DIFFERENTIAL OBJECT MARKING IN SPANISH, HINDI, AND ROMANIAN AS HERITAGE LANGUAGES

Silvina Montrul  
University of Illinois at Urbana-Champaign

Rakesh Bhatt  
University of Illinois at Urbana-Champaign

Roxana Girju  
University of Illinois at Urbana-Champaign

Erosion of differential object marking (DOM)—the overt morphological marking of animate direct objects—has been observed in Spanish heritage speakers who are second-generation immigrants in the United States (Montrul 2004, Montrul & Bowles 2009). We investigated whether DOM is similarly vulnerable in heritage speakers of Hindi and Romanian, two other languages that also exhibit DOM, as well as in first-generation immigrants, adults who are presumably the main source of input to heritage speakers. We report the results of three experimental studies testing acceptability of DOM through a bimodal judgment task in first- and second-generation Spanish, Hindi, and Romanian speakers in the US and native speakers in Mexico, India, and Romania matched for age and socioeconomic status. Our results show structural changes with DOM in all of the heritage speaker groups to different degrees. Acceptance of nontarget DOM omission was more extensive in Spanish than in Hindi and Romanian. First-generation Hindi and Romanian immigrants did not differ in their grammatical proficiency and acceptance of DOM omission from the Hindi and Romanian speakers tested in India and in Romania. However, the first-generation Mexican immigrants displayed similar performance to the Spanish heritage speakers, suggesting that Spanish DOM is prone to L1 attrition in the first generation as well. We discuss linguistic and experiential factors relevant to the three languages and the three immigrant communities to explain these findings.*

Keywords: heritage languages, case, Spanish, Hindi, Romanian, attrition

1. INTRODUCTION. One of the goals of the language sciences is to understand the nature of the language faculty as represented in adult native-speaker knowledge and how that knowledge develops in children. Language development and native-speaker knowledge can be profoundly shaped by the linguistic environment, and this is especially significant in bilingualism, where context of use and amount of input in each language vary. It is, in fact, common for bilingual individuals to experience fluctuations in fluency and proficiency in their languages throughout their lifespans, particularly if language development takes place in the context of large-scale immigration, as in the cases of the heritage speakers discussed in this article. HERITAGE SPEAKERS is a term that has gained ground in the United States to refer to early bilingual speakers of sociopolitically minority languages (Cummins 2005, Valdés 2000), including immigrant languages, such as Spanish, Chinese, Korean, Arabic, Russian, and Hindi, among many others. Heritage speakers are children of immigrants or immigrant children born to one or two parents who are native speakers of the family language or heritage language.

If being exposed to the language in early childhood and beyond is a defining characteristic of a native speaker, then heritage speakers are native speakers because they were

* 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 authors and do not necessarily reflect the views of the National Science Foundation. We are grateful to all of the participants in the study, as well as to Kirsten Hope, Laura Romani, Archna Bhatia, Vandana Puri, Natalie Toomey, Noelia Sánchez-Walker, Vanesa Hernández, Francisca Medrano, Raluca Kim, Luminita Marcus, Andreea Faur, Bogdan Buricea, and Daniela Raducanu, who helped with data collection and transcriptions in the three languages. Versions of this study were presented at the 5th Heritage Language Summer Institute (UCLA, 2011), a plenary talk at GALANA (October 2012), the workshop on Heritage Languages (the Netherlands, 23–25 January 2013), and the University of Cambridge (February 28, 2013). All remaining errors are our own.

Printed with the permission of Silvina Montrul, Rakesh Bhatt, & Roxana Girju. © 2015.
exposed to their family language at home since birth. Yet what sets heritage speakers apart from monolingual or monolingually raised native speakers is that eventual mastery of their heritage language in early adulthood can vary significantly from the end-state achieved by their own parents and from other native speakers of the same language in their linguistic community of origin. Many heritage speakers display non-target structures and patterns of simplification or innovations that could be related to transfer effects from the dominant language. The structural changes common to many heritage language grammars and the potential causes behind them have been enduring interests in the linguistic study of heritage languages. The study of heritage grammars intersects naturally with research on contact linguistics, first language (L1) acquisition, bilingual acquisition in childhood, child and adult second language (L2) acquisition, and L1 attrition.

Much research on heritage languages to date has focused on understanding the properties of individual heritage languages (e.g. Arabic, Spanish, Russian, Korean) by comparing the linguistic behavior of heritage speakers and native speakers of the full variety (Benmamoun et al. 2014, Kim et al. 2009, Montrul 2002, 2004, Polinsky 2006, 2008). Because many heritage speakers make errors similar to those made by L2 learners, another prominent area of research has endeavored to compare heritage language systems to the interlanguage systems of L2 learners (Montrul 2008, 2012). Much less work exists on comparing the language of heritage speakers to the L1 development of children of the same language (for a recent study see Chung 2013), and studies of this sort are critical for understanding the nature of potential delay and arrested development of the language acquired as a minority language. With a few exceptions (Hulsen 2000, Nagy et al. 2011, Pascual y Cabo 2013, Silva-Corvalán 1994), the relationship between the grammars of heritage speakers and those of adult immigrants who may be undergoing attrition has not received much attention. The study we present in this article adds new answers to this question.

Highly relevant to linguistics is what different heritage languages may have in common at the structural level. And if different languages share a given structural property, the question that arises is whether that same property is similarly affected in heritage speakers of these languages. This is indeed the central question addressed by the present study, a theoretically informed large-scale experimental investigation of the potential erosion of differential object marking (DOM) in Spanish, Hindi, and Romanian as heritage languages in the US. DOM is the overt morphological marking of some direct objects, and it is a widespread phenomenon in many languages of the world (Bossong 1991), including Hebrew, Farsi, Turkish, and Bantu languages, among many others, but it does not exist in English. The novelty of our study, therefore, lies in its comparative design because it examines the same linguistic domain (DOM) across different heritage languages (Spanish, Hindi, and Romanian) using the same methodology (bimodal acceptability-judgment tasks).

The study is innovative in several other ways. First, it addresses quantity of early input, although indirectly, by including heritage speakers with different ages of onset of intense exposure to and use of English. Second, it addresses (indirectly) potential changes in the quality of the input heritage speakers may be exposed to by including first-generation immigrants, who have been in the country for several years, to investigate whether their language exhibits attrition in the domain of DOM. Third, the study is also transnational because it included native speakers in Mexico, India, and Romania matched for age and socioeconomic status (SES) in order to evaluate the nature of changes, if there are any, observed in the first- and second-generation immigrant groups
in the US. Finally, we consider the potential effects of several external factors specifically related to each immigrant community—including age at immigration, degree of language use throughout the lifespan, beliefs and attitudes toward the value of transmitting their language to the next generation, and vitality of the language—on the structural changes observed in each language.

Our results show that the heritage speakers of the three languages seem to adopt the grammar of English, which does not overtly mark direct objects, and accept nontarget sentences with animate, specific direct objects without DOM. Yet, DOM omission is affected in the three heritage languages to different degrees. DOM appears to be vulnerable to attrition in the first-generation Mexican immigrants, but the first generation of Hindi and Romanian speakers do not exhibit attrition of DOM. The sociolinguistic information collected to supplement the results of the acceptability-judgment tasks indicates that age of immigration of the sequential bilingual Mexican and Romanian heritage speakers may underlie their differential degree of incomplete acquisition of DOM. At present, Spanish has more vitality in the US than Hindi and Romanian, and these heritage speakers report high use of the language. Therefore, the degree of attrition attested in Spanish immigrants is unlikely due to patterns of language use. We consider several external and internal factors that may account for our results, and we suggest that the degree of DOM erosion in the three languages is better explained by the formal properties of the DOM marker and the syntax of definiteness and specificity in each language. We end by addressing the nature of heritage language grammars with respect to DOM.

Before presenting the details of each experimental study, we describe the dynamic patterns of bilingual dominance and proficiency in immigrant families, followed by a basic description of how DOM works in Spanish, Hindi, and Romanian.

2. Bilingualism in immigrant families. Immigrants bring their own languages into the host country. The family members’ degree of bilingualism and language proficiency in the heritage language is intimately related to their life experience with the majority language. With time and contact with the majority language, their heritage languages undergo changes as they are transmitted to the next generation. First-generation immigrants arrive at the host country as adults. They are typically monolingual speakers of their native language, and most of them learn the majority language as a second language late in life (and imperfectly). Command of the heritage language is strong in this adult group upon arrival, although there can be some perceived changes in retrieval, processing, and use after more than ten years of intense exposure to the majority language, as these immigrants gradually become fluent and frequent users of the majority language (Keijzer 2007, Köpke 2007, Seliger 1996, Schmid & Köpke 2004). The study of these changes falls under the purview of L1 attrition.

Language attrition is characterized as the gradual forgetting of aspects of the native language by a healthy native speaker (Schmid 2011, Sorace 2000), and it takes place at the individual level in contexts where the native language begins to be used less often. It happens essentially when a native speaker becomes, as judged by peers, a nonnative speaker of his/her own language due to difficulties with lexical retrieval, code-switching to fill out lexical gaps, changed pronunciation, morphological errors, and avoidance of certain structures and overuse of others due to influence of the dominant language.

Heritage speakers are the children of the first-generation adults or the second-generation speakers. This group may include children born in the host country to at least one first-generation parent. It also includes immigrant children who come to the host country before the age of five. Some heritage speakers are simultaneous bilinguals,
those exposed to the heritage and the majority language before the age of five; others are sequential bilinguals or child L2 learners, those exposed to the heritage language exclusively at home until age five, and to the majority language once they start school; and there are also some who are late child L2 learners, that is, children monolingual in the heritage language who received some elementary schooling in their home country before immigrating between the of ages of seven and thirteen (generation 1.5 according to Silva-Corvalán 1994).

The study of heritage speakers has focused primarily on second-generation speakers; it is precisely in the second generation when active bilingualism is introduced in the home, giving rise to the wide variability observed in this particular group. Although the children may be reared almost exclusively in the heritage language in early childhood, they are later schooled in the majority language and are drawn to adopt the linguistic and cultural etiquette of the new society. Naturally, the majority language begins to be used more than the home language, especially at school and with peers, and the amount of input in and use of the heritage language decreases. At this time, if not earlier, there is language shift from the heritage language to the majority language, and with the shift come associated changes in the bilingual balance throughout crucial developing years.

Thus, as the majority second language becomes the primary and dominant language of the heritage speakers, their first or native language—the heritage language—gradually becomes their secondary and weaker language. Furthermore, without adequate academic support of the heritage language during the school years, heritage speakers miss the chance to acquire academic literacy skills or enhance the registeral range in the language. By young adulthood, their heritage language resembles a second, imperfectly acquired language, in the sense that it displays lexical and grammatical patterns common to L2 development and falls short of the typical full ultimate attainment reached in a first language acquired in childhood. In general, many (but by no means all) adult heritage speakers may possess good speaking and listening abilities, large vocabulary, native-like levels of pronunciation and fluency, and familiarity with the cultural norms of the language and culture.

The study of heritage speakers raises critical questions about current understanding of language development in a dual-language context and the role of input and experience in both learning and maintaining a language. It also leads us to question what aspects of language develop fully and which ones end up incompletely mastered, as well as how reduced input and use interact with the age of the bilingual speaker in determining patterns of acquisition and language loss. Variability in language development under these conditions also raises fundamental questions about the nature of language. For example, which aspects of grammatical competence are specifically vulnerable and prone to simplification and erosion under limited input conditions in heritage grammars, and which aspects are more resilient, and why?

Research to date has shown that phonetics and phonology seem to be the most robust and best-preserved aspect of heritage language grammars (Au et al. 2002, Lukyanchenko & Gor 2011), whereas inflectional morphology, semantics, and the syntax-discourse interface are quite vulnerable to simplification and loss. Several studies of different heritage languages that used different methodologies have shown that heritage speakers do not properly master case (Montrul et al. 2012, Polinsky 2006, Sekerina & Trueswell 2011, Song et al. 1997), gender agreement in nouns (Montrul et al. 2008, Polinsky 2008), plural morphology (Benmamoun et al. 2014), article semantics (Montrul & Ionin 2010, 2012), verbal morphology (Montrul 2009, Pires & Rothman 2009, Rothman 2007, Silva-Corvalán 1994), relative clauses (O’Grady et al. 2001, Polinsky
and discourse distribution and interpretation of subject object pronouns (Montrul 2004, Otheguy & Zentella 2012, Silva-Corvalán 1994), to name just a few. Many of these areas are acquired by age five in monolingual children. From the perspective of language acquisition, many of the linguistic changes and patterns exhibited by heritage speakers in several linguistic domains (but not all), most notably in morphosyntax and semantics, have been characterized as arising from incomplete, partial, or interrupted acquisition of the heritage language under reduced input conditions (Montrul 2002, 2004, 2008, O’Grady et al. 2011, Polinsky 2006, Silva-Corvalán 1994, 2003). But they could also be due to language attrition in childhood or later on (Polinsky 2011).

Incomplete acquisition and attrition due to insufficient input are two viable potential causes behind the structural properties exhibited by individual heritage language grammars, coupled with, in many instances, strong influence from the dominant majority language (or L2 transfer). It is easier to document actual attrition in adults, who are assumed to have reached linguistic maturity, than in children because, depending on the grammatical property under consideration, children may still be developing their language. This is because attrition would entail full learning, and then loss. Incomplete acquisition of specific aspects of the grammar occurs primarily in childhood due to input that is insufficient for developing the full grammar, and may cooccur simultaneously or sequentially with attrition, depending on the age of the child and the grammatical properties under investigation, as documented by existing longitudinal studies of children (Anderson 1999, Merino 1983, Silva Corvalán 2003, 2014).

Another potential cause of truncated knowledge is different input, either because heritage speakers are exposed to vernacular registers that do not exhibit grammatical properties found in formal registers (e.g. inflected infinitives in Brazilian Portuguese studied by Rothman (2007) and Pires and Rothman (2009)) or because the speakers who provide the main input to the heritage speakers (i.e. parents and older relatives) may exhibit structural changes in their own grammars as well (Sorace 2004, 2011). For example, Otheguy and Zentella (2012) attribute high rates of overt subject pronoun expression in the Spanish spoken in New York City by Spanish speakers of several generations and regions of Latin America to influence from English and dialect leveling. Finally, Silva-Corvalán (1994) proposed that the language-contact situation accelerates diachronic changes already in progress in monolingual varieties (e.g. clitic placement in Spanish). Our study supports the conclusion that transfer from English, insufficient input and use of the heritage language by the second generation, potential attrition in the first generation, and other language-specific factors appear to drive DOM erosion with animate, specific direct objects in Spanish, Hindi, and Romanian as heritage languages.

3. Differential object marking. In languages like Spanish, Hindi, and Romanian some direct objects are marked with overt morphology. DOM marks overtly objects that must be distinguished from subjects on semantic and pragmatic prominence scales (Aissen 2003). DOM is regulated by animacy and specificity (or definiteness) (Comrie 1975), properties of the subject, and semantics of the verb (Torrego 1998, von Heusinger 2008).

In Spanish, animate and specific (definite) direct objects are marked with the preposition a, as in 1. All other direct objects, such as animate, nonspecific (2), inanimate, specific (3), and inanimate, nonspecific (4), typically receive no marking.

