2019

The processing of input with differential objectmarking by heritage Spanish speakers

Jill Jegerski
University of Illinois at Urbana-Champaign

Irina A. Sekerina
CUNY College of Staten Island

Bilingualism: Language and Cognition

cambridge.org/bil

Research Article

Cite this article: Jegerski J, Sekerina IA (2020). The processing of input with differential object marking by heritage Spanish speakers. *Bilingualism: Language and Cognition* **23**, 274–282. https://doi.org/10.1017/S1366728919000087

Received: 1 December 2018
Revised: 3 January 2019
Accepted: 17 January 2019
First published online: 1 March 2019

Key words:
Spanish; heritage language; differential object marking; Visual World Paradigm; eyetracking

Author for correspondence:
Jill Jegerski, jegerski@illinois.edu

The processing of input with differential object marking by heritage Spanish speakers

Jill Jegerski[1] and Irina A. Sekerina[2,3]

[1]Department of Spanish and Portuguese, University of Illinois at Urbana-Champaign, USA; [2]Department of Psychology, College of Staten Island, and Ph.D. Program in Linguistics, The Graduate Center of the City University of New York, USA and [3]National Research University Higher School of Economics, Russian Federation

Abstract

Heritage Spanish speakers and adult immigrant bilinguals listened to wh-questions with the differential object marker *a* (*quién*/*a quién* 'who/who$_{ACC}$') while their eye movements across four referent pictures were tracked. The heritage speakers were less accurate than the adult immigrants in their verbal responses to the questions, leaving objects unmarked for case at a rate of 18%, but eye movement data suggested that the two groups were similar in their comprehension, with both starting to look at the target picture at the same point in the question and identifying the target sooner with a quién 'who$_{ACC}$' than with quién 'who' questions.

Introduction

HERITAGE SPEAKERS are early bilingual acquirers of a home language that is not the majority language of the broader society (e.g., Montrul, 2016). The outcome of heritage language (HL) acquisition is variable and potentially complicated by divergent attainment, attrition, cross-linguistic influence in the individual, and language contact in the broader society. Spanish DIFFERENTIAL OBJECT MARKING (DOM; Bossong, 1991) is a telling example of how these factors might contribute to variability with HL speakers. Montrul (2014) argues that divergent attainment occurs with DOM because exposure to input in childhood is limited in quantity. Pascual y Cabo (2013) additionally suggests that input from older Spanish-speaking family members may be inconsistent in quality, due to first language (L1) attrition.

In this article, we argue for another factor that contributes to variability with DOM, namely, input processing, and focus on two fundamental mechanisms that underlie sentence comprehension: integration and prediction. First, integration is the incremental incorporation of linguistic forms into an existing representation of the input in the mind of the comprehender. It creates cognitive links between form, grammatical function, and meaning that are thought to be a basic mechanism of acquisition (e.g., Ellis & Collins, 2009; Gass, 1997; VanPatten & Cadierno, 1993). Second, integration that is sufficiently rapid can enable the comprehender to predict what comes next in the input, before it appears (DeLong, Urbach & Kutas, 2005; Federmeier, 2007; Kamide, 2008; Van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005).[1] Prediction, like integration, has been proposed as a mechanism for language learning during processing, specifically through predictions that prove inaccurate (Chang, Dell & Bock, 2006; Jaeger & Snider, 2013). Because integration and prediction are potential mechanisms of acquisition (Phillips & Ehrenhofer, 2015), populations of language users who struggle with processing might show divergent attainment. This proposal was originally made for second language (L2) learners (VanPatten, 1996, 2015), but here we explore its explanatory potential with heritage speakers.

The linguistic phenomenon under investigation in the present study is DOM. Spanish consistently marks direct objects that are both human and specific with the particle *a* and variably marks those that are either human or specific but not both, as well as those that are animate nonhuman and specific (Aissen, 2003). Animacy can thus be conceived of as a multilevel hierarchy: human > animate nonhuman > inanimate. In addition to basic declarative sentential contexts, DOM also occurs in subject (1a) and object *wh*-questions (1b), the latter of which show OVS (object-verb-subject) word order, as opposed to the canonical SVO for declarative sentences.

[1]It is worth noting that much of the experimental data commonly cited as illustrating predictive processes could equally be explained by a prediction-free account, in which items that were claimed to have been anticipated are instead integrated more easily into the existing representation of the stimulus (DeLong et al., 2005; Federmeier, 2007; Kamide, 2008; Van Berkum et al., 2005). This is because of timing: effects are often observed during or after the stimulus segment that is to be predicted. Nevertheless, Kamide (2008) points out that the visual world eyetracking paradigm can show effects that occur BEFORE the onset of the segment that is purportedly anticipated and which are difficult to explain without prediction. In any case, the present study was not designed to address this question or contribute to the debate, especially since we did not find any evidence of prediction.

(1) a. *¿Quién visita* ***a*** *la vecina?*
who visits ACC the neighbor
'Who visits the neighbor?'

b. *¿**A** quién visita la vecina?*
ACC who visits the neighbor
'Who does the neighbor visit?'

