Lexical and Morphosyntactic Variation in Persian Heritage Language Outcomes

Khadij Gharibi, Fatih Bayram and Gustavo Guajardo
UiT The Arctic University of Norway

Abstract

Individual variation in heritage language (HL) outcomes does not seem to be random. Instead, this variation can be related to the specific exposure and use patterns heritage speakers (HSs) have with their languages in the contexts they reside. In this study, we present data from 38 child HSs of Persian in English dominant contexts (in New Zealand and the UK), their mothers as well as a control group of age-matched monolinguals in Iran. All participants completed a film-retelling task from which their lexical sophistication (LS) and clausal density (CD) were measured. In addition, the HSs' mothers completed a sociolinguistic questionnaire for their children which was used to calculate proxies for language experiences. Out of the two linguistic measures, the HSs differed from monolinguals only in LS scores. Regarding the relationship between HSs' linguistic scores and language experiences, Random Forest analyses showed HL literacy to be the most important variable for the CD scores; while it was the HSs' age-at-testing for LS. The mothers' scores were only important for the HSs' LS scores. This study contributes to the ongoing discussions on the nature of HL development, outcomes and individual variation.

Keywords: Heritage Language Bilingualism, Vocabulary, Morphosyntax, Random Forest analysis, Literacy, Persian

Introduction

Heritage language speakers (HSs) are individuals who grow up learning (at least) one home language naturalistically that is not the majority language in the wider society. Understanding the nature of the heritage language (HL) grammars has become increasingly popular in the last 2-3 decades (see, e.g., Montrul, 2016; Polinsky, 2018, for reviews).

A large body of studies suggests that HL competence/performance outcomes typically wind up following alternative outcome paths, on a continuum, from (age and SES matched) each other as well as monolinguals, in almost all domains of grammar (e.g., lexicon, phonology, morphology, syntax, semantics and discourse-pragmatics) in both offline behavioral/production as well as online processing experiments (e.g., Montrul, 2008, 2016; Polinsky, 2018; Polinsky & Scontras, 2020, for a comprehensive review). For many studies the reason for this variation has been the shift of dominance (in exposure and use) from the home language to the majority societal language, generally corresponding with the onset of formal education in the majority language in early childhood. However, exactly what is meant by the (potential) change in HL exposure and use has received relatively limited attention— at least until recent years (see for a review Bayram et al., 2021)— compared to documenting differences between HSs and monolinguals. Not all HSs (must) differ from monolingual baselines. In fact, some cannot reliably be differentiated from monolinguals anyway (see, e.g., Kupisch & Rothman, 2018; Bayram et al., 2019).

Considering that HSs are native speakers of their home/heritage language (Rothman & Treffers-Daller, 2016), the above findings have offered challenges as well as opportunities in understanding, and more importantly, accounting for the underlying nature of HL grammars. This is so precisely because “heritage speakers constitute an outcome often assumed to be impossible outside of pathology or trauma: children exposed to a language from birth who nevertheless appear to deviate from the expected native-like mastery in pronounced and principled ways" (Polinsky & Scontras, 2020, p.5).

In an attempt to shed light on the above *impossible* outcome, over the past few years in particular, more attention has been paid to environmental/sociolinguistic factors that contribute to, if not determine, the path of HL development and its outcomes. The main argument is that HSs obtain fully developed, albeit alternative, grammatical systems based on what is available

to them in the linguistic input. This line of research has focused on qualitative input differences and effects of L1 attrition in the first generation of immigrants who provide input of the heritage language to HSs (i.e., cross-generational attrition) (e.g., Sorace, 2004; Rothman, 2007; Rothman, 2007; Pires & Rothman, 2009; Pascual y Cabo & Rothman, 2012; Bayram et al., 2019; Karayayla, 2020). Similar research also investigated the role that access to literacy and formal HL training play in HL competence (e.g., Kupisch, 2007, 2013; Kupisch & Rothman, 2018; Bayram et al., 2019).

Following the footsteps of the recent trends in the HL literature, the current study offers new insights on two fronts: (i) there is an inherent value to measuring HL grammatical competence more holistically; that is, looking at HSs' overall performance in vocabulary and morphosyntax at the same time, and (ii) alternative statistical methods such as random forests could capture sources of variation and trends in language background data in a way that might not be visible in commonly used inferential methods in the existing literature.[i]

Studies on HL vocabulary, morphosyntax and individual variation

The one universal characteristic defining HL grammars is (individual) variation both in terms of development and ultimate attainment (see Montrul, 2016; Polinsky, 2018; Polinsky & Scontras, 2020, for a review). This variation is not random; instead, it is to a large extent context-dependent (Kupisch & Rothman, 2018). The HL development outcomes are systematic just as the outcome of any other naturalistic (typical) child language acquisition scenario, and seem to correlate to, if not predicted by, the specific experiences HSs have with their languages such as, but not limited to, how often, in what contexts and by/with whom they are exposed to and use their HL. For instance, in a study with English/French bilingual children, Paradis et al. (2011) have shown that variation in HL input and use play a deterministic role in vocabulary and morphosyntactic development. More specifically, factors such as parents' proficiency in

the HL and the dominant societal language (e.g., De Houwer, 2007; Chondrongianni & Marinis, 2011), parents' attitudes towards HL use and maintenance (e.g., Nesteruk, 2010; Gharibi & Boers, 2017) and the socioeconomic status of the HL family (e.g., De Cat, 2021) seem to have an impact on HL outcomes to various degrees.

The growing body of research looking in more detail at the role of (home) language use and exposure patterns on HL development shows that the HL context lends itself to a very dynamic cumulation of experiences in which the weighting of specific factors on HL performance varies in different periods and contexts. All these studies have one thing in common: there is no one common characteristic of the HL input/use, quantitative and/or qualitative, that can account for all HL linguistic variation across all groups (e.g., Sorace, 2011; Montrul et al., 2015; Kupisch & Rothman, 2018; Lohndal et al., 2019; Lloyd-Smith et al., 2020; Daskalaki et al., 2020). Such experiences inevitably shape the quality and quantity of input they receive in the HL, which impacts the HL ultimate attainment. However, the potential effect of variation in HL input and use does not manifest itself the same way across all HL languages, or in different domains of grammar (e.g., morphology vs syntax vs lexicon) or even in different features of grammar within the same domain (e.g., word order scrambling vs pro-drop within syntax).

In a study looking at vocabulary and input effects, for instance, Mori and Calder (2017) tested the vocabulary knowledge of 82 school-aged (15-18) Japanese HSs in the United States both in their HL and in English as their dominant language. They used multiple choice tests with 120 Japanese vocabulary items and 220 English vocabulary items ranging from elementary school level to high school level where they had to choose the correct definition or synonym of the given test item. They also collected language background information from their parents to find out what factors played a role in the HSs' vocabulary knowledge. In line with the trends in the HL literature, their results showed that Japanese HSs' vocabulary

knowledge in Japanese lagged behind that of the expected norms (i.e., the school grade norms in the Japanese system) while their L2 English vocabulary was within the expected grade level range. Mori and Calder's (2017) further analysis with the vocabulary data juxtaposed against the parental language background data revealed that reading for pleasure in Japanese, the degree of HL dominance and parents' expectations for their children's success positively predicted the HL vocabulary knowledge.

In another study, Willard et al. (2015) looked at vocabulary knowledge of Turkish HS preschoolers (N=119) and fourth graders (N=121) of Turkish in Germany using an adapted version of the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007). They also collected background information from those children's mothers through an interview and a questionnaire. Their results showed significant links between HS children's Turkish vocabulary and their parents' education and generational status. More importantly, home literacy environment (i.e., exposure to reading materials, learning activities, such as book reading, and parents' literacy habits) and being exposed to Turkish, especially by the mother's use of Turkish, were strong predictors for HSs' vocabulary.