(1) Marina vio a la mujer. (+animate, +specific)
Marina saw DOM the woman
‘Marina saw the woman.’
(2) Marina vio una mujer. \((+\text{animate}, -\text{specific})\)  
‘Marina saw a woman (any woman).’

(3) Marina vio la casa. \((−\text{animate}, +\text{specific})\)  
‘Marina saw the house.’

(4) Marina vio una casa. \((−\text{animate}, −\text{specific})\)  
‘Marina saw a house.’

The DOM marker in Spanish is the same preposition—\(a\)—that appears with indirect objects, as in 5, and verbs with experiencer subjects, as in 6. In these cases, the preposition \(a\) marks dative case, and it is obligatory regardless of the animacy, definiteness, or specificity of the argument. Example 7 shows that \(a\) is also a locative/directional preposition.

(5) Roberto le dio un libro \(a\) Patricia.\(\)  
Roberto \textit{dat.cl} gave a book to Patricia  
‘Roberto gave Patricia a book.’

(6) A Juan le gusta la música.\(\)  
to Juan \textit{dat.cl} like the music  
‘Juan likes music.’

(7) Mario fue a México.  
‘Mario went to Mexico.’

However, there are several counterexamples to the generalization that only specific and animate direct objects are marked with the preposition \(a\) in Spanish. First, animate objects with indefinite determiners, as in 2, can be marked with \(a\) if they are construed as specific, and nonspecific quantifiers like \textit{alguien} ‘somebody’ and \textit{nadie} ‘nobody’ always require \(a\) (\textit{Conocí a alguien} ‘I met somebody’, \textit{No vi a nadie} ‘I didn’t see anybody’). Second, for disambiguation purposes inanimate objects can be marked with the preposition \(a\) if the subject is also inanimate (\textit{El submarino hundió al barco} ‘The submarine sank DOM-the ship’). Third, some verbs of motion and location subcategorize for the preposition (\textit{La calma sigue a la tormenta} ‘Calm follows the storm’). Fourth, with animate nonhuman direct objects, \(a\)-marking is variable (\textit{Mató el/al mosquito} ‘He/she killed the mosquito’). Although for Torrego (1998), definiteness, specificity, aspect, topicality, agentivity, and affectedness determine when objects are marked in Spanish, the exact semantic, syntactic, and pragmatic conditions regulating when accusative objects should be marked with the element \(a\) are quite complex and not entirely clear (Fábregas 2013, Leonetti 2004, 2008, Torrego 1998, Weissenrieder 1990, Zagona 2002).

Spanish, Romanian, and Sardinian are the three Romance languages that exhibit DOM, which was an innovation from Latin. Diachronic studies of Spanish DOM (Company Company 2002, Laca 2006, von Heusinger 2005, 2008) show that the evolution of \(a\)-marking started with personal pronouns, then continued with proper names, and further extended along the definiteness scale to definite, specific animate objects and, lastly, to indefinite, specific animate direct objects.\(^1\)

In several Latin American varieties, \(a\)-marking has been spreading to definite and specific inanimate objects, as in 8 (Company Company 2002, Dumitrescu 1997, von Heusinger 2005), and to topicalizations with clitic doubling in Argentina (Montrul 2013), as in 9.

---

\(^1\) Animacy scale: human > animate > inanimate; definiteness scale: pronoun > name > definite > indefinite specific > nonspecific
(8) Cosecharon el maíz.  
(they)harvested the corn  
‘They harvested the corn.’

(9) A la casa la vio Mariela.  
dom the house it saw Mariela  
‘The house, Mariela saw it.’

The generalization that emerges is that animate, specific (definite) direct objects are at present obligatorily marked in all varieties of Spanish, whereas a-marking with inanimate and nonspecific indefinites is variable (López 2012). While a-marking is further extending to inanimate objects in some monolingual varieties, the most salient pattern in the Spanish spoken in the US, as we will see, is the omission of the marker with animate objects.

Another language we investigated is Hindi/Urdu, an SOV split-ergative language that also exhibits DOM marked with the postposition -ko. Hindi/Urdu does not have definite articles, unlike Spanish. Hence, when -ko appears with direct objects, it signals specificity (Mohanan 1993, 1994). Human animate, specific direct objects must be overtly marked with -ko, as in 10, and are ungrammatical without -ko marking. Nonhuman animate, specific direct objects may be optionally marked with -ko, as in 11.

(10) Mira-ne ramesh-ko dekhaa.  
Mira-ERG Ramesh-DOM saw  
‘Mira saw Ramesh.’

(11) Mira-ne us kutte-ko dekhaa.  
Mira-ERG that dog-DOM saw  
‘Mira saw that dog.’

Inanimate, specific direct objects can also be optionally marked with -ko, as in 12.

(12) Mira-ne ghar(-ko) dekhaa.  
Mira-ERG house(-DOM) saw  
‘Mira saw the house.’

Human animate, nonspecific objects are marginally acceptable without -ko marking, as in 13. Nonhuman animate, nonspecific objects are unmarked, as in 14, and are unacceptable with overt -ko marking. Inanimate, nonspecific objects, as in 15, have no -ko marking either.

(13) Neha kiimaaN-ne us-ke liye laRke dekhe.  
Neha of mother.m.pl-erg her-for boys.m.pl saw.m.pl  
‘Neha’s mother saw boys for her (for marriage).’

(14) Tum vahaaN ek saaNp dekhoge.  
Tum vahaaN ek saaNp-ko dekhoge.  
you there one snake will see  
‘You will see a (nonspecific) snake there.’

(15) Mira-ne ek ghar dekhaa.  
Mira-ERG a house saw  
‘Mira saw a (nonspecific) house.’

Additionally, dative -ko can mark indirect objects (goals, beneficiaries), as in 16, and dative subjects (experiencers) (Narisimhan 1998), as in 17. With indirect objects and dative experiencers -ko marking is obligatory, irrespective of the specificity of the marked argument.

(16) Ramesh-ne Sita-ko kitaab dii.  
Ramesh-ERG Sita-ERG book gave  
‘Ramesh gave a book to Sita.’
Differential object marking in Spanish, Hindi, and Romanian as heritage languages

The third language we are considering in our study is Romanian, which, like Spanish, belongs to the Romance family. DOM in Romanian is marked with the preposition pe, which is homophonous with the locative preposition pe. Although the classification of pe is not straightforward, there is some consensus on the most important parameters of DOM in Romanian: animacy, definiteness, and specificity (Dobrovie-Sorin 1994, Farkas 1978, Mardale 2002, Pană-Dindelegan 1997). As with Spanish, pe-marking applies to animate direct objects only if they are definite or specific, as in 18. Animate nonspecific objects can have optional marking, as in 19, and specific and nonspecific inanimate objects, shown in 20 and 21, are not marked.

(18) Ileana a văzut pe Beatrice. (*Ileana a văzut Beatrice.)
Ileana has seen dom Beatrice ‘Ileana saw Beatrice.’

(19) Claudia a vizitat un prieten. (Claudia a vizitat pe un prieten.)
Claudia has visited a friend ‘Claudia visited a friend.’

(20) Ileana a văzut casa. (*Ileana a văzut pe casa.)
Ileana has seen the.house ‘Ileana saw the house.’

(21) Claudia a văzut o casă. (*Ileana a văzut pe o casă.)
Claudia has seen a house ‘Claudia saw a house.’

Furthermore, Romanian exhibits accusative clitic doubling, by which the object NP is doubled by a clitic (see 9 for Spanish). DOM is required with clitic-doubled animate, specific objects, as in 22, and with animate, nonspecific objects, as in 23. Clitic-doubled inanimate objects are ungrammatical, with or without pe-marking, as in 24 and 25.2

(22) Angelica a văzut-o pe Madonna. (*Angelica a văzut-o Madonna.)
Angelica has seen-cl dom Madonna ‘Angelica saw Madonna.’

(23) Elisabeta a văzut-o pe o femeie. (*Elisabeta a văzut-o pe o femeie.)
Elisabeta has seen-cl dom a woman ‘Elisabeta saw a woman.’

2 Accusative clitic doubling with animate objects is also common in some varieties of Spanish (e.g. Argentinian), as in (i), and like in Romanian, it must cooccur with DOM, or else it is ungrammatical. Inanimate objects are typically ungrammatical with accusative clitic doubling, as in (ii) (cf. Suñer 1988).

(i) Lo vi a Juan/al niño/al perro/el. (*Lo vi Juan/el niño/el perro/el.)
CL.sg.m saw DOM Juan/the boy/the dog/him ‘I saw Juan/the boy/the dog/him.’

(ii) *Juanlo rompió al vaso. (*Juan lo rompió el vaso.)
Juan cl.3sg.m broke dom the.glass ‘Juan broke the glass.’

These Spanish and Romanian facts led Kayne (1975) to the generalization that an object NP may be doubled by a clitic only if the NP is preceded by a preposition. However, subsequent studies have revealed that DOM and clitic doubling are related in Spanish and Romanian but diachronically developed independently and are not exactly the same phenomenon (von Heusinger & Gaspar 2008). Although regulated by similar animacy and specificity constraints, accusative clitic doubling appears in a subset of DOM semantic contexts (Bleam 1999, Leonetti 2004).
(24) *Iulia a văzut-o casa. (*Iulia a văzut-o pe casa.)
Iulia has seen-cl house
Iulia saw the house.’

(25) *Ioana a văzut-o o casă. (*Alexandra a văzut-o pe o casă.)
Ioana has seen-cl a house
Ioana saw a house.’

In addition to the main parameters of pe-marking, there are additional triggering conditions such as metonymical shifts and metaphorical transfers (Gramatica Limbii Române 2005) or discourse prominence (Chiriacescu & von Heusinger 2010). Because these were not the focus of our analysis, however, we do not discuss them here.

The Romanian preposition pe is also the locative preposition, as in 26. It appears with some indirect objects, as in 27, and with dative experiencer subjects, as in 28. In all of these cases the preposition is obligatory.

(26) Cartea este pe masă.
book.DEF is on table
‘The book is on the table.’

(27) Claudia se bazează foarte mult pe mama sa.
Claudia REFLE relies very much on mother.DEF her.
‘Claudia relies very much on her mother.’

(28) Pe Camelia o interesează medicina.
to Camelia CL interest medicine.DEF
‘Camelia is interested in medicine.’

Table 1 summarizes the distribution of DOM in Spanish, Hindi, and Romanian.

<table>
<thead>
<tr>
<th>Spanish</th>
<th>Hindi</th>
<th>Romanian</th>
</tr>
</thead>
<tbody>
<tr>
<td>preposition</td>
<td>postposition</td>
<td>preposition</td>
</tr>
<tr>
<td>animate specific direct objects</td>
<td>animate specific direct objects</td>
<td>animate specific direct objects</td>
</tr>
<tr>
<td>dative (and locative)</td>
<td>dative</td>
<td>locative (and dative)</td>
</tr>
<tr>
<td>experiencer subjects</td>
<td>experiencer subjects</td>
<td>experiencer subjects</td>
</tr>
</tbody>
</table>

A difference between the markers in the three languages is their acoustic salience. In Spanish it is just a vowel /a/, in Hindi it is a postposition /ko/, and in Romanian it is a preposition /pe/. Acoustically, /ko/ and /pe/ are easier to perceive than /a/. Spanish a may be difficult to perceive especially when the verb ends in the vowel [a] as well, as in the third-person singular present (see 29), where the two /a/ sounds are practically collapsed in speech, producing a long vowel [aː]. In the preterite, even if the verb ends in [o], as in 30, the [o] from the verbal ending and the DOM marker [a] diphthongize in rapid speech to [oa] or [ua]. If the verb is in the plural, this problem does not arise: the DOM marker [a] in 31 is more acoustically salient because the coda consonant [n] or [s] in the verb becomes the onset of the next syllable, taking the DOM marker [a] as its nucleus. Thus, the marker may be more difficult to perceive in Spanish than in Hindi and Romanian.3

3 Note that in some sentences, the presence of a is crucial for understanding who is doing what. Compare the minimal pair in (i).

(i) a. Llamó Juan.
(V-S)
‘Juan called.’

b. Llamó a Juan.
(Spro, V-DOM-O)
(s/he)called DOM Juan
‘He/she called Juan.’
(29) Marisa llama a Juana. [a:]
    Marisa calls dom Juana
    ‘Marisa calls Juana.’

(30) Marisa llamó a Juana. [oa], [ua]
    Marisa called dom Juana
    ‘Marisa called Juana.’

(31) Llamamos/llamaron a Juana. [sa], [na]
    (we/they)called dom Juana
    ‘We/they called Juana.’

Aside from acoustic salience, DOM in Spanish and Romanian is sensitive to animacy (animate nonspecific objects can be marked with *a* or *pe*), whereas in Hindi specificity is more relevant than animacy, since specific inanimate can be marked with *-ko*. The DOM markers in the three languages are also the dative case markers appearing with indirect objects and dative experiencer subjects regardless of the animacy, definiteness, or specificity of the argument. In the case of Romanian, *pe* is also a locative preposition.

Several syntactic accounts within generative grammar (Lidz 2006, López 2012, Rodríguez-Mondoñedo 2007, Torrego 1998) propose that DOM objects check specificity or accusative case, depending on the analysis, in a higher functional projection that is different from the projection for unmarked objects. Hence, overtly marked objects are structurally more complex than unmarked objects because they involve movement and phonetic form (PF) realization of morphology (Woolford 2008).

To take one example of Spanish, Torrego (1998) analyzes the dative preposition *a* as an instance of marked accusative case encoded in a functional category specific to Spanish, and different from the functional category for accusative clitics. Consider the clause structure in 32 for transitive verbs. It is in the specifier of vP (base subject position) where subjects typically originate before moving to the tense phrase (TP) in the higher functional layer.

(32) 

While unmarked objects may be assumed to receive accusative case inside the VP in the direct object position, DOM objects move up outside the VP in the overt syntax. In Torrego’s analysis, *v* has a D-feature that can attract the object to raise overtly. Once the object is raised, its case feature is licensed in a specifier position of *v* (to which *V* has already adjoined). If only objects that are semantically similar to subjects are marked in Spanish (Aissen 2003), Torrego’s analysis captures this observation, because marked accusative

---

Example (ia) has a postverbal subject and no object (V-S), whereas example (ib) has the structure S-V-O, except that the subject is a null pronoun. The *a* in front of *Juan* (DOM) indicates that *Juan* is the object and not the subject. (See Montrul 2014 for a comprehension-based experiment that included stimuli sentences of this sort with the same Spanish-speaking participants reported in the present study.)
objects check their features in a subject-like position, probably below the specifier position for the agential subject in these constructions. Torrego further proposes that the preposition in DOM contributes its own D-feature and can be analyzed as an additional functional category (1998:25). Hence, while unmarked direct objects have structural accusative case, marked direct objects have inherent case in addition to the structural one. In another account of Spanish, Rodríguez-Mondoñedo (2007) proposes that animate, specific nominals move through [Spec, vP] to an additional head with dative case where the preposition a checks its features. A similar additional, noncore functional category is assumed by Lidz (2006) for Kannada, a South Asian language that also has DOM. This additional functional projection is not part of the functional projection of the object DP: rather, it is introduced by the DOM marker vanну/yannу in Kannada, so that the case features can be morphologically realized. In attempting to capture the facts of DOM in structural terms, what all of these accounts have in common is that because marked objects have [+specific, +definite], and in the case of Spanish and Romanian also [+animate], semantic features, they are placed in a structural position higher than the basic position for unmarked objects by means of object raising or some similar formal device.