Previous research on DOM has shown that HL speakers use it inconsistently. Specifically, they (a) leave 10 to 50% of animate direct objects unmarked in speech (Montrul, 2004; Montrul & Sánchez-Walker, 2013) and 20% in written language (Montrul, 2014), (b) accept declarative sentences with unmarked animate direct objects as grammatical (Montrul & Bowles, 2009) and (c) misinterpret about 20% of marked direct objects as subjects in picture-matching tasks (Montrul, 2014). Montrul and colleagues have attributed this to a representational difference stemming from the reduced quantity of input in childhood, along with exposure to variable DOM from the older generation of L1 attriters (Montrul & Sánchez-Walker, 2013).

Here we propose that, in addition to the quantity and quality of input, the way the input is processed may also play a role. Two previous studies investigated the online integration of Spanish DOM by heritage speakers and found no effects when direct objects in declarative SVO sentences were left unmarked versus marked (Jegerski, 2018b, self-paced reading), and significant effects with stimuli that had non-canonical VSO word order (Arechabaleta-Regulez, 2016, eyetracking with text). Thus, there was already some evidence that the *a* form is not always integrated during HL processing.

The present study builds on this prior work by using the Visual World eyetracking Paradigm (VWP) to test auditory comprehension of DOM, which may be particularly appropriate for HL users because of underdeveloped literacy in the HL (Benmamoun, Montrul & Polinsky, 2010; Bolger & Zapata, 2011; Jegerski, 2018a). In addition to basic online integration of the form itself, we also examined the potential for DOM to be used to predict later segments of a sentence. Predictive processing had not previously been investigated with Spanish DOM, either with heritage speakers or other populations of Spanish users, and the VWP is ideal for the study of prediction because looking behavior can be observed before the anticipated part of the stimulus is encountered (Kamide, 2008). To this end, we focused on *wh-* questions with DOM at the beginning of a sentence, which could be used to predict a noun phrase (NP) that came three words later. This also allowed us to include stimuli with non-SVO order, which had been proposed as a factor in whether DOM gets processed (Arechabaleta-Regulez, 2016). The comparison group was comprised of adult immigrant bilinguals because our stated definition characterizes heritage speakers by EARLY bilingualism acquired in a minority language setting, rather than bilingualism in general (following Montrul, 2016). Finally, we included verbal responses to the questions as a production measure to maximize comparability with prior studies.

Hence, the three research questions for this study were: i) Do Spanish HL speakers (vs. adult immigrant bilinguals) consistently use DOM in their verbal responses to *who* questions?, ii) Do they show basic integration of the *a* marker during spoken language processing of *who* questions?, and iii) Do they use DOM to generate predictions during spoken language processing of *who* questions?

Table 1. Participant Background Information by Group

	Heritage Bilinguals (*n* = 24)		Adult Immigrants (*n* = 24)	
	M	*SD*	*M*	*SD*
Age	21.0	5.6	26.1	6.2
Age of arrival	4.5	5.9	22.6	5.6
DELE proficiency score (max. 50)	38.1	6.9	47.7	1.7
Self-rating of accent (1 'none' – 5 'very strong')				
English	1.9	.8	3.2	1.2
Spanish	1.8	1.1	1.2	.5

Method

Participants

The experimental group comprised 24 HL speakers of Spanish, childhood bilinguals who were raised in New York. Thirteen were born in the U.S. and 11 were born in seven Latin American countries and immigrated to the U.S. as children (see Table 1). Their sociolinguistic generation would be G1.5, G2, or G3.

The comparison group of adult immigrants (term from Montrul & Sánchez-Walker, 2013) was comprised of 24 bilingual Spanish speakers who were raised and educated in Spanish, in Latin American countries and Spain. Unlike the experimental group, they had all immigrated to the U.S. and become proficient in English as adults. Sociolinguistically, they would be first generation (G1).

Materials

The 20 spoken *wh-* questions in Spanish were each prefaced by a preamble with three sentences (Figure 1). The first sentence introduced two animate referents and a location, the second presented one transitive event with the two main referents, and the third introduced a distractor referent. Next came the stimulus, a *who* question in one of two conditions (Figure 1), querying either the subject or the object of the transitive event. As is typical in VWP studies, it was expected that as participants listened to the question, they would look more over time at the picture corresponding to their response choice. In both stimulus conditions, it was possible to predict the second, post-verbal NP and respond correctly to the question after hearing just the initial *wh-* word and the verb, based on the single transitive event from the preamble (Hopp, 2015). There were 10 items in each of the two conditions, interspersed with 20 similar fillers that had *where* and *where to* questions in pseudorandomized order. For the experimental stimuli, the two animate referents that participated in the transitive event in the preamble were human in 15 sentences and animal in 5 sentences (animal characters were also used by, e.g., Montrul, 2004; Montrul & Bowles, 2009; Montrul & Sánchez-Walker, 2013). Two counterbalanced presentation lists were used. All materials are available upon request from the authors.

