### Contents

**Special Issue–Heritage language studies and early child bilingualism research: Understanding the connection**

**Guest Editors: Suzanne Aalberse and Aafke Hulk**

**Introduction**

Introduction to the special issue – Heritage language studies and early child bilingualism research: Understanding the connection 491

*Suzanne Aalberse and Aafke Hulk*

**Articles**

Simultaneous bilingualism: Early developments, incomplete later outcomes? 497

*Carmen Silva-Corvalán*

Spanish heritage speakers in the Netherlands: Linguistic patterns in the judgment and production of mood 513

*Brechje van Osch and Petra Sleeman*

Heritage language development: Connecting the dots 530

*Silvina Montrul*

Bilingual children and adult heritage speakers: The range of comparison 547

*Maria Polinsky*

Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers 564

*Tanja Kupisch and Jason Rothman*
Visit journals.sagepub.com/home/ijb
Free access to tables of contents and abstracts. Site-wide access to the full text for members of subscribing institutions.
Introduction to the special issue – Heritage language studies and early child bilingualism research: Understanding the connection

Suzanne Aalberse and Aafke Hulk
University of Amsterdam, The Netherlands

Introduction

This issue is the result of a colloquium with the same title organized at the ISB 10 conference at Rutgers, USA, in May 2015. All the presenters, the discussant and one additional colleague have contributed to this collection of articles, which brings together linguists from the domain of heritage studies with those working on early child bilingualism (ECB).

A language qualifies as a heritage language (HL) if ‘it is a language spoken at home or otherwise readily available to young children, and crucially this language is not the dominant language of the larger (national) society’ (Rothman, 2009, p. 156). HLs are learned early in life, either simultaneously with the dominant language or prior to the acquisition of the dominant language of the country; heritage speakers (HSs) are thus early bilinguals. There is, however, a gap between linguists studying ECB and linguists studying HSs. Linguists working on ECB mainly look at the early development of both languages in children growing up bilingually from birth or shortly afterwards and tend to report on similar developmental patterns in monolingual and bilingual children, and on (temporary) delay, acceleration or cross-linguistic influence in this development, mostly in Western Europe and Canada (De Houwer, 1990; Meisel, 1990; Paradis & Genesee, 1996). Linguists studying HSs, on the other hand, mainly look at young adults’ competence in their HL and tend to focus on signs of incomplete acquisition, mostly in the USA (Benmamoun, Montrul, & Polinsky 2013; Montrul, 2008; Polinsky, 2006).

This special issue brings together linguists from both fields to find out how to make the connection: to what extent can we observe similarities in the reported results on child bilinguals and HSs and to what extent do we observe differences? What motivates these similarities and differences? The contributions in this special issue all shed light on the comparison and thereby create new questions. We will briefly discuss some general factors that may influence different outcomes of early bilingualism over the lifespan (see also Hulk & Marinis, 2011).

Sociolinguistic circumstances

The sociolinguistic environment in which bilingualism develops affects its outcome. The social background of HSs and early child bilinguals, as defined above, is not (always) the same. The
effect of sociodemographic context on the type of variation that is likely and not so likely to occur has been investigated in contact-linguistics under the label scenario-approach (Muysken, 2013, and references therein). Although much work on contact-linguistics focuses on adult speakers, the general message also applies to child bilingualism. Muysken (2013, p. 710) defines this as follows: ‘Languages do not interact in a single way, but rather in many different ways, depending on the social setting of bilingualism.’

One such external factor that influences the outcome of bilingual acquisition is the attitude towards the minority language: the reaction of educators and peers on the multilingualism of children in their school years can hinder or facilitate the acquisition process and can cause the results of multilingualism to be different across societies and across social groups. Lambert (1977) reports on the negative impact of feeling ashamed of the HL, which often starts when children enter school. Abrupt disuse of the home language sometimes leads to reduced communication between parents and children and is likely to have a negative impact (De Houwer, 2015). Lambert (1977) refers to the situation where learning a new language leads to loss and negative feelings towards the first language as ‘subtractive bilingualism’ and contrasts it to ‘additive bilingualism’. Taking into account the effect of these two different situations helps to connect the dots between ECB and a heritage language.

### Moving targets

This special issue also investigates the relation between the input bilinguals receive and the output they produce. Some researchers, such as Silva-Corvalán (this issue) and Montrul (this issue), observe differences between the input to the HL and the output bilinguals produce and relate these differences to what they refer to as ‘incomplete acquisition’. Apparently, not all that was offered was acquired. Other researchers, such as Kupisch and Rothman (this issue), building on previous argumentation in Rothman (2007) and Pascual y Cabo and Rothman (2012), stress that the input the HL receives is quantitatively and qualitatively different from that of monolingual learners and this may explain the difference in outcomes.

It is possible, however, that the target is moving (cf. Backus, 2012). We know from the study of world English that new standards become established (Schneider, 2007) and we know from second language studies that children of migrants might use learner forms as identity markers (Nortier & Dorleijn, 2008). Moreover, bilinguals might want to create their own variety. Li (2011) illustrates the creativity of school-aged bilingual children in their bilingual language games. It is possible that during the school-age the linguistic target changes. For example, bilingual constructions that were first perceived as errors might be becoming a new norm.

### Heritage language terms

What is the effect of bilingualism on the path of acquisition and of its outcome? To shed light on this question a baseline is needed. Various baselines are possible: (monolingual) peers in the home country; the first generation of immigrants; other bilinguals; or monolinguals with a similar educational level (cf. Kupisch & Rothman, this issue). Apart from the problem of establishing a fair baseline, there is the problem of labelling the groups we compare HSs to. Some people compare HSs to ‘native speakers’. However, HSs are also native speakers, because they have spoken their HL since birth (Rothman & Treffers-Daller, 2014).

Other researchers contrast HSs to ‘homeland speakers’ or ‘monolingual speakers’. These labels are not completely unproblematic either. For example, for speakers of Spanish in the USA, Mexico might no longer be considered as the ‘homeland’ and speakers in the homeland are not all
monolingual. The most neutral reference to the baseline is probably a reference to the country where the language is spoken such as ‘German in Italy’ (Kupisch & Rothman, this issue).

If variation is attested between HSs and the selected baseline, the next question is how to label these differences. Just stating that there is a difference is not very informative (Silva-Corvalán, this issue), while other more contentful labels, such as ‘incomplete acquisition’ might be used too quickly and can be harmful to the group they describe (Kupisch & Rothman, this issue). In order to understand the connection between early acquisition and adult attainment, more neutral and precise labels are desirable. Polinsky (this issue) gives an overview of possible connections between child language acquisition and adult outcomes and, as such, facilitates the creation of more informative labels. These new labels could focus on HL grammars as the result of processes of bilingual acquisition (cf. Putnam & Sánchez, 2013), where different ‘bilingual optimization strategies’ are used (cf. Muysken, 2013).

Changes across the lifespan

When comparing child bilingual HSs to adult HSs it is clear that many changes in the development of the linguistic system happen across the lifespan. Different processes make it possible that monolingual and bilingual children start out similarly but begin to diverge later on in life. Work by Dąbrowska (2012) supports the idea that schooling in a language affects the outcome of language acquisition. In the present issue this is illustrated by Kupisch and Rothman as well as by Van Osch and Sleeman: both papers show that factors such as literacy and schooling in the home language, community size and training of metalinguistic awareness may affect outcomes of bilingual acquisition (cf. Lynch, 2014). This emphasizes the need to know more about what happens in linguistic performance and knowledge during the lifespan in bilinguals as well as in monolinguals (cf. Pires & Rothman, 2009). It can then be asked what type of learning leads to what kind of linguistic development over the lifetime.

Relative vulnerability or stability of linguistic features

Languages change overtime; this is true for languages spoken in monolingual and in multilingual situations. We know some aspects of language are more prone to change than others. We also know that some domains appear to be vulnerable in all circumstances, whereas others are vulnerable in special contexts only (cf. Müller & Hulk, 2001). This special issue also investigates vulnerability and shows that, in many cases, there is a close connection to vulnerable domains in ECB and adult bilingualism. Van Osch and Sleeman (this issue) show that the use of the subjunctive in Spanish interface domains (as defined by Sorace & Serratrice, 2009) are more vulnerable than that in non-interface domains and that often the indicative is used as the default. Montrul (this issue) and Silva-Corvalán (this issue) show that young and old bilingual speakers struggle with pragmatic constraints on the (non) realization of overt subjects.

Polinsky (2011, this issue) shows that object relative clauses are acquired early, but they are lost overtime in Russian HSs in the USA, even though English has object-relative clauses as well. Interestingly, subject relative clauses, which are cross-linguistically most common, appear to be more stable. Polinsky delves deeper into the issue of stability and shows, for example, that case marking after a number in heritage Russian in the USA is stable despite the fact that it is cross-linguistically rare and sensitive to errors in young learners.

We know from studies on language contact that languages in contact develop differently than languages that develop in isolation (see Trudgill, 2011, for an overview of the literature). Moreover, in bilingual situations the language pair is relevant: (partial or typological) overlap affects relative vulnerability.
In short, in all situations language shows variation, but not necessarily in the same manner. The question is to what extent vulnerability and stability overlap in monolingual and multilingual situations (cf. Aalberse & Moro, 2014) and to what extent vulnerability in young and older learners relate. Moreover, the question is if vulnerability is present across language pairs or only in a restricted set. The articles in this issue all illustrate the relevance of these questions.

Summaries

Silva-Corvalán explicitly bridges the gap between early bilingual acquisition and end stages in second- and third-generation HSs by comparing the grammar of two bilingual children in a longitudinal study with results on second- and third-generation adult Spanish immigrants. The longitudinal study follows the bilingual children until six years of age, and shows that these children start to converge more with English after the age of 4, as they enter school and English becomes the dominant language for them. This is evidenced by an increase in the use of overt subjects around that age, specifically in the youngest child: the developmental patterns match the linguistic behaviour of second- and third-generation adult speakers. This article shows the importance of studying bilingual development over a longer period, before and after the so-called ‘dominance shift’.

Van Osch and Sleeman discuss young adult Spanish HSs in the Netherlands who are tested on their knowledge of mood with an oral production task and an acceptability judgement task, in variable and obligatory contexts. The Dutch HSs show similar vulnerability to other adults and child learner groups reported in the literature: all groups are vulnerable in the variable contexts and the subjunctive appears to be more vulnerable than the indicative. A second result, however, is that the Dutch HSs perform better on the judgement task than on the production tasks. This task effect differs from outcomes in earlier research on Spanish HSs in the USA. This difference is argued to be related to the schooling of the HSs in foreign languages in general in the Netherlands, and shows the importance of taking social dimensions into account when investigating multilinguals.

Montrul reports on her studies investigating the knowledge of pragmatic constraints on subject expression in Spanish in young adult HSs in the USA. She compares these to school-aged monolingual children in Mexico and bilingual children in the USA. Moreover, she looks at the competence of adult immigrants in the USA. Interestingly, she finds that the linguistic behaviour of the bilingual children and the young adults was not much different: both groups show a high use of overt pronouns. Moreover, she also observes that the developmental errors are similar across mono- and bilingual children, although at the same time she finds that all bilinguals use more overt subjects than monolinguals and that they are less sensitive to the coreferentiality constraint. Her contribution shows that linking research on bilingual acquisition in pre-school children with research on young adult HSs not only adds a much needed perspective to understand HL development, but also allows one to raise new questions.

Kupisch and Rothman bridge the gap between studies on ECB and HSs in several ways. Firstly, they describe the turn in the recent past from a deficiency perspective on ECB to a more positive perspective and explore how HL studies can also benefit from this perspective. The insight that controlling for the social situation is key in understanding possible outcomes of bilingualism holds a central position in the second part of their article, where they compare two groups of German-Romance bilinguals in two different social situations: a group of French HSs who went to a French school in Germany versus a group of Italian HSs in Germany who went to a German school. Although both groups of Romance HSs were dominant in German, those who attended formal schooling in their HL performed like monolinguals, whereas those who only had access to the HL at home were different from monolingual peers. The paper by Rothman and Kupisch shows the
importance of factors such as socioeconomic status (SES) and access to schooling to better understand the relation between early and late bilingualism.

Polinsky focuses on the question of vulnerability and stability of linguistic features across the lifespan. She discusses theoretically possible outcomes of the three-way comparison between child bilinguals, heritage bilinguals and adult monolingual speakers, and the underlying reasons for these outcomes. She illustrates this by case studies, for example from American HSs of either Russian or Spanish. Her contribution shows that whereas some aspects of language appear to get lost over the lifespan, other features keep developing over time. In short, multiple outcomes are possible. By making these different possible scenarios explicit and by presenting expected and unexpected outcomes, Polinsky facilitates a meaningful discussion on changes across the lifespan and on notions of relative stability.

In short, the articles in this issue all, in one way or another, contribute to ‘connecting the dots’ between research on early child bilinguals and young adult HSs, and allow us to address new questions in different ways. They all show that it is important to focus on the process of bilingual acquisition across the lifespan and to take both internal and external factors into account.

References


Simultaneous bilingualism: Early developments, incomplete later outcomes?

Carmen Silva-Corvalán
University of Southern California, USA

Abstract

Purpose: Research on the language of heritage speakers has shown that in situations of societal bilingualism the functionally restricted language evidences the simplification of some grammatical domains. A frequent question is whether this stage of grammatical simplification is due to incomplete or interrupted acquisition in the early years of a bilingual’s life, or a result of processes of attrition of acquired knowledge of the underused language. This article considers the issue of incompleteness through an examination of the relationship between bilingual children’s developing grammars and the more or less changed bilingual systems of adult second and third generation immigrants (“heritage speakers”) in the USA.

Methodology: The issue of incompleteness is examined in two corpora: (1) Recordings of 50 Spanish-English adult Mexican-American bilinguals; and (2) Longitudinal data obtained during the first six years of life of two Spanish-English bilingual siblings.

Data analysis: Qualitative and quantitative analyses of the grammar of subjects, verbal clitics, and verb tenses of the Spanish of the bilinguals under study.

Findings: The outcome of reduced exposure and production of a minority language in simultaneous bilingual acquisition reflects the incomplete acquisition by age 6;0 of some aspects of the input language. The bilingual siblings’ unequal control of the minority language is shown to parallel the range of proficiencies identified across the adult heritage speakers.

Significance: Some linguists argue that heritage speakers’ grammars are less restrictive or “different” in some respects but not incomplete. In contrast, this article demonstrates that at least some of the reduced grammars of heritage speakers result from a halted process of acquisition in the early years of life. Furthermore, while difference is not an explanatory construct, incomplete acquisition due to interrupted development caused by restricted exposure and production offers an explanation for the range of proficiencies attested among adult heritage speakers.

Keywords
Incomplete acquisition, simultaneous Spanish-English bilingualism, grammatical subjects, clitics, tenses, heritage speakers

Corresponding author:
Carmen Silva-Corvalán, Department of Spanish and Portuguese, THH-156, University of Southern California, Los Angeles, CA 90089-0358, USA.
Email: csilva@usc.edu
Theoretical preliminaries

Research on the language of heritage speakers has shown that in situations of societal bilingualism the functionally restricted language evidences, among other phenomena, the simplification of some grammatical domains (Montrul, 2009; Silva-Corvalán, 1996; Zentella, 1997, among many). In this context, a frequent question is whether this stage of grammatical simplification is due to incomplete or interrupted acquisition in the early years of a bilingual’s life, or a result of processes of attrition or loss of acquired knowledge of the underused language (Cuza, 2010; Montrul, 2005; Polinsky, 2011).

The notion of “incomplete acquisition” has recently been criticized. The argument is either that the concept itself refer to the language of heritage speakers is flawed, or that the term is not adequate to describe the grammars of these bilingual speakers (e.g. Otheguy, 2013; Otheguy & Zentella, 2012; Putnam & Sánchez, 2013). In contrast, this article will show that the outcome of reduced exposure and production of a minority language in simultaneous bilingual acquisition indeed reflects the incomplete acquisition by age 6:0 of some aspects of the input language. I consider that acquisition has not been completed when a grammatical domain lacks aspects, elements or features present in the learner’s input language (Meisel, 2014). Less exposure to the minority language and fewer opportunities for use of this language, lead to more differences between the speech of a second or third generation speaker and that of their parents or other first generation immigrants who have offered the language input to the succeeding generations.

This paper considers the issue of incompleteness through an examination of the relationship between bilingual children’s developing grammars and the more or less changed bilingual systems of adult second and third generation immigrants (“heritage speakers”, HSs) in the USA. I look at the grammar of subjects, verbal clitics, and verb tenses in two corpora consisting of recordings of the following sociolinguistic conversations:

1. Corpus ELA: 50 Spanish-English adult bilinguals, all Mexican-Americans living in the eastern section of greater Los Angeles (Silva-Corvalán, 1996).
2. A corpus of longitudinal data obtained during the first six years of life of two Spanish-English bilingual siblings (Silva-Corvalán, 2014).

I show that it is possible to identify parallels between the children’s language characteristics and those of the adult bilinguals. This finding lends support to the idea that at least some of the reduced grammars of heritage speakers appear to result from a halted process of acquisition and development in the early years of life rather than attrition or loss of knowledge acquired in childhood. Furthermore, given that the siblings’ input in Spanish is not attrited or reduced, the results suggest that some of the changes that have been identified across generations of bilinguals develop naturally in the acquisition of the heritage language, seemingly regardless of the type of input.

Importantly, my findings should be interpreted with reference to an English-Spanish bilingual community characterized by the shift from Spanish to the socially-dominant language, English, across generations but also by some degree of maintenance thanks to continuous contact with newly arrived Spanish speakers. This sociolinguistic situation gives rise to considerable variation across speakers, captured in the term “bilingual continuum.”

The data

The data from the adults has been selected from recordings of 50 men and women of different ages classified into three immigrant groups according to the length of time that the speakers’ families have lived in the USA (Silva-Corvalán, 1996).
Group 1 includes speakers born in Mexico, who immigrated to the US after the age of 11. Group 2 encompasses speakers born in the US or who have immigrated from Mexico before the age of six. Group 3 also comprises speakers born in the US; in addition, at least one parent responds to the definition of those in Group 2.

The separation into three discrete groups does not correlate directly with three discrete groups in terms of Spanish oral language proficiency. Indeed, because of their social histories, and even though everyone in the sample acquired Spanish from birth, some speakers across groups 2 and 3, the HSs, have similar levels of higher or lower proficiency in Spanish. Overall, however, groups 2 and 3 represent a continuum of less to more distance from the grammars of Group 1.

I have examined bilingual acquisition primarily in a corpus of data obtained during the first six years of life of two English–Spanish developing bilingual siblings, my grandsons Nico and Brennan (see Silva-Corvalán, 2014, for further details). Nico and Brennan acquired greater proficiency in English and use this language significantly more than Spanish. They grew up in a dual-language home: the mother speaks to them exclusively in English; the father used Spanish with Nico almost exclusively until the child was three-and-a-half years old, but use of Spanish decreased from that age on. Brennan, who is two years and nine months younger than Nico, thus heard relatively less Spanish from his father, but in child-directed speech the father used Spanish almost exclusively with Brennan as well, until he was about 3;0 years old. I spoke with the children almost exclusively in Spanish; they used English very infrequently with me. The children spoke almost only English with each other, occasionally responded in English to their father before age 3;0, but addressed him in English more and more frequently after this age.

I observed and recorded the siblings regularly using Spanish, English, or both in a variety of natural and uncontrolled discourse contexts and with different interlocutors, and kept detailed diary notes to age 3;0 for Nico. Diaries and recordings also include adults’ speech addressed to the children.

The children’s degree of proficiency in English by age 6;0 is comparable to that of monolinguals. By contrast, their developing proficiency in Spanish is unequal. I suggest that this inequality results from differences in the amount of Spanish language input and production, more reduced for the younger sibling.

Overall, Nico was exposed to Spanish about one-third of his waking time; Brennan, less than one-third (about 25%). From about age 4;0, exposure to and use of Spanish are further reduced for both children to at most a quarter of the time. It is indeed remarkable that with such limited input the children are able to develop conversational proficiency in Spanish.

Subject realization

We may expect grammatical subjects in Spanish to be vulnerable to English influence, especially so since Spanish is the weaker language and offers options constrained by semantic and pragmatic factors (cf. Müller & Hulk, 2001). The learnability question for the child should be easier for English, which presents a highly regular model of overt, preverbal subjects, but harder for Spanish, since the child needs to learn the semantic and discourse-pragmatic constraints that regulate the expression of subjects.

Studies applying a variationist methodology have shown that the variable expression of a pro-nominal subject is responsive to cognitive, semantic, and discourse factors (Shin, 2014; Travis & Torres Cacoulllos, 2012, among many). An overt subject is required under two conditions: (1) when it is focal, either because it is new information or the focus of contrast, and (2) when it is needed to identify its referent. Overt subjects are favored probabilistically by a number of variables, including subject switch reference (Example 1); verbs in the first person singular; verbs of volition, of saying or speaking, and of mental processes (e.g. pensar ‘think’, creer ‘believe’).
Discontinuity of reference favors the expression of the subject. Continuity of reference (or coreferentiality), as in 2a and 2b, has consistently been shown to disfavor overt subjects. On average, in various Spanish dialects over 40% of non-coreferential pronominal subjects are overt, while only about 25% of coreferential subjects are expressed (Shin & Smith Cairns, 2012).¹

Researchers incorporate quantification to find out similarities and differences in subject expression rates across individuals and across dialects. The assumption is that in a large corpus of data the various factors that constrain subject expression become neutralized and so overall percentages of overt subjects may reveal different stages of acquisition, dialectal differences, and possible processes of change.

Spanish varieties differ with respect to the rate of subject expression. In the siblings’ case, the average percentage for the Santiago (Chile) variety, 38% (Cifuentes, 1980–81), is of interest because this is the variety spoken by their family.

I have examined about 5000 declarative utterances containing a verb that had or could have had a subject. The utterances were coded for several factors, but I discuss only coreferentiality here. The siblings’ use of subjects in English does not differ from what is typical for monolinguals. Subjects are expressed as expected. In what follows I focus only on Spanish.

Quantitative results by age 1;11 indicate that the siblings realize at a very early age that the null subject is grammatical in Spanish but ungrammatical in English: they express pronominal subjects in English over 86% of the time, while in Spanish the rate of expression is below 18% (Silva-Corvalán, 2014). But use of subjects in Spanish appears to be affected as exposure to English increases and exposure to and production of Spanish become more reduced.

The results in Table 1 appear to confirm the hypothesis that a lower amount of exposure to the weaker language makes a child more susceptible to influence from English.² Brennan uses a much higher proportion of overt subjects than his brother from early on. Beyond age 4;0, the older child also increases the percentage of use of pronouns.

The increase in the speech of Brennan and Nico suggests that as English patterns become more entrenched, they may subconsciously replicate the [subject + verb] pattern of English onto Spanish. But Brennan surpasses his brother’s rate of expression by 27 percentage points between the ages of 4;0 and 6;0, thus showing a more intense effect from reduced exposure to Spanish.

It is of course possible that the increased rate of overt subjects might be justified. But the contexts where a subject is clearly new information or required to identify a referent in the children’s data are very rare, so these factors could not explain the increased rate of overt subjects. Thus, I compute overt and null subjects by a frequent objective factor, subject reference (see examples 1 and 2 above).

---

### Table 1. Subject pronoun realization in Spanish at ages 2;0–2;11, 3;0–3;11 and 4;0–6;0.

<table>
<thead>
<tr>
<th>Age</th>
<th>Nico</th>
<th></th>
<th>Brennan (younger)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>2;0–2;11</td>
<td>36</td>
<td>329/912</td>
<td>51</td>
<td>397/775</td>
</tr>
<tr>
<td>3;0–3;11</td>
<td>28</td>
<td>178/636</td>
<td>71</td>
<td>503/711</td>
</tr>
<tr>
<td>4;0–6;0</td>
<td>42</td>
<td>428/1030</td>
<td>69</td>
<td>411/598</td>
</tr>
</tbody>
</table>

1. (a) Tu hermano quería saber (b) si (yo) enseño en USC. [yo “I” is “optional”]

   (a) “Your brother wanted to know if [I] teach at USC.”

2. (a) Pepe es un escritor muy conocido. (b) Es mi vecino.

   (a) “Pepe is a well-known writer. (b) [He]’s my neighbor.”
The frequent expression of coreferential subjects not justified by another favoring factor, as in Example 3, is interpreted as redundant by native speakers. Example 3 is selected from Brennan’s data. It includes five finite verbs, three of which could have had a non-overt subject, but Brennan expresses all five subjects.

(3) Brennan, 4;1: [Talking about making a robot when he grows up]
B: Y mi robot, y mi robot, va a, a ser muy, muy bueno. Sabes que cuando (a) yo dijo “para atrás” y no es parte de mi familia, él, dijo él (b) “para atrás” y él (c) empuja para atrás, tan, tan atrás porque él (d) tiene esos [antenas] que son tan, tan largas que está en outer space y él (e) te tira para outer space, en tu auto.

“B: And my robot, my robot’s gonna, gonna be very, very good. You know when (a) I say ‘back’ and is not part of my family, he, (b) he said ‘back’ and (c) he pushes back, so, so much back because (d) he has these [antennas] that are so, so long they reach outer space and (e) he pulls you to outer space, in your car.”

The subjects in examples 3(a) and 3(b) may be validated for various reasons, but the third person subjects in (c), (d), and (e) are not validated by the discourse context: they are activated, coreferential, non-contrastive, and unnecessary to identify their referent.

Interestingly, an informal survey of 14 adult speakers, nine from Spain and five from South America, did not indicate overall agreement about which subject pronouns should be expressed in a text I asked them to consider. It is no surprise, then, that the so-called syntax-pragmatic interface is vulnerable and open to change. The child does not receive a consistent input that would facilitate learning subjective discourse-pragmatic conditioning factors in situations of reduced language exposure.

A quantitative analysis of coreferentiality in the last age period, when the siblings evidence the highest rates of subject pronouns, shows that the coreferentiality constraint was weak in Nico’s speech, and absent in Brennan’s (see Table 2). Nico expresses coreferential subjects more frequently than the 25% average in adult speech, but he replicates the monolingual trend by expressing switch reference subjects even more frequently. Brennan evidences absence of this constraint: he does not differentiate coreferential from non-coreferential subjects with respect to frequency of expression; the difference between these subjects is not significant.³

In sum, the simultaneous acquisition of a non-null-subject and a null-subject language does not affect the acquisition of subjects in either language at the earliest age stage. But after age 3;0 Brennan’s rates of subject expression diverge further from those reported for monolinguals, and for balanced bilingual children. A possible mechanism accounting for the high rate of pronouns is the child’s copying of the English [subject pronoun + verb] string.

I identified a somewhat similar increase of subject pronouns in the data from some of the adult bilinguals: 32% of subject pronouns in data from A9 in Group 1; 42% in data from L28 from Group

<table>
<thead>
<tr>
<th></th>
<th>Nico</th>
<th>Brennan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overt/Overt + null</td>
<td>Overt/Overt + null</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Coreferential</td>
<td>200/520</td>
<td>38.5</td>
</tr>
<tr>
<td>Non-coreferential</td>
<td>205/461</td>
<td>44.5</td>
</tr>
<tr>
<td></td>
<td>(p&lt; .05)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Subject pronoun realization by coreferentiality (4;0–6;0).
Table 3. Nominal and pronominal subject realization by coreferentiality (adult bilinguals).