Similar to the findings and trends regarding HL vocabulary development, various studies have reported variation in HL morphosyntax as well. While some of these studies used more universal measures of morphosyntactic competence (e.g., grammatical complexity, clausal density, MLU) (see e.g., Gharibi & Boers, 2019; Lahmann, Steinkrauss, & Schmid, 2019; Lloyd-Smith et al., 2020), others looked at more specific linguistic features (e.g., relative clauses, passives, clitics) (e.g., Unsworth, 2013; Montrul, 2016; Montrul & Sánchez-Walker, 2013; Pascual y Cabo, 2020; Pires & Rothman, 2009; Polinsky, 2018).

For instance, Montrul and Sánchez-Walker (2013) compared the production of Differential Object marking (DOM) (a non-existent feature in English) in first generation immigrants, child and adult HSs of Spanish in the US and monolingual speakers in Mexico. To

identify how qualitatively different input affects this specific feature in HSs' developing grammars, they looked at whether variation in their performance can be traced back to input from the first-generation immigrants (in addition to cross-linguistic influence of English). Montrul and Sánchez-Walker (2013) found that both adult and child heritage speaker groups as well as first generation immigrants display significant omission of this property in their production, highlighting the potential effects of cross-generational language attrition. As they indicated, the fact that the adult HSs also displayed some omission of DOM could be related to the attrited input they have received mainly from their parents (first generation immigrants) and other heritage speaker peers.

In another study, Gutierrez-Clellen and Kreiter (2003) used a story narration task and a parent/teacher report to investigate the relationship between input factors (amount of language input at home and at school, and amount of exposure to reading and other literacy activities) and grammatical performance (as measured in T-units) in child HSs of Spanish (age=7-8) in the US in both of their languages. Multiple regression analyses for each language showed that Spanish language use within home was a significant predictor for HSs' grammatical performance in Spanish but none of the input factors were related to their performance in English.

Karayayla and Schmid (2019), on the other hand, investigated vocabulary and morphosyntax in adult HSs of Turkish in the UK using various tasks (e.g., picture narration task, verbal fluency task, C-test) as well a background questionnaire. Overall, their results were in line with the general trends in the HL literature; that is, HSs of Turkish in the UK differed, on average, from age-matched monolinguals in all domains of linguistic performance they measured (lexical access, morphosyntactic complexity, and formal accuracy). As a next step, Karayayla and Schmid (2019) carried out a Structural Equation Model analysis with the data from the personal background questionnaire and the linguistic tasks and showed that factors

such as the amount of exposure and use of the HL as well as the age of onset of the exposure to English played a significant role in predicting HSs' performance in their HL.

In addition to how much HSs are exposed to and use their HL, under which conditions and settings they use their HL also seems to matter. Being exposed to the HL from different sources such as TV and social media (e.g., Jia & Fuse, 2007), having opportunities to use the HL with a variety of speakers in a variety of activities and contexts (e.g., Place & Hoff, 2011; Scheele et al., 2010) and receiving some form of (formal) literacy training in the HL (Unsworth, 2013; Bayram et al., 2017) all seem to matter for the quality of HL input and thus HL development, albeit at different levels for different HL speakers (see e.g., Lloyd-Smith et al., 2020).

As seen in the above-mentioned studies, the HL context is no exception to input effects in language acquisition (e.g., Montrul, 2004; Sorace, 2004; Rothman, 2007, 2009). Nonetheless, the relationship between the qualitative and quantitative nature of input and the HL competence/performance outcome variation is still underexplored, especially in the early years of life when HSs go through major linguistic/life experiences (e.g., starting school in the dominant societal language and become more and more dominant in that societal language). More specifically, there is a growing need for studies that aim to uncover the specific relationship between language experience and HL competence during childhood. This study aims to fill that gap investigating one of the relatively understudied heritage languages in the literature, namely Persian as a heritage language. It looks at both vocabulary and morphosyntactic competence globally of child HSs of Persian juxtaposed against their language background experiences.

This study

The main objective of this study is to better understand the dynamic nature of the relationship between HL performance outcomes and the contexts in which these outcomes are obtained.

The data for the current study come from a film re-telling task where speakers "integrate all areas of linguistic knowledge in real time" (Schmid, 2011, p. 194). We measured their vocabulary and grammatical competencies as two main indicators of HL linguistic competence. In bilingual populations, grammatical knowledge and lexical knowledge have been shown to be strongly correlated (Polinsky, 1997, 2007), and they can serve as fairly reliable indicators of overall language proficiency (Montrul, 2009).

In addition to looking at HSs' own performance, we also looked at their mothers' performance since they are generally seen as the source of primary linguistic input in the HL. We used the same measures to see whether differences in the child HSs' outcome can also be associated with the quality and quantity of input they receive from that primary source.

In doing so, this study attempts to answer the following questions:

1) Do performance outcomes (vocabulary and morphosyntax) of child HSs of Persian differ from that of age and SES matched monolinguals?
2) If so, what experiential factors (demographic and sociolinguistic alike) correlate with, if not account for, the individual variation within the HS group? That is, how do factors such as age at testing, HL literacy, and HL experience (exposure and opportunity for use) play a role in their HL outcome?

Methodology

Participants

There are three groups of participants in this study: (1) child HSs of Persian (New Zealand and the UK combined[ii]); (2) mothers of the same HS participants in group 1; and (3) child monolingual speakers of Persian in Iran (for more details, see Table 1). The HS group comprised 38 Persian-English simultaneous bilingual children (age=5 to 13, mean age: 8.9) who were either born in or immigrated to their host countries before the age of three (see De Houwer, 1990; Paradis, 2007 on classification of simultaneous and sequential bilinguals)[iii]. All

HSs had two parents born in Iran and moved to their respective host countries as adults.

Monolingual controls were matched with the HS group in terms of age, gender and family socioeconomic status (SES). The participants' socioeconomic status was based on the families' education level[iv], the status of their occupation and their address relative to the area/city they live in the country.

The mother group consists of 14 mothers in New Zealand and 16 mothers in the UK (age= 31-49 years, mean age: 39.5 years old). All participants in this group left Iran in adulthood ((Age of Arrival) AoA > 25 years old), except one who left Iran at the age of 13, and have lived in their respective host countries for 3.5 to 35 years (mean Length of Residence (LoR): 10.8 years) (see Table 1). While half of them were housewives, the other half had full time or part time jobs in English speaking environments, except one who was working as a teacher in a Persian weekend school in London.

Table 1: Participants

	N	Mean/Age	SD	Min.	Max.	Mean/LoR	Min	Max
All HSs	38	8.9	1.9	5.6	12.6	_	_	_
HSs in NZ	17	8.1	1.6	6	12	_	_	_
HSs in UK	21	9.6	1.9	5.6	12.6	_	_	_
Controls	38	8.8	1.9	5.6	12.5	_	_	_
All Mothers	30	39.6	4.7	31	49	12.2	3.5	35
Mothers in NZ	14	37.8	4.3	31	47	8.2	3.5	15
Mothers in UK	16	41.3	4.3	33	49	15.7	9	35

Procedure

A film-retelling task was used to collect speech samples from all groups of participants. This task allows extracting data with a controlled content with fairly homogenous choice of vocabulary and style across the samples (Schmid & Beers Fägersten, 2010). Following the first author's studies on Persian HSs in New Zealand (Gharibi, 2016; Gharibi & Boers, 2019), an episode of "Tom and Jerry"[v] was presented.

All participants were asked to watch the episode of "Tom and Jerry" first and immediately after watching it, they were asked to tell the story in Persian as much as they could

recall. The task was not timed; participants took as much time as they desired. All film-retellings were recorded and transcribed according to CHAT conventions (McWhinney, 2000). Subsequently, the transcriptions were checked by two other native speakers of Persian for consistency. The data were used to approximate clausal density and lexical sophistication scores for each participant.