Each stimulus, recorded by a female native speaker of Caribbean Spanish (the dominant local dialect; Pew Hispanic Center, 2016), was presented simultaneously with a visual display

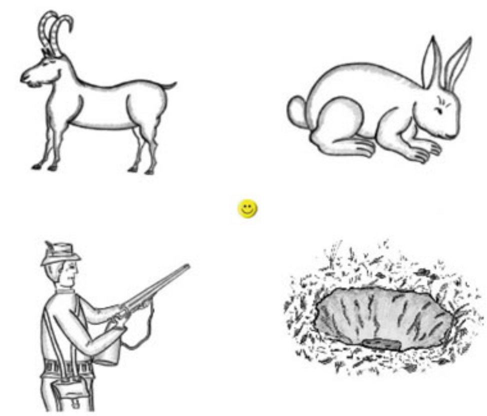
Visual Display		
Preamble		*Una noche, el chivo y el conejo se cayeron a un hoyo.* *El conejo salvó al chivo.* *Por la mañana, el cazador encontró el hoyo vacío.* "One night, the goat and the rabbit fell into a hole. The rabbit saved the goat. In the morning, the hunter found the hole empty."
Who-subject Condition	Stimulus Question	IP1 IP2 IP3 IP4 *¿Quién / salvó / al chivo / en el hoyo?* "Who saved the$_{ACC}$ goat in the hole?"
	Expected Response	Greater fixations on rabbit picture over time Click on rabbit picture
Who-object Condition	Stimulus Question	IP1 IP2 IP3 IP4 *¿A quién / salvó / el conejo / en el hoyo?* "Who$_{ACC}$ did the rabbit save in the hole?
	Expected Response	Greater fixations on goat picture over time Click on goat picture

Fig. 1. Example of a stimulus item, in two conditions and with expected responses for each. Slashes indicate the four interest periods (IPs) for eye movement data.

containing black line drawings of the four referents introduced in the preamble. A central fixation point was indicated with a smiley face.

Additional materials included a language background questionnaire and a 50-item written test of Spanish proficiency adapted from the *Diploma of Spanish as a Foreign Language* and typically used with Spanish heritage speakers (e.g., Montrul, 2004).

Procedure

The VWP for this experiment integrated the visual presentation of four referents with the playing of a recorded stimulus (Figure 1). Calibration of the eyetracker, written instructions, and seven practice trials preceded the 40 trials (20 experimental, 20 filler) for the experiment. For each trial, eye movements were recorded as the participant heard the preamble, a cue to look at the smiley face, and the question. Finally, the participant provided a verbal response and selected a picture response with a mouse click. The experiment was run on an Eyelink 1000 (2005) using Experiment Builder (2015) and on an ISCAN ETL-500 (2003) using DMDX (Forster & Forster, 2003). Afterwards, participants completed the language background questionnaire and written Spanish proficiency test.

Statistical analysis

Data were analyzed with mixed-effects linear and logistic regression using R (R Development Core Team, 2014) with the *lme4* package (Bates, Maechler, Bolker & Walker, 2015). Random effects structures were maximal with random slopes as well as intercepts wherever possible, following the approach outlined by Barr, Levy, Scheepers, and Tily (2013); more detail is provided with each specific model below. *P*-values were obtained using Satterthwaite's approximation for degrees of freedom (Kuznetsova, Brockoff & Christensen, 2014). Alpha was set at .05 and *p*-values less than .10 were approached as potentially significant, to minimize Type II error likelihood (Larson-Hall, 2010). For the eye movement

Table 2. Offline data: Verbal responses and mouse click responses (SDs in parenthesis)

	Verbal Responses		Mouse Clicks	
	Target case marking (%)	Target picture (%)	Accuracy (%)	RTs[1] (ms)
Heritage Bilinguals				
Who-subject	97.5 (14.3)	96.3 (19.0)	96.5 (18.5)	4545 (895)
Who-object	81.6 (38.8)	92.5 (26.4)	97.3 (16.3)	4619 (865)
Adult Immigrants				
Who-subject	100 (0)	98.3 (12.8)	94.6 (22.7)	4753 (1644)
Who-object	97.9 (15.6)	99.2 (9.1)	93.8 (24.3)	4719 (1565)

[1]RTs above 10000 ms were winsorized to the cutoff value.

data, proportions of fixation durations were transformed using the empirical logit function (Barr, 2008). Track loss affected 5.15% of the data.

Results

The mouse click accuracy was high (Table 2, right panel), which suggests that the participants understood the task. Accuracy and RTs did not vary according to Group (HL vs. adult immigrant) or Condition (*who*-subject vs. *who*-object; all *t*s and *z*s < 1).

Verbal responses

Verbal responses to the experimental questions (Table 2, left panel) were scored for accuracy of both case marking (i.e., object NPs were marked with *a* and subject NPs were not marked) and picture selection.

The accuracy of case marking production was partially at ceiling for the adult immigrants because they never marked a subject response with *a* (100% *who*-subject; 97.9% *who*-object), so the effect of stimulus condition could not be analyzed within this group. The HL bilinguals were near ceiling in the subject *wh*-questions (97.5%) but not with the *who*-object questions (81.6%). Their data were therefore analyzed independently, via a logit mixed-effects model (MEM) with Condition (*who*-subject, *who*-object) as a fixed effect and participant and item as random effects (slopes and intercepts). Target marking with DOM was significantly higher in the *who*-subject condition than in the *who*-object condition (estimate = .868, *SE* = .308, *z* = 2.817, *p* = .005). The groups were then compared for just the *who*-object condition, with a logit MEM with Group as a fixed effect and participant and item as random effects (intercepts only). The HL group showed a significantly lower rate of marking with DOM than the adult immigrants (estimate = 1.233, *SE* = .354, *z* = 3.486, *p* < .001).

In addition, following a suggestion from an anonymous reviewer, we examined whether proficiency (DELE test score) and age of arrival played a role in this difference. We ran two logit MEMs on the *who*-object verbal responses from both groups with either Proficiency or Age as a continuous fixed effect (separately to avoid the problem of multicollinearity) and participant and item as random effects (intercepts only with Proficiency; intercepts plus slope for item with Age). Both Proficiency (estimate = .144, *SE* = .047, *z* = 3.093, *p* = .002) and Age (estimate = .136, *SE* = .038, *z* = 3.540, *p* < .001) emerged as significant predictors of production of DOM.