<table>
<thead>
<tr>
<th></th>
<th>Coreferential</th>
<th>Non-Coreferential</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>A9 (Group 1)</td>
<td>10/48</td>
<td>21</td>
<td>40/80</td>
</tr>
<tr>
<td>L28 (Group 2)</td>
<td>10/39</td>
<td>26</td>
<td>64/91</td>
</tr>
<tr>
<td>H48 (Group 3)</td>
<td>19/57</td>
<td>33</td>
<td>59/101</td>
</tr>
</tbody>
</table>

2; and 38% in data from H48 from Group 3 (Silva-Corvalán, 1996). I noted then that the categorical restrictions on subject expression remained intact in second and third generation immigrants, but there was a decrease in the strength of the coreferentiality constraint, as is the case for the siblings, and especially for Brennan (see Shin & Otheguy, 2009, for the decrease of sensitivity to continuity of reference in New York Spanish).

Table 3 shows that the adult bilinguals maintain the difference between same reference and switch reference subjects, but this difference is not statistically significant for the Group 3 speaker. Furthermore, the percentage of coreferential subjects increases steadily from the Group 1 speaker to the Group 2 and the Group 3 speaker. The rate of expressed coreferential subjects in the speech of the Group 3 speaker is 12 points higher than that of A9 (Group 1) and seven points higher than that of L28 in Group 2.

Brennan and Nico’s linguistic behavior is comparable to that of the Los Angeles bilingual adults, even though they have been exposed to a different variety of Spanish. Nico differentiates coreferential from non-coreferential subjects like the HSs in groups 1 and 2; Brennan is closer to the speaker in Group 3.

Clitics in bilingualism

Spanish clitics are unstressed pronominal elements also referred to as “verb clitics,” because they must occur either before or after a host verb. Clitics constitute a complex system encompassing phonological, morphological, syntactic and semantic properties. The set of clitics in Latin American Spanish includes first person singular me; second singular te; third person unmarked for number and case se; third person dative singular and plural le, les; third person accusative singular and plural: feminine la, las, masculine lo, los, and neuter lo (to refer to non-entities); first person plural unmarked for case nos. Third person dative and accusative clitics are co-indexed with the object (object-related), me, te, and nos may be co-indexed with the object or with the subject (subject-related). Se is subject-related in reflexive and reciprocal clauses. Reflexive constructions include true reflexives and a number of other types of “reflexives” with se-unaccusative or se-intensifier (Cuervo & Pérez Leroux, 2015), some “optionally” and others obligatorily reflexive (lexical reflexives). There is also a “spurious se” which stands for le or les in clitic clusters. Examples 4 to 8 illustrate some of these clitic functions.

Subject-related

(4) (Él) _se_, quemó la mano.  
   “He, _se_, burnt his hand.”

(5) ¡No _te_, mojés!  
    not _yourself_, wet  
    “Don’t get wet!”
In addition to contrasts by person, gender and case, there are position constraints: clitics must be preverbal with tensed verbs, and postverbal with infinitives, gerunds and participles. The internal ordering of clitics in clusters is also fixed: se must be in initial position, followed by any other clitics ordered by person: second, first, third person. When clitics refer to an argument of an infinitive or a present participle in a verbal periphrasis with a finite (semi)-auxiliary verb, they may vary after the lexical verb or before the finite auxiliary. This is the only variable context for clitic placement in Spanish.

Previous studies on the acquisition of clitics

The complexity of the clitic system poses a significant learning problem to children. Despite this complexity, studies of the acquisition of clitics by monolingual and bilingual children in Spain (Cuervo & Pérez-Leroux, 2015; Domínguez, 2003; Ezeizabarrena, 1997) show that clitics appear before the age of 2;0 and are used productively with few, mostly morphological errors (e.g. singular for plural clitic, masculine for feminine clitic), no position errors with single clitics and rarely with clitic clusters. Domínguez reports a few cases of lo when no clitic is required, and a few cases of “ungrammatical” clitic-doubling, that is, the object and the clitic are used in the same sentence.5 Cuervo and Pérez-Leroux study the acquisition of clitic clusters by two Spanish monolinguals to age 3;0 and note that clusters are productive from age 2;0. These authors’ Table 4 shows the total number of clusters realized from 2;1 to 3;0 and the types of errors per child. The percentage of errors is 14.60% (20/137) for one child, and 9.38% (9/96) for the other. Omission errors are more frequent than feature and internal ordering errors: 14 of 29 errors correspond to omissions (48.28%). Cuervo and Pérez-Leroux conclude that children produce clusters “with extremely low error rates” (p. 163) and “that before the age of 3 children have fully acquired the syntactic and morphological operations responsible for multiple cliticization” (pp. 163–164).

From these few studies of Spanish clitic acquisition with non-elicited data, I conclude that clitics have been acquired by age 3;0, but the system is not completely stabilized in the grammars of either monolinguals or bilinguals. In production, children do better with single clitics, but not so well yet with clitic clusters (14.60% is not an extremely low error rate). A similar situation of instability characterizes the speech of the siblings and adult bilinguals focus of this article.

Realization of clitics by child and adult bilinguals

I examined the siblings’ spontaneous data between the ages of 4;0 and 6;4 and coded the utterances with realized clitics and those where clitics should have been used. This resulted in a total of 544 sentences, 358 for Nico from ages 5;0 to 5;11, and 184 for Brennan from ages 4;1 to 6;4.6 Brennan
produced 125 clitic contexts from 4;1 to 4;5, but I did not include these in the counts that follow because the objective was to examine clitics at a time close to the children’s start of their first grade, when exposure to English increases.

As reported for monolinguals, the siblings’ data also evidence some errors. In Nico’s data, there is one case of lo instead of la to refer to la serpiente “the snake,” seven sentences with a missing clitic, and four with overproduced clitics. The total percentage of clitic errors in his data is 3.35% (12/358), well below the percentage reported for monolinguals at a younger age, and 1.98% omission errors.7 It may be assumed, therefore, that there is a process of stabilization of the grammar of clitics beyond age 3;0 if input and production are not hampered, but Nico has not reached complete acquisition by age 6;0. The overproduced clitics are all subject-related; three correspond to constructions that frequently occur with an intensifier reflexive clitic; for instance, with verbs of consumption (Example 9), but a clitic is not validated when this type of verb is used with a non-specific object, as Nico does in Example 10. The fourth case of overproduction is with the verb temblar “tremble” (Example 11). That this is the only lexical error is remarkable given the idiosyncratic reflexives typical of Spanish, such as quebrarse “break” and romperse “tear,” which Nico uses correctly.

Table 4. Clitics omitted in required contexts, siblings and heritage speakers.

<table>
<thead>
<tr>
<th></th>
<th>Subject-related</th>
<th>Object-related</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nico</td>
<td>3/131</td>
<td>2.29%</td>
<td>4/223</td>
</tr>
<tr>
<td>Brennan</td>
<td>1/34</td>
<td>2.94%</td>
<td>1/23</td>
</tr>
<tr>
<td>Group 2</td>
<td>4/317</td>
<td>1.26%</td>
<td>6/749</td>
</tr>
<tr>
<td>Group 3</td>
<td>35/527</td>
<td>6.64%</td>
<td>26/954</td>
</tr>
</tbody>
</table>

The seven omissions do not follow a clear pattern: two omissions of an accusative clitic (nos “us,” los “them”), two of the dative clitic le “to her/him,” and one each of a reflexive with a reciprocal construction, an inchoative reflexive, and a lexical reflexive with an inchoative meaning. Examples 12 to 14 illustrate. Note that los (Example 12) and se (Example 14) are omitted in clitic clusters, but Nico produces clusters correctly most of the time (see Example 9 above).

(9) Nico, 5;10: Yo, me, lo voy a comer todo.7
    I, me,Refl go to eat everything
    “I’m going to eat it all.”

(10) Nico, 5;10: Sí, siempre yo ~me como mucho.8
    yes, always I me,Refl eat much.
    “Yes, I always eat a lot.”

(11) Nico, 5;5: Esto también ~se va a temblar.
    “This also se is going to shake.”

(12) Nico, 5;10: Mira, se ~0 comió todos. [zero los = the zucchini]
    look, se ~[them] ate-3psg all
    “Look, [he] se, ate them all.”

(13) Nico, 5;7: Él y Amanda ~0 pusieron a llorar. [zero se, inchoative reflexive]
    “He, and Amanda, [se] started to cry.”

(14) Nico, 5;10: Se ~0 comió todos. [zero se = the zucchini]
    se ~[them] ate-3psg all
    “Se, ate them all.”
(14)  Nico, 5;6: ~0 me cayó un poquito de estos, de la pasta. [zero se, inchoative reflexive]
[\(\text{se}\_\text{refl}\)] to-mey fell a little of thesei, of the pasta.
“A bit of these, of the pasta fell down.”

Between the ages of 5;0 and 6;4, Brennan produces only 59 contexts for the study of clitics, with a total percentage of clitic errors of 6.78% (4/59), about three percentage points higher than Nico’s. Like his brother, he has not reached complete stabilization of clitic usage by age 6;0. Brennan overproduces the clitic se twice with the same verb, poner “put,” which he uses mistakenly in the inchoative reflexive form ponerse “become, begin” (among many other possible translations) in Example 15 (see also Example 13 above). He omits two clitics: me with a lexical reflexive (resbalarse “slip”), and le with pegar “hit” (Example 16). This verb belongs to a group of verbs that can be used with two complements (Juan le, pegó un puñete a Pepe, “Juan hit a punch to Pepe”), although frequently used only with a single complement (Juan le, pegó a Pepe, “John hit Pepe”). It is, thus, an ambiguous type of structure, since when expressing only the “goal” (an indirect object), the child may wrongly consider it the “patient” (a direct object), which as such does not require a clitic co-referent.

(15)  Brennan, 6;4: El perro ~se puso su cabeza en el jarro.
“The dog ~se refl put his head in the jar.”

(16)  Brennan, 5;5: Él 0 pegó a mi papá aquí con su codo. [zero le = him]
[me] hit-3psg to my dadk here with his elbow
“He hit my dad here with his elbow.”

In sum, by age 6;4 the children show full proficiency in clitic placement, gender (with only one exception) and case. The errors concern mainly clitic omissions, 56.25% of total errors, and overproduction of subject-related clitics, 37.50%. This latter type of error is understandable given the idiosyncratic nature of lexical reflexives, illustrated in an unaccusative construction (Example 17) with a non-reflexive type verb, hervir “boil,” and (Example 18), with quemar(se) “burn (itself)” in an unaccusative/middle construction with se.

(17)  El arroz hirvió.
“The rice boiled.”

(18)  El arroz se quemó.
“The ricei se refl burned.”

There are similarities between the children’s use of clitics and the adults’ use in Groups 2 and 3. For the study of clitics I analyzed data from 20 adult speakers, who represent various degrees of Spanish and English proficiency (Silva-Corvalán, 1996). Regarding the omission of clitics, in a subset sample of 13 speakers in Groups 2 and 3, of a total of 2822 contexts for the occurrence of an obligatory clitic, 71 are missing (2.52%). Six speakers in Group 1 produced all required clitics. Table 4 summarizes omissions of object-related and subject-related clitics in the siblings’ data and in the data from five speakers in Group 2 and eight speakers in Group 3.10

Regarding clitic omissions, Group 2 speakers’ behavior is close to that of Group 1, but Group 3 speakers evidence much higher percentages of omissions, closer to Brennan’s percentages, while Nico appears to be in between the two groups. The instability of the clitic system by age 6;0 is thus shown to be also a feature of the speech of some adult heritage speakers.
The verb system

The Spanish verb system includes forms labelled preterite, pluperfect, future, imperfect subjunctive, etc. that are normally used to convey tense, aspect and mood (TMA) distinctions. In this section, I consider the perfective-imperfective (preterite-imperfect) aspectual opposition, and the indicative-subjunctive mood opposition in data from the siblings and from some of the adult bilinguals. To determine if a verb form is used according to the norms of first immigrant adult generation I consider obligatory contexts and discourse tasks which favor one or another form.

Obligatory syntactic contexts include those that impose consecutio temporum constraints in a number of subordinate clauses including adverbial clauses that refer to unrealized situations. For instance, a pluperfect subjunctive in the protasis limits the choice of verb form in the apodosis to pluperfect subjunctive or indicative or to conditional perfect; clauses introduced by antes que “before” or después que “after” require the use of a subjunctive form.

Spanish distinguishes perfective and imperfective aspect in the past. Very simply stated, the preterite conveys the perfective aspect meaning of completed event/situation and implies change, while the imperfect conveys the imperfective meaning of absence of a specific beginning or end point and implies stativity or continuity. Thus, the preterite is required in perfective discourse contexts when the beginning and/or the end of a situation is in focus, a reading that is obtained, for instance, with some punctual adverbial expressions (at 5 o’clock), with complements that force an interpretation of the situation as singular and perfective (see Example 19), etc.

(19) Compró un libro / una casa / un auto.
   “He bought a book / a house / a car.”

The imperfect is required in imperfective environments, most frequently when informing the time and location of an event. For instance, orientation clauses in narratives are coded in the imperfect (Example 20), but narrative abstracts, action clauses and statements which orient or evaluate the narrative events as a whole must be coded in the preterite.

(20) Estábamos jugando con una pelota [Imp, orientation] y alguien me pateó [Pret, action] y me hizo una herida grande [Pret, action]. Eso fue terrible [Pret, overall evaluation].

Mood may be defined briefly as the grammaticalization of the speaker’s subjective attitude toward the degree of assertiveness or certainty of the situation encoded in language. Tenses in the indicative mood communicate assertiveness and factuality. Subjunctive tenses refer to non-experienced and hypothetical situations; they are used mainly in subordinate clauses, which adds to the cognitive complexity of these tenses. Predictably, subjunctive tenses will be more difficult to acquire.

Silva-Corvalán (2014, Ch. 7) shows that in English no differences are observed between the siblings and English-speaking monolingual children. In Spanish, by contrast, the siblings evidence some off-target behavior. Most of the simple indicative mood tenses are not problematic. The siblings are able to talk about the present, the past and the immediate future with little difficulties. But the less frequent and more complex tenses that refer to non-experienced and hypothetical situations were either unstable or not acquired by the age of six. Reduction in exposure to Spanish led to not learning or weakening the on-line command of compound and irrealis tenses, as shown in examples 21 and 22.11

(21) Nico, 5;6: Porque él quería que la gente mala encontrara-ImpSub [on target] pedazos de él para que ellos ~hacia-Imp más robots. [required: hicieran- ImpSub]
“Because he wanted the bad people to find pieces of him so that they made more robots.”
[required: could make]

(22) Brennan, 5;4: ~No cuando ya tengo PresInd dieciséis años. [required: tenga-PresSub]
“Not when [I] am sixteen years old.”

Table 5 presents information about the verb tenses used by the children during the last age period studied. Besides Nico and Brennan, this table includes two Mexican-American children: Daisy and Mike. Daisy is from a Spanish-only home; her English is at an incipient stage. Mike speaks both Spanish and English at home, but he is dominant in English. The data reported in Table 5 are based on recordings done when all the children except Brennan were attending kindergarten; Brennan was in preschool. For purposes of comparison across the children, Table 5 includes data corresponding to the same age range: 5;0 to 6;0. The table also includes an adult from Group (G) 1, S2 (female, 25 years old) to facilitate reference to an unreduced verb system, an adult from Group 2, V21 (female, 16 years old), and one from Group 3, A46 (female, 30 years old). V21 and A46 were exposed to Spanish and English from birth, but have from an early age preferred to speak English. The information in this table is based on a minimum of three hours of recording of each speaker.

Table 5 shows visually that the simpler and more frequent tenses are used in the same manner by all the speakers in the table. Differences among the children and adults correlate with home language and consequent amount of exposure to Spanish. Daisy, who speaks only Spanish at home and attends an English kindergarten, is the only one who uses preterite, imperfect, and present subjunctive as S2 does, and she is beginning to use the imperfect subjunctive. She is also the only child who uses stative verbs with preterite morphology consistently in perfective contexts. Mike, Brennan and A46 additionally evidence unstable preterite morphology. Mike, Brennan and the
adults in Groups 2 and 3 do not use imperfect subjunctive. Brennan and the bilingual adults do not use the present subjunctive either. Note that Mike, Nico and V21 use one more tense than Brennan and A46, the present perfect (he visto “I’ve seen”).

Not only the amount but also the quality of the input received by the children is largely responsible for the patterns observed. Toddlers and preschool children talk more frequently about concrete situations in their immediate environment, they are read or told stories about past events, they talk about what they will do rather than what they would do, and they are given affirmative and negative commands. Children soon acquire the tenses and discourse rules associated with these tasks: they are more frequent and they are also cognitively simpler. These tenses are not a source of difficulty for any of the children.

By age 5;0 to 6;0, then, when bilingual US-born children start kindergarten, they have not yet acquired the complete system of tense, mood, and aspect in Spanish. It is also noteworthy that at this age the English-dominant bilingual children with more reduced exposure to Spanish at home display the same interesting use of the imperfect instead of the preterite (Example 23) with a closed list of stative verbs (ser “to be,” estar “to be,” tener “to have,” haber “there to be,” and poder “can”), just as adult bilinguals in Groups 2 and 3 (marked with * on Table 5), even though at least Nico and Brennan are not exposed to this bilingual variety of Spanish.

By age 5;0 to 6;0, then, when bilingual US-born children start kindergarten, they have not yet acquired the complete system of tense, mood, and aspect in Spanish. It is also noteworthy that at this age the English-dominant bilingual children with more reduced exposure to Spanish at home display the same interesting use of the imperfect instead of the preterite (Example 23) with a closed list of stative verbs (ser “to be,” estar “to be,” tener “to have,” haber “there to be,” and poder “can”), just as adult bilinguals in Groups 2 and 3 (marked with * on Table 5), even though at least Nico and Brennan are not exposed to this bilingual variety of Spanish.

By age 5;0 to 6;0, then, when bilingual US-born children start kindergarten, they have not yet acquired the complete system of tense, mood, and aspect in Spanish. It is also noteworthy that at this age the English-dominant bilingual children with more reduced exposure to Spanish at home display the same interesting use of the imperfect instead of the preterite (Example 23) with a closed list of stative verbs (ser “to be,” estar “to be,” tener “to have,” haber “there to be,” and poder “can”), just as adult bilinguals in Groups 2 and 3 (marked with * on Table 5), even though at least Nico and Brennan are not exposed to this bilingual variety of Spanish.

(23) Nico, 2;8.15: Estaba pateando la pelota a un hombre y alguien pateó la pelota a mi pierna y (a) fue-Pret un awi grande y (b) eso ~estaba-Imp terrible. [expected: estuvo-Pret]

“He was kicking the ball to a man and someone kicked the ball to my leg and (a) it was-Pret a big boo boo and (b) that was-Imp terrible.”

It is remarkable that despite reduced exposure to Spanish, however, Nico has developed a tense system which is very close to that of Daisy. Brennan, on the other hand, has developed a system that is almost the same as that of the Group 3 speaker.12

Further similarities

Further similarities between the siblings and second and third immigrant generation adult bilinguals may be identified. This is interesting because the children had very rarely, if at all, been exposed to the speech of these speakers.

For instance, the children, just as adults, are lexical innovators. They adapt English nouns and verbs to Spanish morphology and use them quite naturally (e.g. lipo from lip “labio,” pular from pull “tirar”). They also equate word combinations in English with parallel combinations in Spanish but assign to these the meaning of the English idiomatic combination (e.g. no puedo esperar “I can’t wait” gets the meaning of “to be very eager for X,” a new form-function in Spanish). These idiomatic combinations are also attested in the Spanish of adult bilinguals.

Conversely, some off-target syntactic constructions are transitory and no longer attested beyond the age of four years and one month; in particular, the placement of adjectives before the noun (verde hoja “green leaf”), preposition stranding (¿Qué es esto para? “What is this for?”), and copies of the English genitive (Kiko’s escuela). In this regard, the siblings’ grammars are closer to those of second generation rather than third generation speakers, whose syntax shows more evidence of convergence with English. I find no evidence of a new mixed language. Structural copies that violate the typological patterns of Spanish constitute qualitative changes that are perceptually salient (e.g. the realization of complementizers, negation, question formation, auxiliary verbs; cf. Hickey, 2010). Therefore, the divergent structures are in time discarded, because the bilingual children either self-correct or are corrected by adults.


Discussion and conclusion

I have discussed some concrete connections between bilingual acquisition in the early years and some aspects of the grammars of adult heritage speakers; in particular, the increased production of overt subject pronouns; object clitic omissions; and a reduced TMA system.

Both children and adults evidence an increased rate of subject pronouns accompanied by a decrease in the strength of the coreferentiality constraint. This characterizes speakers in Group 3 in Los Angeles, but the decrease of sensitivity to continuity of reference is also attested in a large sample of Group 2 speakers in New York (Shin & Otheguy, 2009). This suggests that the weakening of the coreferentiality constraint results from reduced exposure to Spanish in a situation of intense contact with a dominant language (see endnote 2).

I have shown that by age 6;0 the children’s use of clitics shows similar patterns of instability as those attested in the speech of adult HSs. In regard to clitic omissions, overall quite rare, Group 2 speakers’ behavior is close to that of Group 1, who produce all required clitics, but Group 3 speakers evidence a higher percentage of omissions, similar to Brennan’s percentage, while Nico appears to be in between Groups 2 and 3. Reduced exposure to Spanish is thus closely related to the occurrence of clitic omissions.

Lessened exposure is also responsible for the absence of some tenses in the siblings’ Spanish and in one more US-born bilingual child. It is clear that by age 5;0 to 6;0, the children, including a near Spanish monolingual child, have not yet acquired the adult system of tense, mood, and aspect in Spanish. Some features of the children’s TMA system are also characteristic of that of the adult heritage speakers. Notably, the absence of subjunctive and compound tenses, and the use of the imperfect instead of the preterite with stative verbs are features that attest to a halted process of development rather than attrition of knowledge. This interpretation agrees with Polinsky’s (2011, p. 320) statement that the incomplete acquisition of a feature can be assumed “if a child and an adult deviate from the baseline in the same way.”

The parallels identified between the siblings’ linguistic behavior and that of adult heritage speakers have two implications: first, they imply that some aspects of the heritage language of adults are the outcome of an interrupted process of acquisition of this language between the ages of 3;0 and 5;0, when more intensive exposure to the L2 reduces exposure to the heritage language and diminishes the opportunities to use it; and second, given that the siblings’ input in Spanish is not attrited or reduced, the results suggest that some of the changes that have been identified across generations of bilinguals (e.g. by Gutiérrez, 2003; Lynch, 1999; Silva-Corvalán, 1996; Zentella, 1997) develop naturally in the acquisition of the heritage language in childhood, regardless of whether the innovative features occur in the input or not.

The hindered process of acquisition and development of Spanish underlies incompletely or partially acquired grammatical domains (Montrul, 2008; Polinsky, 2006). I do not argue that adult HSs have acquired the entire system of their heritage language incompletely, or that they do not have communicative competence in this language (although those at the lower ends of the bilingual continuum may not, see Silva-Corvalán, 1996), but that some HSs, including many at the higher ends of the continuum, evidence some reduced or incompletely acquired grammatical aspects compared with these (unreduced) aspects in the grammars of their adult models in the preceding generation.

Some linguists argue that in heritage speakers’ grammars variability between expected and non-expected forms reflects “complete acquisition,” that their grammars are less restrictive or “different” in some respects but not incomplete, and that the notion of “incompleteness” should be dismissed (Carreira & Potowski, 2011; Otheguy & Zentella, 2012; Pascual y Cabo & Gómez Soler, 2015; Putnam & Sánchez, 2013, among others). The notion of “difference” requires a point of
comparison to be able to say in what respect a language variety X is different from Y, and in addition it should be necessary to explain what language features differentiate X from Y. This is normal practice in dialectology.\textsuperscript{15} The notion of “incompleteness” as an outcome also requires a comparison with what is “complete.” Thus, if some domains of the grammars of X (Group 1) are “complete,” then in comparison with these grammars the absence of required subjunctive forms, the omission of clitics in required contexts, and the weaker sensitivity to correferential constraints in Y are an indication of difference but also incompleteness with respect to X.\textsuperscript{16} Finally, while difference is not an explanatory construct, incomplete acquisition due to interrupted development caused by restricted exposure and production of Y offers an explanation for the range of proficiencies attested among adult heritage speakers.

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

**Funding**
The author(s) received no financial support for the research, authorship, and/or publication of this article.

**Notes**

1. Other linguistic factors that have been argued to affect subject expression include priming, discourse genre, verb tense; see, e.g. several in Carvalho, Orozco, and Shin, 2015.
2. I use “exposure” as a shorthand for amount of exposure to and production of a language.
3. Based on the Pearson Chi-Square statistic.
4. For a complete analysis of se constructions see Mendikoetxea, 1999; Butt and Benjamin, 2004.
5. Clitic-doubling occurs in all varieties of adult Spanish, but Domínguez does not provide information about its frequency in the children’s language input.
6. Brennan was not recorded when he was 4;5.2 to 5;0.28 because of his reluctance to speak Spanish after a period of very reduced exposure and use of this language (Silva-Corvalán 2014). I resumed recordings from 5;0.28 to 6;4, but he produced only 59 utterances with clitic contexts during this period.
7. In an elicited task, 12 4-year-old children studied in Los Angeles had an omission mean of 2.08% (Mateu, 2015). This similar result underlines the validity of the current case study.
8. The symbol “~” marks an element that is unacceptable in adult monolingual Spanish.
9. At the earlier age, from 4;1 to 4;5, Brennan’s percentage of errors is a bit higher: 8.8% (11/125).
10. Unwarranted clitics are not included in the totals.
11. Abbreviations used in Table 5 and in examples: Imp = imperfect; Ind = indicative; Periphr = periphrastic; Pres = present; Pret = preterite; Sub = subjunctive.
12. The lack of irealis forms in Spanish is a linguistic question and not a cognitive one. In English, Mike, Nico, and Brennan can easily converse about non-factual past and future situations.
13. Brennan produces one English genitive at 5;6 (el niño’s cabeza “the boy’s head”), but the rest of his genitive constructions follow the Spanish pattern from age 4;1.
14. Meisel (2014) suggests that a period of stabilization up to age 11–12 might be needed to ensure that grammatical knowledge remains permanently accessible.
15. I do not ignore the fact that change may and does occur from one generation (1) to the next (2), but in predominantly monolingual communities and under conditions of normal transmission (Thomason & Kaufman, 1988), generation 2 does not differ from generation 1 to the same extent as in the bilingual communities I have studied in Los Angeles.
16. While many speakers in Group 2 and most in Group 3 never use subjunctive, all speakers in Group 1 never fail to use a subjunctive form in a required context, e.g. in the complement of the verb querer “want”
(Es que yo quiero que mis hijos conozcan su idioma “It’s that I want that my children may know their language”).