Semi-structured interviews were conducted with the mothers only to obtain information on the HSs and their families' language background and use patterns. The questionnaire, designed and used in previous studies (Gharibi, 2016; Gharibi & Seals, 2019, 2020), was expanded and applied in the current project with additional questions on parental language use as well as their attitudes towards HL use in the host country in general.

Measures of HL proficiency

Vocabulary

There are a number of methods to calculate the vocabulary knowledge of bilinguals (see for a review Jarvis & Daller, 2013). For the current study, lexical sophistication (LS) was used. As a measure, lexical sophistication aims to quantify the variety and the depth of vocabulary used in a speech sample (Laufer & Nation, 1995). This is done by calculating the frequency of each lexical item from a reference corpus. The word frequency is taken as an index for lexical sophistication in which low LS scores indicate the use of lower frequency words which are considered to be more difficult/complex and therefore higher lexical sophistication. Compared to more traditional methods such as type/token ratio, LS measures have been shown to be more sensitive to capturing variation and neutralizing the effect of data related issues such as text length (e.g., Baese-Berk et al., 2021; Gharibi & Boers, 2019).

To calculate the LS scores in our study, a corpus of raw speech samples of all participants in the three groups was first lemmatized. Lemmatization involves excluding function words and stripping content words of their inflectional morphemes (i.e., tense,

number, person, case, etc.). Items that share the same root are counted as one lemma. For instance, in utterances like "*The cat goes home quickly.", "Jerry is going to outside now."* and "*They went to wake him up.",* "go" would be taken as the root for *goes*, *going* and *went*, and thus would be counted as one lemma. After lemmatizing all participants' film-retelling data, a lemmatized corpus was created for the HSs and their monolingual controls. Following Yilmaz & Schmid (2012) and Schmid & Jarvis (2014), the types of lemmas from the speech samples were divided into five frequency bands. Each of these frequency bands comprised 20% of all tokens (also see Schmid, Verspoor & MacWhinney, 2011). LS score for each participant was measured by calculating the portion of lemmatized items from each frequency band in their specific group corpus. For instance, in a group of participants X, Y and Z, the participant X's LS score is measured from within the lemmatized corpus of the participants X, Y and Z put together. It is important to note again that lower LS scores indicate the use of less frequent words and thus a more sophisticated lexical repertoire compared to those with higher LS scores and thus less sophisticated vocabulary. The same steps were taken to measure the LS scores of the mothers, where the corpus was comprised of lemmatized speech samples only (see the data analysis for rationale on this method) as well as another independent corpus with the data only HSs themselves (see the Findings section below as to why we created two corpora of speech samples by child participants).

Morphosyntax

Traditionally, the morphosyntactic complexity of the narratives was measured by the mean length of utterance (MLU, Brown 1973). However, MLU, despite its ease to compute, has been criticized for lacking validity and reliability, especially in older children and adults, since it has been shown to become less sensitive to individual differences in older populations (e.g., Blake, Quartaro & Onorati, 1993). Instead, clausal density (CD) was found to be a more reliable measure in the older child and adult populations (Nippold, 1993; Mimeau, Plourde, Ouellet, &

Dionne, 2015; Scott, 2004; Scott & Stokes, 1995). This measure is the proportion of embedded/subordinate clauses to independent clauses in a given speech sample. In other words, clausal density is "the extent to which utterances/sentences contain subordinate (dependent) clauses […] a ratio of the total number of clauses (main and subordinate) summed across sentences and divided by the number of sentences in a sample" (Scott & Stokes, 1995, p. 310). For instance, a clausal density of 2.0 would indicate that sentences in the speech sample contain two clauses on the average (one main clause and one subordinate clause), whereas a ratio of 1.50 would mean that a fair number of sentences in the sample were containing one or more subordinate clauses.

Different from the individual LS scores that are calculated within a larger corpus of all the participants' data, the CD score of each participant is measured from their own data. That is, the CD score for a participant X is measured from the dataset consisting of only the participant X's data. Higher scores indicate a higher number of embedded clauses and consequently higher morphosyntactic complexity.

Social and Language background

Semi-structured interviews were conducted with the mothers of HSs following the structure of a sociolinguistic questionnaire based on the family language policy model (Spolsky, 2009). This model has language ideologies, language practices, and language management as its three main components (see King, Fogle, & Logan-Terry, 2008).

The questionnaire consisted of six sections. The first section included 17 questions on demographics of the participating families including their age, their children's age, their length of residence in the host countries, their highest level of education and their occupation at the time of the study. There were also questions on the reason of their immigration and if they have lived in any other countries before. The second section, with 26 questions, was on language background and proficiency, in which the parents were asked if they know any other languages

than Persian and English. They were also asked to report self-evaluation of their own English proficiency before emigration and at the time of the study. This was followed by their reports on their children's Persian proficiency at the time of the study. They were also asked to report if their children know literacy in the HL and how long they have been developing it.

The third section dealt with beliefs on HL development and maintenance (in 5 questions). They were asked how they would feel if they did not successfully pass Persian onto the next generations. The fourth section of the questionnaire included 10 questions on language practices of the families. The questions were on Persian exposure and use in by the HSs, their parents and their siblings if they had any, where the parents were asked how often their children are exposed to (i.e., input) and speak Persian (i.e., output). They were asked to report the frequency of speaking Persian by each parent to the HSs and by the HSs to each parent. There were also questions on how often the HSs practice Persian literacy if they know any. In addition, there were questions about the use of HL outside of the home environment too such as communication with Persian friends in the host and home country, visits to the home country, communication with grandparents, HL literacy/education practices as well as the use of social/visual media. The fifth section asked questions about language management which refers to the families' efforts in helping their children to develop and maintain Persian. This section with 7 questions asked parents about ways they apply to encourage the HSs to develop Persian conversational fluency as well as their HL literacy. The final section addressed parental attitudes on HL development and maintenance in 3 open questions where they could add any comments they had in regards to their experiences with raising bilingual HSs.

The answers in all the questionnaire items that required scaled quantification of values (proficiency, time-spent doing an activity, beliefs, etc.) were based on a backend calculation of a Likert scale of 5, from 0 to 1; 0 being the lowest and 1 being the highest value. The LS and CD scores as well as HL experience factors were brought together in the final stage of the

analysis to look at the sources of individual differences in the HS group. It should be highlighted that the HSs from the two communities (NZ and the UK) were found not to differ significantly regarding their HL experience as the sociolinguistic questionnaire data showed (the raw data can be found in https://osf.io/dsfcm).

Findings

Comparing the heritage speakers to their control groups and their mothers

Lexical Sophistication

It should be recalled that lower LS scores indicate better performance with participants' using less frequent words. As noted above, the three corpora we used for the analyses are: (1) a corpus of data from HSs and monolingual controls only; (2) a corpus of data from the HSs' mothers only for an independent factor that can be used as a proxy for HL input/experience; and (3) a corpus of data from the HSs only. The reason why we created two different HS corpora is two-fold: (i) in HS vs Control comparisons: the LS scores for HSs and their controls were obtained from the corpus that included data from both groups; and (ii) in HS individual variation analysis: the LS scores were calculated from the corpus that only included data from the HSs themselves. This was done to achieve two interrelated goals: (i) to move away from the traditional HS-to-monolingual baseline comparisons predominantly, if not always, disfavoring the HS competence, as also reported in the results herein; and (ii) to have a more ecologically valid approach to the HS individual variation analysis with measures that are better representatives of HS linguistic contexts and competencies enabling us to understand and appreciate HL grammars in their own right (see Bayram et al., 2021; Fuchs, 2021; Polinsky & Scontras, 2019; Rothman et al., forthcoming for a discussion on the ecological value of HS-to-ML comparisons). The descriptive results for HSs and controls from the first corpus are shown in Table 2 below:

Table 2: Lexical Sophistication of the heritage speakers and their controls

	Mean	SD	Min.	Max.
HSs	3.1	.28	2.6	3.7
HSs in the UK	3.1	.29	2.6	3.7
HSs in NZ	3.1	.27	2.6	3.7
Monolingual Controls	3	.23	2.5	3.5

In line with the trends in the literature, the HSs were outperformed by the monolingual controls. While the mean LS score for HSs was 3.1, it was 3 for the monolingual group. A t-test confirmed that the difference between the HSs and the monolingual controls was statistically significant **(**t (74) = 2.6, p = .01).