The accuracy of picture selection in verbal responses was also high overall, though not at ceiling (heritage: 96.3% *who*-subject, 92.5% *who*-object; adult immigrant: 98.3% *who*-subject, 99.2% *who*-object). The data from the two groups were therefore analyzed together in a logit MEM with Group (comparison, HL) and Condition as fixed effects and subject and item as random effects (slope and intercept for subject, intercept only for item). The model showed a main effect of Group (estimate = .820, *SE* = .322, *z* = 2.548, *p* = .011) with no effect of Condition (estimate = .617, *SE* = .554, *z* = 1.115, *p* = .265) and no interaction (estimate = .392, *SE* = .324, *z* = 1.210, *p* = .226). Thus, the adult immigrant bilinguals showed higher accuracy with picture selection in the verbal responses. As with the case marking scores in the verbal responses above, additional MEMs run with a continuous predictor (either Proficiency or Age, plus participant and item as random effects, intercepts only with Proficiency; intercepts plus slope for item with Age) revealed significance for both Proficiency (estimate = .129, *SE* = .037, *z* = 3.528, *p* < .001) and Age (estimate = .071, *SE* = .034, *z* = 2.063, *p* = .039).

Eye Movements

Coarse-grain eye movements

Figure 2 illustrates the mean target advantage for four Interest Periods (IPs; seen in Figure 1) plus one second of silence after the end of the question.

To determine the stimulus segment in which the target picture was first identified, we ran a linear MEM with IP (IP2, IP3), Group, and Condition as fixed effects. Table 3 reveals the main effect of IP: the target advantage was significantly higher overall in IP3 than in IP2. The nearly significant Group × IP interaction was explored with separate analyses by group. These showed a significant effect of IP with the HL group (estimate = .508, *SE* = .175, *t* = 2.903, *p* = .008) but not with the adult immigrant group (*t* < 1). Thus, the main analysis at the IP level showed that the target was generally looked at more starting in IP3, but the borderline interaction and the independent analyses by group suggested a potential difference between the two groups, in which the effect could be traced to the HL group.

Fine-grained eye movements

To more closely examine the time course of processing in a moment-by-moment fashion, the proportions of fixations were calculated in 200-ms bins from the onset of the question (Figure 3). Visually, the divergence points of the two lines suggest that participants began fixating on the target picture more than

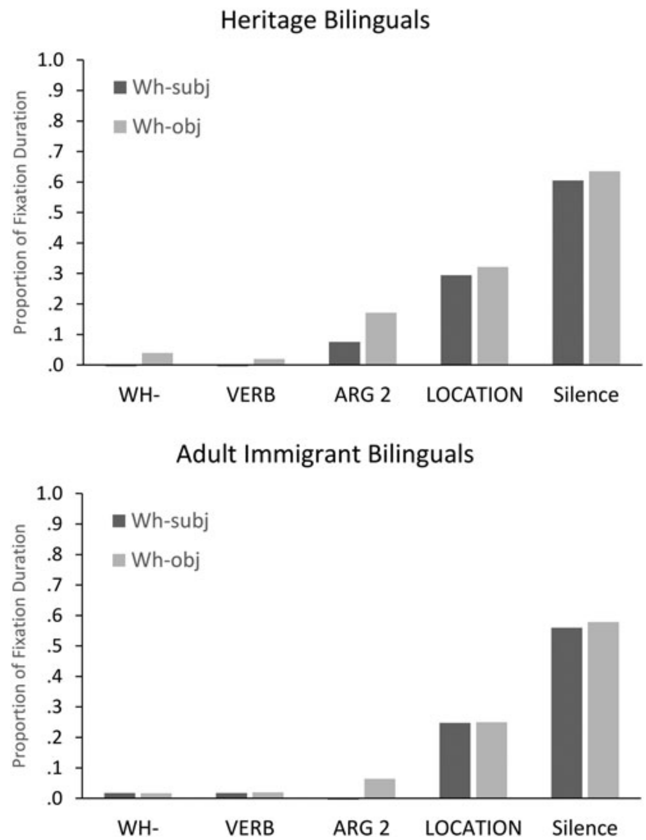


Fig. 2. Coarse-grain eye movements for both groups: Target advantage by stimulus interest period. Target advantage is the proportion of total dwell time within an IP that was on the target picture minus the proportion on the competitor picture. The actual durations of IPs in the individual stimuli varied because they were comprised of different lexical items. For analysis, each IP was delineated by the specific onsets of individual words.