References


**Author biography**

Carmen Silva-Corvalán is Professor Emerita of Spanish and Portuguese linguistics at the University of Southern California, where she taught since she obtained her doctoral degree at the University of California, Los Ángeles in 1979. Silva-Corvalán has published extensively on bilingualism and language contact, and on the semantic and discourse-pragmatic constraints which condition syntactic variation. Her most recent book is the highly praised *Bilingual Language Acquisition: Spanish and English in the first six years* (Cambridge Univ. Press, 2014). She is in the board of editors of numerous journals, she was co-editor of the prestigious journal *Bilingualism: Language and Cognition* from 2005 to 2013, and she is currently a member of its editorial board. Her publications include *Sociolingüística y pragmática del español* (Georgetown Univ. Press, 2001), *Spanish in four continents: Studies in language contact and bilingualism* (Georgetown Univ. Press, 1995), *Language contact and change: Spanish in Los Angeles* (Oxford Univ. Press, 1996), widely cited in studies of bilingualism and language contact, and over one hundred articles in journals and anthologies.
Spanish heritage speakers in the Netherlands: Linguistic patterns in the judgment and production of mood.

Brechje van Osch
University of Amsterdam, Netherlands

Petra Sleeman
University of Amsterdam, Netherlands

Abstract

Purpose: This study investigates heritage speakers of Spanish in the Netherlands regarding their knowledge of Spanish mood. Previous research has demonstrated that heritage speakers of Spanish in the US have problems with mood, especially subjunctive mood and particularly in contexts where choice of mood is variable and depends on semantic and pragmatic factors. Moreover, heritage speakers are often reported to experience fewer problems with oral production tasks tapping into implicit knowledge than with judgment tasks targeting metalinguistic knowledge. This study aims to investigate whether these patterns can be confirmed for heritage speakers of Spanish in the Netherlands.

Methodology: In all, 17 heritage speakers from the Netherlands and 18 monolingual speakers of Spanish completed a contextualized elicited production task. Each item contained a context targeting either indicative or subjunctive mood. Below each context followed the beginning of a sentence which the participants were instructed to complete. Both obligatory and variable uses of mood were included. The results were compared to findings from a contextualized scalar acceptability judgment task described in an earlier study using the same conditions and the same participants.

Data and analysis: All responses were coded as felicitous or infelicitous given the accompanying context and were analyzed using mixed effects modeling. The results demonstrate that the heritage speakers are less accurate in their choice of mood than monolingual speakers, particularly on subjunctive mood and in variable contexts. Furthermore, heritage speakers deviated more from the monolingual patterns in the production task than in the judgment task.

Findings/conclusion: These results confirm several patterns attested for heritage speakers of Spanish in the US, namely the increased vulnerability of subjunctive mood and in contexts where mood is not obligatorily selected. However, in contrast to previous literature, this study reports better performance on a metalinguistic judgment task than on an oral production task. This finding is attributed to differences in societal circumstances between both heritage speaker populations.

Corresponding author:
Brechje van Osch, Department of Dutch Linguistics, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, the Netherlands.
Email: b.a.vanosch@uva.nl
Implications of the research: This study confirms the heterogeneity of heritage speakers as a population and emphasizes the importance of taking societal circumstances into consideration.

Keywords
Heritage speakers, Spanish, mood, task effect, societal differences

Introduction

The study of heritage languages is a fast-growing field of investigation within bilingualism research. The literature reports divergent results concerning heritage speakers’ ultimate attainment. An interesting discrepancy exists between studies conducted in the US and studies conducted in Europe. In the US, heritage language acquisition has been on the research agenda for several decades and typically points towards a divergence from the monolingual norm (e.g. Au, Knightly, Jun, & Oh, 2002; O’Grady, Lee, & Choo, 2001; Montrul, 2009; Polinsky, 2011). In Europe heritage language research started more recently and has rendered mixed results. While some studies report diverging linguistic behavior (e.g. Doğruoz & Backus, 2007; Treffers-Daller, Daller, Furman, & Rothman, 2016, and several studies mentioned in Benmamoun, Montrul, & Polinsky, 2013), others demonstrate high levels of proficiency (Flores, 2015; Irizarri van Suchtelen, 2014; Kupisch, Akpınar, & Stöhr, 2013; Kupisch, Belikova, Özçelik, Stangen, & White, 2016; Kupisch et al., 2014; ), although most of these studies also note minor divergences in some constructions or for some participants. This disparity between heritage speakers (sometimes even of the same language) in different parts of the world raises the question about which way societal differences may affect their ultimate attainment.

A robust finding in the US-based studies is that heritage speakers perform better in tasks measuring implicit knowledge, like oral production and online processing tasks, than in more explicit, metalinguistic tasks such as grammaticality judgment tasks (GJT) (Montrul, 2012 and studies mentioned there). The account suggested for this discrepancy relates to the type of input received. Heritage speakers acquire their heritage language in a naturalistic setting in which they are offered plenty of experience with hearing and speaking the language throughout the life span (although the amount of input and activation relatively decreases at school age). On the other hand, they do not receive much (if any) explicit instruction that might trigger them to deal with their heritage language in a metalinguistic way. Whether this pattern can be confirmed for European heritage languages cannot yet be clearly deduced from the existing literature. Kupisch, in several studies, compared different task types within the same group. In her 2014 study, performance is better on oral production tasks than on more explicit tasks, but elsewhere (Kupisch et al., 2014) the opposite pattern is attested. And in Kupisch et al. (2013) ceiling effects are reported for both task types.

The present paper attempts to add a small piece to the puzzle of putative differences between Spanish heritage speakers in the US and Europe by investigating Spanish as a heritage language in the Netherlands and comparing judgment data to production data within the same group on a single phenomenon, namely Spanish mood.

The paper is organized as follows: in the second section, we explain the contexts in which Spanish mood is investigated in the present study. Next, a summary of the previous research is presented, followed by the research questions. The fourth section describes the methodology of the study and in the fifth section he results will be reported. Finally, we offer a detailed discussion of the data and draw conclusions.
Spanish mood

Spanish has two moods: the indicative is generally used in main clauses and is associated with new information and certainty. The subjunctive almost exclusively occurs in subordinate clauses and denotes presupposed or non-factive information such as emotions, doubts, opinions, wishes, or future events.

In many subordinate clauses, mood is syntactically selected by the lexico-semantic features of the main verb. Affirmative epistemic verbs, like saber (to know), require the indicative, whereas volitional predicates like querer (to want), select the subjunctive. In these contexts, mood selection is obligatory; violations of the rule lead to ungrammaticality, as displayed in Examples (1) and (2).

(1) Sé que tú trabajas / *trabajes.  
I know that you work.IND / work.SUBJ

(2) Quiero que tú *trabajas / trabajes.  
I want that you work.IND / work.SUBJ

Following Kempchinsky (2009), we assume that the mechanism with which the uninterpretable feature of mood in the embedded clause is selected is syntactic, although semantics of the matrix verb are also involved.

In variable contexts both moods are grammatical, but one is usually more felicitous than the other. Two types of variable contexts are of interest to the present study. The first type consists of relative clauses. According to most analyses (Borgonovo, Bruhn de Garavito, & Prévost, 2015 and references therein), mood selection in these contexts is related to the specificity of the antecedent it refers back to. Indicative mood corresponds to a specific antecedent (i.e. the speaker knows its identity) and subjunctive mood to a non-specific antecedent, as illustrated in Examples (3) and (4).

(3) Buscamos un hotel que tiene piscina.  
We are looking for a hotel that has.IN a swimming pool (and I know there is one).

(4) Buscamos un hotel que tenga piscina.  
We are looking for a hotel that has.SUB a swimming pool (but I don’t know whether there is one).

The second variable context consists of negated predicates with an embedded clause, also known as polarity subjunctive (Borgonovo & Prévost, 2003). According to Quer (2001), the subjunctive indicates a shift in the model of evaluation from the epistemic model of the speaker to the epistemic model of the matrix subject. The speaker’s commitment to the truth of the embedded proposition determines which mood is more felicitous. The subjunctive can be used with both a [+commitment] interpretation and a [–commitment] interpretation, as Example (5) illustrates. The indicative on the other hand is only felicitous in the [+commitment reading], i.e. when the speaker thinks the proposition is true, as demonstrated in Example (6).

(5) Lucía no cree que Julio esté enfermo.  
Lucia does not believe that Julio is.SUB sick (but I think he is / and neither do I).

(6) Lucía no cree que Julio está enfermo.  
Lucia does not believe that Julio is.IND sick (but I think he is / # and neither do I).

The three contexts with both conditions and their respective expected moods are summarized in Table 1. Note that the [+commitment] condition within the context of negated sentences is the only condition where both moods are acceptable according to the theory.
Apart from a few archaic fossilized expressions, Dutch does not have a subjunctive mood, and is thus similar to English in this respect.3

**Previous research**

*Spanish mood by heritage speakers*

To our knowledge, the Spanish subjunctive has been investigated only for heritage speakers in the US, not in Europe. Various scholars have examined mood in different contexts. Silva-Corvalán (1994), who looked at oral production data from both second and third generation heritage speakers, attested a gradual decline in the use of the subjunctive with every generation, which was especially pronounced in variable contexts, such as temporal clauses and sentences expressing some degree of uncertainty. Montrul (2009), using a more experimental design, compared obligatory contexts to variable contexts, such as relative clauses, sentences with *cuando* (when) and sentences with *de manera que* (so that). Oral and written production tasks were used to test obligatory subjunctive and a sentence conjunction judgment task was used to test mood in variable contexts. Heritage speakers performed more similar to monolingual controls on the tasks targeting obligatory subjunctive than on the task that tested mood in variable contexts. The authors attribute this difference to the more metalinguistic nature of the latter task. Pascual y Cabo, Lingwall, and Rothman (2012) used a scalar felicitousness judgment task to compare advanced heritage speakers’ judgments of mood in obligatory contexts and in a pragmatically constrained context, namely: negated epistemic predicates such as *creer* (to believe) with an embedded clause. The heritage speakers were able to distinguish between moods in obligatory contexts, but unlike the monolingual controls, they failed to do so in the pragmatically constrained context.

Many of the studies investigating Spanish mood focus on the subjunctive and not the indicative, the underlying assumption being that this mood poses more problems for heritage speakers whose majority language does not contain subjunctive mood. Those studies that compared both moods confirm that heritage speakers are indeed more accurate with the indicative than with the subjunctive, in both obligatory and variable contexts and in both written and oral tasks (e.g. Mira, 2009; Montrul 2009, 2011; Montrul & Perpiñán, 2011).

Together, these studies indicate that US heritage speakers of Spanish have problems with mood, especially with the subjunctive and more so in variable contexts than in obligatory contexts.

**Task type effects in heritage speakers**

As mentioned in the introduction, a general observation in the US literature is that heritage speakers perform better on tasks measuring implicit knowledge, like oral production tasks, than on tasks measuring explicit and metalinguistic knowledge, like written production and GJTs. Several

### Table 1. Expected mood in each condition within contexts.

<table>
<thead>
<tr>
<th>Context</th>
<th>Condition</th>
<th>Expected mood</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBLIGATORY</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Epistemic verbs</td>
<td>Indicative</td>
</tr>
<tr>
<td></td>
<td>Volitional verbs</td>
<td>Subjunctive</td>
</tr>
<tr>
<td>VARIABLE 1: RELATIVE CLAUSES</td>
<td>[+specific]</td>
<td>Indicative</td>
</tr>
<tr>
<td></td>
<td>[–specific]</td>
<td>Subjunctive</td>
</tr>
<tr>
<td>VARIABLE 2: NEGATED SENTENCES</td>
<td>[+commitment]</td>
<td>Indicative/subjunctive</td>
</tr>
<tr>
<td></td>
<td>[–commitment]</td>
<td>Subjunctive</td>
</tr>
</tbody>
</table>
studies compared heritage speakers to L2 speakers. Bowles (2011) examined both these groups using a battery of tests ranging from more to less metalinguistic. Topics of investigation were gender agreement, case marking, tense, aspect, and mood in Spanish. The L2 speakers performed better on the more explicit knowledge tasks, such as an untimed GJT and a metalinguistic knowledge task, whereas the heritage speakers scored highest on more implicit knowledge tasks, like a timed GJT, an oral imitation test, and an oral narration task. In a series of studies, Montrul and colleagues found similar patterns for these two populations for various linguistic phenomena in Spanish, such as gender agreement (Montrul, Davidson, de la Fuente, & Foote, 2014), clitics (Montrul, 2010) differential object marking and tense, aspect and mood (Montrul, 2011; Montrul & Perpiñán, 2011). While L2 learners had relatively more problems in a timed comprehension task and an oral production task, heritage speakers made more errors in the untimed written comprehension task. Mira (2009) attested higher use of the subjunctive by heritage speakers in oral than in written production.

It is not clear whether European heritage speakers demonstrate a comparable pattern concerning task type. Most studies conducted in Europe applied either a production task or a judgment task. Both successful and less successful acquisition has been attested with either task type. Three studies by Kupisch and colleagues are worth mentioning that employed multiple tasks on the same group of participants. Kupisch (2014) compared two groups of German-Italian bilinguals regarding their knowledge of adjective placement in Italian: an Italian-dominant group living in Italy and a German-dominant group living in Germany, the latter of which were considered heritage speakers of Italian. More errors were made in an acceptability judgment task than in a spontaneous production task. However, as the authors themselves note, using completely spontaneous production data makes it difficult to rule out any effects of avoidance. Second, Kupisch et al. (2013) compared two groups of German-French bilinguals: a French-dominant and a German-dominant group (the heritage speakers). The phenomenon of interest was gender assignment and agreement in French. Both an acceptability judgment task and an elicited production task were administered, but the heritage speakers performed at ceiling on both tasks; therefore, not much can be deduced about a possible preference for one task over the other. Third, Kupisch et al. (2014) investigated several different phenomena with the same group of heritage speakers of French, and found that they differed more from the monolingual norm in oral production than in judgment. However, different constructions were tested in each task type: the judgment task targeted morpho-syntactic phenomena like adjective placement, gender and article use, whereas the production data was analyzed for prepositions and phonological elements such as global foreign accent and voice onset time. These results thus do not tell us anything about task type effects within a single phenomenon.

In sum, US-based studies have attested higher vulnerability in tasks that require explicit, metalinguistic knowledge compared to tasks that require fast and implicit knowledge such as oral production. Whether European heritage speakers mirror this pattern has yet to be confirmed by further investigation.

Mood in L1 and 2L1 acquisition

Given that heritage speakers represent the end-state of bilingual acquisition, it would be interesting to look at what knowledge of mood bilingual children possess before the language dominance shift occurs. Unfortunately, to our best knowledge, no studies have been conducted that examine Spanish mood in bilingual acquisition in the context of the Netherlands or any other European country. However, we know that in monolingual acquisition obligatory uses of the subjunctive are acquired early on (age two and three), whereas variable contexts, where mood is constrained by semantic
and/or pragmatic constraints, are acquired at a later age, some as late as eight or nine years (e.g. Blake, 1983; Pérez-Leroux, 1998). It is therefore likely that variable contexts of mood are not fully acquired by the time heritage children go to school. Studies conducted in the US typically attest incomplete knowledge of the subjunctive by Spanish-English bilingual children at school age and later on (e.g. Merino, 1983; Silva-Corvalán, 2003), but no clear distinction is made between the different contexts for mood discussed in the present study.

Research question and hypotheses

The present study examines heritage speakers of Spanish residing in the Netherlands. The specific phenomenon under investigation is mood. Our research question is: Do Dutch heritage speakers of Spanish resemble their US counterparts regarding the patterns in their knowledge of mood?

To answer this question, we distinguish between:

1. obligatory and variable contexts;
2. indicative and subjunctive mood;
3. an implicit task (elicited production) and an explicit task (acceptability judgment).

With respect to items (1) and (2), we expect heritage speakers to resemble their US peers, i.e. we predict that Dutch heritage speakers will perform better on the indicative than on the subjunctive and better in obligatory than variable contexts. Given that these effects are related to the nature of the phenomenon itself, there is no reason to expect differential patterns caused by socio-demographic differences (US vs. Europe) or different dominant languages (Dutch vs. English). Regarding item (3), we do not have specific predictions; even though studies in the US have attested a preference for more implicit task types, societal differences might play a role here.

Method

This paper is part of a larger research project. The heritage speakers’ results on the judgment task are reported in van Osch, Hulk, Sleeman, & Aalberse (forthcoming). In the present paper these results will be complemented with production data to compare between task types.

Participants

In all, 27 adult heritage speakers and 18 monolingual speakers participated in the study. However, 10 heritage speakers were excluded based on criteria that will be discussed below. The participants were of comparable age (19 to 38) and were all university or college students or graduates (not enrolled in Spanish courses at the time of testing). All heritage speakers (three male, 14 female) are second-generation speakers, born in the Netherlands to immigrant parents. Except for one case, all participants grew up in a mixed family where one parent (usually the mother) spoke Spanish and the other spoke Dutch.

Most heritage speakers reported hearing and speaking both languages more or less equally often in the home up until the age of five, when a shift occurred toward more Dutch and less Spanish in the input and output. All reported visiting Spanish-speaking countries frequently from an early age onwards, and 10 had lived in a Spanish-speaking country for a few months during childhood (four participants) or adulthood (six participants). Regarding their language use at the time of the study,
nine participants indicated hearing and speaking mostly Dutch at home, four reported slightly more Spanish and four both languages equally. At work or study and in leisure time, the language use and input is predominantly Dutch for most participants. A total of 11 heritage speakers received instruction in Spanish at some point during primary school, secondary school, or college. Their parents originated from various different Spanish-speaking countries: six from Spain, five from Mexico, three from Colombia, two from Uruguay, and one from Argentina. The parents’ length of exposure to Dutch varied, but the majority had lived in the Netherlands between zero and five years before the participant was born.

The 18 monolingual speakers (five male, 13 female) were tested in the Netherlands, but had lived in the Netherlands for no more than six months and had not yet learned any Dutch. They were all born and raised in Spanish-speaking countries (nine from Spain, four from Mexico, two from Colombia, one from Argentina, one from Nicaragua, and one from Venezuela). To our knowledge there is little evidence for dialectal variety concerning mood in the contexts under investigation here. The only relevant reference we found is from Gallego and Alonso-Marks (2014), who report that monolingual Spanish speakers from Toledo, Spain, used obligatory subjunctive (when giving advice) in oral production more often than speakers in an age-matched group in Rosario, Argentina, who sometimes replaced the subjunctive with other grammatical constructions but never with the indicative. The only Argentinian speaker in the present study did not display a different pattern from the other participants.

Spanish proficiency was measured in three different ways. First of all, the participants judged their own reading, writing, comprehension, and speaking skills. Secondly, they were administered the cloze part and the vocabulary part of the DELE (Diploma Español de Lengua Extranjera), a standardized Spanish proficiency task. Finally, they performed a lexical decision task. All measures correlated significantly with one another, as illustrated in Table 2.

Several inclusion criteria were set to keep the group as homogeneous as possible in terms of proficiency. Only participants who scored 37 or higher on the DELE were included, and this corresponded to a proficiency level ranging from high intermediate to advanced. Furthermore, participants had to get more than 100 out of 149 answers correct on the lexical decision task. Finally, their average self-reported proficiency (calculated over the four language skills) had to be 3.5 or higher, corresponding to a rating half way between intermediate and advanced. Apart from proficiency measures, scores on a mood recognition task (MRT) (taken from Borgonovo et al. (2015)) were used to check whether the participants were able to recognize subjunctive morphology. Only those participants who scored higher than 80% correct on this task were included (following Iverson, Kempchinsky, & Rothman, 2008). A group of 17 out of 27 heritage speakers, and all the monolingual speakers, met all the inclusion criteria. The heritage speakers differed significantly from the monolinguals on all proficiency measures (DELE = t = −4.57, \( p = 7.40 \times 10^{-5} \); lexical decision task = \( t = −7.87 \), \( p \)-value = \( 6.54 \times 10^{-9} \); self-reported proficiency = \( t = −4.28 \), \( p = 0.0004 \)) and on the MRT (\( t = −3.59 \), \( p = 0.0015 \)).

The most important information of the included participants is summarized in Table 3.
Tasks and procedure

The participants started by filling out an extensive paper-and-pencil questionnaire. After this, the lexical decision task, the production task, and the acceptability judgment task (AJT) were carried out on a laptop, with a short break between the production task and the AJT. The MRT and the DELE proficiency test were administered on paper at the very end of the session. For the computerized part of the experiment, all instructions and items were presented both written (on the screen) and aurally. The recordings were made by a native speaker of Colombian Spanish and were understood perfectly by all participants, as indicated during debriefing. The focus of the present paper is on the production task, which contained 54 experimental items and 16 fillers, presented in randomized order and preceded by three practice items. Each item contained a short story followed by an incomplete sentence, which the participants were instructed to read out loud and finish in a way that made sense in relation to the story.

The 54 experimental items were divided into three contexts containing 18 items each: one obligatory context, and two variable contexts. Within each context there were two conditions targeting different moods. In the obligatory context, the correct mood in the embedded clause depended on the lexical features of the matrix verb. In half of the items, the main verb was an affirmative epistemic predicate that obligatorily requires the indicative, like saber (to know). In the other half, the main verb was a volitional predicate such as querer (to want), which selects the subjunctive. An example of an item pertaining to this category is Example (7), in which the predicate querer obligatorily requires the subjunctive form:

(7) Estoy molesto porque mi esposa nunca limpia la casa. Esta noche de nuevo no me ayuda a lavar los platos. Me enojo y le digo:
I am annoyed because my wife never cleans the house. This evening, once again, she doesn’t help me with the dishes. I get angry and say:
Quiero que… Expected answer: “me ayudes”
I want that… (pro) me help.2sg.SUB

The first of the two variable contexts consisted of relative clauses. In half of the items, the antecedent was specific, targeting the indicative, and in the other half it was non-specific, targeting the subjunctive. To rule out any biases within the critical sentence, all antecedents were indefinite and inanimate. Example (8) is an example with a specific antecedent, making the indicative the more felicitous option.

(8) Mi jefe viaja muy frecuentemente a Quebec por el trabajo. Sé que allí existe un hotel de hielo, lo que me fascina, pero no sé si él lo conoce. Le pregunto:

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>DELE score</th>
<th>Accuracy lexical decision task</th>
<th>Self-reported proficiency</th>
<th>Mood recognition task (MRT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heritage</td>
<td>Mean: 25.6</td>
<td>Mean: 42.4</td>
<td>Mean: 110.4</td>
<td>Mean: 5.13</td>
<td>Mean: 94%</td>
</tr>
<tr>
<td>SD: 5</td>
<td>SD: 2.8</td>
<td>SD: 6.4</td>
<td>SD: 0.74</td>
<td>SD: 5.36</td>
<td></td>
</tr>
<tr>
<td>Monolingual</td>
<td>Mean: 26.4</td>
<td>Mean: 46.3</td>
<td>Mean: 130.72</td>
<td>Mean: 5.94</td>
<td>Mean: 99.35%</td>
</tr>
<tr>
<td>(N=18)</td>
<td>SD: 4.5</td>
<td>SD: 8.7</td>
<td>SD: 0.24</td>
<td>SD: 2.23</td>
<td></td>
</tr>
</tbody>
</table>
My boss frequently travels to Quebec for work. I know that there exists a hotel made of ice there, which fascinates me, but I don’t know if he knows it. I ask him:

¿Alguna vez has estado en un hotel que… Expected answer: “consiste en hielo?”

Have you ever been in a hotel that… consists.IND of ice?

The second variable context consisted of negated sentences with embedded clauses. Some of the items were based on the test used in Borgonovo and Prévost (2003). Three types of main verbs occurred: epistemic verbs, like creer (to think); communication verbs, like decir (to say); and perception verbs like ver (to see). The preceding story made clear whether the proposition in the embedded clause was true or not. In half of the items the embedded proposition was true, rendering both moods acceptable. In the other half, the embedded proposition was not true, making the subjunctive the only acceptable mood. In Example (9), the speaker considers the proposition not to be true; therefore, the subjunctive is expected.

(9) Selma camina por la calle y ve a su tía caminando a 20 metros de ella. La llama, pero hay mucho ruido de los coches así que es imposible oír algo.

Selma is walking on the street and sees her aunt walking at 20 meters distance. She calls her name, but there is a lot of noise from the cars, so is it impossible to hear anything.

Selma no cree que su tía la… Expected answer: “oiga”

Selma does not believe that her aunt her… hear.3sg.SUB

After a short break followed the AJT, which is described in detail in van Osch et al. (forthcoming). This task contained similar stories, followed by two sentences that were identical except for the mood of the embedded verb. Both sentences had to be rated on a scale ranging from minus two to two.

**Coding**

All responses were initially coded as “indicative,” “subjunctive,” or “other.” A group of 100 responses were excluded from the analysis for one of the following reasons: (1) the participant did not answer at all; (2) the response did not contain a verb; (3) the response was unintelligible; or (4) the response was completely different from what was targeted content-wise. These cases could not tell us anything about the participant’s knowledge concerning mood. A total of 898 responses were coded as indicative and 847 as subjunctive. The “other” category consisted of 45 responses, which were verbs in future tense, conditional, or infinitive. For the statistical analyses, all “indicative” and “subjunctive” responses were recoded as “expected” or “unexpected,” depending on the condition. The “other” responses were evaluated for felicitousness by two native speaker judges. However, in a number of cases the judges did not agree, for instance on the use of a conditional in a [-commitment] context, as in Example (10):

(10) El hombre no creyó que su propio hijo le robaría.

The man did not think that his own son would rob.COND him.

Therefore, ‘other’ responses were excluded from the statistical analysis.8

Furthermore, the [+commitment] condition within the second variable context (negated predicates) could not be included in the analysis, because no “expected” or “unexpected” mood could be determined. Recall that, according to the theory, both moods are considered acceptable in this condition (see Table 1). Moreover, van Osch et al. (forthcoming) demonstrated that in judgment,
monolingual speakers of Spanish showed differential preference patterns for the three predicate types included in this condition: for perception and communication verbs, the indicative was preferred, whereas for epistemic verbs the subjunctive was more acceptable, a pattern confirmed by the production data in the present study. These findings suggest that this condition is not as homogeneous as it has been considered in linguistic theory and therefore should not be compared to the other conditions in the experiment.

**Results**

Due to the exclusion of the [+commitment] condition, two separate mixed effects models were conducted on the data: (1) one comparing all three contexts, but only the subjunctive-targeting conditions within these contexts, and (2) one including both the indicative and the subjunctive targeting conditions, but excluding the second variable context. P-values were calculated based on estimated degrees of freedom using the Kenward-Roger approximation.

In the first model, both “subject” and “item” were included as random effects. The dependent variable was mood (expected vs. unexpected). Fixed effects were group (heritage vs. control), context (obligatory, relative clauses, and negated sentences) and the interaction between these factors. The results indicate that heritage speakers use the subjunctive less often than monolingual speakers overall ($t=6.40, p=3.17 \times 10^{-8}$), and that the difference between the two groups is larger in the two variable contexts than in the obligatory context ($t=7.41, p=6.49 \times 10^{-10}$). These effects are illustrated in Figure 1.