From the second corpus, we measured the mothers' LS scores as shown in Table 3 below:

Table 3: Lexical Sophistication of the heritage speakers' mothers

	Mean	SD	Min.	Max.
HSs' mothers	3	.15	2.6	3.3

A final set of LS scores was calculated for the HS only from the collection of their speech samples only.

Table 4: Lexical Sophistication scores of the heritage speakers only

	Mean	SD	Min.	Max.
HSs' second score	3.1	.28	2.5	3.7

A t-test showed no significant differences between their scores from the HS-to-monolingual corpus and the HS only corpus (t (38) = .37, p > .05). Despite no differences between the two scores, we will be using the LS scores from the HS-only corpus for the individual differences

analysis (see the section *Input and HL outcomes* below) to adhere to the trends for more ecologically valid measures for HS participants— rather than scores based on comparisons between HSs and monolinguals.

Clausal Density

The descriptive analysis results of clausal density for all three groups are shown in Table 5 below.

Table 5: Clausal density in all participating groups

	Mean	SD	Min.	Max.
HSs	1.30	.17	1	1.7
HSs in the UK	1.31	.15	1	1.5
HSs in NZ	1.31	.2	1	1.7
Monolingual Controls	1.32	.13	1.08	1.5
HSs' Mothers	1.51	.14	1.2	1.8

Interestingly, the HSs were not outperformed by the monolingual controls in their CD scores. Most of the HSs and their monolingual controls had about the same number of embedded clauses in their film-retelling speech samples. A follow-up t-test confirmed that there were no significant differences between the HSs and their monolingual controls (t (74) = -.63, p = .5). The HSs' mothers had a clausal density of more than 1.5, which indicates that a fair number of sentences in their samples contained one or more subordinate clauses. Additionally, a t-test showed a significant difference between HS' CD scores and their mothers' CD results (t (74) = - 5.6, $p < .05$).

Input and HL outcomes

As stated above, one of the main goals of this study is to unpack the relationship between HL linguistic outcomes and the linguistic experiential factors. To do so, we use random forests (Breiman, 2001) in the *ranger* package (Wright & Ziegler, 2017) assessing the effect of the

input related factors and the mothers' LS and CD scores on the HSs' LS and CD scores derived only from the corpora that include only HS data.

Random forests are based on decision trees that use a set of binary rules to predict a response variable. Decision trees, which can be used with both numerical or categorical response variables, belong to a class of statistical models that use recursive partitioning as the main algorithm. Informally speaking, the algorithm first tests if any of the independent variables are associated with the response variable. If it finds more than one independent variable that is associated, then the model determines the strength of association of each of the independent variables with the response variable. The variable with the strongest association is selected for the first binary split. For example, if the independent variable is binary with values M and F, then one subset will contain all the observations with value M and the other subset will contain all those with value F. Each subset constitutes a branch in the tree. This procedure is recursively repeated until all independent variables have been evaluated. A random forest is based on the aggregation of a large number of decision trees. To build the trees and ensure that the trees are sufficiently different from each other, random forests use two different procedures: bootstrap aggregating and random predictor subset selection. Bootstrapping involves creating subsamples of the dataset with replacement, that is each observation can be chosen more than once in the subsample. This results in the subsample having two thirds of the observations and the remaining one third is left out (the out-of-bag sample). Each tree of a random forest is fit to a different bootstrapped sample. The random predictor selection refers to the procedure, where the algorithm selects a random subset of the predictor variables to fit each tree in the forest (this parameter is called *mtry*). For categorical predictors, this is normally the square root of the total number of predictor variables, whereas for continuous predictors it is the number of predictors divided by 3 (Hastie et al., 2001; Strobl et al., 2009).

The choice of random forest over more traditional analyses like linear regression was

based on two main reasons. As explained above, the sociolinguistic questionnaire was made up of 100 questions which would entail 49 different predictor variables in one single model. Since we only had 38 participants, this means that there are more predictors than observations, a problem usually known as $p > n$. Linear regression models are not recommended in this scenario (Bühlmann & van de Geer, 2011; Chakraborty et al., 2012). The second reason is the fact that several of the questions were highly correlated (e.g., HSs' age and their own mother's age $r = 0.52$, mothers' English proficiency before the UK and their English proficiency at the time of the study $r = 0.76$, HSs' literacy and attending Persian classes in the UK $r = 0.84$).[vi] The presence of correlated variables would have made the results uninterpretable and inaccurate. Since we were interested in determining the effect of each of the variables targeted in the questionnaire, we did not want to do a principal component analysis, because this type of analysis, while taking care of the correlation among the variables, obscures the effect of the individual predictors.

Random forests are well-known for being able to handle cases where there are more predictors than observations (Boulesteix et al., 2012) as well as correlations among the predictors (Tomaschek, 2018). They can also handle both continuous and categorical predictors and they are not sensitive to variable scaling. The additional appealing aspect of random forests is the possibility of computing the variable importance of each predictor. This means that the algorithm outputs a ranking of the variables from the most to the least important in explaining the outcome. The random forest in the *ranger* package has the additional advantage that it can calculate *p*-values, so we not only get a ranking of the importance of each variable, but we also get a *p*-value associated with each variable to determine whether the contribution of the variable to explaining the outcome is statistically significant.

There are two available methods to calculate the *p*-value in the *ranger* package. We used the altmann method (Altmann et. al., 2010), because the alternative method requires that

a large proportion of the variable importance scores be negative and only a few of the scores in our dataset had a negative value. The altmann method uses repeated permutations of the outcome variable and estimates the distribution of measured importance for each variable in a non-informative setting. The author of the method recommends between 50-100 permutations, but the creators of the *ranger* package recommend a higher number of permutations to calculate more precise *p*-values, thus we ran 1000 permutations.

We fitted two random forests, one for each of the two scores. In one random forest, the dependent variable was clausal density and in the other, the dependent variable was lexical sophistication. Each random forest was fitted with 5000 trees and the default *mtry* value, which is the square root of the number of predictors. Each random forest contained 52 variables, 49 of which came out of the questionnaire[vii] and the three remaining were the lexical sophistication of both mothers and children and the mothers' clausal density. A complete list of the variables is available online[viii].

Results of Random Forest models

As mentioned earlier, one of the main aims of this study was to know if there is any relationship between the mothers' results and their children's. Prior to discussing the analysis on the HSs' lexical sophistication, recall that we entered the second set of lexical sophistication scores of the HSs into the model. Also, their mothers' scores, which were assessed through the corpus made of their individual speech samples, were entered into the model.

The next model in Table 6 shows the results for lexical sophistication. The model determined that age is the most important predictor of lexical sophistication ($p < 0.001$), followed by the child's mother's own lexical sophistication score ($p < 0.05$). The next significant variable ranks 7th in the variable importance ($p < 0.05$) and this variable refers to parental beliefs on the importance attached to acquisition and maintenance of Persian.

In Figure 1, we show partial dependence plots of lexical sophistication on the three

variables selected as significant by the model. Plot (1-A) shows that as HSs' age increases, their vocabulary also increases (recall that the lower the lexical sophistication score, the higher their sophistication). In plot (1-B), we observe that a stronger belief in maintaining the HL results in more complex vocabulary. Plot (1-C) shows that if the HSs' mothers possess a less varied vocabulary, then their children's vocabulary will be affected in the same direction. In other words, a mother's low lexical sophistication results in the child developing less sophisticated vocabulary.