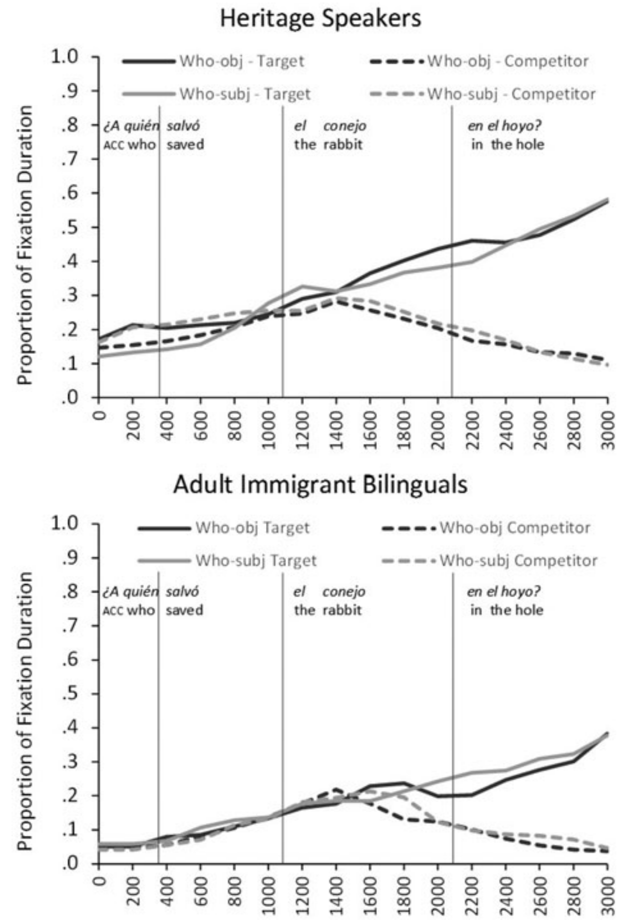


Fig. 3. Fine-grained eye movements for both groups: Proportion of looks in bins of 200 milliseconds. Vertical lines indicate the approximate mean onset for each interest period, since these were different for each individual item. The example stimulus is in the *who*-obj condition and illustrates the words that appeared within each interest period.

Table 3. Coarse-grain eye movements: Output from linear mixed-effects models*

	Estimate	*SE*	*t*	*p*
Group (heritage, adult immigrant)	.478	.337	1.421	.159
Stimulus condition (*who*-subj, *who*-obj)	.010	.275	.036	.972
Interest period (IP2, IP3)	.285	.126	2.271	**.026**
Group × stimulus condition	.184	.275	.669	.505
Group × interest period	.248	.126	1.978	.052
Stimulus condition × interest period	.036	.102	.350	.727
Group × stimulus condition × interest period	.047	.102	.459	.647

*Effects significant at $\alpha = .05$ appear in boldface. Random effects in the model were subject and aggregation ID (intercepts for both, slope for subject only).

the competitor starting in the 1400–1800 ms range, so we conducted planned comparisons only in that window (e.g., Blumenfeld & Marian, 2013; Hanna, Tanenhaus & Trueswell, 2003). As an additional protection against a Type I error, the Bonferroni Correction for two comparisons yields an alpha level of .05/2 = .025.

We ran two linear MEMs with Bin (either 1400–1600 ms or 1600–1800 ms), Group, and Condition as fixed effects. As seen in Table 4, the first model showed a nearly significant main effect of Bin, reflecting an increase in target advantage between 1400 and 1600 ms. The effect of Condition and the Bin × Condition interaction both approached significance, and further analysis showed that the effect of Bin was significant with the *who*-object questions (estimate = 1.850, $SE = .609$, $t = 3.036$, $p = .004$), but not with the *who*-subject questions ($t < 1$). The model comparing the 1600 and 1800 ms time windows showed only a main effect of Bin, indicating that by 1800 ms the target advantage was increasing overall, regardless of the condition. Thus, no differences between the two participant groups were evident in the fine-grained analysis.

Results summary

Only the verbal responses (i.e., the production measure) showed a difference between groups, in the rate of marked objects in the *who*-object condition and in the overall accuracy of meaning, which were both significantly lower for the HL group. No differences according to group or stimulus condition were found in the offline mouse click data or in the eye movements.

Table 4. Fine-grained eye movements: Output from linear mixed-effects models*

	Estimate	*SE*	*t*	*p*
1400–1600 ms				
Group (heritage, adult immigrant)	.150	.704	.213	.832
Stimulus condition (*who*-subj, *who*-obj)	1.230	.704	1.749	.084
Bin	.010	.005	2.253	**.027**
Group × stimulus condition	.750	.704	1.067	.289
Group × bin	.002	.005	.390	.698
Stimulus condition × bin	.008	.005	1.815	.073
Group × stimulus condition × bin	.005	.005	1.100	.274
1600–1800 ms				
Group (heritage, adult immigrant)	.313	.610	.513	.609
Stimulus condition (*who*-subj, *who*-obj)	.160	.608	.263	.793
Bin	.015	.004	4.183	**< .001**
Group × stimulus condition	.283	.608	.465	.643
Group × bin	.001	.004	.281	.779
Stimulus condition × bin	.002	.004	.495	.622
Group × stimulus condition × bin	.001	.004	.371	.711

*Effects significant at $\alpha = .05$ appear in boldface. Random effects in the models were subject and aggregation ID (intercepts only).

Discussion

Regarding the first research question for this study, about the production of DOM, the two groups were different in their production of the *a* marker in the verbal responses to the questions in the *who*-object condition. The heritage speakers left animate direct objects unmarked at a mean rate of 18.4%, which is consistent with the 10–50% rate observed in previous research (Montrul, 2004; Montrul & Sánchez-Walker, 2013). Individual scores correlated with age of arrival, which supports the conclusion of Montrul and Sánchez-Walker that age of acquisition of English and language use patterns at home shape DOM production, because they affect the quantity of input. There was also a correlation with proficiency, which suggests that the patterns of variability observed with a specific linguistic form like DOM can extend to overall knowledge of Spanish as a heritage language (Montrul, 2016), and that the same factors may shape acquisition and maintenance at both levels. In other words, under the right conditions, some individual heritage speakers with higher Spanish proficiency might show more consistent use of DOM.