Spanish, Hindi, and Romanian also exhibit differential subject marking (DSM) (Woolford 2008), which, according to Rodríguez-Mondoñedo (2007), is related to DOM. As we have seen, in the three languages some subjects are also marked with dative case. These subjects lack volitional control, and express a range of thematic roles (experiencer, obligation, possession) that can be categorized broadly under the thematic role of experiencer. Our study included these subjects as well, in Hindi and in Spanish.

4. Previous studies. Given its complexity, DOM certainly presents a challenge since there is significant apparent variability in the system, and both monolingual and bilingual learners need to figure out how to extract the precise syntactic, semantic, and pragmatic constraints that regulate the distribution and polyfunctionality of these markers in Spanish, Hindi, and Romanian. There is very little research on the L1 acquisition of DOM, with the exception of a study by Rodríguez-Mondoñedo (2008) for Spanish. Narasimhan (2005) studied ergativity marking in child Hindi and mentions -ko in relation to ergativity marking, but the study offers no actual analysis of all arguments marked with -ko in Hindi. We are not aware of any study dealing with the acquisition of pe in Romanian.

Rodríguez-Mondoñedo (2008) analyzed the spontaneous production of four Spanish-speaking children (between the ages of 0;9 and 2;11) from the CHILDES database. From a total of 991 V-O exemplars, the children made a total of seventeen errors (98.28% accuracy): nine cases of DOM omission with animate objects out of fifty-three (17% error rate) and eight cases of DOM overgeneralization out of 938 inanimate objects (< 1% error rate). Rodríguez-Mondoñedo concluded that Spanish-speaking children acquire the semantic constraints on the distribution of DOM very early, at least with core cases. But in situations of bilingualism in which the second language has no DOM, the same degree of acquisition by age 3;0 is less likely. Ticio (2015) examined the oral production of six simultaneous bilingual children between the ages 1;0 and 3;0 (five Spanish-English, one Catalan-Spanish) from the CHILDES database. The bilingual children omitted DOM with animate objects 76.62% of the time on average (23.38% suppliance), and Ticio concluded that simultaneous bilingual children do not develop DOM by age 3;0.

Montrul and Sánchez-Walker (2013) tested simultaneous and sequential bilingual children in the US and monolingual children from Mexico (ages six to seventeen) in
story retelling and elicited oral production. The native speakers were equally accurate on marked and unmarked objects (above 90%), whereas the bilingual children made on average more than 30% DOM omission errors with animate objects in the narrative task, and more than 60% omission errors in the elicited production task, suggesting that not all bilingual children master DOM by age three like monolingual children.

Luján and Parodi (1996) and Silva-Corvalán (1994) observed that adult Spanish-English bilinguals from Los Angeles omitted the preposition a with animate, human direct objects in clitic-doubling constructions, as in *Lo veo la niña ‘him/it I see the girl’. Similar examples from Spanish-Quechua bilinguals in Peru are reported by Sánchez (2003). Montrul 2004 and Montrul & Bowles 2009 found significant omission rates of obligatory DOM by adult heritage speakers of Spanish in the US in an oral narrative task (up to 50%, depending on proficiency) and acceptance of ungrammatical sentences with no DOM marking in a written acceptability-judgment task, with mean ratings ranging from 3.5–4 on a five-point scale (where 1 = totally ungrammatical and 5 = perfectly grammatical). The native speakers’ mean for these sentences was 1.2. Montrul and Bowles (2009) confirmed that even advanced heritage speakers are very inaccurate with DOM, not just low-proficiency speakers. If DOM has been incompletely acquired or even lost in some heritage speakers’ grammars, the fact that omissions are also attested in higher-proficiency speakers leads us to entertain the possibility that unmarked animate, specific direct objects may have become a stabilized dialectal feature of Spanish spoken in the US. If so, it is appropriate to ask whether the first-generation Spanish-speaking immigrants show patterns similar to those of the heritage speakers.

Grosjean and Py (1991) documented structural changes in Spanish-speaking immigrants in Neuchâtel, a French-speaking region of Switzerland. French does not have DOM. In their first study Grosjean and Py asked fifteen first-generation Spanish-speaking immigrants from Spain to give acceptability judgments in Spanish. The immigrants accepted grammatical sentences with DOM, as in standard Spanish, as well as ungrammatical sentences with omission of DOM, as attested in their Neuchâtel variant of Spanish. Girard (1995), as reported in Grosjean 2008, tested fifteen second-generation Spanish-speaking immigrants born in Switzerland and living in Neuchâtel using the same test as Grosjean and Py (1991). A comparison of their results with those of the first-generation immigrants showed no difference between the two groups. Grosjean (2008) concludes that under the impact of a dominant language, the competence that adult immigrants have of their native language can change over time, and this change can in turn affect the competence of the second-generation speakers. We agree with Grosjean’s conclusion. Given these findings for Spanish in a bilingual situation, we ask whether heritage speakers of other languages that exhibit DOM show similar omission patterns. The study described next aims to address this important question.

5. COMPARING SPANISH, HINDI, AND ROMANIAN AS HERITAGE LANGUAGES. Our study builds on the experimental findings of Girard 1995, Grosjean & Py 1991, and Montrul & Bowles 2009 by addressing a number of new questions on both methodological and theoretical levels. Our first objective was to obtain more converging evidence for the erosion of DOM in Spanish heritage speakers by expanding the participant pool. Our second objective was to investigate whether DOM is also vulnerable to erosion and loss in other heritage languages. If omission of DOM is attested in Hindi and Romanian heritage speakers in the US, our third objective was to see whether this is likely due to influence from English at the morphosemantic and syntactic levels—since English does not overtly mark DOM—or due to the linguistic properties and distribution of the DOM
marker in each language. In other words, does acoustic salience and syntactic distribution play a role? Are Spanish a, Hindi -ko, and Romanian pe equally omitted when they case-mark indirect objects, dative experiencers, and locatives, or only when they mark animate, specific direct objects?

If the acquisition trajectory of DOM in second-generation heritage speakers lags behind since childhood (see previous studies of bilingual children) and is primarily a result of influence from English, we expect to find similar patterns of DOM erosion in the three populations of heritage speakers studied: Spanish, Hindi, and Romanian. If erosion is related to structural complexity (syntax-semantics), we expect Spanish a, Hindi -ko, and Romanian pe omission to occur more in DOM and dative experiencer contexts than in indirect object and locative contexts. However, if degree of erosion is related to how acoustically salient the DOM marker is in the language, then we expect to find more erosion of DOM in Spanish, where the marker is just a vowel (a), than in Hindi and Romanian, where the marker is syllabic and more acoustically salient (-ko, pe).

Three other issues we address are quantity of input, quality of input, and sociolinguistic characteristics of each immigrant language. Studies have shown that many adult heritage speakers who are simultaneous bilinguals, and who were exposed less to the heritage language than heritage speakers who were raised predominantly monolingually for about the first five years, tend to have weaker command of the heritage language than sequential bilinguals in tense and aspect (Montrul 2002) and in binding interpretations (Kim et al. 2009), but this has not been attested for Spanish, Hindi, and Romanian DOM. If DOM is subject to erosion in the heritage speakers due to potential incomplete acquisition or attrition in childhood, we wanted to find out whether DOM is also subject to attrition in the first generation of adult immigrants, a question that has been raised recently but not investigated directly.

We predict that if cumulative quantity of early input matters, heritage speakers who are simultaneous bilinguals would omit DOM more than sequential bilinguals. If DOM is acquired by age 3;0 in monolingual children but by that age simultaneous bilingual children still omit the marker more than 70% of the time (see previous studies), then it is likely to remain incompletely acquired after that age, when input is more severely reduced during the school-age period. The study by Montrul and Sánchez-Walker (2013) with six- to seventeen-year-old Spanish-English bilingual children suggests that age of onset of bilingualism defined in this way may not be relevant for Spanish DOM, but it may be relevant in the other languages.

If older adult speakers who immigrated in adulthood (after age eighteen) and have been living in the US for several years also omit DOM with animate, specific direct objects, this finding would be a sign of attrition of their native language. Young heritage speakers often report their highest and most frequent use of the heritage language with their parents and relatives, and the participants in our study followed this pattern. Since adult immigrants between the ages of forty and sixty are akin to the parents of the

---

4 We agree with an anonymous referee that this generalization may be too broad, since it depends on how much input simultaneous and sequential bilinguals receive in each language. Naturally, the time and amount of exposure as a simultaneous or sequential bilingual is related to amount of input required to develop linguistic proficiency in the two languages. In a recent study of French-English bilingual children in Montreal, Thordardottir (2013) found that 50% exposure to each language is sufficient to develop the language at monolingual levels in bilingual children. But the reality is that the amount of time or proportion of daily input a bilingual child is exposed to from each of the two languages can range from 0% to 100% depending on circumstances. Some simultaneous bilingual children are exposed to the two languages equally (50% of the time in each), while others may be exposed to one 20% of the time and to the other 80% of the time. Silva-Corvalán (2014) shows that less than 30% exposure to Spanish after age three in simultaneous bilingual children negatively impacts and delays the normal development of the language at age-appropriate levels.
Differential object marking in Spanish, Hindi, and Romanian as heritage languages

young adult heritage speakers (ages eighteen to twenty-five) at the time of testing, it may be possible for some heritage speakers to receive qualitatively different (attrited) input from their parents when they are older, which would also contribute to the apparent arrested development of this feature of their heritage language. (Note that this hypothesis does not apply to child heritage speakers who recently arrived in the US, since the language of their parents would not be attrited upon arrival.) In addition to quantitatively less input in the heritage language than in English, qualitatively different input in the heritage language could further reinforce the heritage speakers’ nontarget attainment of DOM in the heritage language.

These research questions and hypotheses were addressed in three studies—one on Spanish, one on Hindi, and one on Romanian—involving oral and written production, aural and written comprehension, and acceptability-judgment data, in addition to proficiency tests and extensive language-background questionnaires. For each language, the following groups of participants were included: simultaneous and sequential bilingual heritage speakers (age range: eighteen to about twenty-five), first-generation adult immigrants (age range: forty to about sixty), and two groups of age- and SES-matched native speakers in Mexico, India, and Romania: younger native speakers (age range: eighteen to twenty-five) to compare to the heritage speakers, and older native speakers (age range: forty to sixty) to establish whether the first-generation immigrants also exhibited structural changes after several years of living abroad. Due to scope limitations, we focus only on the results from the acceptability-judgment task in each language. The results of other tasks—which largely confirm the trends we report here—have been and will be reported elsewhere (Montrul 2014, Montrul et al. 2012, Montrul & Sánchez-Walker 2013).

5.1. Study 1: Spanish.

Participants. We tested 118 total Spanish speakers: seventy-seven first- and second-generation Mexican-origin immigrants living in the central Illinois and Chicago area, and forty-one Spanish native speakers in Northern Mexico (Guanajuato). The participants were divided into five groups: three in the US—simultaneous bilingual Spanish heritage speakers, sequential bilingual Spanish heritage speakers, and first-generation adult immigrants from Mexico—and two in Mexico: younger adult native speakers, and older adult native speakers.

All of the Spanish speakers completed an extensive language-background questionnaire with sixty-five questions about their bilingual language history, patterns of language use along their lifespan and during schooling, self-ratings of their Spanish and English skills, attitudes toward improving Spanish, and so forth. To complement the self-ratings, which are not usually reliable proficiency measures, all participants completed a standardized written proficiency test in Spanish, consisting of a cloze passage with twenty blanks and thirty multiple choice questions about vocabulary and collocations. The maximum number of points in this task was thus fifty. This is the same task used in several previous studies of Spanish heritage speakers (Montrul et al. 2008).

Simultaneous bilingual Spanish heritage speakers (n = 32). This group consisted of young adults (twenty-one female, eleven male) between the ages of eighteen and twenty-eight (mean 20.6), all born to Mexican parents (mother and father) and residing

5 A study in progress that follows up on the results of the oral tasks reported in Montrul & Sánchez-Walker 2013 tested another group of sixteen bilingual children in the US (ages six to ten) and their mothers on two oral tasks. Preliminary results confirm that the children were significantly more inaccurate than their own mothers, omitting DOM marking with animate, specific direct objects. In fact, the mothers produced the marker 90% of the time or more, showing no signs of attrition.
in the US. They were exposed to Spanish since birth and to English from between birth and age five (mean 3.8). One of them had completed college and the rest were college students at a large university in the US. All of them attended elementary school in the US, where most were schooled exclusively in English and 34.4% received some Spanish instruction through a bilingual program. All of them also attended middle school and high school in the US, where English was the main language of instruction and 80% had instruction in Spanish as a foreign language.

**Sequential bilingual Spanish heritage speakers** (*n = 24*). These heritage speakers were also young adults (thirteen female, eleven male) between the ages of nineteen and twenty-nine (mean 23.6), all born in Mexico to Mexican parents. They immigrated with their parents to the US between the ages of five and thirteen. All but one (95%) were college students at a large university in the US. They were exposed to Spanish since birth and to English as a second language from between the ages of five and thirteen (mean 6.79), and all reported that Spanish was the only language used at home. Except for two, all others in this group of sequential bilinguals attended elementary school in the US (six also attended elementary school in Mexico before immigrating). Only four (16.7%) received some Spanish instruction in elementary school. All of them attended middle school and high school in the US, and English was the main language of instruction. They all had Spanish as a foreign language in high school.

**First-generation adult immigrants from Mexico** (*n = 21*). The adult immigrants (nine male, twelve female) were between the ages of forty and fifty-eight (mean 42.57) and were all from Mexico. They immigrated to the US between the ages of eighteen and forty-seven and had lived here nineteen years on average (range: six to thirty-five years). Their mean age of acquisition of English was 20.76 years (range: eighteen to fifty). They had all received schooling in Mexico. On all self-rating measures, Spanish was their stronger language, and English was their weaker language.

**Younger adult native speakers in Mexico** (*n = 20*). Younger adult native speakers (seven male, thirteen female), ages eighteen to twenty-five (mean 21.05), attending the Universidad de Guanajuato were tested as a comparison group for the young adult heritage speakers. Spanish was their native language, and they spoke it in their daily life. They knew some English as a second language, with age of acquisition ranging from twelve to twenty-four years.

**Older adult native speakers in Mexico** (*n = 21*). Older adult native speakers (seven male, fourteen female) were recruited in Guanajuato as a comparison group for the first-generation immigrants. They ranged in age from forty to sixty-one (mean 46.9). Spanish was their native language and they used it every day. Some had minimal knowledge of English.

Table 2 describes basic information about the five Spanish-speaking groups and includes age at time of testing, age of acquisition of Spanish (AoA Spanish), age of acquisition of English (AoA English), length of residence in the US (LoR US), length of residence in Mexico (LoR Mexico), self-ratings in English (scale from one to five, where 5 = native ability), self-ratings in Spanish, self-ratings by skill (listening, speaking, reading, writing) in each language, and mean, standard deviation (*SD*), and range on the written proficiency test, as well as information about what language they prefer to speak (Spanish, English, both, it depends).