Table 6: Significant predictor variables for lexical sophistication

Variable	Importance	p-value
HSs' Age	1st	< 0.001
Mother's Lexical Sophistication	2nd	< 0.05
Importance of Persian Maintenance	7th	< 0.05

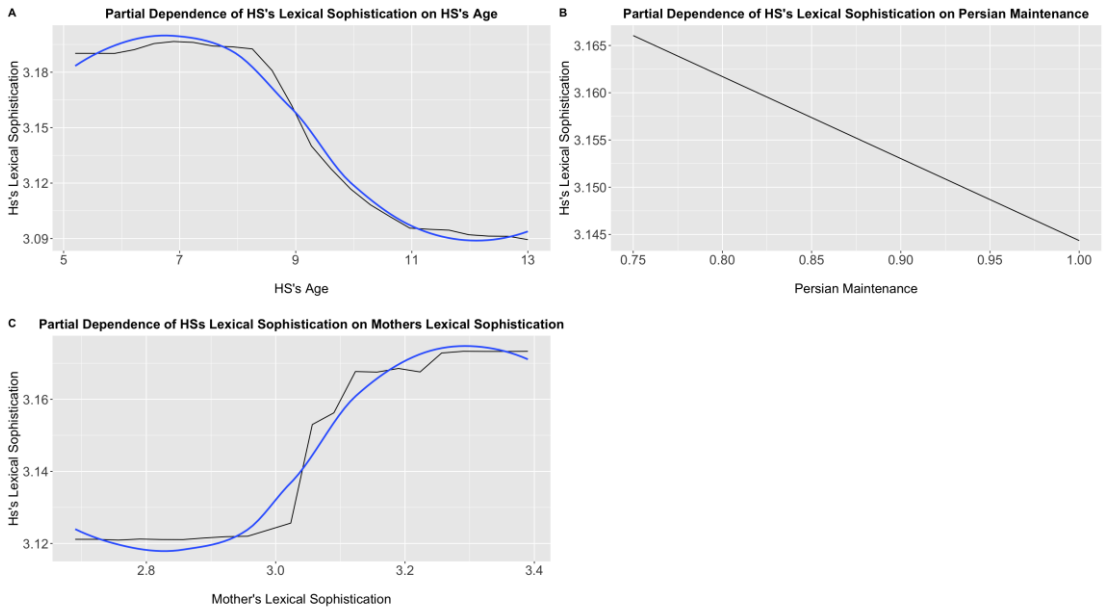

Figure 1. Partial dependence plots of lexical sophistication on the three statistically significant variables. The y-axis represents predicted values of lexical sophistication.

In Table 7, we show the random forest for clausal density. The most important predictor in this case, is whether the participant had any formal literacy training in the HL ($p < 0.05$), followed by whether s/he was encouraged to read and write in Persian ($p < 0.05$). The third most important variable is the level of English of the participant's mother prior to arrival in New Zealand or the UK ($p < 0.05$). The next significant variable ranks 8th and this refers to the language the HSs' parents feel most comfortable in ($p < 0.05$). The last significant variable is the child's sex ($p < 0.05$).

As above, Figure 2 shows the partial dependence plots for clausal density. We observe in plot (2-A) that there seems to be a threshold at which literacy in the HL has an effect on clausal density. It appears that less than 500 hours is not sufficient for the HS child to develop more complex grammar. On the other hand, after 500 hours, there seems to be a relatively steep increase in clausal density. In plot (2-B) we see that encouraging reading and writing in the HL has a positive effect on clausal density such that the more the HS is encouraged to engage in these two activities, the more complex language they will develop. Plot (2-C) shows the partial dependence between clausal density and the mother's English level before arrival. Again, as in the first plot, we observe a threshold effect. Only if the mother possessed a higher English level when she arrived in the new country, does the child get a boost in their clausal density score.

Table 7: Significant predictor variables for clausal density

Variable	Importance	p-value
Literacy in the HL[ix]	1st	< 0.05
Encouragement to read & write in Persian[x]	2nd	< 0.05

Mother's level of English before arrival	3rd	< 0.05

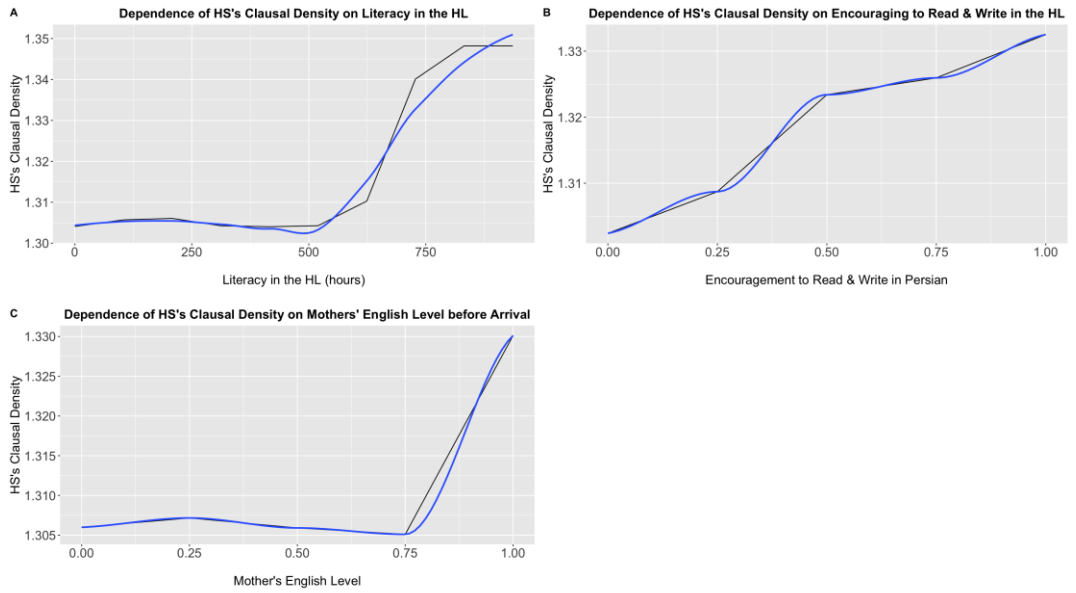

Figure 2. Partial dependence plots of clausal density on the three statistically significant variables. The y-axis represents predicted values of clausal density.

Discussion and Conclusion

In this study, as a first pass, we compared HSs' lexical sophistication (LS) and clausal density (CD) scores to that of their monolingual controls. We found that on average there were no group differences in CD scores while HSs were outperformed by monolinguals in LS scores. At first glance, the asymmetry between morphosyntax and lexicon might look surprising since various studies have shown strong correlations between morphosyntactic and lexical knowledge both in L1 and L2 contexts, especially lexical proficiency being a reliable predictor for morphosyntactic proficiency too (e.g., Polinsky, 1997, 2007). However, we know from child bilingualism studies that despite some delays in developmental trajectories, bilingual children's grammar in general in both of their languages can be robust and similar to that of monolinguals (e.g., De Houwer, 2002; Serratrice, 2013). Given that our HSs' age ranged from 5 to 13 and that age itself was not a predictor for their CD scores (see above for the analysis),

we interpret the results as a sign for the HSs in this particular study to be more resistant to the potential effects of *reduced* input in their morphosyntactic development. On the contrary, our HSs, on average, demonstrated monolingual-like CD performance.