The offline comprehension scores might appear to contradict this trend, because they were high across groups and conditions, but this may have been due to methodological differences. Unlike in the previous study showing variable offline comprehension (Montrul, 2014), DOM was not critical to our experimental task. In other words, it was possible to respond correctly to stimuli without processing the *a* marker, because there was only one transitive event with two NP arguments in the preamble, so once the critical parts of a stimulus question had played (i.e., *who*-V-NP), it was simply a matter of clicking on the other NP that did not appear in the question. For example, for the item for which the preamble stated that *El conejo salvó al chivo* 'The rabbit saved the goat', the image response to the question that mentions the goat is the rabbit and the image response to the question that mentions the rabbit is the goat (i.e., *¿Quién salvó al chivo?* 'Who saved the goat?' > click on rabbit picture, and *¿A quién salvó el conejo?* 'Who did the rabbit save?' > click on goat picture). Hence, a high score on the offline comprehension task does not necessarily mean that DOM was used in any way.

The adult immigrant group showed consistent marking of direct objects (97.9%), a higher rate than the 81.3–87.2% observed by Montrul and Sánchez-Walker (2013). The difference was likely due to length of residence in the U.S., which was 25.9 years in the prior study and 3.5 years in the present study. The negative correlation with age of arrival is also consistent with this account. These observations support the conclusion of the previous study that adult immigrants use DOM consistently when they arrive to the U.S. and then long-term residency leads to attrition. However, attrition is not widespread among those who have only been in the U.S. for a few years. Thus, a potential implication of these findings for heritage speakers and their early and continued exposure to input with DOM is that the length of residence of their older Spanish-speaker family members can be an important factor.

For the second research question, about the integration of DOM during processing, the groups were very similar in when they appeared to distinguish the target picture from the competitor (e.g., more looks to the goat picture while hearing *¿A quién salvó el conejo?* 'Who$_{ACC}$ did the rabbit save?'). Specifically, this occurred 200 ms faster for the *who*-object versus the *who*-subject stimuli (although still not soon enough to show prediction, as discussed below under the third research question), so the participants must have used the *a* marker to their advantage, because it was the only difference between the two conditions. For the heritage speakers, this result differs from one previous study (Jegerski, 2018b), which had found no evidence of online integration of the *a* marker, and is consistent with another (Arechabaleta-Regulez, 2016), which had observed sensitivity to

DOM during the reading of declarative sentences with VSO word order, but not with SVO. It appears that word order may be important for HL bilinguals, because both studies that observed online integration of the *a* particle included a stimulus condition with a word order other than SVO. Thus, the heritage speakers and adult immigrants in our study processed DOM similarly, although such similarity may be modulated by word order (consistent with the L2 theory of VanPatten, 1996, 2015).

The similarity between the two participant groups with regard to the time course of online comprehension might appear to be at odds with the difference observed in their production of the *a* marker in the verbal responses to the questions in the *who*-object condition, discussed above under the first research question. One possible explanation is that there could be a causal relationship between the speed of processing and inconsistent DOM production, along the lines of the relationship between lexical processing efficiency and expressive vocabulary in bilingual Spanish–English 3-year-olds (Hurtado, Grüter, Marchman & Fernald, 2014). The present study was not designed to test such a hypothesis, so we leave it for future research.

The third research question was about using DOM to generate predictions during processing. Because each preamble included one transitive event (e.g., *El conejo salvó al chivo*. 'The rabbit saved the goat.'), it was possible to predict the second, post-verbal NP of a stimulus question and begin to look more at the target response after hearing just the first word (i.e., *quién* 'who' or *a quién* 'who$_{ACC}$') or the first word and the verb. Prior work had shown that case information is used predictively by prototypical native speakers of German (Hopp, 2015; Kamide, Scheepers & Altmann, 2003; but cf. Hanne, Sekerina, Vasishth, Burchert & De Bleser, 2011), but Spanish had not previously been investigated, and neither of the participant groups in our study appeared to use case predictively. Rather than looking to the target as soon as possible, based on a prediction of what came next in the question, they waited to actually hear the second NP argument (e.g., *¿A quién salvó el chivo…* 'Who$_{ACC}$ did the goat save…'), at which time prediction based on the object marker could no longer be observed, because all critical components of the question had already been presented. There are at least four potential explanations for this outcome.

First, there may be cross-linguistic differences that affect all speakers of the language, if the Spanish case marker *a* is generally less likely to be used predictively than German case (perhaps because Spanish direct objects are not all marked, so the form is less reliable than in German). Given that predictive processing with DOM has not yet been investigated with monolingual users of Spanish, there is still quite a bit of uncertainty in this area.

Second, another reason for the apparent lack of predictive processing could be that this occurs with all bilinguals. Some theories suggest that late bilinguals have specific difficulty generating predictions online (Kaan, 2014). However, such claims are based on assumptions of weaker lexical representations and slower integration that occupies more cognitive resources, which might be considered with heritage bilinguals, but does not seem likely to apply equally to adult immigrants with their first language.