The heritage speakers rated their English higher than their Spanish, while the adult immigrants and the Mexican native speakers (younger and older) rated their Spanish higher than their English. The five groups differed statistically on their self-ratings of
English ($F(4,113) = 108.2, p < 0.0001$), their self-ratings of Spanish ($F(4,113) = 13.89, p < 0.0001$), and on the written Spanish proficiency test ($F(4,113) = 6.78, p < 0.0001$). According to Tukey post-hoc tests, the two groups of heritage speakers did not differ on their self-ratings of English, but differed from the three other groups (all $p$’s < 0.0001). The same pattern of significance obtained for self-ratings of Spanish: the two heritage speaker groups did not differ from each other but they differed from the other three groups: adult immigrants, younger native-speaker Mexicans, and older native-speaker Mexicans (all $p$’s < 0.0001). The mean accuracy scores on the written Spanish proficiency test did not differ statistically among the three US groups. That is, the heritage speakers and the adult immigrants obtained similar accuracy scores according to Tukey (all $p$s > 0.6) and were less accurate than the two groups of native speakers (both $p$s < 0.05). These results show that age of acquisition of onset of bilingualism (or AoA English) is not significant for the Spanish heritage speakers recruited for this study. There are apparent signs of attrition in the first-generation adult immigrants from Mexico as compared to the native speakers tested in Mexico.

**Experimental task.** We administered an acceptability-judgment task, where we manipulated suppliance and omission of a with different arguments and predicates. Because reading and writing are the weakest skills in the heritage speakers, the acceptability-judgment task was self-paced and bimodal, with stimulus presentation in visual and auditory modality (cf. Montrul and Bowles (2009), who used a written acceptability-judgment test). It could be argued that this type of task is not ideal for testing knowledge of vernacular varieties in immigrant populations, especially if the speakers are not all exposed to standard varieties through schooling. Furthermore, acceptability-

<table>
<thead>
<tr>
<th>UNITED STATES GROUPS</th>
<th>MEXICO GROUPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIMULTANEOUS</td>
<td>SEQUENTIAL</td>
</tr>
<tr>
<td>HSs</td>
<td>HSs</td>
</tr>
<tr>
<td>N</td>
<td>32</td>
</tr>
<tr>
<td>age at testing</td>
<td>20.6</td>
</tr>
<tr>
<td>AoA Spanish</td>
<td>at birth</td>
</tr>
<tr>
<td>LoR US (years)</td>
<td>20.6</td>
</tr>
<tr>
<td>LoR Mexico</td>
<td>—</td>
</tr>
<tr>
<td>Spanish feels like an L1</td>
<td>46.9%</td>
</tr>
<tr>
<td>Spanish feels like an L2</td>
<td>53.1%</td>
</tr>
<tr>
<td>self-rating English (1–5)</td>
<td>4.7</td>
</tr>
<tr>
<td>self-rating Spanish (1–5)</td>
<td>4.1</td>
</tr>
<tr>
<td>English listening</td>
<td>4.9</td>
</tr>
<tr>
<td>English speaking</td>
<td>4.7</td>
</tr>
<tr>
<td>English reading</td>
<td>4.7</td>
</tr>
<tr>
<td>English writing</td>
<td>4.7</td>
</tr>
<tr>
<td>Spanish listening</td>
<td>4.5</td>
</tr>
<tr>
<td>Spanish speaking</td>
<td>3.9</td>
</tr>
<tr>
<td>Spanish reading</td>
<td>4.0</td>
</tr>
<tr>
<td>Spanish writing</td>
<td>3.3</td>
</tr>
<tr>
<td>Spanish written proficiency test (max = 50)</td>
<td>mean</td>
</tr>
<tr>
<td>SD</td>
<td>8.06</td>
</tr>
<tr>
<td>range</td>
<td>13–49</td>
</tr>
</tbody>
</table>

Table 2. Information about the Spanish-speaking participants.
judgment tasks can require metalinguistic knowledge, and heritage speakers have been shown to possess less metalinguistic awareness than L2 learners, who are typically very literate (Bowles 2011, Montrul 2011). At the same time, other studies have found that many heritage speakers can do grammaticality judgments successfully (Montrul & Bowles 2009, Montrul et al. 2012). We acknowledge that acceptability judgments have their limitations and should not be the only tool for examining linguistic knowledge in variable language environments. Yet, as our results show, we found that our heritage speakers discriminated between grammatical and ungrammatical sentences, accepting some critical ungrammatical sentences and rejecting grammatical ones. In other words, our results do not show a pattern of inability to detect ungrammaticality and overacceptance of grammatical and ungrammatical sentences, as some may fear. At the same time, we note that our study included other oral production measures as well as aural and written comprehension tasks, and that the patterns obtained in the tasks reported here converge on the patterns found with the other measures with Spanish and Hindi analyzed to date.

The bimodal acceptability-judgment task consisted of 140 sentences (seventy-five grammatical, sixty-five ungrammatical) divided into twenty-eight types, with five token sentences per type, depending on the structure. Fourteen sentence types are the focus of the study and the rest were fillers. We included sentences with animate/inanimate, definite/indefinite direct objects, indirect objects, and dative subjects, where the presence and omission of the preposition a were manipulated, as in Table 3.

1. **Animate direct objects**
   a. definite with DOM
      Marina vio a Madonna.
      ‘Marina saw Madonna.’
   b. definite with no DOM
      *Julia vio Shakira.
      ‘Julia saw Shakira.’
   c. indefinite with DOM (optional)
      Mi abuelo conoció a unos pintores.
      ‘My grandfather knew some painters.’
   d. indefinite with no DOM
      Antonia vio una gitana.
      ‘Antonia saw a gypsy.’

2. **Inanimate direct objects**
   a. definite with DOM
      *Josefa eligió al auto rojo.
      ‘Josefa chose the red car.’
   b. definite with no DOM
      Ignacio visitó la universidad.
      ‘Ignacio visited the university.’
   c. indefinite with DOM
      *Adriana siguió a unas direcciones.
      ‘Adriana followed some directions.’
   d. indefinite with no DOM
      Pablo escuchó un concierto.
      ‘Pablo listened to a concert.’

3. **Indirect objects**
   a. clitic-doubled indirect object
      Sara le envió un mensaje a su maestra.
      ‘Sara cl sent a message to her teacher.’
   b. clitic-doubled indirect object with no a
      *Juan le dio su padre dinero.
      ‘Juan cl gave his father money.’
   c. double object
      *Javier dio su padre dinero.
      ‘Javier gave his father money.’

4. **Dative experiencers**
   a. dative experiencer
      A Juan le gusta el fútbol.
      ‘(To) Juan cl likes soccer.’
   b. dative experiencer with clitic and no a
      *Mariela le interesaba la arquitectura.
      ‘Mariela is interested in architecture.’
   c. nominative experiencer
      *Héctor gusta el vóley.
      ‘Héctor likes volleyball.’

Table 3. Target sentences included in the Spanish bimodal acceptability-judgment task.
All sentences were recorded in a soundproof room by a female native speaker of Mexican Spanish. The task was administered through a web interface, with each sentence presented in written Spanish and with an audio player below. Participants were instructed to read each sentence and play the sound file before rating each sentence on a scale of one to four (1 = completely unacceptable, 4 = perfectly acceptable). A zero (0) option meaning ‘I don’t know’ was presented independently of the scale, not as a middle point. This was done to avoid conflating a scale of acceptability with a scale of certainty; some participants may choose the zero option because they are uncertain of their judgments, while others may be fairly certain that the sentence is neither very acceptable nor very unacceptable. The task was self-paced. All tests were conducted in the presence of a Spanish-speaking research assistant.

Results. Mean acceptability ratings for each sentence type listed in Table 3 were entered into a repeated-measures ANOVA, with sentence type and grammaticality (grammatical sentences, ungrammatical sentences) as within-subjects factors, and group (simultaneous bilingual heritage speakers, sequential bilingual heritage speakers, adult immigrants, younger Mexican native speakers, older Mexican native speakers) as between-subjects factor. All main effects were highly significant at $p < 0.0001$ (sentences: $F(6,113) = 247.4$; grammaticality: $F(1,113) = 668.8$; and group: $F(4,113) = 4.14$), and so were the interactions (all $ps > 0.05$): sentences by group ($F(23,113) = 5.18$), grammaticality by group ($F(4,113) = 4.08$), sentences by grammaticality ($F(6,113) = 431.4$), and sentences by grammaticality by group ($F(24,113) = 9.08$).

A main finding is that the three US groups did not differ from each other, and they differed significantly from the two groups of Mexican speakers when collapsed into two groups by region: the US groups vs. the Mexico groups ($F(1,116) = 15.84$, $p < 0.0001$). At the same time, the simultaneous bilingual and the sequential bilingual heritage speakers gave similar acceptability ratings, suggesting that, at least for Spanish, the age of onset of bilingualism is not relevant for DOM. Therefore, the rest of the analyses were based on participants divided by the two main regions, US vs. Mexico.

Figure 1 displays the mean acceptability ratings for sentences with animate direct objects, definite and indefinite, with and without DOM (sentences 1a, 1b, 1c, and 1d in Table 3). The key sentence type is 1b, since animate, specific (definite) direct objects are ungrammatical without DOM. The two native-speaker groups from Mexico assigned acceptability ratings of 1.1 and 1.2 to these sentences, as expected. By contrast, the three US groups accepted these ungrammatical sentences significantly more often, with ratings close to 3 on the acceptability scale. The difference between the US groups and the Mexico groups was significant ($F(1,116) = 138.32$, $p < 0.0001$). The strength of the acceptability ratings of the simultaneous bilinguals, the sequential bilinguals, and the adult immigrants did not differ ($F(2,74) = 0.24$, $p = 0.78$). This result suggests that DOM is subject to attrition in first-generation immigrants as well. All groups were very accepting of all the other grammatical sentences, although the US groups gave slightly lower ratings to indefinite objects than the native-speaker groups (with DOM: $F(1,116) = 6.34$, $p = 0.013$; with no DOM: $F(1,116) = 16.01$, $p < 0.0001$).

Figure 2 shows the acceptability ratings for inanimate direct objects, definite and indefinite, which are typically unacceptable with DOM. For the four sentence types, the difference in acceptability ratings between the two Mexico groups and the three US groups was significant. That is, the two groups of heritage speakers and the adult immigrants rated the grammatical sentences with no DOM as significantly less acceptable than the native speakers: inanimate definite objects with no DOM: $F(1,116) = 9.88$, $p = 0.002$; inanimate indefinite objects with no DOM: $F(1,116) = 9.87$, $p = 0.002$. Re-
call that inanimate objects are unmarked, but a diachronic tendency in Latin America is the gradual expansion of DOM to inanimates. This expansion is evident in the results of the native speakers, who rated inanimate definite and inanimate indefinite objects with DOM at two, slightly higher than definite animate objects with no DOM (1.1 and 1.2 in Fig. 1). Interestingly, the US groups also show this tendency, and assigned higher ratings than the native speakers to inanimate objects with DOM, both definite ($F(1,116) = 11.18, p < 0.0001$) and indefinite ($F(1,116) = 13.8, p < 0.0001$).

Figure 3 shows the results for indirect objects, which are always marked with the preposition a regardless of animacy, definiteness, or specificity. Clitic-doubled indirect objects (3a in Table 3) were compared to ungrammatical sentences with the dative clitic and no preposition (3b) and to ungrammatical sentences with no clitic and no preposi-
tion, following the structure of double objects in English. All groups were equally accepting of grammatical sentences ($F(4,116) = 1.12, p = 0.35$), but the US groups and the Mexico groups differed on their acceptability ratings of the ungrammatical sentences with omitted $a$ ($F(1,116) = 26.37, p < 0.0001$) and their acceptability ratings of double objects ($F(1,116) = 9.71, p = 0.002$). The three US groups were more accepting of sentences with the dative clitic and no preposition than of sentences with double objects (paired-sample $t$-tests, $t(77) = 8.04, p < 0.0001$), perhaps because they perceive the dative clitic as the main marker of dative case rather than the preposition. Without the clitic and the preposition, the sentences were judged to be largely unacceptable.

![Figure 3. Mean acceptability ratings on indirect objects, Spanish bimodal acceptability-judgment task.](image)

Figure 4 shows the results for dative experiencer subjects. The grammatical sentences carry an obligatory dative clitic and the preposition $a$, as in 4b in Table 3. As with indirect objects, the ungrammatical sentences were of two types: sentences with the clitic but missing the preposition, as in 4b, and sentences with nominative experiencers mirroring the structure of English, as in 4c. All groups were equally accepting of the grammatical sentences and largely rejected nominative experiencers, although the mean ratings of the simultaneous bilinguals for nominative experiencers were higher than the ratings of both the younger ($F(1,116) = 3.29, p = 0.49$) and the older Mexican native speakers ($F(1,116) = 4.05, p = 0.014$). All groups, however, were quite accepting of ungrammatical sentences with the clitic but no preposition, including the native speakers in Mexico. However, the ratings of the three US groups were higher than the ratings of the two Mexico groups ($F(1,116) = 8.83, p = 0.004$). This phenomenon in native speakers has been attested in other studies (Pascual y Cabo 2013) and suggests that the preposition $a$ may be in the process of being lost in these constructions in native speakers. We see that the tendency is more advanced in the bilinguals living in the US.

We compared acceptability of ungrammatical sentences with omitted $a$—animate, specific direct objects, indirect objects, and dative experiencer subjects—to see whether $a$-omission occurs regardless of syntactic context. We found that ungrammatical animate, specific direct objects without DOM were rated higher than ungrammatical indirect objects with no $a$ by the three US groups ($t(77) = 6.94, p < 0.0001$). This suggests that the speakers are sensitive to syntactic context, and that the erosion of $a$ cannot be due mainly to its low acoustic salience.
To summarize, this study confirms that DOM omission with animate, specific direct objects is a robust finding in Spanish heritage speakers when compared to monolingual speakers in Mexico. The new finding in this study is that this phenomenon also affects the first generation of immigrants, which suggests that DOM undergoes attrition. Heritage speakers report highest use of the language with their parents and older relatives (first-generation immigrants). Hence, incomplete acquisition in the heritage speakers may be reinforced by emerging patterns in some speakers of the parental generation, especially those who have been in the country for many years. Finally, age of onset of bilingualism in the heritage speakers tested was not a significant factor, confirming the trend attested by Montrul and Sánchez-Walker (2013) with children.

5.2. Study 2: Hindi.

Participants. We tested 101 total Hindi speakers: fifty-nine first- and second-generation immigrants from India living in the central Illinois, Chicago, New Jersey, and Austin, Texas, areas; and forty-two Hindi native speakers in New Delhi and neighboring areas in India. The participants were divided into four groups: two in the US—simultaneous bilingual Hindi heritage speakers, and first-generation adult immigrants from India—and two in India: younger adult native speakers, and older adult native speakers.

All of the Hindi speakers completed the same language-background questionnaire described for Spanish. Because we did not have access to any standardized proficiency measure in Hindi, we developed a cloze-based written Hindi proficiency test. We selected an advanced text in Hindi and extracted every seventh word to yield forty blanks; hence, the maximum number of points was forty. The test was piloted and normed with fluent Hindi speakers before it was used for this study.

Simultaneous bilingual Hindi heritage speakers (n = 26). The Hindi heritage speakers were young adults between the ages of eighteen and twenty-five (mean 21.5). Twenty-six were simultaneous bilinguals exposed to English and Hindi in early childhood, who were born in the US or came to the US before age three. There were only seven heritage speakers who were sequential bilinguals and immigrated to the US later in childhood, after age seven. Because this group is very small and cannot be statistically compared to the group of twenty-six simultaneous bilinguals, we omitted the re-
sults of the sequential bilinguals and focused only on the patterns exhibited by the twenty-six simultaneous bilinguals. As a result, we do not address the issue of age of onset of bilingualism in Hindi.