The vocabulary knowledge of HSs, as measured by their LS scores, however, was less sophisticated/varied than that of the monolingual controls. Bi-/multilingual children grow up learning two languages and they do so through exposure to those languages at different times of the day, in different contexts, at different rates and with different interlocutors, potentially leading to differences in the shape and size of the vocabulary in all their respective languages (see e.g., Gharibi & Boers, 2017; Oller, Pearson, & Cobo-Lewis, 2007; Bialystok, Luk, Peets, & Yang, 2010; Daller & Ongun, 2018; Daller, 2020). Just like other bi-/multilingual children, HSs grow up learning (at least) two languages either simultaneously or sequentially. In many cases, their dominance in their home language shifts toward the societal language as a result of schooling and thus changes in their social environment in early childhood. With this shift generally comes reduction in the patterns of HL use/exposure in contexts— in most cases HL use/exposure gets more and more limited to a small and home-environment driven network of individuals. Limited opportunities/environments may also create challenges for learning new and more varied vocabulary items as well as activation and reactivation of the previously learnt ones (e.g., Unsworth, 2016; Thordardottir, 2011; Hoff, 2003).

Although, on average, the HSs and monolinguals seem to have performed similarly, at least in one of the performance scores (clausal density, CD), the spectrum of individual differences within the HS group is noticeably wider than that of monolinguals and thus is worth discussing in its own right. In the random forest analysis, it was found that the HSs' CD and LS scores are predicted by different factors (see Tables 6 & 7, and Figures 1 & 2 above and also see Lloyd-Smith et al., 2020 for similar results). In CD scores, for instance, more varied use of embedded clauses was strongly associated with exposure to formal literacy practices in

the HL, as indicated by the first two factors in the model output. As well as the increase in the HSs' exposure to formal language training in their HL, in Persian in our case, the increase in their parents' explicit efforts to encourage them to practice literacy in the HL positively influenced their CD performance. Interestingly, this trend may be more pronounced with those mothers who reported to have a higher level of English proficiency before arriving to the host country (see the relevant graph in Figure 2). In a way, the sharp difference between the highest and the rest of the English proficiency graph seem to indicate a *necessity* versus *preference* attitude. That is, for those mothers who had lower English proficiency, they probably felt or still feel more comfortable to use Persian as the language of communication with their children. This is also reflected by some of the mothers, especially those with low English proficiency, that they would not want to use English with their children for fear of being evaluated/corrected all the time for the *mistakes* they might make. After all, for the HL children English is one of their native languages while all the mothers learnt English as an L2 in adulthood. However, the more balanced Persian-English bilingual mothers might prefer to use Persian with their children rather than using English simply because they could, given that they might be more aware of the risk of not supporting home language development for their children and of the advantages of growing up as balanced bilinguals like they themselves are.

All this is also in line with findings regarding the relationship between formal (language) education and ultimate attainment not only in monolingual L1 settings (see e.g., Dąbrowska, 1997, 2012) as well as in the HS communities. For instance, Kupisch and Rothman (2018) and Bayram et al. (2019) show that some differences between HSs and their monolingual controls can be due to differences in exposure to and engagement with (formal) education in the HL. The more formally educated an individual HS is, the more likely that their linguistic competence/performance falls within the expected standard baseline ranges. Engagement in (formal) literacy activities in the HL also impacts overall language acquisition

and maintenance and academic achievement (see e.g., Baker, 2006; Biber & Hared, 1991; Bigelow & Tarone, 2004; Eisenchlas, Schalley, & Guillemin, 2013; Hoff, 2006; Tsimpli, 2014). On the other hand, HSs with no (formal) literacy training in their HL in the early years of development face difficulties in maintaining their HL (see e.g., Cummins, 2005; Eisenchlas, Schalley, & Guillemin, 2013 for bilingual communities in Canada, the United States and Australia).

The other factor that predicted the HSs' CD scores was related to their mothers' proficiency in English before their emigration to their respective host countries. All the mothers reported to have low to intermediate level of proficiency in English before emigration— on a Likert scale of five from none to excellent where most of the responses accumulated toward the lower end of the scale. This is also confirmed by the other items in the questionnaire related to their preferred language when they interacted with their children (the majority of mothers reported that they used Persian with their children regularly). Other studies qualitatively analyzing the same interview data used in this study also showed how parents' relatively low proficiency in the societal language reinforces and motivates their use of the HL as the language of parenting (Gharibi & Seals, 2019, 2020; Gharibi & Mirvahedi, 2021). This in turn creates a context at home where there are more opportunities for children to be exposed to qualitatively and quantitatively richer HL input.

For LS scores, age (at the time of testing) was the strongest factor in determining the choice of words HSs used during the task. Let us remind ourselves that in the LS calculation, use of less frequent words meant a more sophisticated lexical knowledge. In our study, the older a heritage speaker is, the richer their lexicon is too. Similar results have been reported in other studies too (Gharibi & Boers, 2019). The age effect also indicates that the HSs in our study seem to have spent increasingly more time in various HL contexts as they grew older enabling them to learn more words that can be less frequent (and thus more sophisticated) for

those HSs that are much younger (e.g., Gharibi & Boers, 2017, 2019; Golberg, Paradis, & Crago, 2008). The next factor in the model associated with HSs' LS scores was their mothers' LS scores. We found that the mothers' more sophisticated vocabulary use explained the increase in the level of lexical sophistication in their children. The final factor predicting a better LS score for the HSs was their parents' beliefs on the value of maintaining and supporting their children's home language development, suggesting that parental language ideologies— whether they believe HL (home) language maintenance is important or not— form a central component of their family language policies and thus directly impacts their children's HL development and maintenance. These three factors combined highlight the context-dependency of lexical growth in bilingual children and are consistent with previous research documenting the positive impact of mother (parent)-child interaction in bilingual children's home language vocabulary development (e.g., Snow & Zhao, 2010; Limia, Ozcaliskan, Hoff, 2019; Quiroz et al., 2010; Gharibi & Seals, 2019; Gharibi & Mirvahedi, 2021).

Overall, our results offer further support to the previous findings showing that lexical development, especially in the home language of child bilinguals, is more vulnerable to input related factors compared to morphosyntactic development which seems to be more robust (e.g., Armon-Lotem et al., 2021; Pearson et al., 1997; Paradis & Genesee, 1996). More importantly, our findings show that although both morphosyntax (as measured by clausal density) and vocabulary (as measured by lexical sophistication) in child HSs are influenced by input and environmental factors, not all factors interact with both domains at the same time or at the same level (e.g., Lloyd-Smith et al., 2020). Literacy practices in the HL and mother's proficiency in the dominant societal language were the strongest predictors for the HSs' clausal density scores while it was the HSs' age, the quality of their mothers' vocabulary and parental attitudes toward HL maintenance that determined the HSs' level of lexical sophistication. These differences in in the weighting of various input factors affecting different domains of HL outcomes lends

support to (i) the complex and dynamic nature of HL acquisition (Lohndal et al., 2019; Putnam & Sanchez, 2013; Polinsky & Scontras, 2020), and (ii) that the observed differences between HSs, compared to each other as well as monolinguals, are not random, instead differences in the context (HL exposure and use patterns), in which HSs acquire their HL, contribute, if not determine, the development and ultimate attainment in their HL (Montrul, 2016; Polinsky, 2018; Kupisch & Rothman, 2018; Bayram et al., 2021; Karayayla & Schmid, 2019).

Our results also highlight the importance of making well-informed decisions for choosing either lexicon-based or morphosyntax-based global measures while examining the overall HL proficiency in HL studies. As seen from the results above, testing only one (LS or CD alike) would lead to conflicting interpretations and potentially fail to capture the whole gamut of our participants' competence in their home language. The same is also important for investigating the role input-related factors play in HL individual outcomes variation. We would have also missed the forest for the trees had we only looked at the relationship between input factors holistically (i.e., aiming for a HS *quotient*, see Marian & Hayakawa, 2020) and only one of the measures (LS vs CD). Of course, we are in no position to claim that the factors investigated in this study or collected using the tools we did, linguistic or otherwise, used to gather such information can offer a complete understanding of the complexities and the dynamicity within the HL context.