In the second statistical model, only “item” was included as a random effect, because including “subject” did not significantly improve the model. Fixed effects were group (heritage vs. control), context (obligatory vs. relative clauses), and condition (indicative

![Figure 1. Production results on all subjunctive-targeting conditions, split out between groups and contexts.](image)
van Osch and Sleeman

523

Comparison between production and judgment

For a full report of the AJT results please see van Osch et al. (forthcoming). Summarizing those results, we can say that (1) heritage speakers have weaker preferences than monolingual speakers; (2) within the subjunctive targeting conditions, the difference between the groups is larger in the two variable contexts than in the obligatory context; and (3) looking only at the obligatory context and the relative clause context, the heritage speakers differ more from the monolingual speakers in the relative clause context, especially in the [-specific] condition, where the monolinguals have a stronger preference for the subjunctive in the [-specific] condition than for the indicative in the [+specific] condition, whereas heritage speakers’ preferences are equally strong in both conditions. These patterns thus exactly mirror the patterns from the production task, confirming the reliability of both tasks. To be able to compare the two tasks more directly, separate analyses were performed to measure for each group individually whether they significantly distinguished between the two moods in each task. Mixed effects analyses were run on each condition, without including any

Figure 2. Production results on the obligatory context and the relative clause context, split out between groups, contexts and conditions.
fixed effects group, only random effects (“item” and “subject”). The p-values for the intercepts of these models are displayed in Table 4.

Grey cells indicate a deviation from the monolingual pattern. Table 4 shows that in judgment, heritage speakers differ from the monolingual pattern only in the [–commitment] condition within the second variable context containing negated predicates: where monolingual speakers prefer the subjunctive over the indicative as expected, heritage speakers do not significantly distinguish between the two moods.

In production, the image looks quite different. The heritage speakers follow the monolingual patterns only in the obligatory context. In the [+specific] condition within relative clause context, monolinguals do not significantly prefer the indicative to the subjunctive, although there is a strong tendency. Heritage speakers on the other hand do significantly prefer the indicative. In the [–specific] condition it is the other way around: heritage speakers fail to distinguish between the two moods, whereas monolinguals, as expected, prefer the subjunctive. In the negated sentences context, the monolinguals behave in line with predictions: a significant preference for the subjunctive in the [–commitment] condition. Heritage speakers, in contrast, tend to overuse the indicative here, although the difference is not significant. The overall picture that arises from this table is that heritage speakers differ more from the monolingual patterns in production than in judgment.⁹

Discussion

The aim of this study was to examine Spanish heritage speakers in the Netherlands concerning their knowledge of mood in oral production compared to judgment. We expected variable contexts to be more vulnerable than obligatory contexts, and the subjunctive to be more vulnerable than the indicative, mirroring the patterns attested for US heritage speakers. Regarding task type effects, we did not have specific predictions for our heritage speakers. We discuss these three factors separately.

Obligatory vs. variable mood

In both judgment and production, heritage speakers differ more from monolinguals in contexts where choice of mood depends on semantic and pragmatic factors. Similar patterns were found by
Montrul (2007, 2008, 2009) and Montrul and Perpiñán (2011) in written and oral production tasks and in a mood recognition task. In this respect, the Dutch heritage speakers resemble their US counterparts. The increased vulnerability in variable contexts may be accounted for by the Interface Hypothesis (Sorace & Filiaci, 2006), given that the more vulnerable contexts can be argued to be located at the interface between syntax and semantics/pragmatics, whereas the obligatory cases represent a purely syntactic instantiation of mood selection. This line of thought is discussed more extensively in van Osch et al. (forthcoming).

**Indicative vs. subjunctive mood**

In the relative clause context, heritage speakers differ more from monolingual speakers in the subjunctive targeting ([-specific]) condition, than in the indicative targeting (+specific]) condition. The same effect was found in the judgment data. The larger difference with the control group in this condition is caused by the monolinguals having a greater preference for the subjunctive in the [-specific] condition than for the indicative in the [+specific] condition, a finding also reported in Borgonovo et al. (2015) and Giancaspro (2015). In the [-commitment] condition in the negated predicate context, heritage speakers fail to prefer the subjunctive over the indicative in both tasks. They even overuse the indicative in the production task, although the difference does not reach significance. This increased vulnerability with the subjunctive compared to the indicative confirms patterns attested in previous research conducted in the US (Mira, 2009; Montrul, 2009, 2011). The effect may be attributed to the default status of the indicative in Spanish (Sánchez-Naranjo, 2014 and references there), given that the indicative is the more frequent mood, especially in main clauses, where the subjunctive only appears in fixed expressions, like after ojalá (expression of hope). Also, the indicative, denoting factual information, semantically has a less marked interpretation than the subjunctive, which is used to express hopes, dreams, future events, etc.

**Judgment vs. production**

Research in the US has demonstrated that Spanish heritage speakers have more difficulties with tasks measuring explicit knowledge than implicit knowledge. This has been attested for various phenomena (Montrul, 2010; Montrul et al., 2014) including mood, both in obligatory (Montrul, 2011) and variable contexts (Bowles, 2011; Mira 2009). In the present study however, the opposite result was found. The patterns attested for the production task compared to those attested in the judgment task described in van Osch et al. (forthcoming) show that the differences with the monolingual patterns are larger in production than in judgment. Why would Spanish heritage speakers in the Netherlands demonstrate the inverse pattern from their US peers? We suggest that it may be related to differences between the two societies. First of all, Dutch heritage speakers may have an advantage over their US peers when it comes to metalinguistic awareness, because Dutch children have much experience learning foreign languages in school. English lessons start during primary school, and the vast majority of high school students also learn at least one other foreign language (usually French or German, but Spanish and Mandarin are becoming popular electives), and often more than one, for a number of years. A relation between multilingualism and an increased level of metalinguistic awareness has been documented in literature on third language acquisition, as noted in Jessner (2015). Thus, the Dutch heritage speakers’ increased metalinguistic awareness may help them with the more explicit tasks. Secondly, there are important differences between the size and type of heritage speaker communities in both countries. The Spanish-speaking community in the US is large in number and often concentrated in Hispanic neighborhoods, where heritage speakers may interact frequently in Spanish with neighbors, friends, shopkeepers, people in the streets, etc. In contrast, the Spanish-speaking community in the Netherlands is small and dispersed, with little
cohesion among its members. It is thus plausible that the Dutch heritage speakers have less opportunity to activate their heritage language in their day-to-day life than their US counterparts. Consequently, Dutch heritage speakers of Spanish, even though they have accurate knowledge of their heritage language, may have trouble accessing this knowledge in oral production.

However, to make compelling claims about a comparison between the two communities, we need more information about the participants in the cited US studies. We do not know exactly how much contact they had with other members of the Spanish-speaking community. Furthermore, as mentioned by an anonymous reviewer, we have to be aware of possible effects of level of education and socio-economic status (SES). Regarding level of education, the participants in the US-based studies are similar to the participants in the present study: they are all university students. Regarding SES, we do not have information about the participants in the present study, but we have some evidence that at least some immigrant communities from Spanish-speaking countries in the Netherlands tend to have a relatively high SES (see Irizarri van Suchtelen, 2016 regarding Chilean immigrants). In the US studies no information is given about the participants’ SES. To properly investigate the effect of societal differences, studies need to be conducted which test a comparable population in Europe and in the United States using the same test battery and providing detailed information of all participants’ background and current situation. Nevertheless, this study adds a small piece to the heritage puzzle and once again underscores the heterogeneity of heritage speakers and the risk involved in generalizing over different types of heritage speaker communities.

Another problematic issue worth mentioning concerns the baseline. It has been suggested that heritage speakers may receive not only less but also qualitatively different input, due to the fact that their parents, often long-term immigrants themselves, have suffered from attrition (Paradis & Navarro, 2003; Pires & Rothman 2009). Some researchers therefore advocate comparing heritage speakers to first generation immigrants. In the present study, monolingual speakers constitute the baseline, for reasons of comparability with the US-based studies. Moreover, we were interested in the linguistic patterns in the outcome of heritage language acquisition, rather than the underlying factors that may influence its development.

An anonymous reviewer wondered about the possible effects of regularity and tense of the produced verbs. Concerning regularity, both groups produced more regular than irregular verbs but this division was similar across contexts and conditions, and there was no significant effect of regularity on mood for either group ($z=0.50, p=0.62$ for the heritage speakers and $z=-0.86, p=0.39$ for the monolinguals). As for tense, the story is somewhat more complicated. Present and past tense verbs forms were not equally divided over all contexts: in the relative clause context both groups produced almost exclusively present tense verbs, because the verbs in the main clause were always in the present tense as well.11 In the other two contexts both tenses occurred more or less equally frequently. Since this division of the tenses applies to both the judgment and the production task, this problem is irrelevant for the effect of task type. To exclude the possibility that the skewed division confounded the effect of context attested in this study (i.e. the increased vulnerability in the variable contexts), a separate analysis was performed on the present tense items only. The effects remained the same: heritage speakers differed more from monolinguals in the two variable contexts in the subjunctive targeting conditions ($t=-5.84, p =5.31 \times 10^{-7}$), and in the relative clause condition, the difference was biggest for the [–specific] items ($t=-3.73, p =4.36 \times 10^{-4}$). Therefore, it is safe to assume that even the skewed division of the tenses did not considerably bias the results.

**Conclusion**

The present study demonstrated that heritage speakers of Spanish in the Netherlands have problems with Spanish mood, especially with the subjunctive and in variable contexts, where mood depends on semantic and/or pragmatics features. These patterns are in line with previous studies
with Spanish heritage speakers from the US. But contrary to what has been attested in those studies, the Dutch heritage speakers in the present study did not show an advantage for production compared to judgment. In fact, it was the other way around: they behaved more similarly to monolingual speakers in a metalinguistic acceptability judgment task than in an oral production task. We suggest that this result may be accounted for by differential societal circumstances between both heritage speaker populations. Since Hispanic communities in the Netherlands are small and dispersed, Dutch heritage speakers of Spanish may lack the experience of speaking their heritage language on a regular basis that would help them in an oral production task. Furthermore, the multilingual nature of the Dutch educational system may give rise to a higher metalinguistic awareness from which the Dutch heritage speakers would benefit in a more explicit judgment task. This finding once again confirms the heterogeneity of heritage speakers as a population and emphasizes the importance of taking societal circumstances into consideration.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Although Kupisch (2013) emphasizes the research on European bilingual child acquisition, which in many cases can be considered child heritage acquisition.
2. The [+commitment] reading is obtained through an external negation interpretation (i.e. “It is not the case that Lucía believes that Julio is sick”).
3. Except for the fact that English has subjunctive in expression such as “I require that he come,” as well as the for-to-construction with volitional predicates: “I want (for) you to work.”
4. One person reported having lived mostly in Uruguay until the age of three.
5. But see e.g. Molina, 2008 and Serrano, 1995; for regional variation w.r.t. mood in other contexts.
6. Except for one item about a cat.
7. Obtained through personal communication with the authors.
8. For reasons of space, we could not include a qualitative analysis of the “other” category, as suggested by an anonymous reviewer.
9. This pattern is confirmed by the individual data. Individual patterns for each participant, in all conditions and in both tasks were examined. For the heritage speakers, wherever there was a discrepancy between judgment and production (i.e. the expected pattern was demonstrated in one of the tasks, but not the other), it was more often “judgment: yes” and “production: no” than the other way around. For native speakers, such a pattern was not attested.
10. We think the stronger preference for subjunctive in the [–specific] condition than for the indicative in the [+specific] condition is related to the fact that relative clauses of the type “I am looking for…” or “I want to buy…” simply occur more often with non-specific antecedents. It thus might be a frequency effects. For reasons of space, this line of argument has not been included in this paper (but see van Osch et al. (forthcoming)).
11. For reasons of comparability with the judgment task and with previous studies e.g. Borgonovo et al. (2015).

References


**Author biographies**

Brechje van Osch is a PhD candidate at the University of Amsterdam. Her research project, under supervision of Aafke Hulk, focuses on heritage speakers of Spanish in the Netherlands. She is particularly interested in the difference between interface and non-interface phenomena.

Petra Sleeman is a senior linguist at the University of Amsterdam. She has publications in the domain of theoretical linguistics and acquisition. Her main fields of interest are Syntax and the Syntax-Pragmatics Interface. She has mainly worked on Romance languages, but she has also done research on English and Dutch. Recent papers have appeared in Lingua and Glossa.
Heritage language development: Connecting the dots

Silvina Montrul
University of Illinois at Urbana-Champaign, USA

Abstract
To date, the vast majority of research on the linguistic abilities of heritage speakers has focused on young adults whose heritage language is no longer developing. These adults began their journey as bilingual children acquiring the heritage language with the majority language simultaneously since birth or sequentially, as a second language. If longitudinal studies are not always feasible, linking research on the structural development of bilingual pre-school children with research on young adult heritage speakers adds a much needed perspective to understand the initial state and the end state of heritage language development. The purpose of this study is to connect the beginning of heritage language development with its ultimate attainment by comparing the expression of subjects in Spanish in 15 school-age bilingual children and 29 young adult heritage speakers, all of them simultaneous bilinguals with English as the dominant language and Spanish as the weaker language. The oral production of null and overt subjects by child and adult heritage speakers was compared to that of age-matched monolingual speakers in Mexico (20 children, 20 adults). To provide a wider context the study includes a group of 21 adult immigrants, who could also potentially influence the input to the heritage speakers. The results confirm that discourse pragmatic properties of subject expression in Spanish are vulnerable to incomplete acquisition and permanent optionality in child and adult bilingual grammars.

Keywords
Spanish, null/overt subjects, incomplete acquisition, attrition

Heritage language acquisition is a type of early bilingual acquisition that takes place in a specific sociolinguistic environment. The heritage language is a socio-politically minority language, acquired as a first language during the first years of life, as in sequential bilinguals, or simultaneously with the majority language since birth, as in simultaneous bilinguals (Montrul, 2008; Polinsky, 2006). The acquisition of heritage languages deals with the developmental stages and outcome of learning a heritage language from childhood to adulthood, as well as the wax and wane of the heritage language in response to input factors (Montrul, 2016).

Corresponding author:
Silvina Montrul, University of Illinois at Urbana-Champaign, 4080 Foreign Languages Building, MC-176, 707 S. Mathews Ave., Urbana, IL 61801, USA.
Email: montrul@illinois.edu
The vast majority of young adult heritage speakers in the United States are unbalanced bilinguals with stronger command of English than of the heritage language, and in the past two decades there has been keen interest in understanding the linguistic outcome (i.e. ultimate attainment) of heritage language acquisition. The particular emphasis has been on the weaker language of young adults who have achieved a stable level of bilingualism (Benmamoun, Montrul, & Polinsky, 2013a, 2013b; O’Grady, Kwak, Lee, & Lee, 2011; Sekerina & Trueswell, 2011), especially to explain why heritage speakers exhibit systematic gaps in vocabulary, morphological knowledge and in certain discourse pragmatic interfaces compared to fluent bilinguals and native speakers of the same language raised monolingually in a majority language context (Benmamoun et al., 2013a, 2013b; Laleko, 2010; Montrul, 2016; O’Grady et al., 2011; Rothman, 2007). To address these questions, the linguistic competence of heritage speakers has been compared to that of age-matched monolingual and bilingual speakers, monolingual children and older adult bilinguals comparable to the parental generation (Chung, 2013; Montrul & Sánchez-Walker, 2013; Pascual y Cabo, 2013).

Another interest has been to identify the factors that may explain the slower and halted development of the heritage language when compared to the development of the majority language in the same bilingual individuals. Very often, heritage languages exhibit systematic simplification and structural changes. Silva-Corvalán (1994, 2014, 2016), Polinsky (2006), Montrul (2008), and O’Grady et al. (2011), among others, consider that insufficient input and use of the heritage language during childhood contribute to incomplete acquisition, or better yet acquisition without mastery (Montrul, 2016), of several aspects of the language. Although some researchers consider the term “incomplete acquisition” imprecise when applied to heritage language acquisition in general, and unfortunate for possibly inviting value judgments on the speakers and their grammars (Kaltsa, Tsimpli & Rothman, 2015; Otheguy, 2016; Putnam & Sánchez, 2013; Pascual y Cabo & Rothman, 2012), they do not deny that heritage speakers exhibit linguistic characteristics worth explaining.1 Other factors that may also contribute to divergent outcomes in heritage language acquisition and that can co-exist or not with incomplete acquisition or actual attrition, are the quality of the input due to ongoing diachronic change (Lohndal & Westergaard, 2016) as well as majority language transfer (Montrul & Ionin, 2012), as well as frequency and complexity of particular structures (O’Grady et al., 2011). I have stressed (Montrul, 2002, 2008, 2016) that the best way to show that a grammatical property was not mastered by a certain age is to conduct longitudinal studies of bilingual children as they develop their heritage and majority languages from birth to adulthood, a certainly daunting task that nobody has yet undertaken. And to evaluate whether the input is different, the child-directed speech should be examined as well. Except for Silva-Corvalán’s (2014) longitudinal study of two bilingual siblings from approximately ages 1 to 6 in the two languages and the quality of the heritage language input, other existing longitudinal studies (Anderson, 1999; Merino, 1983) have spanned a couple of years in childhood, have focused only on the heritage language, and have not analyzed the parental input. See Montrul (2016) for more extensive discussion.

To date, the vast majority of research on the linguistic abilities of heritage speakers has focused on young adults whose heritage language is no longer developing. These adults began their journey as bilingual children acquiring the heritage language with the majority language simultaneously since birth or sequentially, as a second language. If longitudinal studies are not always feasible, I argue that linking research on the structural development of bilingual pre-school children with research on young adult heritage speakers adds a much needed perspective to understand the initial state and the end state of heritage language development. For example, an area that continues to attract significant attention in Spanish and other pro-drop languages in acquisition and attrition is the expression of null and overt subjects (Carvalho, Orozco, & Shin, 2015; Kaltsa et al., 2015; Otheguy & Zentella; 2012; Paradis & Navarro, 2003, Sorace 2005; Tsimpli, Sorace, Heycock, &...
Filiaci, 2004, among others). Silva-Corvalán (2014) found that the mastery of the pragmatic constraints on subject expression in Spanish lagged behind in the bilingual children studied. Interestingly, the children displayed some of the same patterns attested in the adult heritage speakers reported in Silva-Corvalán (1994), a finding that points to an important connection between bilingual children and adults. By comparing monolingual children and bilingual children, and child and adult heritage speakers, we gain more insight into the developmental route of this particular phenomenon in Spanish in contact with English. At the same time, if early bilingual acquisition is the beginning of the bilingual development of heritage speakers, whereas examining the linguistic characteristics of adult heritage speakers represents their ultimate attainment, the natural question that arises is what happens in between the beginning and the end? To answer this question we must also look at school-age children, since the onset of school marks the beginning of language shift and the change of balance for most bilingual children whose heritage language is not supported at school (Hurtado & Vega, 2004).

In this article I attempt to connect the beginning of heritage language development with its ultimate attainment by comparing the production of null and overt subjects in Spanish in school-age bilingual children and young adult heritage speakers, all of them simultaneous bilinguals with English as the dominant language and Spanish as the weaker language. I investigate whether the pattern of acquisition of the pragmatic constraints on subject expression in adult heritage speakers can be traced back to developmental patterns exhibited by school-age bilingual children. The results of the production of null and overt subjects by child and adult heritage speakers are also compared to those of age-matched monolingual speakers in Mexico. Finally, and to provide a wider context, I also examine whether a group of adult immigrants may exhibit attrition, which could also potentially influence the input to the heritage speakers.

Child bilingualism and heritage language acquisition: What is the connection?

Since the 1980s, important research has been conducted on bilingual first language acquisition in Europe (Deuchar & Quay, 2000; Ezeizabarrena Segurola, 2011; Meisel 2001, 2007; Müller & Hulk, 2001), Canada (Genesee, 1989; Genesee, Nicoladis & Paradis, 1995), the United States (Bolonyai, 2007), Hong Kong (Yip & Matthews, 2006) and Australia (Döpke, 1992). Most of these studies have focused on children growing up in bilingual professional families. Serratrice (2002) studied a child—Carlo—the son of an American father and an Italian mother living in Scotland. Even though Carlo is reported to be bilingual in Italian and English, Italian is a heritage language in this case because it is not the language of the wider speech community. Similar examples are the Hungarian-speaking child studied by Bolonyai (2007), the Spanish-speaking child from the United Kingdom described in Deuchar and Quay (2000), and the German-speaking child from Australia studied by Döpke (1992). The fact that one or the two home languages spoken by these bilingual children is a minority language in the wider community makes them heritage speakers (Montrul, 2016), although they have not been called heritage speakers in the literature. The term is now beginning to be applied to them as well (Kupisch, 2013).

Because early linguistic development by age three or four in bilingual children has been generally found to be balanced and on schedule in each language (Genesee, 1989), it may be implicitly assumed that simultaneous bilingual children end up knowing their two languages very well into adulthood, but this is not always the case. When children realize that their home language is a minority language and it is not spoken beyond the home, they often switch to the majority language spoken by their social group (Kerswill, 1996). De Houwer (2009, p. 3) describes the child studied by Von Raffler-Engel (1965), who grew up in Florence, Italy. His mother always addressed him in
Italian and his father, who was American, in English. The parents spoke English to each other but because the child’s broader environment was Italian-speaking, the child refused to speak English despite the parents’ best efforts. Consequently, the child’s Italian was much stronger than his English.

For theoretical and clinical reasons, the study of simultaneous bilinguals has prioritized the preschool period, tracing the emergence and independent but parallel development of the two languages (Meisel, 2007; Paradis & Genesee, 1996). Although simultaneous bilingual development in childhood may lead to fluent and balanced bilingualism in adulthood when the two languages are supported by input and are used frequently, this pattern of development and ultimate attainment is not the most common in bilingual children of heritage languages. A more typical pattern is what Silva-Corvalán (2014) reported (see Figure 1): some children show parallel and balanced bilingual development until about age 3 or 4, with divergence after ages 5 to 6 when they begin to attend school in the majority language and significant exposure to the majority language increases. After age 5 to 6 the majority language takes over in dominance, and the development of the heritage language slows down, stagnates, and even regresses in specific grammatical areas (inflectional morphology, vocabulary, etc.). These general tendencies, however, do not imply that achieving balanced bilingualism after age 4 is not possible in heritage speakers (see Kupisch, Belikova, Özçelik, Stangen, & White, in press); it is just apparently not very common in the United States.

Although the relationship between young bilingual children and young adult heritage speakers may be obvious—after all, bilingual children are future heritage speakers or heritage speakers who started as bilingual children—only recently have there been attempts to link the early linguistic development of bilingual children with the linguistic abilities of young adult heritage speakers. Silva-Corvalán (2014) is the first study that makes a connection between early development in simultaneous bilingual heritage speakers and the linguistic outcome of young adult heritage speakers (Silva-Corvalán, 1994). Silva-Corvalán’s study of two Spanish–English bilingual siblings living in Los Angeles revealed that from the onset of production and up until age 3, the children showed independent and age-appropriate development of both English and Spanish morphosyntactic structures and did not differ from the development of the same structures in the monolingual acquisition of English and of Spanish. After age 3, when Spanish clearly became the weaker language in the two children, they began to display incomplete development and structural influence from English as compared with monolingual norms in some specific structures (Figure 1). Silva-Corvalán’s analysis focused on the same grammatical areas investigated in Silva-Corvalán’s (1994) study of three generations of Mexican immigrants in Los Angeles—the Tense-Aspect-Mood (TAM)
system, the copulas *ser* and *estar*, pragmatic distribution of null/overt subjects—to allow comparison of the child and adult bilinguals. According to the results of Silva-Corvalán (2014), the copulas exhibited target-like and robust development, while aspects of TAM and the null vs. overt subject expression in topic continuity contexts did not fully develop when input and use of Spanish fell below 25 to 30% after age 3;00. This amount of input was not sufficient for full development of the grammatical areas examined well into age 5;11. The TAM system developed by the siblings was comparable to that of second and third generation Spanish–English bilingual adults who have not received formal education in Spanish (Silva-Corvalán, 1994), and so was the inconsistent mastery and misuse of overt pronouns in topic continuity contexts. Silva-Corvalán’s data add substantial support for the argument that the processes of simplification and loss attested in adult bilingual Spanish are most likely the consequence of an interrupted process of acquisition between the ages of 3;0 and 5;0, when more intensive exposure to English reduces exposure to and use of Spanish.

By comparing young bilingual children and young adult heritage speakers we are comparing the early stages of bilingual development and ultimate attainment, as illustrated in Figure 2. But what we are missing is a closer look at what happens during the school-age period, which is a time of significant language development as well, in both monolingual and bilingual individuals. In fact, the study of adult heritage speakers is highly informative as to the importance of this period in bilingual language development.

In order to understand heritage language development more fully, we must investigate the link between early bilingual acquisition in pre-school and adult heritage speakers of college age. One promising way to connect the dots is by studying in more detail what happens to the heritage language during the entire school-age period, from 5 to 17 years of age, an understudied period in monolingual and bilingual development. If longitudinal studies spanning several years are difficult to impossible to carry out, short-term longitudinal studies and cross-sectional studies comparing school-age bilingual heritage speakers and adult heritage speakers are more feasible. Recent examples are O’Grady et al. (2011), Polinsky (2011), and Montrul and Sánchez-Walker (2013). O’Grady

Figure 2. The early stages and the ultimate attainment of subtractive bilingualism in minority language speaking bilinguals.
et al. (2011) found that the school-age Korean heritage speakers in the United States were more accurate on scope of disjunction under negation in Korean than adult heritage speakers of Korean. Polinsky (2011) also found that school-age heritage speakers of Russian were more accurate on comprehension of relative clauses in Russian than adult heritage speakers. O’Grady et al. (2011) and Polinsky (2011) found evidence of attrition over the lifespan for the adult heritage speakers. Lack of schooling and exposure to advanced language may perhaps underlie the level of language regression observed in the adults in these studies. By contrast, Montrul and Sánchez-Walker (2013), who studied accuracy on production of differential object marking (DOM) in Spanish heritage speakers, found that the school-age children were less accurate than the adult heritage speakers, who were also quite inaccurate. The non-mastery of DOM in young adults was seemingly related to patterns found in earlier stages of bilingual development. Thus, comparing the linguistic knowledge of heritage speakers and school-age bilingual children provides another possibility to address potential attrition or incomplete acquisition of specific grammatical properties of the heritage language. The study presented in this article continues this important line of research by bringing new data from school-age bilingual children, young adult heritage speakers and adult immigrants (the parental generation) on the expression of null and overt subjects in Spanish.

**Null and overt subjects in Spanish**

The expression of null and overt subjects in null subject languages is assumed to be regulated by morphosyntactic and discourse pragmatic factors and has received significant attention in child and adult bilingualism (Paradis & Navarro, 2003; Silva-Corvalán 1994, 2014; Sorace, 2005; Toribio, 2000). For example, the use of null and overt subjects in Spanish is relevant to establish reference in discourse. When there is no switch in reference between a series of sentences, null subjects are appropriate to establish topic continuity and overt subjects are pragmatically infelicitous, as in (1). By contrast, overt subjects are appropriate when there is topic shift and a different referent is introduced, as in (2). Null subjects are pragmatically illicit or infelicitous in these switch reference contexts (examples from Silva-Corvalán, 1994, p. 148).2

(1)  
*Pepe no vino hoy a trabajar. *Pepe/?él/Ø estará enfermo.  
Pepe no came today to work Pepe/?él/Ø will be sick  
‘Pepe did not come to work today. He must be sick.’