All of the heritage speakers were born to Hindi-speaking parents (both father and mother). Due to the multilingual situation in India, all of the parents spoke both English and Hindi and thirteen additionally spoke another regional South Indian language (Punjabi, Gujarati, Marathi, Multani, Sandhi, Maithili, Bangla, Haryanvi, Mewadi, Tamil, Bihari, Himachli, Telugu). Most of the heritage speakers (85%) spoke English and Hindi before age five; 15% spoke only Hindi. Except for one participant who reported hearing only Hindi in early childhood, all others heard both Hindi and English, as well as the other South Indian languages spoken at home. All of them were schooled in English in the US, and none of them received instruction in Hindi as a heritage, second, or foreign language at school. Use of Hindi was mostly with the parents and to a more limited extent with the siblings. All of the heritage speakers had traveled to India.

**First-generation adult immigrants from India (n = 26).** This group consisted of adults from India between the ages of nineteen and sixty-four (mean 38.69). Seven participants were recent arrivals in the US (age nineteen) and the rest were between forty and sixty-four years old and had lived in the US between four and thirty-five years (mean 12.7). On average, they have lived in India for twenty-three years. They were highly educated, all of them bilingual in English and Hindi. Although 65% spoke both Hindi and English before age five because they grew up in India, their reported use of Hindi at present was higher than that reported by the heritage speakers in the US. These adults also spoke other regional languages from India (Punjabi, Gujarati, Marathi, and others).

**Younger adult native speakers in India (n = 22).** This group consisted of young adults between the ages of eighteen and twenty-five (mean 20.72) residing in Delhi who had not lived in the US. Sixteen were college students and six were graduate students. They were also bilingual in Hindi and English, and their parents spoke Hindi, English, and in some cases Punjabi.

**Older adult native speakers in India (n = 20).** Adults between the ages of forty and sixty (mean 46.05) residing in India were tested as a comparison group to the adult Hindi-speaking immigrants. They all knew English as well, and some were speakers of other regional languages (Punjabi, Gujarati, Tamil, and Himachli).

Table 4 gives basic information about the four groups of Hindi speakers. The profiles of the Hindi speakers match the description of the high status of English in India and the characteristics of the Hindi-speaking diaspora in the US described by Gambhir and Gambhir (2013). The four groups were bilingual in English and Hindi, and even the adult immigrants and the two groups of Hindi speakers tested in India were exposed to English in childhood. The four groups self-rated their overall abilities in English very high, but there was a significant difference between the groups ($F(3,93) = 5.46, p = 0.002$), largely because the heritage speakers rated their English significantly higher than the other three groups (all $ps < 0.05$). The four groups also differed significantly in their self-ratings of Hindi ($F(3,93) = 28.78, p < 0.0001$): the heritage speakers rated their Hindi lower than the other three groups (all $ps < 0.05$). Hindi is the weaker language of the heritage speakers as confirmed by a paired-sample $t$-test comparing English and Hindi ($t(25) = 8.17, p < 0.0001$). For the three other groups, by contrast, their abilities in the two languages were balanced, and there were no statistical differences in their own self-ratings of English compared to Hindi according to a paired-sample $t$-test (adult immigrants: $t(25) = 1.44, p = 0.16$; younger native Hindi speakers: $t(21) = 1.19$, etc.)
older native Hindi speakers: \( t(19) = 0.63, p = 0.54 \). Note also that there were differences for the Hindi heritage speakers in their self-ratings of the four skills \( F(3,22) = 50.32, p < 0.0001 \): they rated listening and speaking skills in Hindi significantly higher than reading and writing (pairwise comparisons \( p < 0.0001 \)). Most of them had low literacy in Hindi. Finally, the results of the written Hindi proficiency test showed highly significant group differences \( F(3,93) = 51.70, p < 0.0001 \). According to Tukey post-hoc tests, the heritage speakers scored lower than the three other groups \( (p < 0.0001) \), and the adult immigrants and the two native-speaker groups in India did not differ from each other in Hindi written proficiency. These scores show that although the heritage speakers are less proficient in Hindi than the other three groups, there are no apparent signs of attrition in the first-generation adult Hindi speakers.

**Experimental task.** Because the heritage speakers’ strongest skills in Hindi are listening and speaking, we also administered an acceptability-judgment task with visual and auditory stimulus presentation. All sentences were recorded by a female native speaker of Hindi, following the same procedure described for Spanish.

The bimodal acceptability-judgment task consisted of 216 sentences (108 grammatical, 108 ungrammatical) divided into twenty-four types, with between five and twelve

#### Table 4. Information about the Hindi-speaking participants.

<table>
<thead>
<tr>
<th></th>
<th>UNITED STATES GROUPS</th>
<th>INDIA GROUPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BILINGUAL HSs</td>
<td>ADULT IMMIGRANTS</td>
</tr>
<tr>
<td>( N )</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>age at testing</td>
<td>21.5</td>
<td>38.7</td>
</tr>
<tr>
<td>AoA Hindi at birth</td>
<td>1.88</td>
<td>4.53</td>
</tr>
<tr>
<td>AoA English at birth</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>LoR US (years)</td>
<td>23.3</td>
<td>12.7</td>
</tr>
<tr>
<td>LoR India</td>
<td>—</td>
<td>22.9</td>
</tr>
<tr>
<td>Preferred language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hindi</td>
<td>3%</td>
<td>15%</td>
</tr>
<tr>
<td>English</td>
<td>50%</td>
<td>8%</td>
</tr>
<tr>
<td>both</td>
<td>47%</td>
<td>77%</td>
</tr>
<tr>
<td>Hindi feels like an L1</td>
<td>35%</td>
<td>89%</td>
</tr>
<tr>
<td>Hindi feels like an L2</td>
<td>65%</td>
<td>11%</td>
</tr>
<tr>
<td>self-rating English (1–5)</td>
<td>5.0</td>
<td>4.7</td>
</tr>
<tr>
<td>self-rating Hindi (1–5)</td>
<td>3.3</td>
<td>4.9</td>
</tr>
<tr>
<td>English listening</td>
<td>5.0</td>
<td>4.5</td>
</tr>
<tr>
<td>English speaking</td>
<td>5.0</td>
<td>4.5</td>
</tr>
<tr>
<td>English reading</td>
<td>5.0</td>
<td>4.6</td>
</tr>
<tr>
<td>English writing</td>
<td>4.9</td>
<td>4.4</td>
</tr>
<tr>
<td>Hindi listening</td>
<td>3.5</td>
<td>4.7</td>
</tr>
<tr>
<td>Hindi speaking</td>
<td>2.9</td>
<td>4.7</td>
</tr>
<tr>
<td>Hindi reading</td>
<td>1.9</td>
<td>4.4</td>
</tr>
<tr>
<td>Hindi writing</td>
<td>1.6</td>
<td>4.0</td>
</tr>
<tr>
<td>Hindi written proficiency test (max = 40)</td>
<td>mean 23.7</td>
<td>38.2</td>
</tr>
<tr>
<td></td>
<td>SD 9.01</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>range 10–40</td>
<td>32–40</td>
</tr>
</tbody>
</table>

\(^6\) We have also calculated proficiency from speech samples, taking into account words per minute, lexical diversity, lexical complexity, and so forth, and the differences between the groups are the same as revealed in this task.
token sentences per type, depending on the structure. The task included twelve sentence types (sixty sentences total) with animate and with inanimate, specific and nonspecific direct objects, indirect objects, and dative subjects, where the presence and omission of the case marker \(-ko\) were manipulated. The rest of the filler sentences tested agreement and ergativity. The target sentences are listed in Table 5.

1. ANIMATE DIRECT OBJECTS
   a. specific with DOM  
   b. specific with no DOM  
   c. nonspecific with DOM (optional)  
   d. nonspecific with no DOM

2. INANIMATE DIRECT OBJECTS
   a. specific with DOM  
   b. specific with no DOM  
   c. nonspecific with DOM  
   d. nonspecific with no DOM

3. INDIRECT OBJECTS
   a. with \(-ko\)  
   b. without \(-ko\)

4. DATIVE EXPERIENCERS
   a. with \(-ko\)  
   b. without \(-ko\)

Table 5. Target sentences included in the Hindi bimodal acceptability-judgment task.

The task was self-paced and administered through a web interface, with each sentence presented in Hindi script with an audio player below. Participants were instructed to read each sentence and play the sound file before rating each sentence on a scale of one to four (1 = completely unacceptable, 4 = perfectly acceptable). This was particularly helpful for the heritage speakers. A Hindi-speaking research assistant was present during the testing.

Results. Mean acceptability ratings for each sentence type illustrated in Table 5 were entered into an omnibus ANOVA with repeated measures, with sentence type as within-subjects factor and group (simultaneous bilingual heritage speakers, adult immigrants, younger Indians, older Indians) as between-subjects factor. Because the twelve target sentence types were not fully balanced in terms of grammaticality (some were grammatical, some were ungrammatical, and some varied in grammaticality depending on context), grammaticality was assessed in follow-up analyses. There was a significant main effect for sentences \((F(11,93) = 143.59, p < 0.0001)\), a significant main effect for group \((F(3,93) = 14.05, p < 0.001)\), and a significant sentence-by-group interaction \((F(3,93) = 16.45, p < 0.0001)\) followed up in subsequent statistical analyses.

Figure 5 shows the results of animate (human) direct objects. Specific direct objects are grammatical with DOM \((-ko\) marking) and ungrammatical without it (see sentences
The four groups rated the grammatical sentences with *ko as highly acceptable, although the heritage speakers rated them less acceptable than the other three groups (Tukey, $p < 0.01$). The crucial ungrammatical sentence type is animate, specific direct objects without *ko. We see in Fig. 5 that the heritage speakers differed significantly from the three other groups ($F(3,96) = 42.93$, $p < 0.0001$), assigning higher overall ratings to these ungrammatical sentences (2.7) than the three other groups (1.3, 1.1, and 1.1), as predicted. Nonspecific direct objects are grammatical without *ko and ungrammatical with *ko, but this difference depends on context. Here, our participants showed variability in their judgments rather than more categorical ratings, most likely because our test required the sentences to be judged in isolation, without context, and it is possible that some speakers took into account possible contextual factors in judging these sentences. Still, the difference between nonspecific animate direct objects with and without *ko was significant for all groups ($F(1,93) = 10.60$, $p < 0.0001$). The heritage speakers judged sentences with nonspecific animate objects with DOM as significantly less acceptable than the three other groups (all $ps < 0.05$), whose ratings did not differ as per Tukey post-hoc tests.

![Figure 5](image.png)

**Figure 5.** Mean acceptability ratings on animate direct objects, Hindi bimodal acceptability-judgment task. (SIMBIL HS: simultaneous bilingual heritage speakers, AD INDIA IM: adult immigrants from India, Y INDIANS: younger adult Hindi native speakers in India, O INDIANS: older adult Hindi native speakers in India.)

Figure 6 shows the results of the sentences with inanimate direct objects. A repeated-measures ANOVA comparing the four sentence types by group showed a main effect for sentences ($F(3,96) = 19.56$, $p < 0.0001$) but no main effect for group. Nonspecific inanimate direct objects with DOM (*ko marking) were consistently rated as less acceptable than the other sentences by all groups, although the ratings were not categorical, as in the case of animate, specific direct objects.

Finally, Figure 7 shows the results of *ko omission with indirect objects and dative experiencer subjects, which are ungrammatical. A repeated-measures ANOVA with sentences and grammaticality as within-subjects factors and group as between-subjects factor showed a main effect for sentences ($F(1,93) = 37.03$, $p < 0.0001$), grammaticality ($F(1,93) = 1401.0$, $p < 0.0001$), and group ($F(3,93) = 5.04$, $p < 0.003$), and a sentence-by-grammaticality-by-group interaction ($F(3,93) = 8.93$, $p < 0.0001$). Once again, the heritage speakers were different from all of the other groups (all $ps < 0.05$), especially
with grammatical and ungrammatical indirect objects. The heritage speakers accepted omission of \(-ko\) and were less accepting of \(ko\)-marked indirect objects than the other three groups (all \(ps < 0.0001\)). The heritage speakers also accepted omission of \(-ko\) with dative experiencers, and were less accepting of \(ko\)-marked dative experiencers than the other three groups (all \(ps < 0.0001\)). It is also worth noting that the adult immigrants were more accepting of ungrammatical sentences with no \(-ko\) with dative experiencers than the Hindi speakers in India. In fact, the difference between the four groups for this particular sentence type was significant \((F(3,93) = 23.22, p < 0.0001)\). According to Tukey post-hoc tests, the first-generation immigrants were more accepting of these sentences than the younger \((p < 0.0001)\) and older adult native speakers in India \((p < 0.0002)\). Yet their ratings were lower than those of the heritage speakers \((p < 0.015)\). At the individual level, seven speakers in the immigrant group and fourteen in the heritage speakers group had mean ratings higher than the mean ratings of the native Hindi speakers in India. Although we found no evidence of attrition of \(-ko\) omission with animate, specific objects in the first-generation immigrant group, there is emerging attrition of \(-ko\) with dative subjects.
Comparing the heritage speakers’ acceptability ratings on omission of -ko with animate, specific direct objects, indirect objects, and dative experiencers revealed a significant difference between the sentence types ($F(2,26) = 15.68$, $p < 0.0001$). Pairwise comparisons indicated that the heritage speakers accepted -ko omission with animate, specific direct objects (2.7) and dative subjects (2.4) significantly more than ko-omission with indirect objects (1.9) ($p < 0.0001$). Animate, specific direct objects also received marginally statistically higher ratings than dative experiencers without -ko, suggesting that the heritage speakers do not omit -ko across the board either: they are sensitive to its syntactic and semantic distribution.

To summarize, the heritage speakers display signs of incomplete acquisition or L1 attrition in Hindi in the domain of DOM, as judged by acceptability ratings. (See Montrul et al. 2012 for similar patterns in oral production.) There are no apparent signs of attrition of DOM in the adult immigrant group, unlike what was found for the Spanish-speaking Mexican immigrants in study 1. However, there seems to be some evidence of attrition with dative subjects. We also found that -ko marking is omitted more with direct objects than with indirect objects and dative experiencers, similar to what we found in Spanish.

5.3. Study 3: Romanian.

Participants. We tested 125 total Romanian speakers: seventy-four first- and second-generation immigrants from Romania living in the Central Illinois and Chicago area, and fifty-one Romanian native speakers in Brașov, Romania. The participants were divided into five groups: three in the US—simultaneous bilingual Romanian heritage speakers, sequential bilingual Romanian heritage speakers, and first-generation adult immigrants from Romania—and two in Romania: younger adult native speakers, and older adult native speakers.

All of the Romanian speakers completed the same language-background questionnaire described for the Spanish and Hindi studies. Because we did not have access to any standardized proficiency measure in Romanian, as with Hindi, we developed a cloze-based written Romanian proficiency test. We selected an advanced text in Romanian and extracted every seventh word to yield twenty-seven blanks; hence, the maximum number of points was twenty-seven. The test was piloted and normed with fluent Romanian speakers before it was used for this study.