There is no *catchall* linguistic proficiency component or HL input/use experience factor that has the power to account for all the individual variation across all age groups and/or contexts in any given group of heritage speakers. Addressing these issues with more detail but better precision and rigor is important not only for scientific inquiry, formal linguistic, psycholinguistic or otherwise, but also for making more informed decisions in developing/implementing pedagogical, clinical, political, social policies and practices for the

languages of the very population we work with in our studies. Future research should consider these factors in understanding HSs and their linguistic competencies more comprehensively.

---

[i] The data and R script are available on https://osf.io/xj6nf/?fbclid=IwAR2cSJOKPb1BjRNuqhhnBaoBm1ZPGT7oHlrPHxjPOPBdrx4jn-pmv9ierY0.

ii Although UK and NZ are two different countries, they share the English language as the dominant societal language, which is the main reason behind the merging of the two HS groups into one in this study. The main focus herein is not making comparisons between these two locations in a binary fashion. Instead, we are interested in the granularity of individual/experiential differences our HS participants have and how these map onto their HL competencies. Keeping the dominant societal language constant but looking at two different contexts allows us to capture the richness and variation in HL experiences on a wider spectrum we would not be able to do by looking at either UK or NZ only.

iii It should be noted that sixteen of the HSs (out of thirty-eight) were siblings from eight different families (6 siblings from New Zealand and 10 siblings from the UK), and thirteen were the only child in their families. This sociodemographic pattern was also matched in the control group of children in Iran. However, we did not find any "sibling" effect in our results (see the Results section for more details).

iv The parents' educational level was measured through a five value Likert scale (Highschool diploma or lower, bachelors' degree, masters' degree, PhD or higher)

v This film lasts about six minutes, and unlike the other stories of this series, it is a friendly story about a puppy that was found by Jerry. Jerry tries to take the puppy into the house where Tom lives but Tom keeps throwing them out. He feels bad and goes out to find them, but he falls into the river. Jerry and the puppy save him, and so Tom lets the puppy stay and live with them.

vi A complete correlation table is available online on https://osf.io/xj6nf/?fbclid=IwAR2cSJOKPb1BjRNuqhhnBaoBm1ZPGT7oHlrPHxjPOPBdrx4jn-pmv9ierY0.

vii Keep in mind that majority of the answers to the items in the questionnaire are structured on a Likert scale of 5, from 0 to 1 in which 1 is the highest value. An exception to this is the calculation of hours spent in HL literacy training which is highlighted in the text.

viii They are available on https://osf.io/xj6nf/?fbclid=IwAR2cSJOKPb1BjRNuqhhnBaoBm1ZPGT7oHlrPHxjPOPBdrx4jn-pmv9ierY0.

ix Literacy in the HL was measured as the total number of hours spent in a formal language training class in school (i.e., Saturday schools).

x Encouragement to read and write in Persian was measured by asking parents how often they encourage their children to read and write in Persian.

References

Altmann, A., Tolosi, L., Sander, O. & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure, *Bioinformatics* 26:1340-1347.

Armon-Lotem, S., Rose, K., & Altman, C. (2021). The development of English as a heritage language: The role of chronological age and age of onset of bilingualism. *First Language*, 41(1), 67-89.

Baese-Berk, M. M., Drake, S., Foster, K., Lee, D. Y., Staggs, C., & Wright, J. M. (2021). Lexical diversity, lexical sophistication, and predictability for speech in multiple listening conditions. *Frontiers in Psychology*, *12*, 2328.

Baker, C. (2006). *Foundations of bilingualism*. Clevedon, UK: Multilingual Matters.

Bialystok, E., Luk, G., Peets, K. F., & Sujin, Y. A. N. G. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and cognition*, 13(4), 525-531.

Biber, D., & Hared, M. (1991). Literacy in Somali: linguistic consequences. *Annual Review of Applied Linguistics*, 12, 260-282.

Bigelow, M., & Tarone, E. (2004). The role of literacy level in second language acquisition: Doesn't who we study determine what we know?. *TESOL quarterly*, 38(4), 689-700.

Blake, J., Quartaro, G., & Onorati, S. (1993). Evaluating quantitative measures of grammatical complexity in spontaneous speech samples. *Journal of Child Language*, 20(1), 139-152.

Boulesteix, A. L., Janitza, S., Kruppa, J., & König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(6), 493-507.

Breiman, L. (2001) Random Forests. *Machine Learning*, 45, 5-32. https://doi.org/10.1023/A:1010933404324

Brown, R. (1973). Development of the first language in the human species. *American psychologist*, 28(2), 97.

Bühlmann, P., & Van De Geer, S. (2011). *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media.

Bylund, E. (2009). Maturational constraints and first language attrition. *Language learning*, 59(3), 687-715.

Cabo, D. P. Y., & Rothman, J. (2012). The (il) logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied linguistics*, 33(4), 450-455.

Chakraborty, S., Ghosh, M., & Mallick, B. K. (2012). Bayesian nonlinear regression for large p small n problems. *Journal of Multivariate Analysis*, *108*, 28-40.

Chondrogianni, V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism*, 1(3), 318-345.

Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal*, 585-592.

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219-253.

Daller, M. H. (2020). Turkish heritage speakers in Germany. *Studies in Turkish as a Heritage Language*, 60, 17.

Daller, M., & Ongun, Z. (2018). The threshold hypothesis revisited: Bilingual lexical knowledge and non-verbal IQ development. *International Journal of Bilingualism*, 22(6), 675-694.

Daskalaki, E., Elma, B. L. O. M., Chondrogianni, V., & Paradis, J. (2020). Effects of parental input quality in child heritage language acquisition. *Journal of child language*, *47*(4), 709-736.

De Cat, C. (2021). Socioeconomic status as a proxy for input quality in bilingual children?. *Applied Psycholinguistics*, *42*(2), 301-324.

De Houwer, A. (1990). *The acquisition of two languages from birth: A case study*. Cambridge University Press.

De Houwer, A. (2002). Comparing monolingual and bilingual acquisition. *Alkalmazott Nyelvtudomány*, 1, 5-18.

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, 28(3), 411.

Dunn, L. M., & Dunn, D. M. (2007). PPVT4: Peabody Picture Vocabulary Test (4th ed.). Minneapolis, MN: NCS Pearson.

Eisenchlas, S. A., Schalley, A. C., & Guillemin, D. (2013). The importance of literacy in the home language: The view from Australia. *Sage Open*, 3(4).

Fuchs, Z. (2021). Facilitative use of grammatical gender in Heritage Spanish. Linguistic Approaches to Bilingualism. DOI: https://doi.org/10.1075/lab.20024.fuc

Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied psycholinguistics*, 29(1), 41-65.

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 23(2), 283-294.

Gutierrez–Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied psycholinguistics*, *24*(2), 267-288.

Hastie, T., Tibshirani, R. & Friedman, J.H. (2001) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, New York.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child development*, *74*(5), 1368-1378.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental review*, 26(1), 55-88.

Karayayla, T., & Schmid, M. S. (2019). First language attrition as a function of age at onset of bilingualism: First language attainment of Turkish–English bilinguals in the United Kingdom. *Language Learning*, 69(1), 106-142.

King, K. A., L. Fogle, and A. Logan-Terry. 2008. Family language policy. *Language and Linguistics Compass*, 2(5), 907–922.

Köpke, B., & Schmid, M. S. (2004). Language attrition. First language attrition: Interdisciplinary perspectives on methodological issues, 28(1).

Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564-582.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language and Hearing Research*, 50(5), 1280-1299.

Lahmann, C., Steinkrauss, R., & Schmid, M. S. (2019). Measuring linguistic complexity in long- term L2 speakers of English and L1 attriters of German. *International Journal of Applied Linguistics*, *29*(2), 173-191.

Lanza, E., & Svendsen, B. A. (2007). Tell me who your friends are and I might be able to tell you what language (s) you speak: Social network analysis, multilingualism, and identity. *International journal of bilingualism*, 11(3), 275-300.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics*, 16(3), 307-322.