(2)  
Hoy no fui a trabajar. Pepe/él/*Ø pensó que estaba enferma.  
today I no went to work Pepe/él/*Ø thought that I was sick  
‘Today I did not go to work. Pepe/he thought I was sick.’

Several studies have shown that bilinguals who speak a null subject language tend to overproduce overt subjects (especially pronouns) in topic continuity contexts (Sorace, 2004). One explanation for such overuse of overt subjects emphasizes the influence from English (Montrul 2004, 2006), an overt subject language. Because the use of null and overt pronouns in discourse engages the syntax-discourse interface, an alternative explanation for the overuse of overt pronouns is the complexity of such interface in pro-drop languages: bilinguals resort to the overt subject option because it is linguistically less complex (Silva-Corvalán, 1994; Sorace, 2005). In the context of Spanish in New York, where many Spanish-speaking bilinguals also overuse pronominal subjects, Otheguy and Zentella (2012) proposed that both contact from English and dialect leveling are at play in these changes. Interestingly, comprehension studies of monolingual children in Mexico show that 6- to 12-year-old children tend to overuse null subjects in switch reference contexts, the opposite
pattern found in bilingual grammars, and only when they are about 13 to 14 years old do they use null and overt subjects as adults (Shin & Cairns, 2012). Therefore, monolingual children also experience developmental difficulties with the discourse pragmatic properties of null and overt subjects, at least in comprehension. If the discourse pragmatic properties of null and overt subjects take so long to converge on the adult grammars in a monolingual environment, such a phenomenon is likely to remain underdeveloped in bilingual contexts where children use Spanish less than the majority language (Silva-Corvalán, 2014).

In the next section, I present a study addressing whether potential incomplete acquisition of pragmatic constraints in young adult heritage speakers relates to developmental patterns exhibited by school-age bilingual children.

The study

In this study, oral production data from 35 school-age children (20 Mexican speakers and 15 simultaneous bilinguals) reported in Montrul and Sánchez-Walker (2015) are compared with data from 60 adults (29 simultaneous bilingual heritage speakers, 20 Mexican native speakers, 21 Mexican immigrants), as shown in Table 1 (children) and Table 2 (adults). Because the child and adult heritage speakers were of Mexican origin, the control groups were recruited in Guanajuato, Mexico.

The parents of all the children completed a short language background questionnaire eliciting information about the children’s linguistic abilities, and the parents of the bilingual children were asked to report on use of Spanish and English. In addition, the bilingual children were administered the standardized Peabody Picture Vocabulary Test–Revised (PPVT–IV) and the Test de Vocabulario en Imágenes Peabody (TVIP), standardized with Puerto Rican and Mexican populations. The children in Mexico also took the TVIP in Spanish. The bilingual children were significantly less proficient in Spanish than in English as per parental assessments ($t(15) = 4.8, p < 0.0001$) and PPVT and TVIP scores ($t(15) = 9.54, p < 0.0001$). The bilingual children differed from the Mexican children on TVIP scores ($t(35) = 25.51, p < 0.0001$).

The 29 heritage speakers consisted of young adults ranging from 18 to 28 years (mean 20.6), all born in the United States to Mexican parents (mother and father). They were exposed to Spanish since birth and to English between birth and age 5 (mean 3.8). One of them had completed college and the rest were college students at a large university in the United States. All of them attended from elementary, middle school and high school in the United States. Only 34.4% received some Spanish instruction through a bilingual program; for all the others the main language of instruction was English. Eighty percent had instruction in Spanish as a foreign language in middle school and high school.
The other comparison group consisted of 21 adult immigrants, who emigrated to the United States after age 18 and had been residing in the country for several years (range 8–37 years, $M = 25.9$ years). They were between 25 and 58 years old at the time of testing ($M = 45.4$ years), and they have completed elementary school only ($n = 5$), high school only ($n = 6$), and college ($n = 10$). Finally, a group of 20 Mexican native speakers tested in Guanajuato, matched in age (range 40–61 years) and SES to the adult immigrant group, was included. Although the level of education of this group also ranged from elementary school only to graduate degree, this group was more skewed toward speakers with elementary education only ($n = 13$), while of the remaining seven, three had completed high school, one some college and three had a graduate degree. Thus, if differences between the speakers in Mexico and the immigrants are found, it cannot be due to lower level of education and instruction of the adult immigrants.

The adult participants completed a lengthy language background questionnaire, with questions about their proficiency self-assessments in English and Spanish, and about frequency of language use throughout early and late childhood, adolescence and adulthood with their parents and other relatives. They also completed a written Spanish proficiency test consisting of a cloze passage with 20 blanks and a 30-item multiple choice vocabulary test. The maximum score was 50. Table 2 summarizes basic descriptive information of the adult groups.

According to one-way ANOVAs the three groups differed on their self-ratings in English ($F(2,69) = 124.7, p < 0.0001$), self-ratings in Spanish ($F(2,69) = 19.6, p < 0.0001$), and written proficiency in Spanish ($F(2,69) = 8.39, p < 0.0001$). The heritage speakers rated their English higher than their Spanish (paired samples t-test, $p < 0.0001$), while the two other groups rated their Spanish higher than their English (paired samples t-tests, both groups $p < 0.0001$). The heritage speakers differed significantly from the Mexican speakers and the adult immigrants in their self-ratings of English and Spanish (all $p < 0.0001$). In terms of Spanish proficiency, the adult immigrants did not differ from the heritage speakers ($p > 0.90$), but both groups differed from the Mexican speakers ($p < 0.05$). Thus, for the heritage speakers English is the dominant language. For the immigrants, Spanish is still the dominant language as per self-ratings, but their written proficiency in Spanish did not differ from that of the heritage speakers.

All participants were provided with 14 colored pictures of the children’s tale Little Red Riding Hood and were asked to look at the pictures and narrate the story in Spanish and in the past, with as much detail as possible. The research assistant who administered the task could see the pictures. All the oral narratives were audio recorded, transcribed and coded for analysis. In this article, we focus on the rate of null and overt subjects and the distribution of subjects in two pragmatic contexts: same reference and switch reference.

### Table 2. Information about the adult participants.

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Age (in years)</th>
<th>AoA English (in years)</th>
<th>LoR US (in years)</th>
<th>Self-rating English (max = 5)</th>
<th>Self-rating Spanish (max = 5)</th>
<th>Spanish proficiency score (max = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heritage speakers</td>
<td>29</td>
<td>20.1 (1.4)</td>
<td>2.8 (1.9)</td>
<td>17.3 (2.1)</td>
<td>4.8 (.4)</td>
<td>3.8 (.8)</td>
<td>36.8 (8)</td>
</tr>
<tr>
<td>Mexican immigrants</td>
<td>21</td>
<td>47.1 (6.5)</td>
<td>20.7 (3.8)</td>
<td>25 (7.7)</td>
<td>3.6 (.8)</td>
<td>4.9 (.2)</td>
<td>37.6 (8.7)</td>
</tr>
<tr>
<td>Mexican speakers</td>
<td>20</td>
<td>47 (6.5)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>1.2 (.5)</td>
<td>5</td>
<td>45.7 (2.5)</td>
</tr>
</tbody>
</table>

AoA: age of acquisition; LoR: length of residence. Standard deviations appear in parentheses.
Since the narrative is in the third person, it mostly elicits third person subjects. The few first and second person pronouns produced were not included in the analysis. Counts and percentages of verbs with subjects were calculated for each participant, as well as counts and percentages of types of subjects, including lexical noun phrases (NPs), and null and overt subject pronouns. Licit and illicit or redundant uses of lexical subjects, pronominal subjects and null subjects in same reference and switch reference contexts were also calculated. For example, overt subjects were considered redundant when they referred to the same referent mentioned in the previous sentence (a context where a null pronoun would be felicitous) (see (3)), and they were considered correct if they were used when there was a change of referent in the discourse, or to establish emphasis (see (4)).

(3) . . . el muchacho tienen unas estijeras. El muchacho le está poniendo bolas de piedras en su barriga. (participant 304)

(4) ‘. . . the young man have scissors. The young man is putting balls of stone in his belly.’ El lobo malo atacó al Capurecita Roja. El señor llega. (participant 10)
‘The bad wolf attacked Little Red Riding Hood. The man arrives.’

(5) Similarly, null subjects were considered illicit when there was a change of referent in the story line, and the participant used a null subject, making the context ambiguous or unclear. Ahí el leñador oyó los gritos de Caperucita Roja y Ø [el lobo] se la comió. (participant 5)
‘There the woodcutter heard the screams of Little Red Riding Hood and Ø [the wolf] ate her.’

However, there were some null pronouns that appeared to be licensed by the researcher and participant’s shared knowledge and joint attention to the story (see note 3), as in (6). These task-related null subjects in switch reference contexts occurred in both monolinguals and bilinguals. They were less than 2.4% of all the data and were excluded from the analysis.

(6) Después vino un cazador y vio el lobo. Ø Encontraron a su abuela. then came a hunter and saw the wolf Ø found-3rd-pl her grandmother
‘Then a hunter came and saw the wolf. They found her grandmother.’
(Participant 311) (In the picture the hunter comes in with a dog, thus the third person plural in the second sentence can refer to the dog and the hunter.)

The data from the school-age children and the adult data by the region of the monolinguals (Mexico) and the bilinguals (United States) were compared statistically to test whether the monolingual children differ in development from the monolingual adults on the one hand, and to what extent the bilinguals are similar to the monolingual children on the other. We also compared the child and adult heritage speakers, and the heritage speakers with adult immigrants. The following specific research questions were considered for the analysis.

1. Do child and adult heritage speakers of Spanish produce more overt subjects in Spanish than child and adult speakers in Mexico?
2. Do child and adult heritage speakers exhibit more difficulty with the pragmatic constraints on overt subjects than the pragmatic constraints on null subjects?
3. Do adult immigrants from Mexico produce more overt subjects than native speakers from Mexico, a potential sign of attrition?

Figure 3 displays the rate of overt subjects produced by all the groups.
As can be seen in Figure 3, the school-age Mexican children produced significantly more overt subjects (17%) than the adult Mexican speakers, and the difference was significant ($t(40) = 4.23, p < 0.023$). This suggests that 11-year-olds have not yet reached the adult distribution. A one-way ANOVA comparing the monolingual and bilingual groups was significant ($F(4,100) = 3.69, p < 0.008$), but Tukey post-hoc tests indicated that the main statistical difference was between the Mexican adults and all the other groups. The three bilingual groups produced more overt subjects than the monolingual Mexican children (children 12.1%, heritage speakers 1.6%, and immigrants 6.2%); the four groups did not differ statistically ($F(3,109) = 1.26, p = 0.29$).

Figure 4 (Mexican speakers) and Figure 5 (bilingual speakers) compare pragmatically felicitous and infelicitous uses of lexical, pronominal and null subjects in same reference and switch reference contexts. In same reference contexts null subjects are felicitous and overt subjects are not: the school-age children from Mexico produced more infelicitous uses of overt subjects (4.1% lexical, 6.3% pronominal) than the adults (1.9% lexical, 2.3% pronominal) ($F(1,39) = 3.42, p = 0.04$). In switch reference contexts, lexical and pronominal subjects are felicitous and null subjects are pragmatically illicit. The monolingual Mexican children produced more illicit null subjects than the adults (5.9% vs. 1.4%, $F(1,39) = 3.42, p = 0.04$).

Figure 5 shows the results of the bilingual groups. In same reference contexts, the bilingual groups differed from each other on redundant uses of lexical ($F(2,69) = 6.54, p < 0.0001$) and pronominal subjects ($F(2,69) = 3.54, p < 0.041$), largely due to the fact that the bilingual children produced significantly more pragmatically illicit overt subjects (24.7% and 12.2%) than the heritage speakers (11.8% and 3.6%) and the adult immigrants (4.6% and 5.2%) (Tukey, $p < 0.05$). Compared to the less than 2.3% of illicit overt subjects produced by the Mexican adults in Figure 4, the immigrants in Figure 5 have a higher rate of overt subjects in same referent contexts (more than 4.6%). Yet, it is much lower than the rate produced by the bilingual children (24.7%) and in the heritage speakers (11.8%).

In switch reference contexts, where overt subjects are felicitous, the child and adult heritage speakers produced twice the rate of illicit null subjects (12.7% and 10.5%) compared to 5.9% produced by the monolingual children in Figure 4. The bilingual groups differed on accuracy rates of lexical ($F(2,69) = 6.54, p < 0.0001$) and pronominal subjects ($F(2,69) = 6.54, p < 0.001$), largely due to the lower accuracy of the bilingual children, who differed from the heritage speakers and the
The adult immigrants produced 5% of illicit null objects, more than half the rate of the child and adult heritage speakers. The 5% overuse of overt subjects by the adult immigrants was largely due to six individuals, and not to the entire group. (The range of redundant pronouns in these six individuals ranged from 4% to 23%).

**Discussion**

The aim of this study was to attempt to connect the dots between the bilingual development of simultaneous bilingual children and adult heritage speakers. To place the results in a wider context, I examined the rate and the distribution of null and overt subjects in an oral narrative task administered to school-age bilingual children, young adult heritage speakers, adult Mexican

---

**Figure 4.** Mexican speakers’ rate of lexical, pronominal and null subjects in same reference and switch reference contexts.

**Figure 5.** Rate of lexical, pronominal and null subjects in switch reference contexts.
immigrants and age-matched school-age children and adults from Mexico. Comparing the linguistic behavior of child and adult bilinguals allows us to see whether many of the apparent grammatical differences found in young adult heritage speakers and native speakers in the country of origin can be traced back to protracted development in childhood. Furthermore, comparing heritage speakers to adult immigrants, who are akin to the parents of the heritage speakers and the main source of input, helps us address, at least indirectly, whether the type of input heritage speakers receive at present exhibits changes due to attrition.

Shin and Cairns (2012) found in their experimental study of sentence comprehension that monolingual Mexican children did not establish reference at adult levels until after 12 years of age. The results of the oral narratives in the present study show that the expression of subjects in Spanish is subject to developmental effects in monolingual children from Mexico when compared to adults from the same region. In general, the Mexican school-age children produced more overt subjects in same reference contexts than adult native speakers. Our first research question was whether child and adult heritage speakers of Spanish produce more overt subjects in Spanish than child and adult speakers in Mexico. According to our results, child and adult heritage speakers of Spanish indeed produced more overt subjects than the adult Mexican speakers, but their rates did not differ statistically from the rates of the monolingual children. The bilingual children did produce more overt subjects than the monolingual children, although this difference did not reach statistical significance. If the differences between the monolingual children and adults relate to the acquisition of reference, it looks like the bilingual school-age children exhibit similar developmental difficulties, which are further enhanced in a bilingual context. These tendencies are also found in young adult heritage speakers, suggesting a developmental link between the two groups.

Monolingual children have been shown to have problems with null subject reference in comprehension tasks until age 12 (Shin & Cairns, 2012), while the interpretation of overt subjects seems to be the main problem in bilingual grammars. Our second question was whether child and adult heritage speakers would exhibit more difficulty with the pragmatic constraints on overt subjects than the pragmatic constraints on null subjects. If bilingual children and young adult heritage speakers produce higher rates of overt subjects and of pragmatically illicit subjects than their age-matched monolinguals, it can be assumed that the bilinguals did not have a chance to develop the pragmatic features of Spanish subjects fully, especially if child and adult bilinguals do not differ from each other.

The results showed that the linguistic behavior of the child and adult bilinguals was not very different: they all produced more redundant null and overt subjects than the Mexican native speakers. The child and adult bilinguals produced pragmatically illicit null subjects like the monolingual children, but they produced much higher rates of redundant overt subjects as attested in bilingual children and adults (Montrul, 2004; Paradis & Navarro, 2003; Silva-Corvalán, 1994, 2014). Yet, of all the bilingual groups, the bilingual school-age children displayed the highest rates of redundant lexical and pronominal subjects. The fact that the young adult heritage speakers also show high use of overt pronouns when compared to the adult Mexican group suggests that difficulty with the discourse pragmatic distribution of subjects continues well into the adult years. Although indirectly, we were able to establish a link between patterns found in school-age children and patterns found in adult heritage speakers (cf. Silva-Corvalán, 1994, 2014).

Finally, this study also sought to establish whether there is a relationship between the changes observed in the heritage speakers and in the parental generation. If there is deviation from monolingual performance, is it similar in both groups? We asked whether adult immigrants from Mexico produce more overt subjects than native speakers from Mexico. Redundant uses of overt subjects have been portrayed as a key vulnerable area in adult immigrants undergoing attrition (Sorace, 2000; Tsimpli et al., 2004; Kaltsa et al., 2015), although the magnitude of this effect is much
smaller in L1 attrition than in heritage speakers (Montrul, 2008). Montrul (2004) reported illicit uses of null and overt subjects in Spanish heritage speakers and Sorace (2004) speculated that the results could be due to intergenerational attrition, although Montrul’s study did not include data on immigrants. The present study, which used the same narrative task as Montrul (2004), included adult immigrants to verify this possibility and to see whether attrition in adults could also contribute to the incomplete acquisition of pragmatic features of Spanish subjects attested in the heritage speakers. The results show that adult immigrants from Mexico, who have been in the United States for an average of 25 years, produced more overt subjects than adult speakers in Mexico. Although the heritage speakers produced higher error rates than the immigrants, the statistical analyses were not significant.

The question is whether the rate of null and overt subject production by the heritage speakers is related to the input they may receive from the parental generation. Heritage speakers often report their highest and most frequent use of the heritage language with their parents and older relatives (Hurtado & Vega, 2004). Since adult immigrants are akin to the parents of the young adult heritage speakers at the time of testing, this would suggest that some heritage speakers may also receive at present (but not while they were growing up) qualitatively different (attrited) input, with occasional misuse of null and overt subjects in discourse. Residual optionality in the input would reinforce the grammar resulting from the delayed development of their Spanish during the school-age period, which exhibits non-target mastery of the pragmatic features of subjects. Thus, while bilingual school-age children show developmental and bilingualism effects, the adult heritage speakers’ grammars are the result of bilingualism, reduced input in childhood, and possibly input effects from relatives who may undergo attrition.

In conclusion, subject expression is vulnerable in child and adult bilingual grammars, and there is a natural connection between the grammatical development of bilingual children and the ultimate attainment of school-age and young adult heritage speakers. Consistent with what Silva-Corvalán (1994, 2014) found in child and adult heritage speakers, this study further confirms that discourse pragmatic properties of subject expression in Spanish are vulnerable to incomplete acquisition and permanent optionality in child and adult bilingual grammars (Sorace, 2000, 2005, 2011).

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This material is based in part upon work supported by the National Science Foundation under Grant Number BCS-0917593, ARRA. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

**Notes**

1. Kaltsa, Tsimpli and Rothman (2015) state: “Our issue with the term incomplete acquisition to refer to heritage language differences is two-fold. First, it does not take into account possible input differences that heritage language speakers may be exposed to compared to monolingual children. In this case, the acquired grammar is complete insofar as the input properties are similar in quality and quantity for native grammatical development, since age of onset issues and age considerations with respect to cognitive development are similar. Secondly, we believe the term to be descriptively inaccurate. A truly incomplete grammar would be one that is not systematic for the domain that is claimed to have been incompletely acquired. This is not the case of heritage grammars, as might be the case for some properties in L2 grammars, labeled as such under the origin of the term from Schachter’s (1990) Incompleteness Hypothesis. If it were to be stable and ‘incomplete’ then it is, in principle, a different grammar, (p. 268).” In response
to these assertions, I submit that empirical evidence for the potential effects of input quality on heritage grammars are given by Montrul & Sánchez-Walker (2013) and Pascual y Cabo (2013). Montrul, Bhatt & Girju (2015), who also compared heritage speakers and first generation immigrants, did not always find a link between the two. Thus, input differences cannot be at play, at least in these cases. Furthermore, in my understanding, Montrul (2008), Polinsky (2006), O’Grady et al. (2011) and Silva-Corvalán (2014) have never equated “incomplete acquisition” with lack of systematicity in heritage language grammars nor have they ever used the term to denote or imply a value judgment on the speakers. Benmamoun et al. (2013a, 2013b), Polinsky (2008), Kim, Montrul and Yoon (2009) and Montrul (2016) have stressed how heritage grammars, even when they exhibit simplification when compared to the source language, are systematic and not random, following strikingly similar simplification patterns in the same grammatical areas in language after language. See Silva-Corvalán, 2016, for a similar position.

2. Overt subjects can also be contrastive (Ordóñez & Treviño, 1999; Frascarelli, 2007), and involve another layer of functional projection in the syntax (topic).

3. Sorace (2004) questioned the suitability of this methodology to investigate the pragmatic distribution of subjects for two reasons: 1) the number of tokens produced by each subject differs, and 2) because the story is universally known and there is no information gap between the participant and the researcher, perhaps making the participant produce more null subjects in shared knowledge context than expected. I am aware of this criticism and acknowledge that using a narrative of this sort limits the ability to control for context and draw clearer conclusions about the participants’ knowledge of null and overt subjects. However, this task is valuable for comparing children and adults and for identifying emerging trends that can be the focus of a future experimental task. Finally, I share Sorace’s concern that this narrative and the way it was administered may have prompted some participants to use null subjects in switch reference contexts because they were licensed by the shared context of the narrative. I considered this possibility in the data analysis.

References


**Author biography**

Silvina Montrul is Professor of Spanish and Linguistics at the University of Illinois at Urbana-Champaign, where she is also Director of the University Language Academy for Children and Director of the Second Language Acquisition and Bilingualism Lab. She is also co-editor of the journal *Second Language Research.*
She is author of *The Acquisition of Spanish* (Benjamins, 2004), *Incomplete Acquisition in Bilingualism, Re-examining the Age Factor* (Benjamins, 2008), *El bilingüismo en el mundo hispanohablante* (Wiley-Blackwell, 2013), and *The Acquisition of Heritage Languages* (Cambridge University Press, in 2016), as well as numerous articles in journals such as *Bilingualism: Language and Cognition*, *The International Journal of Bilingualism*, *Language Learning*, *The Heritage Language Journal*, *Studies in Second Language Acquisition*, *Language Acquisition*, *Second Language Research*. Her research focuses on linguistic and psycholinguistic approaches to adult second language acquisition and bilingualism, in particular syntax, semantics and morphology. She also has expertise in language loss and retention in minority language-speaking bilinguals, or heritage speakers.
Bilingual children and adult heritage speakers: The range of comparison

Maria Polinsky
University of Maryland, USA

Abstract
This paper compares the language of child bilinguals and adult unbalanced bilinguals (heritage speakers) against that of bilingual native speakers of their home language (baseline). We identify four major vectors of correspondence across the language spoken by these three groups. First, all varieties may represent a given linguistic property in a similar way (child bilinguals = adult heritage speakers = bilingual native speakers of their home language). This occurs when either (i) the property in question is highly robust and is acquired by learners without difficulty or (ii) the property is already in decline in the baseline. We illustrate scenario (i) with data from Russian count forms, which are morphologically quite complex. The preservation of these forms in child bilinguals and adult heritage speakers suggests that simplicity of encoding is not the only factor determining robustness of retention. Second, child and heritage speakers may share a linguistic structure that differs from the one found in the baseline (bilingual native speakers of their home language ≠ child bilinguals = adult heritage speakers). This scenario occurs when incipient structural changes in the baseline become amplified in the language of next-generation bilinguals, or when a given structure is rare, confined to a specific register, and/or reinforced through literacy. Third, a structure may be acquired by bilingual children faithfully, but undergo reanalysis/attrition in the adult heritage language (bilingual native speakers of their home language = child bilinguals ≠ adult heritage speakers). Russian relativization illustrates this scenario; child bilinguals show native-like performance on relative clauses but adult heritage speakers show an exaggerated subject preference in the interpretation of gaps. Finally, a structure that is not fully learned by child speakers may be reanalyzed by adult heritage speakers following general principles, thus bringing the adult heritage representation closer to that of the baseline (bilingual native speakers of their home language = adult heritage speakers ≠ child bilinguals). Heritage speakers’ production and comprehension of psychological predicates in Spanish illustrates this possibility.

Keywords
Attrition, count forms, heritage language, numerical expressions, psychological predicates, reanalysis, relativization
Introduction

Linguistic research over the last two decades has uncovered a significant number of properties that, taken together, allow us to identify heritage language as a particular phenomenon within bilingualism. Heritage language (HL) is typically understood as the language that early bilinguals grew up exposed to at home, before becoming dominant in the main language of their society (Valdes, 2000; Polinsky & Kagan, 2007, among others). The recognition of HL speakers as a special speaker group has inspired a ream of new empirical, experimental, and theoretical studies (see Benmamoun, Montrul, & Polinsky 2013a, 2013b; Montrul, 2016; Scontras, Fuchs, & Polinsky, 2015, for overviews). To date, the bulk of this research has focused on comparisons between adult heritage speakers and adult L1 speakers of the baseline, i.e. the language that constitutes the main input to heritage learners.

Although this research has uncovered a number of structural properties characteristic of HL, the origins of these recurrent properties remain underexplored. In order to fully understand adult HL, it is imperative to consider the language of ‘future heritage speakers’: childhood bilinguals who are still receiving daily input in the home language but who operate under similar sociolinguistic conditions to those reported for adult heritage speakers. These conditions include residence outside the country where the HL is dominant, lack of formal education in the HL, an (upcoming or ongoing) rapid switch to the dominant language of their society, and decreasing input from the home language as they spend more time in school (He, 2011, 2014).

In comparisons of future and current heritage speakers, two possible approaches seem pertinent: longitudinal studies that follow the same subjects from childhood to adulthood, and experimental studies that compare children acquiring the HL (‘future heritage speakers’) to HL adults with closely matched linguistic biographies. The work of Silva-Corvalan (2014; this volume) is an example of the former type of study; the author carefully monitors the language development of two siblings growing up in a bilingual Spanish-English environment, showing how English eventually becomes the subjects’ stronger language. Longitudinal studies are by necessity limited to a small number of subjects, and while they contain valuable observations, they need to be supplemented by controlled comparisons among larger pools of speakers. Such comparisons should also incorporate the language of first-generation immigrants (the baseline), which provides input for bilingual learners.

In much existing research on HL, differences between a baseline language variant (spoken by first-generation immigrants) and the standard language variant (spoken in the immigrants’ country of origin, i.e. the homeland) are disregarded; however, evidence from recent studies suggest that such differences can be significant. Languages with large, diasporic populations do not remain homogeneous. The Spanish of first-generation US immigrants from Chiapas, Mexico cannot be assumed to be the same as Peninsular Spanish or even a different variety of Latin-American Spanish. Recent studies suggest that the language of first-generation immigrants undergoes minor changes that persist and spread rapidly in the HL of subsequent generations. The propensity to use overt rather than null subjects is a well-known feature of HL, for example; crucially, this tendency emerges regardless of the presence of pro-drop in the dominant language of the society (cf. immigrant Italian in a Spanish-dominant environment, as discussed in Sorace and Filiaci, 2006; immigrant Russian in a Hebrew-dominant environment, as discussed in Dubinina and Polinsky, 2013; and immigrant Spanish in an English-dominant environment, as discussed by Montrul in this volume, and by Otheguy and Zentella, 2012, p. 9).2

Assembling data on baseline speakers, childhood bilinguals, and heritage speakers is a challenging task, and one function of this article is to highlight the need for additional data collection on these crucial three-way comparisons. Once such data are available, it becomes possible to
conduct a structural comparison across the three varieties. When comparing a particular structure across three related linguistic variants, several logical possibilities arise: all three variants may treat the structure in the same way, each may treat it differently, or any two variants may pattern against the third. In what follows, I discuss and illustrate several of these logical possibilities in the context of HLs, using data from recent studies. An important goal of this kind of work is to be able to predict, for a particular phenomenon or population, which logical scenario will develop. Although the present paper is not able to answer this question, it is my hope that, by pulling together disparate study results into a coherent preliminary discussion, we can lay the groundwork for future analyses.