Simultaneous bilingual Romanian heritage speakers ($n = 23$). The simultaneous bilingual heritage speakers were young adults between the ages of eighteen and twenty-three (mean 20.34) living in the US who had been born to Romanian parents. Ten of them (43.5%) were born in Romania and immigrated with their parents before the age of five, and the rest (56.5%) were born in the US. One participant had a postgraduate degree, and the rest were completing or had completed college. Except for one mother who spoke Farsi and Romanian, all other parents spoke Romanian and some English, or Romanian and another European language as a second or third language (Hungarian, Russian, German, Flemish, French, Italian). Fifteen (65%) of the simultaneous bilinguals spoke English and Romanian before age five, the rest only Romanian. Only eight participants (35%) lived at home with a Romanian-speaking grandparent, and all of them reported speaking only Romanian with their grandparents. All of them attended elementary, middle, and high school in the US with instruction in English. None of them had access to Romanian as a foreign or second language at school.

Sequential bilingual Romanian heritage speakers ($n = 19$). This group consisted of Romanians between the ages of eighteen and thirty-five (mean 23.4) who immi-
grated to the US with their Romanian parents between the ages of seven and fourteen. They had been living in the US for an average of 12.2 years at the time of testing. Their mean age of acquisition of English as a second language (in the US) was nine, so this group had a substantial period of monolingualism in Romanian before becoming bilingual in the US. Like the parents of the simultaneous bilinguals, the parents of the sequential bilinguals also spoke mainly Romanian and one or more European languages as second or third languages, such as Hungarian, German, Russian, and French. Seven heritage speakers (37%) attended elementary school in Romania, and 63% attended elementary school in the US. All of them attended middle and high school in the US with no access to Romanian as a second or heritage language beyond the home.

**First-generation adult immigrants from Romania (n = 32).** This group consisted of Romanians between the ages of twenty-two and fifty-eight (mean 39.09) who immigrated after age eighteen and had been living in the US between four and twenty years. They all completed their education in Romania and had knowledge of other European languages, the same as the ones reported by the heritage speakers about their parents.

**Younger adult native speakers in Romania (n = 30).** The younger adult native speakers tested in Romania were between the ages of eighteen and twenty-seven (mean 22.7). They were recruited from Transylvania University of Brașov in Brașov, Romania. They all had some knowledge of English and some other European language, which they had learned as a second language in Romania.

**Older adult native speakers in Romania (n = 21).** The older adult native speakers were also recruited from the city of Brașov, Romania. They were between the ages of forty-one and sixty (mean 49.42). They were all schooled and educated in Romania and spoke some English and other European languages.

Basic descriptive information about the five Romanian-speaking groups is presented in Table 6. All groups were asked to self-assess their linguistic ability in English and in Romanian, since most participants, including the native speakers tested in Romania, had been exposed to some English.

There were significant differences between the groups on their ratings of their English ability ($F(4,124) = 26.54, p < 0.0001$) and their ability in Romanian ($F(2,124) = 10.53, p < 0.0001$) because the two groups of heritage speakers rated their English higher and their Romanian significantly lower than the other three groups. The simultaneous bilinguals and the sequential bilinguals did not differ on their self-ratings of English, but they differed on their self-ratings of Romanian: the simultaneous bilinguals rated their Romanian much lower than the sequential bilinguals, and the difference was significant ($p < 0.0001$). The sequential bilinguals and the adult immigrants did not differ from each other on their Romanian or English self-ratings. As with the other heritage speakers described in this study, the weakest skill in the Romanian heritage speakers was writing and the strongest aural comprehension (listening). Compared to their English, Romanian is the weaker language of the simultaneous and sequential bilingual heritage speakers according to their self-ratings (simultaneous bilinguals: $t(23) = 10.5, p < 0.0001$; sequential bilinguals: $t(19) = 7.39, p < 0.0001$), while for the adult immigrants their abilities in the two languages were balanced and there were no statistical differences in their self-ratings of English compared to Romanian ($t(32) = 1.24, p = 0.22$). For the two Romanian native-speaker groups, their Romanian was stronger than their English, as expected (younger Romanians: $t(30) = 5.93, p < 0.0001$; older Romanians: $t(21) = 7.17, p < 0.0001$).
The results of the written Romanian proficiency test showed highly significant group differences \((F(4,124) = 16.71, p < 0.0001)\). According to Tukey post-hoc tests, the simultaneous bilingual heritage speakers scored lower than the four other groups \((p < 0.0001)\), and the sequential bilinguals scored lower than the adult immigrants and the two native-speaker Romanian groups (all \(ps < 0.05\)). The adult immigrants and the two native-speaker groups in Romania did not differ from each other in their Romanian proficiency as measured by the Romanian cloze test (all \(p s > 0.05\)). These results show that age of onset of bilingualism is significant for the Romanian heritage speakers and that there is no apparent sign of attrition in the first-generation adult immigrants from Romania. We now examine whether these proficiency scores relate to the results of the main task.

**Experimental task.** The main task was a bimodal acceptability-judgment task with stimulus presentation in auditory and visual modality. The sentences were recorded in a soundproof room by a female Romanian native speaker. The task consisted of 150 sentences, seventy grammatical and eighty ungrammatical, divided into thirty sentence types. Twenty sentence types were target and the rest were fillers. The sentences were presented in a randomized list form with a scale ranging from one to four given underneath. The target sentences tested animate and inanimate, specific and nonspecific direct objects where the presence of the preposition \(pe\) was manipulated. Because accusative clitic doubling is the preferred option in Romanian, the same sentence types also appeared with accusative clitic doubling, explaining why there were twenty target sentence types for Romanian. We also tested the use of the preposition \(pe\) with locatives and with indirect objects. The remaining filler sentences included indirect objects and

<table>
<thead>
<tr>
<th>United States Groups</th>
<th>Romania Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous bilinguals</td>
<td>Sequential bilinguals</td>
</tr>
<tr>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>23</td>
<td>19</td>
</tr>
<tr>
<td>AoA Romanian at birth</td>
<td>3.7</td>
</tr>
<tr>
<td>AoA English</td>
<td>17.9</td>
</tr>
<tr>
<td>LoR US (years)</td>
<td>10.6</td>
</tr>
<tr>
<td>Romanian feels like an L1</td>
<td>43.5%</td>
</tr>
<tr>
<td>Romanian feels like an L2</td>
<td>56.5%</td>
</tr>
<tr>
<td>Self-ratings in English (1–5)</td>
<td>4.9</td>
</tr>
<tr>
<td>English listening</td>
<td>3.7</td>
</tr>
<tr>
<td>English speaking</td>
<td>4.8</td>
</tr>
<tr>
<td>English reading</td>
<td>4.8</td>
</tr>
<tr>
<td>English writing</td>
<td>4.7</td>
</tr>
<tr>
<td>Romanian listening</td>
<td>4.1</td>
</tr>
<tr>
<td>Romanian speaking</td>
<td>3.5</td>
</tr>
<tr>
<td>Romanian reading</td>
<td>3.7</td>
</tr>
<tr>
<td>Romanian writing</td>
<td>2.9</td>
</tr>
</tbody>
</table>

**Table 6. Information about the Romanian-speaking participants.**

<table>
<thead>
<tr>
<th>Romanian written proficiency test (max = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
</tr>
<tr>
<td>(SD)</td>
</tr>
<tr>
<td>range</td>
</tr>
</tbody>
</table>

The sentences tested animate and inanimate, specific and nonspecific direct objects where the presence of the preposition \(pe\) was manipulated. Because accusative clitic doubling is the preferred option in Romanian, the same sentence types also appeared with accusative clitic doubling, explaining why there were twenty target sentence types for Romanian. We also tested the use of the preposition \(pe\) with locatives and with indirect objects.
Dative experiencer subjects marked with dative case and clitic-doubled. The target sentences analyzed in the present study are listed in Table 7.

### Table 7. Target sentences included in the Romanian bimodal acceptability-judgment task.

1. **ANIMATE DIRECT OBJECTS**
   - a. specific with DOM Roxana a văzut pe Madonna.  
     ‘Roxana has seen/saw Madonna.’
   - b. specific with no DOM  
     *Marina a văzut Madonna.  
     ‘Marina has seen/saw Madonna.’
   - c. nonspecific with DOM (optional) Claudia a văzut pe o femeie.  
     ‘Claudia has seen/saw a woman.’
   - d. nonspecific with no DOM Georgiana a văzut o femeie.  
     ‘Georgiana has seen/saw a woman.’

2. **CLITIC-DOUBLED ANIMATE DIRECT OBJECTS**
   - a. specific with DOM Angelica a văzut-o pe Madonna.  
     ‘Angelica has seen/saw Madonna.’
   - b. specific with no DOM  
     *Anca a văzut-o Madonna.  
     ‘Anca has seen/saw Madonna.’
   - c. nonspecific with DOM (optional) Elisabeta a văzut-o pe o femeie.  
     ‘Elisabeta has seen/saw a woman.’
   - d. nonspecific with no DOM  
     *Ecaterina a văzut-o o femeie.  
     ‘Ecaterina has seen/saw a woman.’

3. **INANIMATE DIRECT OBJECTS**
   - a. specific with DOM Ileana a văzut casa.  
     ‘Ileana has seen/saw a house.’
   - b. specific with no DOM  
     Roxana a văzut casa.  
     ‘Roxana has seen/saw the house.’
   - c. nonspecific with DOM (optional) Doina a văzut o casă.  
     ‘Doina has seen/saw a house.’
   - d. nonspecific with no DOM  
     *Ecaterina a văzut-o o femeie.  
     ‘Ecaterina has seen/saw a woman.’

4. **INDIRECT OBJECTS**
   - a. with pe Claudia se bazează foarte mult pe mama sa.  
     ‘Claudia relies very much on her mother.’
   - b. without pe  
     *Cezara se bazează foarte mult mama sa.  
     ‘Cezara relies very much on her mother.’

5. **LOCATIVES**
   - a. with pe Cartea este pe masă.  
     ‘The book is on the table.’
   - b. without pe  
     *Cartea este masă.  
     ‘The book is table.’

As with Spanish and Hindi, the task was presented through a web interface. Participants were instructed to read each sentence and play the sound file before rating each sentence on scale of one to four (1 = completely unacceptable, 4 = perfectly acceptable). Romanian-speaking research assistants conducted the testing.

**Results.** Mean acceptability ratings for the target sentences were submitted to a repeated-measures ANOVA, with sentences as within-subjects factor and group as between-subjects factor. The analysis revealed a main effect for sentences ($F(15,118) = 475.01$, $p < 0.0001$) and group ($F(4,118) = 13.75$, $p < 0.0001$), and a sentences-by-group interaction ($F(4,118) = 14.31$, $p < 0.0001$). The post-hoc Tukey tests confirmed that the ratings of the simultaneous and sequential bilinguals differed from each other ($p < 0.002$), and they differed from those of the adult immigrants (both $ps < 0.05$). The adult immigrants and the native speakers in Romania did not differ from each other (all $p$s $> 0.5$).
Figure 8 shows the results of animate, specific direct objects, and Figure 9 animate, specific direct objects with accusative clitic doubling. A repeated-measures ANOVA focusing on these four sentence types without clitic doubling (Fig. 8) confirmed the significant sentences-by-group interaction ($F(3,118) = 14.21, p < 0.001$). The simultaneous and sequential bilinguals did not differ from each other ($p = 0.50$), but the two groups differed from the other three groups at $p < 0.0001$. Sentences with animate direct objects marked with pe received higher acceptability ratings from the two groups of heritage speakers than from the three other groups. Figure 9 shows that the native speakers prefer animate, specific direct objects with pe-marking when they have accusative clitic doubling, which explains why the acceptability ratings for grammatical animate, specific direct objects with pe are lower in Fig. 8 than in Fig. 9 for all groups, except for the two groups of heritage speakers (simultaneous bilinguals: $t(23) = 0.40, p = 0.68$; sequential bilinguals: $t(19) = 1.8, p < 0.088$). But the sentences that are critical to our study—animate, specific direct objects without pe-marking or DOM (ungrammatical)—showed significant differences at the $p < 0.0001$ level among the groups. That is, the simultaneous bilinguals were more accepting of these ungrammatical sentences than the sequential bilinguals ($p < 0.0001$), confirming an effect of age of onset of bilingualism for the erosion of DOM in Romanian. At the same time, ungrammatical sentences with animate, specific direct objects without DOM with clitic doubling in Fig. 9 received lower acceptability ratings than the sentences without clitic doubling in Fig. 8 (simultaneous bilinguals: $t(23) = 5.14, p < 0.0001$; sequential bilinguals: $t(19) = 2.95, p < 0.0001$). This pattern suggests that accusative clitic doubling may favor DOM retention with animate, specific direct objects.

With respect to animate, nonspecific direct objects (Fig. 8), the two heritage speaker groups accepted unmarked objects to the same extent and significantly more than the other three groups (all $p s < 0.0001$); the same pattern of significance obtained for nonspecific direct objects marked with DOM, which are ungrammatical depending on context. Comparison of indefinite direct objects with no DOM shows that when these sentences have clitic doubling (Fig. 9) they are clearly ungrammatical for the first-generation adult immigrants and the two groups of Romanian native speakers. By contrast, the two groups of heritage speakers assigned statistically higher acceptability ratings to these sentences than the three other groups (all $p s < 0.0001$).

Figure 10 shows the results for inanimate direct objects, which are ungrammatical with pe marking, and the acceptability ratings are very clear for the first-generation adult immigrants and for the native speakers tested in Romania. By contrast, the simultaneous bilingual heritage speakers assigned higher acceptability ratings to pe-marked specific and nonspecific inanimate objects than all of the other groups, including the sequential bilingual heritage speakers (all $p s < 0.0001$). This pattern suggests some overgeneralization of pe-marking to inanimates as well.

Finally, Figure 11 shows omission of pe with indirect objects and with locatives to see if it occurs to the same extent as omission of pe as an instance of DOM with animate, specific direct objects. The results show that the bilingual heritage speakers gave higher acceptability ratings to the two ungrammatical sentences than all of the other groups (all $p s < 0.0001$). Yet the heritage speakers assigned higher acceptability ratings to animate, specific direct objects without pe (Fig. 8) than to indirect objects without pe (simultaneous bilinguals: $t(24) = 5.53, p < 0.0001$; sequential bilinguals: $t(19) = 6.04, p < 0.0001$) and than to locatives without pe (simultaneous bilinguals: $t(24) = 3.23, p < 0.004$; sequential bilinguals: $t(19) = 3.53, p < 0.002$). Thus, the heritage speakers
discriminate between the semantic and syntactic contexts for *pe*, and *pe*-omission is not accepted across the board.

To summarize, the results of the Romanian groups showed significant acceptance of DOM omission with animate, specific direct objects. The simultaneous bilingual heritage speakers, many of whom were exposed to Romanian and English from an early age, are more affected than the sequential bilinguals who came to the US in middle and late childhood. There was evidence of incomplete acquisition of this phenomenon in the simultaneous bilingual heritage speakers and attrition in the sequential bilingual heritage speakers. In all sentences tested, the first-generation immigrants did not differ in their acceptability judgments from those of the native speakers tested in Roma-
nia. There are no signs of attrition of this phenomenon in the first-generation adult immigrants.