Limia, V., Özçalişkan, Ş., & Erika, H. O. F. F. (2019). Do parents provide a helping hand to vocabulary development in bilingual children?. *Journal of child language*, 46(3), 501-521.

Lloyd-Smith, A., Bayram, F., & Iverson, M. (2020). The effects of heritage language experience on lexical and morphosyntactic outcomes. *Studies in Turkish as a Heritage Language*, 60, 63.

Lohndal, T., Rothman, J., Kupisch, T., & Westergaard, M. (2019). Heritage language acquisition: What it reveals and why it is important for formal linguistic theories. *Language and Linguistics Compass*, *13*(12).

MacWhinney, B. (2000). The CHILDES Project: Tools for analyzing talk. transcription format and programs (Vol. 1). Psychology Press.

Marian, V., & Hayakawa, S. (2021). Measuring bilingualism: The quest for a "bilingualism quotient". *Applied Psycholinguistics*, *42*(2), 527-548.

Mimeau, C., Plourde, V., Ouellet, A. A., & Dionne, G. (2015). Comparison of measures of morphosyntactic complexity in French-speaking school-aged children. *First Language*, 35(2), 163-181.

Montrul, S. (2004). Convergent outcomes in L2 acquisition and L1 loss. *First language attrition: Interdisciplinary perspectives on methodological issues*, 259-279.

Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *International Journal of bilingualism*, 13(2), 239-269.

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

Montrul, S. A. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor* (Vol. 39). John Benjamins Publishing.

Montrul, S., & Sánchez-Walker, N. (2013). Differential object marking in child and adult Spanish heritage speakers. *Language Acquisition*, *20*(2), 109-132.

Mori, Y., & Calder, T. M. (2017). The role of parental support and family variables in L1 and L2 vocabulary development of Japanese heritage language students in the United States. *Foreign Language Annals*, *50*(4), 754-775.

Nesteruk, O. (2010). Heritage language maintenance and loss among the children of Eastern European immigrants in the USA. *Journal of Multilingual and multicultural Development*, 31(3), 271-286.

Nippold, M. A. (1993). Developmental markers in adolescent language: Syntax, semantics, and pragmatics. *Language, Speech, and Hearing Services in Schools*, 24(1), 21-28.

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28(2), 191-230.

Paradis, J. (2007). Second language acquisition in childhood. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 387–405). Blackwell Publishing.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent?. *Studies in second language acquisition*, 18(1), 1-25.

Paradis, J., Nicoladis, E., Crago, M., & Genesee, F. (2011). Bilingual children's acquisition of the past tense: A usage-based approach. *Journal of child language*, 38(3), 554.

Pascual y Cabo, D. (2020). Examining the role of cross-generational attrition in the development of Spanish as a heritage language: Evidence from gustar-like verbs. *Linguistic Approaches to Bilingualism*, *10*(1), 86-108.

Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied psycholinguistics*, 18(1), 41-58.

Pires, A., & Rothman, J. (2009). Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 211-238.

Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2- year- olds' bilingual proficiency. *Child development*, *82*(6), 1834-1849.

Polinsky, M. (1995). American Russian: Language loss meets language acquisition. In *Formal Approaches to Slavic Linguistics*. Cornell Meeting. Michigan Slavic Publications.

Polinsky, M. (2007). Reaching the end point and stopping midway: Different scenarios in the acquisition of Russian. *Russian Linguistics*, 31(2), 157-199.

Polinsky, M. (2011). Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in second language acquisition*, 305-328.

Polinsky, M. (2018). *Heritage languages and their speakers* (Vol. 159). Cambridge University Press.

Polinsky, M., & Scontras, G. (2020). A roadmap for heritage language research. *Bilingualism: Language and cognition*, 23(1), 50-55.

Putnam, M. T., & Sánchez, L. (2013). What's so incomplete about incomplete acquisition?: A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 478-508.

Quiroz, B. G., Snow, C. E., & Zhao, J. (2010). Vocabulary skills of Spanish—English bilinguals: impact of mother—child language interactions and home language and literacy support. *International Journal of Bilingualism*, 14(4), 379-399.

Rothman, J. (2007). Heritage speaker competence differences, language change, and input type: Inflected infinitives in Heritage Brazilian Portuguese. *International Journal of Bilingualism*, 11(4), 359-389.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International journal of bilingualism*, *13*(2), 155-163.

Scheele, A. F., Leseman, P. P., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*(1), 117-140.

Schmid, M. E., Köpke, B. E. (2019). *The Oxford handbook of language attrition*. Oxford University Press.

Schmid, M. S. (2011). *Language attrition.* Cambridge University Press.

Schmid, M. S., & Fägersten, K. B. (2010). Disfluency markers in L1 attrition. *Language learning*, 60(4), 753-791.

Schmid, M. S., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, 17(04), 729-748.

Schmid, M. S., Verspoor, M., Verspoor, M. H., & MacWhinney, B. (2011). Coding and extracting data. In *A dynamic approach to second language development* (pp. 39-54). John Benjamins.

Scott, C. (2005). Learning to write. In H. Catts & A. Kamhi (Eds.), *Language and reading disabilities* (2nd ed., pp. 233–273). Boston: Pearson.

Scott, C. M., & Stokes, S. L. (1995). Measures of syntax in school-age children and adolescents. *Language, Speech, and Hearing Services in Schools*, 26(4), 309-319.

Serratrice, L. (2013). Acquisition of features in the nominal domain in bilingual acquisition. *International Journal of Bilingualism*, 17(5), 657-664.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism Language and Cognition*, 7(2), 143-145.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic approaches to bilingualism*, *1*(1), 1-33.

Spolsky, B. (2009). *Language management*. Cambridge University Press.

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological methods*, *14*(4), 323.

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, *15*(4), 426-445.

Tomaschek, F., Hendrix, P., & Baayen, R. H. (2018). Strategies for addressing collinearity in multivariate linguistic data. *Journal of Phonetics*, *71*, 249-267.

Treffers-Daller, J., Daller, M., Furman, R., & Rothman, J. (2016). Ultimate attainment in the use of lexical collocations among heritage speakers of Turkish in Germany and Turkish-German returnees. *Bilingualism: Language and Cognition*, 19(3), 504-519.

Tsimpli, I. M. (2014). Early, late or very late?: Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism*, 4(3), 283-313.

Unsworth, S. (2013). Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics*, *33*, 21-50.

Unsworth, S. (2016). Early child L2 acquisition: Age or input effects? Neither, or both?. *Journal of child language*, *43*(3), 608-634.

Willard, J. A., Agache, A., Jäkel, J., Glück, C. W., & Leyendecker, B. (2015). Family factors predicting vocabulary in Turkish as a heritage language. *Applied Psycholinguistics*, *36*(4), 875-898.

Wright, M. N., Ziegler, A. (2017). *ranger*: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, 77(1): 1-17. DOI:10.18637/jss.v077.i01

Yilmaz, G., & Schmid, M. S. (2012). L1 accessibility among Turkish-Dutch bilinguals. *The Mental Lexicon*, 7(3), 249-274.

Data Availability Statement:

The data, R script and a complete correlation table are available at https://osf.io/xj6nf/?fbclid=IwAR2cSJOKPb1BjRNuqhhnBaoBm1ZPGT7oHlrPHxjPOPBdrx4jn-pmv9ierY0.

Address for correspondence

Khadij Gharibi, Khadij.Gharibi@uit.no, UiT The Arctic University of Norway, Department of Language and Culture (ISK), Tromsø 9037, Norway

Co-authors information

Fatih Bayram, Fatih.Bayram@uit.no, UiT The Arctic University of Norway, Department of Language and Culture (ISK), Tromsø 9037, Norway

Gustavo Guajardo, Gustavo.a.Guajardo@uit.no, Stensberggata 10 Oslo 0170, Norway