Before I proceed, a clarification is in order concerning the data discussed in this paper. Most of the studies I discuss compare adult speakers to children above the age of three. The underlying assumption in these studies is that children’s mastery of the particular structures under investigation has had sufficient time to develop by this age. I adopt the same assumption in my discussion below. This focus on slightly older children allows researchers to consider both simultaneous and sequential bilinguals. Ideally, these two groups should be considered separately, but in practice, they are often collapsed into one group.

In the next four sections, I outline four hypothetical acquisition scenarios describing possible patterns of mastery among bilingual native speakers (BIL), adult heritage speakers (AHL), and child bilinguals (future heritage speakers, CHL).

1. All three groups show equal mastery (or non-mastery) of the property in question (BIL = CHL = AHL).
2. The property is present in the baseline but undergoes change in child language and the adult HL (BIL ≠ CHL = AHL).
3. Child bilinguals do not differ from the baseline with respect to the property; adult heritage speakers differ from both groups (BIL = CHL ≠ AHL).
4. The adult HL patterns with the baseline with respect to property X, while child language differs (BIL = AHL ≠ CHL).

It’s language all the way down

First, let us consider scenario (1a); all three groups show equal mastery (or non-mastery) of a given property.

This scenario suggests two possibilities: first, that the property is quite robust and is acquired unproblematically, or second, that the property is undergoing change and restructuring in the baseline and the other varieties (Meisel 1986, 1989, 2001). Below, I illustrate this with data from Russian numerical expressions.

In Russian numerical phrases, the noun varies in form depending on the numeral. Nouns accompanied by numerals five and higher or the words *mnogo* ‘many’ and *malo* ‘few’ appear in the genitive plural; nouns occurring with the numerals one-and-a-half, two to four and the word *oba* ‘both’ require a special paucal form. For the overwhelming majority of Russian nouns, the paucal coincides with the genitive singular.

(2) a. odin velosiped
   one bike.NOM
b. tri velosiped-a
   three bike-PAUCAL/GEN.SG
c. vosem’ velosiped-ov
   eight bike-GEN.PL
Russian acquisition of case-forms is laborious and strewn with errors (see Gvozdev, 1961; Slobin, 1966; Gagarina & Voeikova, 2009; Eisenbeiss, Narasimhan, & Voeikova, 2008 for monolinguals; Schwartz & Minkov, 2014 for bilingual children). Acquisition of case-forms for numerical expressions may appear equally challenging, although the data are inconsistent. On the one hand, Russian Child Language Data Exchange System (CHILDES) data show no errors with numerical expressions. On the other hand, Gagarina and Voeikova remark that ‘[e]rrors in [the acquisition of number in the nominal domain] are frequent with all Russian children and are found until the age of five years’ (Gagarina & Voeikova, 2009, p. 199).

Russian morphology has a great number of syncretisms; note that, in the discussion below, the crucial question is whether or not children use the correct case-form, not the particular allomorph. The Russian genitive is actually an ideal domain for investigations of case substitution, as such substitution is particularly common with the genitive of negation (Babyonyshev, Ganger, Pesetsky, & Wexler, 2001; Modyanova, 2006) and genitive of possession. Even so, case substitution does not appear to take place in numerical expressions. Aside from two examples cited by Gagarina and Voeikova (2009), no case substitution errors with count forms are reported elsewhere in the literature. In their examination of case-form production in the speech of Hebrew-Russian bilingual children (ages three years three months to five years three months), Schwartz and Minkov (2014, p. 76) observe 14 case substitutions in 127 contexts where the dative singular is warranted, and only three substitutions in 269 contexts calling for the genitive singular. Even if all three incorrectly used genitives were in numerical expressions (which seems unlikely: see Schwartz & Minkov, 2014, p. 72), that’s a remarkably low number compared to the respective number of substitutions of the dative.

To investigate count-form errors among adult heritage speakers, we conducted an analysis of heritage narratives produced by American Russian (n = 31) and German Russian (n = 19) speakers (average age 23 years and seven months) (Ivanova-Sullivan, 2013; Denisova-Schmidt, 2014). All the subjects produced a narrative based on the same video clip, so the data were largely comparable. Despite the cross-subject variation in proficiency level, and a number of systematic errors (Ivanova-Sullivan, 2014), there were no errors in the use of count forms.

In another study, 19 American English-dominant heritage speakers (average age 18 years and nine months) were asked to read stimuli with matching or mismatching case forms in the numerical expression (see Xiang, Harizanov, Polinsky, & Kravtchenko., 2011 for the details of the experiment), and rate the examples on a seven-point Likert scale (1: completely unnatural, 7: completely natural). (Note that, since this study included reading tasks, we had to limit participation to heritage speakers literate in Russian). Filler sentences included ungrammatical conditions, in particular ones where the animate accusative appears in place of the dative, as shown below.

(3) Ženja podaril syn-u/*syn-a …
    Zhenya gave.as.gift son-DAT/son-ACC
    nov-ujу igrušк-u.
    new-ACC toy-ACC

‘Zhenya gave his son a new toy …’

Figure 1 shows the ratings obtained for conditions where the noun in the numerical expression appeared in the correct vs mismatched form. Heritage and native speakers perceived the differences between correct and mismatched forms equally clearly; the differences in ratings between the grammatical and ungrammatical conditions were statistically significant for both groups.

Conversely, the ratings for fillers that included the mismatched case of the indirect object were different across speaker groups (Figure 2). Native speakers were equally discriminating of matches
Polinsky and mismatches, but heritage speakers’ ratings for both conditions were not significantly different, which suggests that their sensitivity to case mismatches is not global and is much sharper with numerical expressions.7

The findings of these studies suggest that Russian speakers from all walks of life use count forms correctly. If so, we have a clear illustration of scenario (1a). This possibility, in turn, raises an interesting question for future research: What are the universal principles governing acquisition of numerality in monolingual and bilingual children?

**Difference between the baseline and all heritage speakers**

In scenario (1b), a given property in the baseline undergoes a parallel change in the language of child bilinguals and heritage speakers.

Such a pattern may arise under a number of different conditions. For example, a language property that is rare and only available in certain registers may be naturally absent in the language of
speakers whose input is restricted to informal and familiar contexts. Consider the Spanish non-
finite absolute construction. The absolute construction is syntactically similar to an adverbial sub-
ordinate clause, but lacks any sort of subordinating conjunction that overtly specifies its logical
connection to the clause it adjoins to. The predicate of the absolute construction occurs in the par-
ticipial form, followed by any nominal complements.

(4) [Llegado el sacerdote], la boda podía comenzar.
    arrive.PARTICIPLE the priest the wedding could begin.
    ‘The priest having arrived here, the wedding can begin.’ (Bruno, 2011, p. 264)

The Spanish absolute construction is sensitive to the external/internal argument distinction: its verbal
argument must be interpreted as internal, which rules out the use of unergatives (Alcázar & Saltarelli,
2008; Bruno, 2011; Hernanz, 1991; López, 2001). Spanish-learning children, heritage speakers, and
L2 speakers show equal difficulties in producing and evaluating this construction, and especially in
enforcing the unaccusative constraint (see Montrul, 2006, p. 44-51, for Spanish heritage speakers,
and Montrul 2004, p. 329, for L2 Spanish). These findings suggest that the acquisition of the absolute
construction is limited to more formal registers and may happen relatively late in development.

Scenario (1b) may also arise when a particular language structure or property is in the incipient
stages of change in the baseline grammar. In this case, HL acquisition of the baseline may amplify
this change and make it more apparent. Consider the development of differential object marking
(DOM) in Spanish. DOM is a generalization over a set of facts about Spanish grammar – in par-
ticular, the requirement that objects which are [+animate] and [+specific] be marked with the
preposition a (‘a personal’), whereas other types of objects appear unmarked.

(5) a. María quiere a un abogado. [+animate, +specific]
    ‘Mary wants a (specific) lawyer.’

b. María quiere un abogado [+animate, -specific]
    ‘Mary wants a lawyer (any lawyer).’

(6) Juan destruyó (*a) la ciudad [-animate, +specific]
    ‘Juan destroyed the city.’

Rodríguez-Mondonedo (2006) shows that monolingual Spanish children acquire adult-like DOM
by the age of three. Montrul and Sánchez-Walker (2013) and Montrul, Bhatt, and Girju (2015) tested
DOM in English-dominant heritage speakers of Spanish, first-generation immigrants (whose lan-
guage serves as the input to the heritage speakers), and L1 speakers from different age cohorts in
Mexico, and found that child and adult heritage speakers of Spanish omit a personal on a regular
basis. The situation is complicated, however, by the fact that the same pattern of DOM omission has
been observed in the language of first-generation immigrants. Leaving aside the possible reasons for
such a change, it appears that heritage Spanish has picked up an incipient change in the baseline and
generalized it, yielding scenario (1b).

Reanalysis in adult language

Reanalysis in (adult) HL may produce opposing results: attrition in the HL versus the baseline and
child language (scenario (1c), discussed in the following subsection) or realignment of the HL with
the baseline against child language (scenario (1d), discussed in the subsection “Getting better with age?”). Theoretically, HL reanalysis may also lead to a situation in which all three varieties (the baseline, the language of child bilinguals, and HL) differ; I am not aware of examples instantiating such a situation. The crucial observation underlying all these scenarios is the fact that agents of change—heritage speakers—follow a number of universal principles that are not unique to HL.

Attrition?

In this section, I consider a scenario in which bilingual children (future heritage speakers) pattern with bilingual native-speaker adults with respect to a given property, whereas heritage speakers differ from both groups (scenario (1c)).

In this scenario, a particular structure or language property is initially acquired by bilingual children, but undergoes attrition (in the proper sense of the word) in adulthood. Attrition of this kind is observed in the treatment of A-bar dependencies in adult heritage Russian (Polinsky, 2011).

As a subtype of A-bar dependency, relative clauses offer a window onto structural preferences in language. Consider the subject-extracted relative clause in (7a) and the object-extracted relative clause in (7b). In both cases, the gap and the relative pronoun reference the subject of the matrix clause, the reporter.

(7)  a. The reporter, [who, t, harshly attacked the senator] admitted the error.
    b. The reporter, [who, the senator harshly attacked t] admitted the error.

It is a well-known empirical generalization that, if a language can relativize at a given position in the Accessibility Hierarchy (Keenan and Comrie, 1977), it can also relativize at every position above it.

(8)  subject > direct-object > indirect-object > oblique-object > …

Development of relative clauses more or less follows the scale in (8); children tend to acquire subject and object relatives first, and to acquire them relatively early (Hamburger & Crain, 1982).

Assuming that relative clauses are linguistically robust and acquired relatively early, we might expect them to be resistant to change in heritage grammars. If relativization does not undergo the same processes of attrition that other areas of heritage grammars do — that is, if heritage speakers and native speakers perform equally well in comprehending and producing relative clauses — this finding would support the notion that competence in relativization is independent of quantity or quality of exposure. If, however, heritage speakers do diverge from native speakers in their performance with regard to relative clauses, then the observed differences may inform the trajectory of heritage grammars.

Like English, Russian allows relativization at any point in (8), as the following examples demonstrate.

(9)  a. deti, [kotor-ye ti polučili podarki ot babuški] children.NOM.PL REL-NOM.PL received gifts.ACC.PL from grandma.GEN ‘(the) children that/who received gifts from Grandma’
    b. podarki, [kotor-ye deti polučili ti ot gifts.NOM.PL REL-ACC.PL children.NOM.PL received from babuški] grandma.GEN ‘(the) gifts that the children received from Grandma’
The similarity between the two systems makes the production of Russian relative clauses by English-dominant heritage speakers a particularly interesting phenomenon: since both languages have parallel restrictions on relative clauses, the likelihood of transfer effects is low. Where the two languages differ, however, is in their approach to scrambling. Unlike English, Russian has widespread scrambling (Bailyn, 2004), from which relative clauses are not exempt: the non-extracted noun phrase in a relative clause may occur either pre-verbally (10a) or post-verbally (10b).

(10) a. deti, [kotor-ye ti polučili podarki]  
   children,REL-NOM.PL received gifts,ACC.PL 

b. det, [kotor-ye ti podarki polučili]  
   children,REL-NOM.PL gifts,ACC.PL received 

‘(the) children that received (the) gifts’

Polinsky (2011) used a picture-matching task to investigate the following questions: (i) does heritage Russian allow for the same expressivity in relativization structures observed in the baseline language? (ii) does the presence of scrambling in the baseline Russian grammar (but not in the dominant English grammar) affect the grammar of relative clauses in the corresponding HL?

To answer these questions, speakers were presented with relativization structures that crossed two types of relative clause gaps (subject vs object) with two orders of arguments in the relative clause (noun-verb vs verb-noun). Given the well-established preference for subject gaps in relativization, subject-extracted relative clauses were predicted to be easier for heritage speakers to process than object-extracted structures. In addition to this universal tendency to favor subject gaps, it was also expected that the speakers would show effects of their dominant language. Specifically, Polinsky (2011) predicted that correspondences of surface order between certain Russian and English constructions would lead to differences between heritage speakers and native speakers in the processing of scrambling within relativization structures.

The participants included two paired groups of speakers: age-matched Russian-monolingual and English-bilingual children (average age six) and age-matched Russian-monolingual and heritage-speaking adults. The participants were asked to choose between two pictures in response to an auditory question containing a relative clause. The stimuli all featured reversible actions: for example, ‘chasing’ in Figure 3. The questions varied in terms of subject vs object extraction and scrambled vs unscrambled argument order in the relative clause.

The monolingual speakers, both adults (n=26) and children (n=15), found the task almost trivial, choosing the correct picture with ceiling-level accuracy. Bilingual children (n=21; average age six) performed equally well. Adult heritage speakers (n=29), however, exhibited a stark asymmetry in their performance between subject- and object-extracted relative clauses. These participants performed quite well in subject-extracted identification tasks, but performed at chance in response to questions involving object extraction. In other words, they reanalyzed the Accessibility Hierarchy in (8) as the contrast between subject gaps and gaps in all other positions.

Since both monolingual and bilingual children performed essentially at ceiling, the reanalysis in the adult heritage grammar cannot be attributed to a fossilized child-language variant. Rather, these findings suggest that over their lifespan, the heritage speakers’ competence with respect to relative clauses has degraded, leaving the adults speakers capable of comprehending only the easier subject-extracted relative clauses. Thus, it appears that relativization is a less robust area of linguistic competence than we tend to think: with reduced input and insufficient maintenance, competence in
this area can decline. The observed attrition undoubtedly relates to a loss of morphological knowledge. If the heritage speakers did not process the nominative vs accusative distinction, no further clues were available to clarify the orientation of the relative clause; they simply observed a clause with a transitive verb, a single overt argument, and a gap. In the absence of morphological cues, the default preference of these speakers was to treat such clauses as subject-extracted relatives.

At first blush, it is natural to attribute the observed attrition to pressure from the dominant language, in this case English. However, if English were to blame, then relative clauses without scrambling (i.e. those mapping directly onto the word order of the analogous English sentences) should have been easier for heritage speakers to process. The results of the study showed that this was not the case: heritage speakers performed equally well in identifying both scrambled and unscrambled subject-extracted configurations, and equally poorly in identifying both scrambled and unscrambled object-extracted configurations. Thus, we must conclude that attrition in Russian heritage grammar, at least in the domain of relative clauses, is not the result of English transfer effects. Instead, it is most likely a result of restructuring in the absence of sufficient maintenance or input.

Getting better with age?

Our final scenario, (1d), inverts the pattern just discussed: here, the adult HL patterns with the baseline, while the child language differs.

U-shaped learning — where a child makes developmental errors and then recovers, eventually developing an advanced, target-like knowledge — is a clear instance of (1d). While we will ultimately need to compare U-shaped learning in monolingual and bilingual child language acquisition, such a comparison is beyond the scope of the current discussion. Clear cases of scenario (1d) in the context explored in this study — comparisons between the performance of later child bilinguals and their adult counterparts — have yet to be identified. The case I will discuss in this section, which to my mind offers the best illustration of this scenario to date, also reflects innovations acquired from the baseline language, and thus bears some resemblance to scenario (1b), discussed above.

The development I will discuss here pertains to Spanish psych-verbs such as gustar ‘like’, which Pascual y Cabo investigates in his dissertation (2013). Psych-verbs are rarely uniform within a given language (e.g. Belletti & Rizzi, 1988; Landau, 2010); in Spanish, they fall into at least three
classes. Pascual y Cabo concentrates on Spanish class III psych-verbs, among which _gustar_ is the most common. This class of verbs has several properties relevant to the discussion below. First, for all these verbs, the experiencer precedes the verb (_Katherine_ in (11a)), but it is the post-verbal theme ( _los kiwis_ in (11a)) that serves as the syntactic subject of the sentence. Second, verbs of this type necessarily receive a stative reading. As statives, they resist passivization, (11b).

(11) a.  A Katherine le gustan
    |PRP    | 3SG.DAT.CL
    |los kiwi-s. | like.PRESENT.3PL
    ‘Katherine likes kiwis.’

b. *Los kiwi-s son gustad-os
    |the kiwi-PL | be.PRESENT.3PL
    |por Katherine). | like.PARTICIPLE-M.PL

The argument structure of stative psych-verbs has been the subject of much discussion in the literature on L1 and L2 acquisition of Spanish. Gómez Soler (2011) analyzes spontaneous child speech and shows that children start producing target-like _gustar_ constructions quite early, at approximately one year and ten months. In a subsequent comprehension study, Gómez Soler (2012) shows that children as young as three are able to comprehend this class of psych-verbs, but that their performance tends to vary with the specific verb used. Specifically, children perform remarkably well (at 79% accuracy) with _gustar_, but at chance (52%) with less common stative-only psych-verbs such as _faltar_ ‘lack’. As is so often the case, different tasks yield different findings; Torrens, Escobar, and Wexler (2006) report that children do not have adult-like comprehension of these psych-verb constructions until the age of six. Although the exact timing of the acquisition of stative-only psych-verbs in Spanish is still up for debate, the evidence at hand supports the modest claim that these verbs are acquired _later_ by monolingual Spanish children than agentive predicates.

Assuming that the atypical argument structure of psych-verbs such as _gustar_ contributes to their encumbered acquisition, Pascual y Cabo (2013) examines English-dominant heritage speakers of Spanish, who often lack formal schooling in their home language. Based on a comprehension study of class III psych-verbs in Heritage Spanish, Pascual y Cabo hypothesizes that heritage speakers may reanalyze the psych-verb _gustar_ as optionally agentive, rather than strictly stative. If this reanalysis has taken place in the grammar of heritage speakers (turning class III psych-verbs into class II psych-verbs), we should find evidence of _gustar_ appearing in passive constructions such as example (11b) above. The results of an acceptability judgment task confirm this prediction: as expected, native speakers in this study judge passive constructions of stative-only psych-verbs as categorically unacceptable, while heritage speakers at varying levels of proficiency are more likely to accept these constructions. Pascual y Cabo takes this result as confirmation of his hypothesis that heritage speakers (at least sometimes) reanalyze class III psych-verbs as agentive.

In the next stage of his research, Pascual y Cabo considers the possible trajectory of this reanalysis, comparing the performance of the original population of heritage speakers to child bilingual and monolingual speakers. He posits that, if the reanalysis of _gustar_ were due to attrition, then at some earlier point in HL development, we should find more target-like behavior, which is later lost. Concretely, we should expect monolingual (and bilingual) children to perform better than heritage adults at judging passive _gustar_ constructions unacceptable. However, this prediction is not borne out: both monolingual and bilingual children perform worse at this task than the (adult) heritage speakers. The fact that heritage speakers behave more like adult native speakers than child
monolingual speakers suggests that heritage speakers actually improve their mastery of these psych-verbs over time.

Following Lightfoot (1991, 1999, 2012), Pascual y Cabo argues that ‘superficial performance innovations provided in the input from the immigrant generation contribute to the changes in H[eritage] S[peakers’ ] grammars’ (Pascual y Cabo, 2013, p. 131). In other words, he attributes the heritage speakers’ poor performance to flawed input from the original source: L1 monolingual immigrants. These immigrant parents sometimes produce target-like gustar constructions, and sometimes do not; the next-generation immigrant speakers (i.e. HL learners) receive this already-non-standard input from their parents and create an ambiguous mental representation of the syntax of the constructions at issue. This ambiguity forces heritage speakers to (economically) reanalyze the psych-verb construction to allow the otherwise off-limits agentive structure. Note that an additional reanalysis trigger could come from English. All the Spanish heritage speakers surveyed by Pascual y Cabo were dominant in English, which lacks quirky subjects such as the class III psych-verbs. Thus, structural transfer from English may emerge as an additional factor forcing reanalysis in the heritage grammar.

Conclusions and predictions for future work

The speech of childhood bilinguals offers several important insights for HL research. On the one hand, child bilingual speakers, unlike true heritage speakers, are still receiving daily input in the HL; on the other hand, they are receiving that input under the same sociolinguistic conditions as are reported for adult heritage speakers: absence of bilingual education, rapid switch to the dominant language of their society, and decreasing input in the home language as they spend more time in school. Information about the baseline language – the language that serves as input to bilinguals during their acquisition of the home language – is a critical ingredient in any study that seeks to understand childhood bilingualism and ensuing HL competence. Comparing the language of childhood bilinguals exclusively to the language of monolingual adults may lead to missed generalizations.

Comparisons between L1 and L2 under second language acquisition are inspired by the need to understand how language learners deploy universal grammatical principles in adapting their L1 knowledge to the L2. Two questions raised by the discussion presented here are: (i) Which universal grammatical principles are of primary importance in language acquisition and attrition? and (ii) Which configurations of a given language survive ‘under stress,’ that is, in the context of restricted or noisy input? The HL development scenarios I have explored in this paper provide windows onto general properties of language design. Heritage speakers’ reliance on universal language principles speaks directly to Plato’s problem in language: data from heritage speech reveals how grammar is acquired under conditions of reduced input and usage.

Altogether, this paper has identified four major possible vectors of correspondence among the language of baseline speakers, childhood bilinguals, and heritage speakers. These possibilities include cases where all three language varieties represent a given linguistic property in a similar way (1a) and cases where two of the three varieties show a parallel representation, but the third variety differs (1b)-(1d). A fifth scenario, in which all three varieties differ in the representation of a given property, is theoretically possible, but (to my knowledge) is not attested.

All three varieties will show the same treatment of a given property (BIL = CHL = AHL) when the property in question is highly robust in the language and is acquired by learners without difficulty. Identifying commonalities in the language of baseline speakers, childhood bilinguals, and heritage speakers can offer important insight into fundamental configurations or properties of natural language. Note that the example I offer in this paper – the case of Russian nominal forms in numerical
expressions (count forms) – is morphologically quite involved. On a broad level, this suggests that general simplicity of encoding is not necessarily a defining characteristic of a robust pattern.

A situation in which a particular structure in the baseline differs from the structure in child and adult heritage speech (BIL ≠ CHL = AHL) may arise when incipient structural changes taking place in the baseline become amplified in the language of next-generation bilinguals. Patterns of this kind may therefore serve as a window into changes in the baseline. Such a scenario may also arise when a given structure is rare, confined to a specific register, and/or reinforced through literacy. Child bilinguals may simply not have access to such a structure, and adult heritage speakers, whose literacy is often subpar, may not be exposed to it either.

Finally, two mirror-image possibilities exist in which one heritage-speaker group (childhood or adult speakers) patterns with the baseline group, while the other does not. When a structure is acquired by bilingual children faithfully, but undergoes later reanalysis in the adult HL (BIL = CHL ≠ AHL), this process of attrition can expose universal principles of language structure. We saw in this paper that, despite expectations based on language transfer, the reanalysis of relative clauses by Russian heritage speakers consistently follows the subject preference in the interpretation of gaps; this finding supports the robustness of a general subject-extraction preference cross-linguistically. Conversely, a structure that is not fully learned by child speakers may be reanalyzed by adult heritage speakers following general principles, thus bringing the adult heritage representation closer to that of the baseline (BIL = AHL ≠ CHL). Table 1 below summarizes the different scenarios considered in this paper.

<table>
<thead>
<tr>
<th>Table 1. Correspondences between baseline, language of child bilinguals, and heritage language.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correspondence type</td>
</tr>
<tr>
<td>Baseline = language of child bilinguals = adult heritage language</td>
</tr>
<tr>
<td>Baseline ≠ language of child bilinguals = adult heritage language</td>
</tr>
<tr>
<td>Baseline ≠ language of child bilinguals = adult heritage language</td>
</tr>
<tr>
<td>Baseline = language of child bilinguals ≠ adult heritage language</td>
</tr>
<tr>
<td>Baseline ≠ language of child bilinguals ≠ adult heritage language</td>
</tr>
<tr>
<td>Baseline = adult heritage language ≠ language of child bilinguals</td>
</tr>
</tbody>
</table>

Identification of these possible scenarios allows us to track and categorize developmental trajectories in the language of bilingual children and adults. The patterns of correspondences can also serve as tools in mapping out robust and vulnerable units of language design.

The next major question to be addressed is more challenging: is it possible to predict what scenario will arise for a given language property within a given bilingual population? The answer to this question is far from clear, but I would like to offer some initial thoughts.

First, consider age-of-acquisition effects. Cognitive development, proficiency at a given age, biological age, and socioeconomic status of the learner together constitute significant factors in
language development. In monolingual speakers, acquisition is ultimately successful, but it may take some time and is prone to errors (Montrul, 2008). In the HL of child learners, a relatively late age of acquisition for certain phenomena or structures may mean that those structures are never acquired perfectly and remain weak in the adult grammar. We can thus anticipate that phenomena or structures associated with a relatively late age of acquisition will be vulnerable under unbalanced bilingualism. However, it remains impossible to predict at this point whether such vulnerable domains will undergo transfer, loss, or reanalysis. On the face of it, all these possibilities should all be available and should be examined on a case-by-case basis.

Age of acquisition is related to amount of exposure/input, although the two phenomena should be kept separate (Johnson & Newport, 1989). Indeed, significant exposure to a language does not guarantee that all the received material will be internalized by the learner; in other words, the intake (what is internalized and processed by the learner) and the input do not necessarily match (see Gagliardi, 2012, for a recent discussion of this issue in L1 acquisition). The role of exposure in determining intake is significant in monolingual acquisition; when two languages are in competition, the exposure factor may be even more salient. The precise interaction of input and intake in early HL acquisition is still to be determined, but in my opinion, this interaction is one of the crucial factors affecting the outcome of bilingual acquisition.