6. Discussion. The study of heritage language speakers, their patterns of language use, and their linguistic varieties, as well as issues related to language policies that affect their language identity, have long been the realm of sociolinguistics (Fishman 1964, 1965, 1966), although heritage language speakers and heritage languages have not always been referred to in this way (Dorian 1981, Gal 1979, Silva-Corvalán 1994, Zentella 1997). We approached our study from a formal linguistic and experimental perspective with the goals of obtaining more converging evidence for the erosion of DOM with animate, specific direct objects in Spanish heritage speakers by investigating whether erosion of DOM is also a feature of other heritage languages that exhibit this property, such as Hindi and Romanian. If DOM is vulnerable in all of these languages to different degrees, another objective was to trace the root of the differences by

![Figure 10. Mean acceptability ratings on inanimate direct objects, Romanian bimodal acceptability-judgment task.](image1)

![Figure 11. Mean acceptability ratings on indirect objects and locatives with and without *pe, Romanian bimodal acceptability-judgment task.](image2)
considering the structural properties and syntactic distribution of the markers in each language: namely, whether Spanish a, Hindi -ko, and Romanian pe are omitted more in cases of DOM than when they case-mark indirect objects, experiencers, or locatives. We also consider external sociolinguistic factors that may account for differences in the three immigrant populations.

Two major innovations in this study were the inclusion of heritage speakers who differ in age of onset of bilingualism (simultaneous vs. sequential) and the inclusion of three comparison groups: one group of first-generation immigrants to the US and two groups of native speakers in the countries of origin closely matched in age, education, and SES to the heritage speakers and the immigrants. DOM in Spanish and ko-marking in Hindi have been shown to be mastered before age three in monolingual children. Investigating this phenomenon in heritage speakers with different onsets of bilingualism allows us to address, although indirectly, the impact of predominantly monolingual and bilingual input and use of the heritage language in early language development and the relative stability of language acquisition at that early age for the grammatical property investigated. Sequential bilinguals who enjoyed a longer period of monolingualism or dominance in their heritage language were hypothesized to have better command of DOM than heritage speakers who were exposed to English and the heritage language since or closely after birth. If DOM is subject to incomplete acquisition or attrition in the heritage speakers as compared to age-matched peers in the countries of origin, testing adult immigrants and age-matched controls in the countries of origin also allowed us to establish whether DOM is subject to attrition in adults as well.

This question, which has not received much attention in the emerging literature on heritage speakers so far, is critical because adult immigrants are typically the main interlocutors and sources of input to heritage speakers. What may look like the result of incomplete acquisition or attrition in childhood in the heritage speakers may, in fact, be a stabilized feature in a new variety of Spanish, for example, similar perhaps to what Sharma (2005) showed for the distribution of definite and indefinite articles in Indian English. Shin and Otheguy (2013) found that current changes taking place in the Spanish of New York with respect to overt subject-pronoun expression are driven by affluent female adult immigrants with several years of residence in New York, who have high levels of English proficiency, more contact with other Spanish-speaking immigrants, and—most relevant to our purposes—more extensive contact with US-born bilinguals, that is, young adult heritage speakers. Hence, at least with increased use of overt subject pronouns in the Spanish of New York, the mothers of many heritage speakers seem to be leading the change. Because attrition in adult immigrants, if it happens, occurs after more than ten years of living in the host country, the language of the parents of young bilingual children who are relatively recent arrivals cannot be attrited, but that of the parents and grandparents of young adult heritage speakers can.\(^7\) We tested our hypotheses by administering a bimodal acceptability-judgment task in each language, in addition to other oral and written tasks reported elsewhere (Montrul 2014, Montrul et al. 2012, Montrul & Sánchez-Walker 2013).

Our findings can be summarized as follows. We confirmed high acceptance of ungrammatical sentences with DOM omission with animate, specific direct objects by the heritage speakers of Spanish, Hindi, and Romanian. Therefore, DOM with animate,
specific direct objects is vulnerable to simplification or loss not only in Spanish but also in other languages that exhibit DOM. We found that acceptance of omission of the markers a, -ko, and pe was higher with animate, specific direct objects than with indirect objects in the three languages and with locatives in Romanian. In Spanish and Hindi there was also significant omission of a and -ko with dative subjects. Thus, the omission of these markers is not blind to structural distribution or the type of case they represent (inherent and meaning-based in DOM and dative experiencers vs. structural case in indirect objects, for example) (Butt 2006, Woolford 2006).

We also found that age of onset of bilingualism, which we were able to address in Spanish and in Romanian, was not significant for Spanish, but it certainly made a difference in Romanian. The simultaneous bilingual heritage speakers of Romanian, exposed to English much earlier than the sequential bilingual heritage speakers who immigrated as children later, were significantly more accepting of ungrammatical sentences, especially those omitting pe-marking with animate, specific direct objects, than the sequential bilingual heritage speakers. At the same time, the sequential bilinguals exhibited higher acceptability ratings of ungrammatical sentences than the adult immigrants, suggesting attrition in the sequential bilingual heritage speakers.

Comparison of the adult immigrant groups in the three languages revealed interesting findings. The immigrants from Mexico showed higher acceptability ratings of ungrammatical sentences with DOM omission in Spanish, compared to the native speakers from Mexico, which we take as a sign of attrition. In fact, there were no statistical differences between the two groups of Spanish heritage speakers and the adult immigrants from Mexico. By contrast, there was no evidence of attrition of DOM in the adult Hindi-speaking immigrants and the adult Romanian immigrants (although a few Hindi speakers showed some attrition patterns with the -ko marking of dative experiencer subjects). The Hindi-speaking immigrants had on average spent fewer years in the US and were five years younger than the Mexican immigrants; these differences could greatly impact their level of attrition. However, the Mexican immigrants and the Romanian immigrants were more comparable in age and length of residence, yet the Romanians did not exhibit attrition.

Table 8 summarizes these overall patterns by counting how many participants in each group accepted ungrammatical sentences with DOM omission more than the native-speaker baselines, as established by the higher end of their mean acceptability ranges (1 = totally unacceptable, 4 = perfectly acceptable). The mean acceptability range for native speakers in Mexico was between 1 for the lowest and 2.17 the highest, for the Hindi speakers in India between 1 and 2.16, and for native speakers in Romania between 1 and 2.17.

<table>
<thead>
<tr>
<th>Language</th>
<th>Simultaneous Bilingual HS</th>
<th>Sequential Bilingual HS</th>
<th>Adult Immigrants</th>
<th>Younger Native Speakers</th>
<th>Older Native Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>25/32 (78%)</td>
<td>15/24 (62%)</td>
<td>13/21 (62%)</td>
<td>0/20 (0%)</td>
<td>0/21 (0%)</td>
</tr>
<tr>
<td>Hindi</td>
<td>19/26 (73%)</td>
<td>—</td>
<td>0/26 (0%)</td>
<td>0/22 (0%)</td>
<td>0/20 (0%)</td>
</tr>
<tr>
<td>Romanian</td>
<td>17/23 (74%)</td>
<td>8/19 (42%)</td>
<td>0/32 (0%)</td>
<td>0/30 (0%)</td>
<td>0/21 (0%)</td>
</tr>
</tbody>
</table>

Table 8. Number and percentage of individuals in each group whose mean acceptability ratings for ungrammatical unmarked DOM animate, specific direct objects was above the highest individual mean acceptability rating for native speakers of the languages in each country.

These individual results mirror the group results: many heritage speakers within each group are affected, and so are 62% of the first-generation adult immigrants from Mex-
Overall, Spanish is more affected than Hindi and Romanian. What could explain the erosion of DOM in the three languages in this manner?

### 6.1. Complexity and Dominant Language Transfer

Let us first focus on the similarities among the three languages, since DOM is subject to erosion in all three: it is thus a vulnerable structure. Two factors that may have contributed to the eradication of DOM in the grammars of many speakers (and which are probably working together, we believe) are the structural complexity of DOM on the one hand, and the influence from English on the other.

Consider grammatical complexity first, which we admit is a very controversial and elusive construct in linguistics. To give an example, McWhorter (2007) measures grammatical complexity along three dimensions: overspecification (degree of grammaticalized semantic distinctions), structural elaboration (number of rules or foundational elements required to generate surface forms), and irregularity (degree of morphological regularity and suppletion). Extending this measurement to animate, specific direct objects, languages that mark DOM are more overspecified and present more grammatical elaboration than grammars that do not mark DOM.

As we have seen, DOM encodes semantic notions such as animacy, definiteness, and specificity, among others, and involves both structural accusative and inherent dative case, according to Torrego’s analysis (see the tree in 32 (§3) and related discussion). Inherent case is handled by the syntax-semantics interface. In contrast to nonmarked objects, DOM objects must move outside of the VP to be overtly case-marked by the DOM marker in the specifier of the vP. Instead of staying in situ close to the vP, they move out due to their referential properties, which likens them to nominative and dative subjects (see also López 2012). Dative indirect objects, by contrast, receive only structural case in the VP. Alternatively, under more current theories, the object may move because it is affected and marks specificity—as in languages that have object shift—and not necessarily for case reasons. (Argumental accusative and dative clitics in Spanish and in Romanian are morphological manifestations of accusative and dative structural case.) Carrying the semantic features of animacy and specificity, DOM is both semantically and syntactically more complex, involving more structure than unmarked objects and indirect objects, in addition to movement. Unmarked inanimate objects and non-specific objects receive case structurally lower in the VP, like all objects in English. O’Grady and colleagues (2011) provide cognitive explanation for the complexity of DOM, capitalizing on the nature of form-meaning mappings. The DOM markers in the three languages studied are not just DOM markers, but have other functions as well. So, we have one form (a, -ko, pe) with multiple meanings and functions (direct object, indirect object, dative subject, locative), or what O’Grady and colleagues call an opaque form-meaning mapping. At the same time, while all indirect objects and dative experiencers are obligatorily marked in these languages, only a subset of direct objects (animate, specific) is obligatorily marked. Such opaque mappings require a very high frequency of occurrence in the input to be reliably acquired and mastered, and adequate amounts of input are precisely what are missing in heritage speakers. This would explain why DOM, which lies at the morphology-syntax-semantics interface, is so vulnerable in heritage languages under the pressure from English (a language with no DOM).

Transfer from the dominant contact language cannot be denied, and a move away from linguistic complexity is further reinforced by English. English has both prepositional datives (Juan gave a book to Maria), analogous to the Spanish dative construction (except for the lack of a dative clitic), and double object datives (Juan gave Maria a book). According to Torrego (1998), the English double object construction and Span-
ish DOM are roughly equivalent in their semantic and syntactic characteristics, with the exception that Spanish marks the affected object overtly while English does not. Furthermore, although English has topicalizations, it does not have clitic left dislocations like Spanish and Romanian, or scrambling like Hindi, structural options that signal specificity and reinforce the presence of DOM in these languages. Finally, English has nominative-marked experiencers with stative psych verbs (Mary likes John).

At the individual, psycholinguistic level, the cognitive structure of the majority language appears to be imposing itself on the structure of Spanish, Hindi, and Romanian in the heritage speakers. Significantly more input and use of English than of Spanish, Hindi, and Romanian throughout their lifespans have reinforced both the structures that the two languages share and those that are linguistically less complex, such as English-based nominative experiencers for dative experiencers and unmarked and unscrambled direct objects. Thus, the structural changes going on in the grammars of heritage speakers are most likely due to the influence from English, since the change happens to be in the retreat rather than in the advancement of overtly marked case. Similar loss of inherent case as a result of language contact and incomplete acquisition in bilinguals is reported by Lightfoot (1991, 1999), with the loss of inherent dative case in Middle English, and by Polinsky (2006), with the loss of the genitive of negation and semantically based instrumental case in Russian heritage speakers.

Admittedly, direct support for the dominant-language transfer hypothesis would have to be sought by examining heritage speakers of Spanish in contact with a language that exhibits DOM: if in this case heritage speakers do not move in the direction of losing DOM, then dominant-language influence gets strong support. Although we do not have data from Spanish heritage speakers in Romania or Romanian heritage speakers in Spain, we did administer our Spanish tests to Romanian native speakers learning Spanish as a second language in Romania (Montrul 2015) and to Turkish speakers learning Spanish in Turkey (Montrul & Gürel 2015). Turkish exhibits DOM with specific objects, like Hindi, and animacy is less relevant. The results showed that the Romanian- and Turkish-speaking L2 learners were highly accurate in judging Spanish sentences with DOM; in fact, the Romanians did not differ statistically from any of the Spanish native speakers tested in Mexico. These results contrast sharply with those of English-speaking L2 learners of Spanish, who have been shown to experience significant difficulty with DOM in Spanish, even at advanced levels of proficiency (Bowles & Montrul 2009, Farley & McCollam 2004, Guijarro-Fuentes 2012, Guijarro-Fuentes & Marinis 2007, Zyzick & Marqués 2012).

Although transfer and simplification are viable explanations for the DOM omission patterns attested in all of the heritage speaker groups in this study, erosion of DOM in Spanish, Hindi, and Romanian under the pressure of English in the US is apparently more extensive and extended in Spanish as an immigrant language than in Hindi and Romanian, affecting not just the second generation of immigrants but also the first. Similar DOM omission patterns were found in the Spanish spoken by Spanish immigrants and their children in Neuchâtel, Switzerland (Grosjean & Py 1991, Girard 1995 reported in Grosjean 2008), where Spanish is in contact with French, another language with no DOM.

6.2. DIFFERENTIAL DEGREES OF EROSION IN THE THREE LANGUAGES. We now consider potential internal and external factors that may explain the differences between the three languages. By internal factors we mean specific linguistic properties of the languages. External factors are sociolinguistic characteristics of each community, patterns of language use, and vitality of the language, to name just a few.
Language-internal, structural factors. An initial consideration underlying the high degree of erosion in Spanish was that the preposition *a* is less acoustically salient than the Hindi and Romanian DOM markers *ko* and *pe*. If acoustic salience were the main factor, then the *a*-marker should be similarly omitted in indirect object contexts, for example, because the same acoustic issues would arise in this case (*Juan llamó a Susana* ‘Juan called Susana’ vs. *Juan preguntó a Susana* ‘Juan asked Susana (something)’). However, we found that the ratings for ungrammatical indirect objects without *a*, especially when there was no clitic in Spanish, were much lower than the ratings of animate, specific direct objects without *a*, suggesting that the acoustic salience of *a* cannot be the main factor behind our findings. Furthermore, similar asymmetric acceptability ratings for animate, specific direct objects and indirect objects were obtained for Hindi *-ko* and Romanian *pe*, and the DOM and dative markers in these languages are more acoustically salient than Spanish *a*.

Another important consideration has to do with other structural properties of the languages. For example, Spanish and Romanian have definite articles, and definite articles mark definiteness and specificity in these languages. By contrast, Hindi does not have definite articles. The indefinite determiner in Hindi is the numeral *ek* ‘one’. Specific and nonspecific definite NPs, including generics, have no articles. Definiteness and specificity can be marked by case marking (*-ko*) or word order, such as scrambling, in Hindi. This means that, structurally speaking, *-ko* is needed more in Hindi than perhaps *a* and *pe* are needed in Spanish and Romanian to mark definite and specific direct objects.