Third, a more plausible explanation lies in the experimental task. The two prior VWP studies that had shown predictive processing with case (Hopp, 2015; Kamide et al., 2003) employed simple declarative sentences as stimuli and simple *yes/no* comprehension questions. Plus, the argument roles of the pictures were not plausibly reversible (e.g., a fox eats a hare rather than vice versa), so it would be quite easy to predict after hearing just a few words of a sentence like *The fox will eat* _____. In the present study, the demands of predictive processing were much greater because the stimuli were contextualized questions and because the argument roles were reversible (e.g., a rabbit saves a goat or vice versa), so participants had to recall the preamble narrative in detail. Without the preceding context, there is no way to predict what comes next in a question like *Who saved* _____, even if the case marker is efficiently processed. On the other hand, the task became much easier after hearing the second NP, because the target was then just the other of the two protagonists from the narrative and it could be identified by a simple process of elimination.

Thus, we reason that the target was identified at that point and not sooner because the demands of the experimental task made the target difficult to identify early on in the stimulus question, but easy to identify later on, regardless of whether the case marker was integrated rapidly during processing. The results therefore do not necessarily represent a failure to use case marking predictively. This account is supported by the mouse click responses, which reflected a high level of accuracy, despite the apparent lack of prediction. It is possible that similar bilingual participants could show evidence of predictive processing using DOM with another task, and monolingual Spanish users might as well, so further investigation of both questions is warranted. We also might predict, based on the overall observation from this investigation and two prior studies that heritage bilinguals can integrate DOM during online comprehension but appear to do so inconsistently (Arechabaleta-Regulez, 2016; Jegerski, 2018b), that they might be less likely to use DOM predictively than adult immigrant bilinguals or monolinguals.

Fourth and finally, another potential contributor to the apparent lack of predictive processing in our results has to do with the NPs in the stimuli and the animacy hierarchy that plays a fundamental role in Spanish DOM (i.e., human > animate nonhuman > inanimate). Marking with the particle *a* is consistent on direct objects that are human, but it is variable on animates that are nonhuman. For this reason, the *wh*- expression *a quién* 'who$_{ACC}$' might be less expected to query a nonhuman NP like *chivo* 'goat' than a human one like *cazador* 'hunter' and this weaker association might reduce the potential of the *a* marker to be used predictively in sentences with nonhuman arguments. In the present study, only 5 of the 20 experimental stimuli contained nonhuman arguments, so it was not possible to examine this question via a post-hoc analysis of the potential effects of the animacy hierarchy. Nevertheless, similar animal characters have been used successfully in previous studies to elicit DOM production among Spanish heritage speakers and other Spanish users (Montrul, 2004; Montrul & Bowles, 2009; Montrul & Sánchez-Walker, 2013), under the assumption that nonhuman animate direct objects that are ascribed human characteristics are typically marked with *a* (Aissen, 2003). Because it has not been an issue in previous research, it seems unlikely that the inclusion of 25% nonhuman animate NPs in our stimuli played any major role in the results of this study.

Taken as a whole, the outcome of the current study provides evidence in support of the input-based, multifactorial account of DOM among heritage speakers proposed by Montrul (Montrul, 2014; Montrul & Sánchez-Walker, 2013) and builds on this previous work by adding a new factor to the mix, based on the theory of input processing (VanPatten, 1996, 2015). Our verbal response data show that variability in production is associated with general proficiency in Spanish and with the age of onset of bilingualism which, like previous research (Montrul, 2014; Montrul & Sánchez-Walker, 2013), suggests that the quantity of input in childhood is important for heritage speakers. Montrul's work has also suggested that the QUALITY of input can be a factor, after observing that adult immigrant family members also begin to produce DOM less consistently, although production data from the adult immigrant bilinguals in the present study suggest that this takes time, occurring after more than 3.5 years of residency.

Beyond the quantity of input in childhood and the quality of the input later on, we propose here that the way the input is processed might further limit the acquisition and maintenance of DOM among heritage speakers. In other words, even when the input contains animate direct objects marked with *a*, it will not always be processed in a way that promotes acquisition and maintenance of the form (i.e., via integration and prediction). Our eye movement data show that basic online integration occurs in at least some contexts, which may be sufficient to maintain DOM in the grammar to the extent that it is produced most of the time (81.6% in the present study). Nevertheless, when the results of previous research (Arechabaleta-Regulez, 2016; Jegerski, 2018b) are taken into consideration, the overall trend is that online integration is variable, so not every instance of DOM available in the input will reinforce the heritage speaker's knowledge of the form. Moreover, our eyetracking results suggest that predictive processing based on DOM is inconsistent at best, as it was not observed in the current study and had not been investigated previously. Based on these observations, we argue that variable processing of input (i.e., integration and prediction), along with the quantity and quality of input, may help to explain why the production of DOM among heritage speakers is not as consistent as that of recent adult immigrants.

Acknowledgements. This research was supported in part by a Research Grant and by a Humanities Released Time award to Jill Jegerski, both from the Campus Research Board of the University of Illinois at Urbana-Champaign. Irina Sekerina was partially supported by the Center for Language and Brain NRU Higher School of Economics, RF Government grant, ag. № 14.641.31.0004. We would also like to thank our Research Assistants, Juan Cardona, Krista Evensen, Sara Fernández Cuenca, Sandra Flores Hernández, and Eimy Saldaña, and three anonymous reviewers.

References

Aissen J (2003) Differential object marking: iconicity vs. economy. *Natural Language and Linguistic Theory* **21**, 435–83.