Bilingual acquisition data are not available for all languages. In particular, exemplary studies such as Kupisch (2006), which illustrates a much-needed balance between the input in the dominant language and HL, are rare. As a result, linguists often rely on monolingual L1 data when estimating the amount of input and exposure received by HL learners. Empirical gaps of this sort serve as a reminder that we need more data on the input bilingual speakers receive in both their languages, at different stages of linguistic development. In addition, corpora of child-directed speech and child speech in bilingual contexts can also serve as a resource that would allow us to track incipient changes in the baseline, such as those observed for baseline Spanish by Pascual y Cabo (2013) or by Montrul and Sanchez-Walker (2013).

Age of acquisition and input both relate to what might be considered the ‘ecological’ aspects of language. Both factors have a direct bearing on the reorganization of language structure in vulnerable domains. Many factors can make a language domain vulnerable; here, I will highlight just two general trends. The first has to do with the difficulty of going from one module of grammar to another, a phenomenon captured under the rubric of the Interface Hypothesis (Sorace, 2011; Sorace & Filiaci, 2006). Assuming that grammatical operations that intersect two components of grammar increase cognitive load for speakers, we can anticipate that such operations will pose some difficulty both in initial acquisition and in ongoing maintenance. This area has so far been best investigated with respect to the interface between discourse and syntax (especially in the licensing of null pronominals), and we need to explore other interfaces (for example, between syntax and prosody) with similar vigor if we are to make more progress in understanding developmental trajectories in bilingual speech. The prediction is clear: if a phenomenon at an interface is challenging to a bilingual child learner, learning will deviate from the trajectory of the baseline, and this difference will persist in adult heritage speech. So far, this prediction has been confirmed in two notable cases: the persistent emergence of pro-drop in child and adult heritage speech (see the discussion of Spanish above) and the divergent use of topic marking by heritage speakers of Japanese and Korean (Laleko & Polinsky, 2016). The difficulty in the acquisition of these arguably interface-related phenomena is apparent in child learners and seems to linger on in adult heritage speakers. Whether such lingering effects can be completely ascribed to insufficient input, however, remains an open question.

The second structural trend that deserves special consideration has to do with low tolerance for ambiguity in language (Benmamoun et al., 2013b). This intolerance for ambiguity seems to be a general heuristic, leading to an across-the-board desire for one-to-one mapping between meaning and structure; assuming that lack of ambiguity allows heritage speakers to communicate more
effectively, we can predict that certain ambiguous structures in a language are likely to be reanalyzed by heritage speakers of that language. As an example, consider the well-known ambiguity between strict and sloppy readings in ellipsis, illustrated in the following English example.

(12) The linguist promoted his new theory and the logician did too.
(i) The linguist promoted his new theory and the logician promoted the linguist’s new theory. [STRICT READING]
(ii) The linguist, promoted his, new theory and the logician, promoted his, new theory. [SLOPPY READING]

The sloppy reading, in which the linguist and the logician each promoted their own separate theories, is attributed to the presence of a bound variable dependency. The strict reading, on the other hand, represents a simple coreference: the pronoun picks up as its antecedent the linguist and this referent is carried over into the unpronounced VP. Ambiguity under ellipsis has received significant attention in both theoretical linguistics and processing. Within processing, a number of researchers have shown that monolingual speakers prefer the sloppy (bound variable) reading and process it more quickly than the strict reading (Frazier & Clifton, 2000; Koornneef, 2008; Shapiro, Hestvik, Lesan, & Garcia, 2003; Vasić, 2006). For heritage speakers, one might then predict that they will generalize this processing preference and eliminate the ambiguity altogether, assigning only the sloppy reading to the elliptical structures (as in example 12). This is a testable prediction that can be explored in child and adult HL, with a general expectation that whatever ambiguity is allowed in the baseline will be reduced in the HL.

Acknowledgements
I am grateful to Suzanne Aalberse, Aafke Hulk and to anonymous reviewers for comments on this paper. All errors are my responsibility.

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

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by the NSF grant SBR-1619857.

Notes
1. Not all researchers consider a shift in dominance from the home language to the societal language to be an essential feature of HL (see Polinsky & Kagan, 2007; Cantone, Müller, Schmitz, & Kupisch, 2008, for a discussion of different approaches). The studies reviewed in this article all adopt this definition; for the sake of consistency, I consider “heritage speakers” to include specifically those speakers who have shifted substantially away from their home language to the societally dominant language.
2. Baseline variants may also preserve an older version of the standard language while speakers in the country of origin innovate, which may lead to the impression that HLs are more conservative.
3. As a reviewer quite rightly points out, it is certainly not the case that all linguistic structures are fully mastered by the age of three. However, most of the specific structures examined in the studies I discuss are acquired relatively early in L1 acquisition.
4. Abbreviations: ACC—accusative, AHL—adult heritage language, BIL—baseline input language, CBL—child bilinguals’ language, DAT—dative, HL—heritage language.
5. The entire numerical expression also carries its own case. The salient pattern for our purposes arises when the numerical expression is in the nominative or accusative form.

6. Other cases prone to substitution include the dative case and cases licensed by prepositions (Gagarina & Voeikova, 2009; Lepskaja, 1997; Polinsky, 2006).

7. The ratings of the indirect object condition, as in (3), confirm the observation that heritage Russian speakers often replace the dative with the accusative (Polinsky, 2006).

References


**Author biography**

Maria Polinsky is Professor of Linguistics, Associate Director of the Language Science Center, and Director of the Guatemala Field Station at the University of Maryland, College Park. Her research focuses on cross-linguistic variation in syntactic phenomena and on differences between monolingual and bilingual populations, heritage speakers in particular.
Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers

Tanja Kupisch
Universität Konstanz, Germany
UiT, the Arctic University of Norway, Norway

Jason Rothman
University of Reading, UK
UiT, the Arctic University of Norway, Norway

Abstract
This paper integrates research on child simultaneous bilingual acquisition more directly into the heritage language acquisition literature. The child simultaneous bilingual literature mostly focuses on development in childhood, whereas heritage speakers are often tested at an endstate in adulthood. However, insights from child simultaneous bilingual acquisition must be considered in heritage language acquisition theorizing precisely because many heritage speakers demonstrate the adult outcomes of child simultaneous bilingual acquisition. Data from child simultaneous bilingual acquisition raises serious questions for the construct of incomplete acquisition, a term broadly used in heritage language acquisition studies to describe almost any difference heritage speakers display from baseline controls (usually monolinguals). We offer an epistemological discussion related to incomplete acquisition, highlighting the descriptive and theoretical inaccuracy of the term. We focus our discussion on two of several possible causal factors that contribute to variable competence outcomes in adult heritage speakers: input and formal instruction in the heritage language. We conclude by offering alternative terminology for heritage speaker outcomes.

Keywords
Heritage language, heritage speaker bilinguals, simultaneous bilinguals, incomplete acquisition, input, schooling

Corresponding author:
Jason Rothman, University of Reading, UK.
Email: j.rothman@reading.ac.uk
Introduction

Prior to the early 1960s, the commonly accepted view of bilingualism—held even by many psychologists and linguists—was that it was subtractive for normal development. Bilingualism was considered a ‘bad thing’, at least with regards to cognitive and linguistic development. Bilingualism was thought to confuse the child, presumably due to some type of cognitive overload related to the complex task of acquiring, maintaining and balancing two languages. As a result, it was often held that exposing children to more than one language would likely result in so-called imperfect acquisition of both languages. In the extreme, it was believed that bilingualism conferred disadvantages for learning more generally. Many studies between 1920 and 1960 supported this general view (e.g. Saer, 1923), indicating that bilingualism correlated with lower IQ, cognitive deficiencies and even mental retardation.

It is not a coincidence that most of the above-mentioned research emerged from monolingual societies, particularly the United States and the UK. On the positive side, such research was fuelled by an earnest interest in understanding possible consequences of bilingualism during a period of sharp increase in immigration in these same societies. On the negative side, it was indiscriminately accepted as a result of a political climate that needed to account for immigrant children’s scholastic underperformance.

Given what we know about bilingualism 50 years on, it seems inconceivable that the prevailing view once aligned bilingualism with being a burden for general development. Consider the fact that more than 50% of the world’s population found itself then—as it does today—in naturalistic bilingual situations, that is, where more than one language prevails at the societal level. In such environments, bilingualism is the default state. If it were true that bilingualism was harmful for general development, bilingual societies would place their citizens at significant risk relative to monolingual ones. So, the question is how to reconcile the claims from the early 20th-century literature with the fact that bilingualism as a default state has existed for millennia with great success in multilingual societies, in countries such as Luxemburg, Switzerland, India and South Africa.

It was suggested that developmental disadvantages did not apply to bilingualism per se, but rather reflected something about bilingual development outside of bilingual societies. But can it really be the case that bilingualism is either neutral (neither facilitates nor subtracts) or advantageous for development in multilingual societies, while being subtractive in monolingual societies? Although it is true that bilingualism at the societal level has positive consequences for individual bilinguals’ language development and ultimate acquisition outcomes, it is not clear how having access to a (significantly) larger bilingual community could ward off the detrimental outcomes said to be caused by the ‘confusion’ inherent to bilingualism. The task of the bilingual child in and outside of a bilingual country is one and the same. If confusion obtains because of having two languages in a single mind, it applies to all bilingual scenarios. Societies that are themselves bilingual likely provide children with greater quantities and richer qualities of input relative to what bilinguals receive in monolingual societies. We would expect then that environment has much to bear on linguistic development and outcomes, but would not be a potential cause of mental retardation in one extreme and possible cognitive advantages in the other.

Seminal work by Peal and Lambert (1962) effectively changed the way in which bilingualism was viewed. They highlighted how previous studies did not control for deterministic intervening variables that obscured any meaningful conclusions about bilingualism. Prior to this exceedingly few studies imposed careful selection criteria for the inclusion of bilingual participants. Bilinguals with various degrees of proficiency, from various socio-economic statuses, across a range of...
language pairings, with diverse ages of onset and amounts of exposure, were haphazardly grouped together as if none of these variables mattered. To make matters worse, these heterogeneous groups were then compared to homogeneous monolinguals of considerable socioeconomic advantage. Peal and Lambert (1962) imposed meticulous criteria for participant selection and demonstrated that, once deterministic variables were properly controlled, their highly proficient bilinguals were not only equal to, but indeed outperformed appropriately matched monolingual counterparts for both measures of verbal and non-verbal intelligence.

Since the early 1960s virtually nobody has claimed that bilingualism is subtractive in any way, although the effects of bilingualism are still being discussed. According to some researchers, there are no qualitative differences between monolinguals and simultaneous bilinguals in terms of language development (Meisel, 2011). Research has further shown that bilingualism might confer some cognitive advantages to executive functions (e.g. Bialystok, 2009; Bialystok, Craik, & Luk, 2012). Others argue in favour of limited advantages depending on degree of language usage (Luk & Bialystok, 2013), but determining the extent of these advantages at present is difficult (Valian, 2015). Whether or not bilingualism results in any cognitive advantages hardly matters; what we know for sure today is that bilingualism is not a ‘bad’ thing by any measure. Knowing that bilingualism is not ‘confusing’ to the child is a dramatic change from the past. Yet despite the fact that this has been known for decades within the language sciences, cognitive science and psychology, the traces of injudicious claims made in the early part of the 20th century have not disappeared outside of academia. Today many childcare providers (from medical to educational), government agencies, educational institutions, individual practitioners (teachers, tutors, etc.), parents and more still operate under the idea that bilingualism creates disadvantages. The footprints of previously misguided research still have huge implications for policy and practice related to supporting bilingualism and bi-literacy, offering counsel and advice to parents, maintaining heritage languages, creating language policy and enacting language education.

To be sure, the notion that bilingualism can cause general mental incapability is no longer believed valid within academia. However, the notion that bilingualism, at least under certain conditions, can result in imperfect linguistic acquisition is still very much alive. It is our view that the latter notion is as misguided as the former at the conceptual, theoretical and practical levels. In this article, we will address the notion of so-called incomplete acquisition in the context of heritage language (HL) bilingualism (e.g. Montrul, 2008, 2016). Our position is that incomplete acquisition as a term to describe differences between monolingual controls and heritage speaker (HS) bilinguals is theoretically flawed and misleading, if not unintentionally insensitive. No one denies that the typical endstate grammars of HSs are, on a continuum, different from appropriately matched monolinguals. However, we will argue that different and incomplete, in this domain, are not potential synonyms. Moreover, we will introduce to this general discussion the literature on simultaneous child bilingualism (2L1), which, as a whole, adds a unique perspective and especially strong support against the construct of incomplete acquisition. We focus our discussion on two of several possible causal factors that contribute to variable competence outcomes in adult HSs as compared to monolinguals1: (i) qualitative input factors (e.g. Pascual y Cabo & Rothman, 2012; Rothman, 2007; Sorace, 2004) and (ii) general inaccessibility to formal education/literacy (e.g. Tsimpli, 2014) in the HL. To the extent that we succeed in highlighting the conceptual and theoretical imprecision of the term incomplete acquisition, we will argue that it is time for the field to replace it with more accurate and less potentially evaluative labels that capture not only the description of differences, but inch us towards explanatory adequacy of how and why these differences obtain (as has been attempted in recent years, e.g. Kupisch, 2013; Meisel, 2013; Pascual y Cabo & Rothman, 2012; Pires & Rothman, 2009; Putnam & Sánchez, 2013; Scontras, Fuchs, & Polinsky, 2015).
Setting the context: Who are heritage speakers and child bilinguals?

Given the topic of this article, it seems fitting to first define what we mean by the HS label. Reasonably, one might expect that there is one, unambiguous definition to which all researchers working on HSs ascribe. Under such a scenario, determining which bilinguals qualify as HSs would be straightforward. However, this is not clear in practice. Indeed, all HSs are bilingual by definition, but certainly not all bilinguals are HSs. Agreeing on characterizing parameters for HS-ness is non-trivial, especially because the term *heritage speaker* itself is relatively new to the field of language acquisition and still does not have ubiquitous use across researchers around the world. The term originated in North America and has over the past two decades been used there increasingly. In the past few years, the term has become more globally used, for example, in Europe and beyond. For researchers primarily interested in formal linguistic issues related to the process, development and ultimate attainment of HL grammars, it is crucial that one makes a difference *a priori* between those that are naturalistic bilingual acquirers/speakers of an HL and those that are current adult learners of a second language (L2) that is a language of familial heritage. While L2 HL learners, unlike typical novice L2 learners, might have different (higher) motivations, cultural connections, access to (family) native speakers outside the classroom and have had some limited exposure prior to the start of L2 learning, they are not early childhood naturalistic acquirers of the HL. Therefore, they are not native speakers of the HL in the same way as HS bilinguals (Rothman & Treffers-Daller, 2014). Interesting to study in their own right as a potential subset of L2 learners, L2 HL learners are different from HS bilinguals. Herein, we follow Rothman’s (2009) definition of what an HL is. This definition is relatively uncontroversial, according in several domains with other available definitions (e.g. Benmamoun, Montrul, & Polinsky, 2013; Montrul, 2008, 2016; Polinsky & Kagan, 2007), but one crucial difference consists in its purposeful avoidance of the term *incomplete acquisition*.

A language qualifies as a *heritage language* if it is a language spoken at home or otherwise readily available to young children, and crucially this language is not a dominant language of the larger (national) society […] the *heritage language* is acquired on the basis of an interaction with naturalistic input and whatever in-born linguistic mechanisms are at play in any instance of child language acquisition. Differently [from monolingual acquisition], there is the possibility that quantitative and qualitative differences in *heritage language* input, influence of the societal majority language and differences in literacy and formal education can result in what on the surface seems to be arrested development of the *heritage language* or attrition in adult bilingual knowledge (Rothman, 2009, p. 156).

Importantly, this definition differs from others, according to which the HL is a minority language, acquired naturalistically but for a speaker whose first language ‘did not develop fully at age-appropriate levels’ (e.g. Benmamoun *et al*., 2013, p. 133, based on Valdés, 2000) or ‘often does not reach native-like attainment during adulthood’ (Benmamoun *et al*., 2013, p. 133). Rothman’s definition does not imply anything with regard to potential acquisition outcomes. An HS is a native-speaker bilingual of a minority language spoken at home and either also a native speaker (in the case of 2L1) or a child L2 learner of the majority language of the society in which she/he lives and is educated. Under either scenario, it is virtually inevitable that the HS will wind up being dominant in the societal majority language. However, it seems incongruous to talk about HSs as not reaching ‘native-like attainment during adulthood’ precisely because they are native speakers by definition. Accepting that HSs are a subset of native speakers (Rothman & Treffers-Daller, 2014) presupposes that one defines a ‘native speaker’ with respect to age of onset in a naturalistic context, rather than by some dubious proficiency levels.
that monolinguals supposedly have and/or holding the view that dominance is a necessary deterministic factor. Confusing nativeness with dominance and/or having convergence on a standard monolingual variety as a benchmark only serves to highlight the comparative fallacy of monolingual vs native bilingual comparisons. Such a practice also has serious implications for determining ‘nativeness’ for subgroups of monolinguals as well as adults who migrate and shift dominance towards an L2 much later in life. HSs do reach native attainment in adulthood because they are native speakers by definition. It may just happen to be the case that the HS native outcomes differ from other sets of natives in ways potentially similar to subgroups of monolinguals with limited exposure and training in the standard language and adult bilinguals under first language (L1) attrition.

A child bilingual (2L1) is typically described as an individual exposed to two languages from birth or very early on in childhood (before three or four years of age, see Meisel, 2011). Research within 2L1 acquisition is often based on *longitudinal* case studies of developing children between the ages of 18 months and five to six years where children are observed in both languages. This research, which has mostly shown qualitative similarities between 2L1 and monolingual development alongside various degrees of protracted developmental delay or acceleration, has largely been carried out in Western Europe or Canada (e.g. De Houwer, 1990; Hulk & Müller, 2000; Genesee, Nicoladis, & Paradis, 1995; Meisel, 1990, 1994a,b; Paradis & Genesee, 1996, among many others). In contrast, HS research normally looks at acquisition outcomes, it is often based on *cross-sectional* experimental studies with young adults who have been observed only in their minority language. Bringing the trends of available studies together, the picture that emerges shows HSs to be communicatively competent, fluent speakers of their HL. However, HSs display endstate variation in their knowledge and use of the HL that does not normally define other sets of native speakers, specifically monolinguals. Such differences between HL grammars and the monolingual standard can transcend virtually all aspects of the language system, from morphosyntax to phonetics/phonology and especially the lexicon (see Montrul, 2008, 2016). As they are observable and replicated, highlighting such differences should not be controversial. What is apparently contentious is hypothesizing how difference obtains and what it should mean.

To our knowledge, nobody has ever explicitly stated that child 2L1 learners are not HSs — in fact, Benmamoun *et al.* (2013, p. 133) explicitly include them — and clearly Rothman’s (2009) definition would very much include them too. It just so happens that most of the existent research examining HS bilingual competence has examined HS adults (but see Flores, Santos, Marqes, & Jesus, under review; Montrul & Potowski, 2007; Pascual y Cabo, 2013; Polinsky, 2011; Rodina & Westergaard, 2015). This means that much HS research has, in fact, examined 2L1 acquisition, but at a very different stage of development than the majority of research falling unambiguously under the 2L1 label. Whereas historically the 2L1 literature has focused almost exclusively on developing bilingual grammars in childhood, HS studies have focused on adult endstate grammatical knowledge of either child 2L1 or child L2 speakers. The relevance of highlighting this is at least two-fold. First, we wish to bring to the attention of researchers who work on child bilingualism the idea that HSs are simply the adult outcome versions of the children they typically study. Second, we wish to highlight how bringing what is known from the child bilingualism literature into the domain of HS studies raises serious questions about claims made in the adult HS literature. We take it as a given that 2L1 children who grow up outside bilingual communities are HS bilinguals. The main difference between these two groups is terminological and differentiated primarily by age.

To the extent that it is fair to conceive of child 2L1ers as a subtype of HSs, the fact that the majority of 2L1 research concludes that 2L1 development is not qualitatively different from monolinguals’ complicates the observation of ubiquitous qualitative differences in the endstate grammars of HSs compared to monolinguals. By bringing the 2L1 literature to bear on HS theoretical
epistemology, we hope to highlight that the explanation for differences in HS outcomes requires more than conjecture about specifically what happens between the timeframe of six to eight years and young adulthood, the former being the age 2L1 studies usually stop and the latter the age where HS studies typically pick up.

**Studies of ‘successful’ heritage speakers**

In this section, we summarize studies of adult HSs in Germany, comparing two populations, HSs of Italian and HSs of French. The latter perform monolingual-like, in spite of their variable and quantitatively reduced input throughout the lifespan. The eight studies reported were all carried out as part of the project E11 at the Research Centre of Multilingualism in Hamburg, precisely where in the late 1980s and 1990s many of the early systematic studies with 2L1 children were carried out, based on projects such as Deutsch und Französisch–Doppelter Erstspracherwerb (DUFDE), Baskisch und Deutsch–Doppelter Erstspracherwerb and others (see Schmidt & Wörner, 2012, for an overview). The aforementioned projects traced 2L1 bilinguals in their two languages over several years, finding evidence for separate language development and qualitative similarities with monolingual children, especially in morphosyntax (see De Houwer, 1995, for an overview). More than two decades later, E11 looked at equivalent populations (German-French, German-Italian), but this time during **adulthood**. In what follows, we examine the motivation behind the selection of participants in E11, then summarize a number of studies.

The research in E11 stands out among HS studies in restricting the data collection to bilinguals with simultaneous exposure to two languages from birth. This is most easily guaranteed in speakers who grew up in binational families, where parents have different native languages which they consistently use with their children. The reason for including only bilinguals with exposure to both languages from birth was to exclude age of onset (AoO) in the majority language as a potential confound. Now how is AoO in the majority language relevant to the acquisition of the HL, which is acquired from birth? With simultaneous exposure from birth HL and majority language are always in contact, including during the very early years, which are generally deemed to be crucial for language acquisition. In the case of successive acquisition, i.e. when AoO in the HL precedes AoO in the majority language, the HL has more time to develop independently. Thus, the acquisition of some phenomena in the HL may already be completed before exposure to the majority language begins. In this respect, simultaneous bilinguals from birth might be seen as facing the hardest task, since their HL evolves under the potential influence of the majority language at all times. There is indeed evidence that the earlier presence of the majority language at home (e.g. through mixed marriages) may have a negative impact on the development of certain domains of the HL (see Rodina & Westergaard, 2015, for children and Van Suchtelen, 2014, for adult HS), suggesting that the two populations, simultaneous and sequential bilinguals, should be studied separately.

Moreover, E11 used predominantly bilingual rather than monolingual controls. For example, instead of comparing French/Italian HSs to French/Italian monolinguals, the controls were HSs of German from France/Italy who had acquired French/Italian and German from birth, i.e. the mirror image of French/Italian HSs in Germany. Majority-language speaking bilinguals are better controls than monolinguals because the (same) two languages are always in contact during acquisition, so that the possibility of mutual language influence is constantly given. Moreover, they have the same AoO in their two languages, and are thus preferable to L2 learners with the same language combinations, who often served as groups of comparison in previous research (e.g. Montrul, 2008). Coincidentally, most bilingual speakers participating in the studies also had high school degrees (corresponding to German ‘Abitur’).
In summary, E11 investigated relatively homogeneous populations: bilinguals with the same AoO in their two languages and a relatively high socio-economic status. In anticipation of the studies summarized below, it is worth noticing that most participants in the German-French group even attended the same school, where the HL was not the target of instruction but the medium.

The target phenomena were gender marking, adjective placement, determiner use in generic expressions and voice onset time. All phenomena are sufficiently different between French/Italian and German, so that cross-linguistic influence (CLI) is likely to occur. In fact, all these phenomena had been studied previously in developing bilinguals and were shown to be vulnerable to CLI.

**Heritage speakers of French attending French schools in Germany**

Most speakers who eventually took part in the study, coincidentally, had attended the Lycée Français de Hambourg, a school that is part of a network of Francophone schools outside of France that follow the curriculum of the French ministry of education. Besides French, the school attempts to integrate German language and culture and allows for the simultaneous attainment of the French Baccalauréat and the German Abitur (‘Abibac’), facilitating university access in both countries. Despite attending this school, most of the participants felt more at ease using German than using French and it was typical for them to speak German outside school, e.g. even when talking to each other in the school yard.

The first study (Kupisch, Akpinar, & Stöhr, 2013) was concerned with the acquisition of gender in French. Gender is an early acquired property in German-French bilinguals, with children choosing articles correctly in more than 90% of all contexts by the age of three. The acquisition of gender may be protracted in unbalanced bilinguals but generally follows the patterns of monolingual development in qualitative terms, i.e. regarding the type of errors made. German and French both have grammatical gender, but German has one additional gender and gender assignment and agreement patterns differ across the languages, creating a high potential for errors and CLI. The speakers were tested via an acceptability judgment task (AJT) and in elicited production. In both cases, particularly challenging conditions were included, e.g. gender agreement with adjectives whose equivalents in German are not gender marked, or nouns representing conflicting gender assignment rules, e.g. the noun *sentinelle* ‘nightguard’ was contextualized as a man despite being grammatically feminine (the ending –elle is associated with feminine gender). Results showed successful assignment and agreement with over 95% accuracy. Minor disadvantages were observed in gender assignment as compared to agreement, especially with nouns that are infrequent and/or represent conflicting assignment rules. This seems to indicate disadvantages with linguistic information that has to be memorized (i.e. intrinsic lexical information that cannot be derived from any productive rule) as compared to rule-based linguistic knowledge, as relevant for syntactic gender agreement and rule-based gender assignment.

The second study (Kupisch et al., 2014), was concerned with adjective placement. In this domain too, French differs from German in having two possible positions for adjectives. Some adjectives can be used in both positions, often with a difference in meaning (*un grand homme* ‘a famous man’ vs *un homme grand* ‘a tall man’). German, like English, only has prenominal adjectives. It is known that simultaneously bilingual children acquiring a Romance language and a Germanic language use both positions from early on but often overuse the prenominal position, which is present in both languages, especially if their Germanic language is dominant. The speakers in Kupisch et al. (2014) were tested via an AJT, again, including particularly challenging conditions. For some items, only the context gave a clue to the correct adjective position; in other cases, the adjective was part of a fixed expression and its position could not be derived by any rule or
even contradicted existing rules. On average, participants performed with 90% accuracy. Speakers were highly accurate when the position of the adjective was dependent on contextual information. By contrast, fixed expressions were relatively error prone. Cases of incorrect use mostly resulted in overusing postnominal adjectives. The results of this study too suggest that when proficient bilinguals perform differently from monolinguals, this may be due to lexical gaps rather than CLI from the dominant language.

The third study was concerned with the use of articles in generic plural determiner phrases (DPs, Barton, 2015). Generic DPs generalize over a class of individuals or objects, e.g. Sunflowers are yellow. In French, these must be formed with a definite article, while the absence of the article leads to ungrammaticality (see 1(a)). In German, article use is optional (see 1(b)).

1. (a) Fr. *(Les) tournesols sont jaunes. 
   (b) Ge. (Die) Sonnenblumen sind gelb
   the sunflowers are yellow
   ‘Sunflowers are yellow.’