A difference between Spanish and Romanian is that accusative clitic doubling is preferred in Romanian and unacceptable in Mexican Spanish. We saw that omission of *pe* in Romanian was less acceptable in non-clitic-doubled direct objects (Fig. 9) than in clitic-doubled direct objects (Fig. 8). Kayne’s generalization states that clitic doubling needs the preposition marking the object NP. This would explain why, perhaps, clitic doubling protects DOM from being eroded in Romanian. Now, accusative clitic doubling is not grammatical in Mexican Spanish, the variety tested in this study, but it is very common in Argentinian Spanish. A prediction this explanation for Romanian makes is that heritage speakers from an Argentinian background and first-generation adult immigrants from Argentina should omit DOM with animate direct objects less than heritage speakers of Mexican origin or other varieties that do not admit accusative clitic doubling, behaving more like the Romanian heritage speakers. In short, it is possible that the particular structural properties of the languages, and the syntax of definiteness and specificity in each language, play an important role in the degree of erosion and retention of DOM in different heritage languages.

External sociolinguistic factors. Let us now turn to other potential explanations related to characteristics of the immigrant communities, to see to what extent they could also explain the patterns of results found. In terms of sociolinguistic characteristics of the three immigrant groups, with over fifty million Spanish speakers, the Spanish speakers have a wider speech community in the US than the Hindi and Romanian speakers (Carreira 2013, Gambhir & Gambhir 2013). Although not uniformly, of all heritage languages in the US, the Spanish-speaking community has more access to Spanish in media and advertising, and in instruction in elementary school through bilingual programs and in middle and high school as a foreign language (Carreira 2013). For example, according to their answers to the extensive language-background question-

---

8 Sentences with accusative clitic doubling were included in the Spanish test as well, and all of the speakers rated these sentences as highly unacceptable.
naire administered to supplement the results of the experimental tasks, all of the heritage speakers tested had access to Spanish as a second language in middle school and high school, whereas the Hindi and the Romanian heritage speakers did not, unless instruction of some sort was provided by their parents or religious organizations/groups. Table 9 shows some estimates of language use with family members, siblings, and friends in the five groups of heritage speakers.

Table 9 shows some estimates of language use with family members, siblings, and friends in the five groups of heritage speakers.

<table>
<thead>
<tr>
<th></th>
<th>SPANISH HERITAGE SPEAKERS</th>
<th>HINDI HERITAGE SPEAKERS</th>
<th>ROMANIAN HERITAGE SPEAKERS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SIMULT</td>
<td>SEQUENT</td>
<td>SIMULT</td>
</tr>
<tr>
<td>Parents spoke heritage</td>
<td>81.3%</td>
<td>91.7%</td>
<td>15.4%</td>
</tr>
<tr>
<td>language only at home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoke heritage language</td>
<td>68.8%</td>
<td>70.8%</td>
<td>15.4%</td>
</tr>
<tr>
<td>when talking to their</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parents in early childhood</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoke heritage language</td>
<td>6.3%</td>
<td>33.3%</td>
<td>7.7%</td>
</tr>
<tr>
<td>only with siblings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoke heritage language</td>
<td>18.8%</td>
<td>12.5%</td>
<td>26.9%</td>
</tr>
<tr>
<td>with grandparents living</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at home</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoke heritage language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exclusively with friends</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>early childhood</td>
<td>21.9%</td>
<td>62.5%</td>
<td>0% (65.4% Eng)</td>
</tr>
<tr>
<td>elementary school</td>
<td>19.4%</td>
<td>25.0%</td>
<td>0% (92.9% Eng)</td>
</tr>
<tr>
<td>middle school</td>
<td>6.3%</td>
<td>25.0%</td>
<td>0% (100% Eng)</td>
</tr>
<tr>
<td>high school</td>
<td>5.1%</td>
<td>12.5%</td>
<td>0% (100% Eng)</td>
</tr>
<tr>
<td>Watched TV in heritage</td>
<td>90.6%</td>
<td>100%</td>
<td>46.2%</td>
</tr>
<tr>
<td>language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were read to in heritage</td>
<td>53.1%</td>
<td>70.8%</td>
<td>50%</td>
</tr>
<tr>
<td>language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were encouraged to speak</td>
<td>81.3%</td>
<td>75.0%</td>
<td>53.8%</td>
</tr>
<tr>
<td>heritage language</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9. Estimated reported use of the heritage language with parents, grandparents, siblings, and friends by the heritage speakers.

All speakers reported using English and the heritage language, but Table 9 reports the percentage of speakers who spoke the heritage language exclusively with parents, grandparents, siblings, and friends. In terms of language use throughout the lifespan, the Hindi speakers are the ones who use the heritage language less frequently, preferring mostly English. The Spanish speakers reported higher use of the heritage language in the family than the Romanian speakers, including more interactions with friends who speak the heritage language. The Romanian families appear to provide more literacy opportunities to their children at home than the Spanish-speaking families, while the Hindi speakers were the most disadvantaged in terms of literacy. They were the least literate of the three groups largely because their language uses a different script, whereas the Spanish and Romanian scripts are easier to read. Although all of the heritage speakers seem to use both languages to some extent, the Hindi and the Romanian speakers reported higher use of English-only than the Spanish heritage speakers. Thus, in terms of overall heritage language use and availability, Spanish has an advantage over Hindi and Romanian, yet it shows the highest degree of erosion of DOM.

We also found an age-of-acquisition effect in the Romanian heritage speakers but not in the Spanish heritage speakers. This pattern was evident from the proficiency measures (self-ratings and written cloze test) as well: in Spanish, there were no overall proficiency differences between simultaneous and sequential bilingual heritage speakers, while in Romanian there were. One reason for this difference between Spanish and Romanian may be related to the patterns of immigration of the two communities. The sequential bilingual heritage speakers from Romania immigrated to the US later in
childhood (between the ages of five to fourteen, mean age of nine) and had a longer length of residence in Romania (average 10.63 years) than the sequential bilinguals who immigrated as children from Mexico between the ages of five and twelve (mean 6.8) and lived in Mexico for an average of 6.8 years. All in all, the sequential bilinguals from Romania had longer exposure to their native language (3.8 years more) than the sequential bilinguals from Mexico, explaining, in part, the differences in their proficiency scores relative to the other groups and the differences in DOM omission between the Spanish-speaking and Romanian-speaking sequential bilinguals.

Table 9 also shows that the sequential bilingual Romanian heritage speakers speak more Romanian with their parents and friends than the sequential bilingual Spanish heritage speakers speak Spanish to their parents and friends. They reported using Romanian often or seldom, and it was mostly when speaking with their parents and siblings, although they frequently used it along with English. When asked what language they prefer to speak, none, however, said Romanian. The fact that the adult Romanian group does not seem to show signs of attrition whereas the first-generation adult immigrants from Mexico do could also contribute to these results. When the sequential bilingual Romanian heritage speakers speak Romanian to their parents, the input from their parents is target-like, but when the sequential bilingual Spanish heritage speakers speak to their elders, who also produce unmarked animate, specific direct objects, the input reinforces the grammar of unmarked objects entertained by many Spanish heritage speakers.

Another important difference between the three groups is education and SES. The Hindi heritage speakers and the Hindi-speaking adult immigrants had the highest level of education when compared with the Mexican and Romanian immigrants. They came largely from professional backgrounds (doctors, engineers, scientists, lawyers, etc.), like the profiles described by Gambhir and Gambhir (2013). By contrast, the Spanish and the Romanian heritage speakers came from predominantly working-class backgrounds, as judged by the reported occupation of their parents in the US, although the Romanian immigrants seemed to have higher level of education than the Spanish immigrants. Whether and how working-class and economic experiences affect language development is a matter of controversy. Several studies have shown that working-class children have smaller vocabularies and different profiles of syntactic development than children of professional families (Hart & Risley 1995, Huttenlocher et al. 2002). Recent studies on adults suggest that educational and SES background have deterministic effects on language proficiency and linguistic competence in monolingual and bilingual speakers (Dąbrowska 2012 and commentaries therein, Pakulak & Neville 2010). Yet the studies documenting differences in morphology and syntax are about later, rather than earlier, language development. DOM is a feature of early language development because it is supposedly acquired by children before the age of three. At the same time, we see a marked difference between the Spanish and the Romanian speakers, even when many of the Romanians also came from working-class families. It is consequently not certain that the working-class status of the parents can account for our results. Certainly, SES is an issue that remains to be verified in Spanish by testing Spanish heritage speakers who come from professional families.

Knowledge of other languages is another factor worth considering. Most of the Spanish speakers were monolingual in Spanish when they became bilingual in English. By contrast, all of the Hindi speakers were bilingual and multilingual, speaking English and Hindi from an early age and in most cases a third South Asian regional language. The Romanian speakers also spoke English at higher levels than the Spanish speakers in Mexico, and were all knowledgeable of at least one other European language. Re-
search has shown that additional language learning is facilitated in bilinguals as compared to monolinguals (Bruhnde & Garavito 1997). It may also be the case that knowledge of more than one language enhances language aptitude and protects against attrition of the first language. This would explain why the Spanish-speaking adult immigrants exhibit attrition of Spanish DOM whereas the Hindi and the Romanian speakers do not, although further research is warranted before any firm conclusions can be drawn. Table 10 summarizes the main factors discussed.

### Table 10. Summary of main findings and potential factors affecting DOM erosion.

<table>
<thead>
<tr>
<th>FINDINGS</th>
<th>SPANISH</th>
<th>HINDI</th>
<th>ROMANIAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>incomplete acquisition in heritage speakers</td>
<td>no age of onset of bilingualism effect</td>
<td>attrition in first generation</td>
<td></td>
</tr>
<tr>
<td>attrition in first generation</td>
<td>—</td>
<td>no attrition in first generation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LANGUAGE FEATURES</th>
<th>SPANISH</th>
<th>HINDI</th>
<th>ROMANIAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>low acoustic salience of the <em>a</em> marker</td>
<td>high acoustic salience of the <em>ko</em> marker</td>
<td>high acoustic salience of the <em>pe</em> marker</td>
<td></td>
</tr>
<tr>
<td>articles</td>
<td>no articles</td>
<td>articles</td>
<td></td>
</tr>
<tr>
<td>clitic left dislocations</td>
<td>scrambling</td>
<td>accusative clitic doubling</td>
<td></td>
</tr>
</tbody>
</table>

7. Conclusion. We have shown that differential object marking, which lies at the morphology-syntax-semantics interface, is a vulnerable feature in heritage languages in contact with a language like English that does not mark direct objects. At the same time, the degree and extent of the erosion of DOM differs in Spanish, Hindi, and Romanian as heritage languages. DOM is extensively more affected in Spanish than in Hindi and in Romanian, even when Spanish enjoys more presence in public life and education in the US than Hindi and Romanian, and heritage speakers appear to use their heritage language more than Hindi speakers, for example. While both internal linguistic factors and external sociolinguistic factors play a role in the structural changes observed, we support the conclusion that the structural properties of the DOM markers together with the syntax of definiteness and specificity in each language (more than external factors) seem to account for the degree of DOM erosion in each language. Because DOM is affected in the adult Mexican Spanish immigrants, whereas Hindi and Romanian immigrants are not affected, the degree of erosion of DOM observed in the Spanish heritage speakers is more extensive (in terms of number of individual speakers) than the extent observed in Hindi and Romanian. Therefore, for some linguistic structures and for some populations, incomplete acquisition in heritage speakers is also related to qualitatively different input provided by the parental generation.

We want to end with some thoughts about the nature of the individual grammars of heritage speakers. Throughout this study, we have emphasized the notions of incomplete acquisition and attrition in the face of suboptimal input in the heritage language and extensive and intensive exposure and use of English, the majority language in these cases, as the key reasons for the patterns found. But the validity of these notions is relative to a baseline, which could be the native-speaker groups in the countries of origin.
used in this study or, if this were a longitudinal study measuring the linguistic abilities of heritage speakers and adult immigrants over time, it would be the same individuals at an earlier time. Now, moving away from the monolingual-bilingual comparison, and from the unintentional ‘deficit’ view of bilingual knowledge this comparison unfortunately entails (Bhatt 2002), and considering instead heritage grammars in their own right, it behooves us to address what type of linguistic representations of DOM these speakers have. In other words, what do their grammars look like, and how are they different from the target?

The heritage speakers and adult immigrants who do not differ on their acceptability ratings of DOM-marked objects have a linguistic representation of these objects along the lines of the structure shown in 32, where the DOM object moves to a higher structural projection above the VP. Unmarked objects stay within the vP. When these speakers use English, they do not mark animate, specific direct objects and project a structure where animate, specific direct objects, like all objects, receive structural case within the VP. We can imagine that these bilingual speakers have a grammar of marked objects for Hindi, Spanish, and Romanian and a different grammar for unmarked animate, specific direct objects in English, which does not involve movement or overt marking. In other words, they generate two grammars, and they use one grammar for each language. In contrast, the heritage speakers and adult immigrants who incorrectly accept unmarked animate, specific direct objects in Spanish, Hindi, and Romanian also have two grammars of animate, specific direct objects, except that because English is more dominant than the heritage language in these speakers, the higher cognitive and psycholinguistic activation level of English with respect to the level of activation of the heritage language leads them to recruit the grammar of animate, specific direct objects of English—a representation with no marking and no movement—for the heritage language as well. What we are saying is that all of these bilingual speakers generate multiple grammars in the sense of Amaral & Roeper 2014, Kroch & Taylor 1997, and Yang 2003, and the coexisting grammars of DOM occur at the representational and processing levels (Sharwood Smith & Truscott 2014). As in situations of diachronic change in general, different grammars coexist for an extended period of time until one of them is not generated anymore (Lightfoot 1991). The bilingual speakers who are consistent in their acceptability judgments, either because they rejected all or most sentences with DOM or because they accepted all or most of them, have activated one or the other grammar of DOM (the one of the heritage language or the one of English). Those who consistently accept unmarked objects in the heritage language may have already eradicated the DOM grammar, while those bilinguals who show more inconsistent acceptability ratings are still entertaining and activating both grammatical options in the heritage language (the one for English and the one for the heritage language). Our next step is to investigate the systematicity of these heritage grammars in greater depth by testing, for example, to what extent heritage speakers respect or violate the definiteness scale proposed for DOM (Aissen 2003).

In conclusion, incomplete acquisition and attrition in heritage speakers and the first generation of immigrants lead to structural changes and different end-state grammars that can be described and explained using the available tools of theoretical models developed to account for native, monolingual grammars.

REFERENCES


Benmamo, Elabbas; Abdulkafi Albirini; Silvina Montrul; and Eman Saadah. 2014. Arabic plurals and root and pattern morphology in Palestinian and Egyptian heritage speakers. *Linguistic Approaches to Bilingualism* 4.89–123.


Differential object marking in Spanish, Hindi, and Romanian as heritage languages


KEIJZER, MERel. 2007. Last in first out? An investigation of the regression hypothesis in Dutch emigrants in Anglophone Canada. Amsterdam: LOT.


Montrul, Silvina; Rakesh Bhatt; and Archna Bhatia. 2012. Erosion of case and agreement in Hindi heritage speakers. Linguistic Approaches to Bilingualism 2.141–76.


Montrul, Silvina; Rebecca Foote; and Silvia Perpiñán. 2008. Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. Language Learning 58.3–53.


Department of Linguistics and
Department of Spanish and Portuguese
University of Illinois at Urbana-Champaign
4080 Foreign Languages Building, MC-176
707 South Matthews Ave.
Urbana, IL 61801
[montrul@illinois.edu]
[rbhatt@illinois.edu]
[girju@illinois.edu]

[Received 7 April 2013; revision invited 29 January 2014; revision received 16 March 2014; accepted 1 May 2014]