Arechabaleta-Regulez B (2016) Online sensitivity to DOM violations by Spanish heritage speakers. Paper presented at the International Workshop of the Unity and Diversity in Differential Object Marking, Institut National des Langues et Civilisations Orientales, Paris, France.

Barr DJ (2008) Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language* **59**(4), 457–474.

Barr DJ, Levy R, Scheepers C and Tily H (2013) Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68**, 255–278.

Bates D, Maechler M, Bolker B and Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**(1), 1–48.

Benmamoun E, Montrul S and Polinsky M (2010) *White paper: Prolegomena to heritage linguistics*. National Heritage Language Resource Center.

Blumenfeld HK and Marian V (2013) Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology* **25**(5), 547–567.

Bolger PA and Zapata GC (2011) Psycholinguistic approaches to language processing in heritage speakers. *Heritage Language Journal* **8**, 1–29.

Bossong G (1991) Differential object marking in Romance and beyond. In Wanner D and Kibbee D (eds), *New Analyses in Romance Linguistics: Selected Papers from the XVIII Linguistic Symposium on Romance Languages*. Amsterdam: John Benjamins, pp. 143–170.

Chang F, Dell GS and Bock K (2006) Becoming syntactic. *Psychological Review* **113**(2), 234–272.

DeLong KA, Urbach TP and Kutas M (2005) Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* **8**, 1117–1121.

Experiment Builder (Version 2.1) [Computer Software]. (2015) Mississauga, Ontario, Canada: SR Research.

Ellis N and Collins L (2009) Input and second language acquisition: The roles of frequency, form, and function. *The Modern Language Journal* **93**(3), 329–335.

ETL-500. [Apparatus]. (2003) Woburn, MA: ISCAN.

Eyelink 1000 [Apparatus]. (2005) Mississauga, Ontario, Canada: SR Research.

Federmeier KD (2007) Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology* **44**(4), 491–505.

Forster KI and Forster JC (2003) A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers* **35**, 116–124.

Gass SM (1997) *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Hanna J, Tanenhaus MK and Trueswell JC (2003) The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language* **49**, 43–61.

Hanne S, Sekerina I, Vasishth S, Burchert F and De Bleser R (2011) Chance in agrammatic sentence comprehension: What does it really mean? Evidence from eye movements of German agrammatic aphasic patients. *Aphasiology* **25**(2), 221–244.

Hopp H (2015) Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching* **53**(3), 277–306.

Hurtado N, Grüter T, Marchman VA and Fernald A (2014) Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism: Language and Cognition* **17**(1), 189–202.

Jaeger TF and Snider NE (2013) Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition* **127**(1), 57–83.

Jegerski J (2018a) Psycholinguistic perspectives on Spanish as a heritage language. In K Potowski (Ed.), *Routledge handbook of Spanish as a heritage/minority language* (pp. 221–234). New York: Routledge.

Jegerski J (2018b) The processing of the object marker *a* by heritage Spanish speakers. *International Journal of Bilingualism* **22**(6), 585–602.

Kaan E (2014) Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism* **4**(4), 257–282.

Kamide Y (2008) Anticipatory processes in sentence processing. *Language and Linguistics Compass* **2**(4), 647–670.

Kamide Y, Scheepers C and Altmann GTM (2003) Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research* **32**(1), 37–55.

Kuznetsova A, Brockoff PB and Christensen RHB (2014) lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Available at CRAN.R-project.org/package=lmerTest.

Larson-Hall J (2010) *A guide to doing statistics in second language research using SPSS*. New York and London: Routledge.

Montrul S (2004) Subject and object expression in Spanish heritage speakers: A case of morpho-syntactic convergence. *Bilingualism, Language and Cognition* **7**, 125–142.

Montrul S (2014) Structural changes in Spanish in the United States: Differential object marking in Spanish heritage speakers across generations. *Lingua* **151B**, 177–196.

Montrul S (2016) *The Acquisition of Heritage Languages*. Cambridge, UK: Cambridge University Press.

Montrul S and Bowles M (2009) Back to basics: Differential Object Marking under incomplete acquisition in Spanish heritage speakers. *Bilingualism: Language and Cognition* **12**, 363–83.

Montrul S and Sánchez-Walker N (2013) Incomplete acquisition of differential object marking in child and adult Spanish heritage speakers. *Language Acquisition* **20**, 1–31.

Pascual y Cabo D (2013) Agreement reflexes of emerging optionality in heritage speaker Spanish (Unpublished doctoral dissertation). University of Florida, Gainesville.

Pew Hispanic Center. (2016) *Hispanic Population and Origin in Select U.S. Metropolitan Areas, 2014*. Retrieved from http://www.pewhispanic.org/interactives/hispanic-population-in-select-u-s-metropolitan-areas/

Phillips C and Ehrenhofer L (2015) The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism* **5**(4), 409–453.

R Development Core Team (2014) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at http://www.R-project.org/.

Van Berkum JJA, Brown CM, Zwitserlood P, Kooijman V and Hagoort P (2005) Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **31**, 443–467.

VanPatten B (1996) *Input processing and grammar instruction in second language acquisition: Theory and research*. Norwood, NJ: Ablex.

VanPatten B (2015) Input processing in adult SLA. In VanPatten B and Williams J (eds), *Theories in Second Language Acquisition*. (2nd ed.), New York: Routledge, pp. 113–134.

VanPatten B and Cadierno T (1993) SLA as input processing: A role for instruction. *The Modern Language Journal* **77**, 45–57.