Crucially, bare nominals are ungrammatical in French, while being the standard norm in German. As with adjective placement, the languages overlap partially, but this time it is French that allows only one type of syntactic structure (article+noun), while German allows two structures (article+noun, bare noun). As for their semantics, definite marked French DPs are ambiguous between generic reference and reference to a specific set of individuals or objects. For instance, Les chats sont intelligents (literally, the cats are intelligent) can refer to cats in general or a specific group of cats. German, like English, disambiguates between the two readings by using bare nouns; e.g. Cats (in general) are smart vs The cats (we just saw) are smart, although, as mentioned before, in some varieties of German the definite article is used for generic reference as well. Barton (2015) studied acceptance of bare nominals in French under the potential influence of German. In an AJT, bilingual French speakers rejected bare nouns more than 90% of the time. Furthermore, in a truth value judgment task which assessed the preferred reading of definite plural nominals (e.g. French Les tournesols sont jaunes, German Die Sonnenblumen sind gelb), the bilinguals preferred a generic reading in French and a specific reading in German. In summary, the HSs of French performed on par with monolingual speakers in accepting and interpreting plural subject nominals. With respect to this phenomenon, the result is particularly remarkable since previous studies have reported late acquisition and CLI for developing bilinguals acquiring a Germanic and a Romance language (Kupisch & Pierantozzi, 2010; Serratrice, Sorace, Filiaci, & Baldo, 2009) and incomplete acquisition in adults (Montrul & Ionin, 2012).

The fourth study (Lein, Kupisch & van de Weijer, forthcoming) looked at voice onset time (VOT). VOT is the most salient cue differentiating language-specific realizations of voiced (/b, d, g/) and voiceless (/p, t, k/) plosives. It refers to the interval between the release of the stop and the onset of voicing. In French and German, voiced and voiceless stops are associated with different types of VOT. Lein et al.’s study focussed on voiceless /k/, which in French exhibits a short voicing lag (VOTs ranging between 30-50 milliseconds (ms) across studies) and in German a long voicing lag (VOTs ranging between 35-80 ms across studies). Thus, German voiceless stops have longer VOTs than French ones. In acquisition, short lag is comparatively unproblematic and acquired early while long lag time is relatively more marked. Early bilingual children acquiring languages with different VOT realizations may show delayed acquisition of target values (Kehoe, Lleó, & Rakow, 2004). Lein et al. investigated VOT productions of seven HSs in spontaneous conversations of 20-30 minutes. Their French VOTs were compared to those of French-dominant bilinguals from France as well as their VOT productions when speaking German. Their average VOTs in French were 37.2 ms long, compared to 35.6 ms in French-dominant bilinguals from France. The difference was not significant.
Furthermore, all HSs had clearly distinct VOT ranges in their two languages (mean for German 50.8 ms). Overall, the HSs performed within the range of monolinguals and did not differ significantly from their bilingual controls who spoke German as their HL and French as their dominant language.

Comparison with heritage speakers of Italian attending German schools in Germany

The same four domains that were studied in the French bilingual group were also investigated with the Italian HSs introduced above. The same methodology was used, albeit with minor modifications.\(^4\) With respect to the target phenomena, Italian and German display the same differences as French and German. The bilinguals in this group had all attended monolingual German schools. However, some of them had attended weekly HL instruction classes during their childhood, where they practised reading and writing Italian. At the time of testing, some were in the process of studying Italian at university, or studied or worked in Italy. The group also included participants who did not use Italian regularly after starting primary school. Although all participants could read and write in Italian, they did not necessarily use these skills on a regular basis.

To anticipate the results, the Italian HSs differed more from monolingual norms than their French counterparts in each of the phenomena. Average scores (percentage of monolingual-like answers) are reported in Table 1.

For grammatical gender, there were two experiments, an elicited production task and an AJT (Table 1 shows the combined percentage of correct use). Similar to the French HSs, the Italian HS were better at marking agreement than at assigning the correct gender to the noun. However, the gap between the success rates in agreement vs assignment is larger compared to the French HSs, suggesting that there was a higher number of nouns the participants were not sufficiently familiar with.\(^5\)

Adjective placement was tested by means of an AJT. The difference in results between the French and the Italian HSs is comparatively less pronounced here. In the Italian experiment, problems mostly occurred with adjectives whose meanings varied depending on the position. Again, it was typical for speakers to overuse the postnominal position, i.e. the default and more frequent position for Romance adjectives.

The study on article use with generic plural subjects showed the largest contrast between the French and the Italian HSs. Correcting bare nouns into nouns with definite articles turned out to be particularly difficult for the Italian HSs with correctness scores as low as 33%. When it came to spontaneous interpretations of definite marked plural subjects in a truth value judgments task, Italian HSs preferred the generic interpretation, as: is also typical for monolingual speakers.

The VOT study also focused on the production of the voiceless stop /k/ in naturalistic speech. Italian is similar to French in that /k/ is produced with a short lag. The Italian HS produced average VOTs of 43.8 ms (compared to the 36.5 ms produced by the Italian-dominant bilingual

---

**Table 1.** Comparison of studies with French HSs and Italian HSs, correctness scores and average VOTs (data from Bianchi, 2013; Kupisch 2012, 2014).

<table>
<thead>
<tr>
<th></th>
<th>Gender assignment</th>
<th>Gender agreement</th>
<th>Adjective placement</th>
<th>Article use with generic DPs</th>
<th>VOT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>French HSs</strong></td>
<td>93.2%</td>
<td>95%</td>
<td>90.1%</td>
<td>90.5%</td>
<td>37.2 ms</td>
</tr>
<tr>
<td><strong>Italian HSs</strong></td>
<td>79.7%</td>
<td>92.8%</td>
<td>82.6%</td>
<td>33%</td>
<td>43.8 ms</td>
</tr>
</tbody>
</table>

DP: determiner phrase; HS: heritage speaker; VOT: voice onset time.
controls from Italy). In other words, the HSs’ VOT were comparatively more ‘German-like’, though still within the limits of what has been reported for Italian monolinguals (38.4 ms). The Italian HSs produced significantly longer VOTs in German (59.3 ms), suggesting language separation similar to the French HSs, but a stronger tendency to produce longer, i.e. more German-like, VOTs in their HL.

Summary

Four studies were carried out with French and Italian HSs in Germany, testing the same phenomena with the same methodology. The French HSs had attended French schools during their childhood, while the Italian HSs had attended German schools. Overall, the French HSs outperformed the Italian HSs. At the same time, the two groups appeared to face qualitatively similar problems. For instance, difficulties were more apparent when lexical knowledge was required (e.g. gender assignment) than in the application of grammatical rules (e.g. gender agreement). Moreover, both groups showed clear evidence of language separation. For instance, in the study on adjective placement, HSs overused the postnominal position, which is typical for and unique to their Romance languages, rather than the prenominal position that is common to both languages. Similarly, VOT realizations occurred in clearly distinct ranges in the speakers’ two languages.

The explanation for the French HSs’ advantage may be that they had attended French schools continuously for more than 10 years. They were familiarized with the standard (academic) register and French literature from an early age, and they had experience taking exams and language tests in French. They were thus familiar with formal registers of French, a broader range of lexical domains and task-taking procedures. Some of the Italian HSs were also formally instructed in Italian. However, their experience did not provide them with the same type of linguistic experience that the French HSs had accumulated during their earlier years. Taken together, the data sets emerging from the E11 project lend support to Rothman’s (2009, p. 156) claim that “… differences in literacy and formal education can result in what on the surface seems to be arrested development…”. Our point is definitively not that education is necessary for bilinguals to converge more closely on domains of grammar that are more standard-like, but that consistent educational contexts provide them with quantitatively and qualitatively more opportunities to do so. By extension then, the absence of this might explain some differences that obtain when schooling is not provided in the HL.

Discussion

The term ‘incomplete acquisition’

What does ‘incomplete’ mean? In our view, it should mean that a grammar is somehow unable to fulfil the remit of language because it lacks properties that govern the constitution of natural language, as some have argued for adult L2 acquisition (see Schachter, 1990, and his Incompleteness Hypothesis), or is not sufficiently developed as a system to reliably encode and decode all necessary linguistic information (as might be the case of pidgeons). This is clearly not the case of HS grammars. In our view, naturally acquired native grammars that are sufficiently developed for communication cannot be incomplete, only different—potentially drastically—from one another by comparison. HSs are native speakers of their HL. The fact that they often differ from monolinguals is interesting and potentially revealing for theoretical questions (see Benmamoun et al., 2013), but difference from an arbitrary standard does not suggest that said system is incomplete.

Linguistic completeness of any grammar, heritage or otherwise, cannot be determined by comparison to another grammar. For example, we do not argue that English is incomplete compared to German because the former no longer has morphological case inflections. But it is still relevant to
discuss in the present context. The main difference as it relates to this being a standard practice in HS studies is, in our view, one of not properly managing expectations. In the case of the English-German comparison, no one expects speakers exposed to English to wind up with a system that has morphological case outside of the pronominal system. Of course, English is not incomplete in comparison to German. They are just different, in much the same way as are apples and oranges. But is it fair to expect HSs to wind up with grammars like native monolinguals? If not, then they too are apples and oranges, or at least different species of apples. The comparison is equally unjust. Fortunately, linguistic theory affords us the possibility of avoiding this. Defining completeness on the basis of whether grammars abide by the universal rules of natural language formation and being sufficiently developed for communicative purposes avoids potential pitfalls and should be the only criterion for determining if a language is complete. If so, then HS grammars are complete, full stop.

The state of linguistic science at any particular moment in time has no obligation to be right. In fact, the humble goal of science is merely to be increasingly more accurate over time. History has proven most reasonable theories, hypotheses and constructs to be flawed, at least to some degree. Determining incorrectness, however, is only possible in the first place on the coattails of high-quality proposals and only after they have had sufficient impact on a field. Arguing against a preceding proposal is likely the highest testament to its importance and effectiveness in pushing any field forward. Although science can never be assured to be right, it should be as descriptively and explanatorily precise as possible. This obligation entails accepting \textit{a priori} that specific proposals will be wrong most of the time. The construct of \textit{incomplete acquisition} is a case in point. The pioneering work done by several researchers who assume this construct is not to be understated. To be sure, our argument is not with the empirical basis that shows differences between HSs and monolingual controls, but solely with the precision of the term \textit{incomplete acquisition}. As we outlined in the introduction, the history of studying bilingualism over the past 100 years shows that terminology matters a great deal. Once offered to the public domain, terminology can have far-reaching and long-lasting effects, even—perhaps especially—when these are unintended by their original promoters. Once the ink dries on the words we put to paper, we no longer control how others interpret and use them. It seems illogical to believe that (many/most) scientists in the early part of the 20th century had any ill intention in using inaccurate and inconsiderate labels to describe bilingualism. Nevertheless, some 50 years later these ill-fated labels still echo. From time to time then, it is useful to review the terms we use, especially when such terms relate to real groups of people and can influence policies that affect them.

Potentially inaccurate terminology does not always matter to the same degree. When debating terminological minutiae that has little potential impact, there is no need to enter into arguments. However, where there is potential for impact—at the level of science and beyond—it is worth forcing debates. Although most scientists are principally concerned with the task of doing good science, we also have a commitment to the communities we engage with—scientific peers, our participants and the public we aim to serve via our research. In this light, we submit that the time has come to do away with the term \textit{incomplete acquisition}. Although we have shown exceptions, no one denies that HS grammars are often significantly different from monolingual ones. So, why are we suggesting the term \textit{incomplete acquisition} is universally inaccurate? The short answer to this question is because the term is theoretically flawed and potentially harmful.

\textit{Differences between developing 2L1s and endstate HSs in Europe and North America?}

We have pointed out that many HS studies that label their acquisition outcome as ‘incomplete’ have emerged from the US, while many studies reporting monolingual-like acquisition in
early developing bilinguals have emerged in Europe. These differences may have various explanations:

(i) The conditions for bilingualism vary across countries and continents;

(ii) Studies on early bilinguals often focused on ‘elite bilinguals’, while studies on adult HSs often focused on stigmatized populations;

(iii) Studies on early developing bilinguals were often based on naturalistic speech (where potentially problematic structures can be avoided) and studies on adults HSs on experimental data (where potentially problematic structures are often explicitly targeted);

(iv) Early bilinguals might have been studied ‘too early’, i.e. before their development could have been ‘arrested’ or ‘attrited’.

The motivation for carrying out the research summarized in section 3 was to look at a European population of adults comparable to previous studies on developing 2L1s, but using endstate data. The comparison of the two groups of HSs (French/Italian) in this paper shows that there is nothing special about bilingually educated European HSs. With respect to similar tasks and phenomena, some HSs behave like monolinguals, others don’t. For example, the Italian HSs had noticeable problems using articles with generic DPs, thus performing very similar to Spanish HSs in the US on comparable tasks (e.g. Montrul & Ionin, 2012). By contrast, the French HSs, who developed their French under similar conditions as monolinguals, with all the extra high quality input that formal schooling provides, behaved like monolinguals. It follows that the lack of formal education in the HL could be the primary explanation for HS/monolingual differences.

**Formal education: Why it matters and what for**

The results reported in section 3 suggest that diverging trajectories in language input and output, including additional training of the native language, spoken and written, in a formal context is precisely what distinguishes monolingual and multilingual speakers, as well as multilingual speakers with different proficiencies. Since formal training happens between later childhood and early adulthood, it may explain the ubiquitous differences between monolinguals and HSs noted in the literature. Previous studies on school-aged heritage bilinguals are in line with this conclusion. Relevant examples are Montrul and Potowski’s (2007) study of six- to eleven-year-old English-Spanish children in Chicago and Kupisch and Pierantozzi’s (2010) study of six- to eleven-year-old German-Italian children in Hamburg. The children in both studies attended dual immersion schools, which combine minority language-speaking children with majority language-speaking children in the same classrooms and provide some of the teaching through the minority language. In both studies, the heritage bilinguals performed differently from monolinguals, i.e. Spanish and Italian monolinguals, but in neither study was there a decline in linguistic abilities with increased age, suggesting that schooling of the HL can prevent language loss in such contexts. However, both studies left open whether schooling has long-lasting effects (beyond school) and how much schooling is necessary to attain monolingual-like competence. In a somewhat different context, Bylund and Diaz (2012) investigated the effects of HL classes (once per week) on first language proficiency in HSs of Spanish in Sweden in 12th grade. The results showed that HSs who attended HL classes at the time of testing performed better on grammaticality judgments tasks (GJTs) and cloze tests than HSs who did not attend any classes at the time of testing. However, the latter group had attended HL classes before, and the two groups did not differ with regard to total HS class experience, suggesting that once attendance to HL classes stops, its positive effects on L1 proficiency
become less obvious (p. 605). The authors raise the question of the extent to which their findings are generalizable to HS populations in other contexts of HL education and testing formats where literacy skills are less relevant than in GJT and cloze tests. The studies reviewed in section 3 seem to provide us with a tentative answer:

There are (at least) three different dimensions in which schooling might explain differences between monolinguals and bilinguals. The most obvious dimension concerns properties of the standard language that are typically considered to be indicators of sophisticated language, elaborate style and/or register, e.g. subjunctive in German, or French liaison. Such phenomena are, often taught to monolingual speakers at school, though not necessarily mastered after instruction. Other examples include the genitive case in German, clitic object pronouns in some Romance languages and the inflected infinitive in Brazilian Portuguese, which are subject to regional and diastratic variation. Obviously, speakers with exposure to regional varieties but without additional exposure to the standard variety at school, have fewer chances of acquiring these properties (Pires & Rothman, 2009).

A second dimension in which formal schooling in the HL might foster HL maintenance is in having the HL not as the object of learning but as the medium of instruction, because learners will be familiarized and learn how to deal with scholarly instructions in this language. Anecdotally, HSs often report having problems understanding instructions when taking part in experiments, while the very same instructions seem to be clear to most L2 learners.

A third dimension regards aspects of language that are not explicitly taught in heritage or foreign language classes, but that are nevertheless (positively) affected by language exposure at school. Bylund and Diaz (2012, p. 305) ask whether HL class attendance has its strongest effects on written and more metalinguistic modalities of L1 proficiency as compared to more automatic and implicit L1 knowledge, or whether other modalities are positively affected too. The studies reported in section 3 suggest that formal instruction will also affect non-instructed, implicit linguistic knowledge, since the phenomena discussed there are typically not part of native or foreign language instruction. So, why should there be beneficial effects? Schooling involves working with texts whose grammar and lexicon go beyond colloquial language, e.g. classic literature or texts in formal, technical or scientific language. Thus, learners are automatically exposed to a wider range of grammatical constructions and vocabulary items, even if these are not explicitly addressed. Furthermore, the language spoken by teachers in dual immersion classes is likely to sound different from the language that HSs hear at home, being mostly closer to the standard variety. Of course, dialect speakers too may have their first contact with the standard variety when they go to school, but they differ from HSs in having the opportunity to hear their majority language outside of the school setting. Finally, formal school settings provides significantly more opportunity to use the language and its different registers more authentically with a greater variety of people, especially peers of the same age.

So why was formal training in our studies more beneficial for the French bilinguals than for the Italian bilinguals? One major difference between the two groups was that the French HSs had French as the medium of instruction and exposure (in a school setting) every day with input from various speakers. The Italian HSs, by contrast, (those who attended HL classes) had Italian as the target of instruction, only once per week (if at all) and often by only one single teacher. This suggests that in order for schooling to have long-terms beneficial effects (in morphosyntax, lexicon and pronunciation), input needs to be continuous and come from a variety of different native speakers. Evidence of this type suggests that dual-language bilingual education has benefits to language development that is retained in adulthood.

Another difference between the two groups of bilinguals is that Italian HSs in Germany are typically exposed to first generation speakers from different regions in Italy, who often communicate
amongst each other in a neutral regional variety (if not in German). By contrast, speakers of French (at least in Europe) have a stronger inclination to convergence on the alleged standard language. In addition, the two groups tend to have different attitudes towards language norms. Italian dialects are very much alive today and speakers are somewhat more tolerant towards variation. This means that it might be less important for an Italian HS to sound like a monolingual speaker of the standard language.

A final question to ask at this point is whether French or Italian were acquired at the expense of linguistic abilities in German. The answer is simple. The bilingual populations who participated in the aforementioned studies were also investigated in their dominant language, i.e. German. None of the studies reported any differences from monolingual-like, standard German norms (see Barton, 2015 on generic nominals for an example).

Comparison with monolinguals

HSs often—but not always—have significantly less input in the HL than a monolingual speaker of the same language. Their input is often qualitatively different, they tend to lack formal literacy training, and they use the language in fewer contexts than monolinguals. Furthermore, unlike monolinguals, HSs have another (usually dominant) language that can exercise some influence on the HL. So, why should we expect them to wind up sounding like monolinguals? If the potential for convergence with monolingual systems is unlikely from the start, how can we call bilinguals incomplete for not having achieved something they are unlikely to achieve in the first place? Apparently, as we saw in our review of several studies in the previous section, when some contributing variables are nullified across HSs and monolinguals, such as an education in the HL during early childhood, differences seem to disappear. This adds support to our general position. By extension, this suggests that being a less monolingual-like HSs is more a reflection of not having had the same opportunity to be convergent. The HS grammar is not incomplete, it is just complete and different as a matter of circumstance.

Differences between early bilinguals and monolinguals have been discussed in terms of being quantitative or qualitative. These debates have been more intensive in research on developing bilinguals than in research on adult HSs. Learners count as qualitatively different from monolinguals if they commit errors that are typical for adult L2ers but never observed with monolinguals, or if they acquire certain properties in a different order than monolingual children. By contrast, accelerated or delayed acquisition with respect to monolinguals are seen as quantitative differences.

For a long time, a predominant view was that the syntactic development of early bilinguals is similar to that of monolinguals (e.g. Meisel, 2011; Paradis & Genesee, 1996), provided that AoO happens below a certain age. Over time, diverging opinions have emerged, with some scholars arguing that the weaker language of a bilingual can have L2-like features and thereby be qualitatively different from (monolingual) L1 development (Schlyter & Håkansson, 1994). Further, there is no agreement on what counts as a quantitative or qualitative difference. While qualitative differences have become a matter of debate for syntacticians, phonologist have already accumulated evidence that the simultaneous development of two phonological systems can be qualitatively different from those of monolinguals, e.g. resulting in inverted acquisition orders due to CLI (see Kehoe, 2015; Lleó & Cortés, 2013). Given the undeniable existence of extreme inter-individual differences across HSs, it will be hard to maintain the original view that differences between early bilinguals/HSs and monolinguals are merely quantitative in nature. Clearly, there are factors beyond AoO that determine acquisition outcomes. What matters is not whether the emerging or final systems are different—we know they often are—but how they are different and under what circumstances.
Finally, we wish to point out how HL acquisition, as compared to monolingual acquisition, might relate to language change. For instance, English once had morphological case in the nominal domain, like present day German. Over many generations this changed, presumably because the language of the generation before was not taken up exactly the same way by the subsequent generation (e.g. Lightfoot & Westergaard, 2007). The idea is that HS grammars might simply show us how quickly language change can take place in bilingual contexts with less than optimal support for the development and maintenance of HLs. The difference between significant grammatical language change over many generations and the accelerated case of individual HS grammars might have more to do with context than anything else. In the typical case of language change, the changing language is the dominant societal language. In the case of HSS, we are looking at change to the minority one. The fact that HS grammars at the individual level exist on a continuum, meaning some are more monolingual-like than others, is likely a mere reflection at the individual level of the potential each HS had to become monolingual-like. Related to what we discussed in 4.1. above, one difference between HL research in the US and HL research in Europe might be that the former often focuses on large minority communities with comparatively low prestige, whereas many studies of early bilinguals in Europe have looked at comparatively smaller communities and more prestigious languages (Romance languages in Germany, for instance). This may also play a role in explaining the different results in the two types of research traditions: in large communities with much contact among its members, we might be dealing with language change on a bigger scale, whereas in small communities differences are more likely to occur at the individual level.

Conclusions

We hope to have given sufficient context regarding why we consider the term incomplete acquisition to be theoretically and descriptively flawed and inadvertently insensitive. If we are on the right track, then the term should be replaced without further consideration. Calling a bilingual’s grammar incomplete because it fails to meet an unattainable threshold of comparative similarity to a different group of native speakers unwittingly perpetuates the notion that bilingualism results in imperfect language acquisition of either one or even both languages. Although this most certainly is not the message intended by linguists who use this term, it is often what is understood by non-linguists who access such work but do not have the training to understand the nuances of what is meant. Of course, the standard meaning of ‘incomplete’ assumes the potential of completion, so understood without proper context it is reasonable to deduce that incomplete acquisition entails that bilingualism, at least in a HL context, results in imperfect acquisition. In addition to being misleading, the term unwittingly privileges monolingualism and perpetuates it as the ‘should-be default state’. It could also give the impression that there is something wrong with HS grammars; something that we should endeavour to fix. Of course, HSSs are not broken and education in the standard variety is not meant to fix them, but rather provide them with an opportunity to acquire different registers of their HL that will happen to be closer to what monolinguals speak. Confronting myths related to bilingualism has been an uphill battle for the larger part of a century. Insofar as terminology matters, researchers have an obligation to use accurate terminology that cannot be misinterpreted or potentially harmful to individuals. We cannot imagine that any HS appreciates being told that their grammars are incomplete, although we can imagine that they would hardly mind knowing the non-evaluative reality that their grammars are different from monolinguals.

In summary, there are no good reasons for calling heritage speaker grammars ‘incomplete’ because (i) not all heritage grammars are different from monolinguals and (ii) the reasons for deviance from monolinguals can be manifold, including issues related to quantity and quality of input and access to literacy training, which all apply to monolinguals as well. That a similar degree of
inter-individual variation is not usually seen across monolinguals is likely to be due to two facts. First, the typical monolingual controls are formally educated in that language; when not, there are often significant differences (see Dabrowska, 2012). Second, monolinguals almost always grow up within a majority language context and their L1 is the majority language or a mutually intelligible dialect of the majority language.

Since we have suggested that the term incomplete acquisition is inappropriate to describe any outcome of naturalistic acquisition—bilingual or otherwise—it seems appropriate to offer some alternative terms that capture the differences between monolingual and HS grammars. We submit that differential acquisition is a more appropriate term. The word differential captures the reference to difference compared to monolinguals as well as degrees of difference across individual Hs. Future research that is designed specifically to tease apart the contributing variables that make many HS grammars different from monolingual ones will be well served to consider the epistemological discussion offered in this paper.

At a minimum, this discussion underscores the complexity of the HS situation. It underscores the need to expand the empirical basis of the research we do, for example, replacing monolingual controls with appropriate bilingual ones, undertaking pseudo and actual longitudinal data, getting fine-grained background information and running correlations with input factors. In our view, doing this will make HS research more ecologically valid. It will bring us beyond the current state of describing HS differences towards actual explanation of how and why differences emerge without resorting to backward assumptions regarding the path of development based solely on endstate experimental data. It will also foster the need to explain the composition of HS grammars in their own right, under a view that accepts them as complete entities worthy of formal description.

Acknowledgements

We wish to thank the audiences at ISB10 at Rutgers University at the Cambridge Romance Seminars; Miriam Geiss, Cristina Flores, Anika Lloyd-Smith and two anonymous reviewers for commenting on earlier versions of this paper. Many thanks to Aafke Hulk and Suzanne Aalberse for initiating this timely discussion of research on heritage and early bilinguals, and for giving us the opportunity to present our views.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. When referring to ‘monolinguals’ we mean speakers who grew up in homes where only the majority language was spoken and who attended monolinguals schools in the majority language.
2. It is possible, even likely depending on the HL, that Hs (as defined here) enroll as adults in language classes, either studying their HL in foreign language—alongside L2 HLs (as defined here) and novice L2 learners—or in classes specifically designed for Hs. In such cases, Hs are exposed to the standard variety of the HL, which may even be considered a specific case of third language acquisition (see Polinsky, 2015).
3. Benmamoun et al. (2013) explicitly state that with reference to heritage speakers, ‘non native-like is equivalent to ‘different’ rather than ‘incomplete’ (p. 167, FN 14). In other publications, however, the term ‘incomplete’ is used with reference to heritage speakers (e.g. Montrul, 2008).
4. These minor differences concerned the following aspects: (i) The control sentences were not exactly the same in the two languages. (ii) The Italian study on gender tested particle agreement rather than adjective
agreement. (iii) The number of participants was not exactly the same in all Italian studies because we started to publish first results while additional data was still being collected. (iv) Not all data were suitable for acoustic analyses in the VOT study.

5. The nouns providing difficulties in assignment typically contradicted the general assignment rules of Italian, e.g. *il copra* (‘the copra’), a noun which carries masculine gender although ending in –a, i.e. a consonant normally associated with feminine gender.

References


**Author biographies**

Jason Rothman is a Professor of Multilingualism and Clinical Language Sciences at the University of Reading as well as a Professor II of Linguistics at UiT The Arctic University of Norway. He has conducted research on first, second and third language acquisition as well as heritage language bilingualism, with a particular focus on the acquisition of morphosyntax. His research makes use of a diverse range of behavioral, psycho- and neurolinguistic methodologies to investigate quantitative and qualitative aspects of linguistic competence in different groups of learners.

Tanja Kupisch is a Professor of Linguistics at the University of Konstanz as well as a Professor II of Linguistics at the UiT The Arctic University of Norway. She has conducted research on bilingual first language acquisition, first, second and third language acquisition, bilectal acquisition as well as language dominance. Her work has focussed on the syntax and semantics of Determiner Phrases and pronunciation.