Verb Frequency and DOM in Heritage Speakers of Spanish

|  |
| --- |
| Esther Hur1 |
| 1 Rutgers University |
|  |

# Author note

Correspondence concerning this article should be addressed to Esther Hur, 15 Seminary Place, New Brunswick, NJ 08901. E-mail: [kh624@scarletmail.rutgers.edu](mailto:kh624@scarletmail.rutgers.edu)

Abstract

Heritage Speakers (HS) of Spanish have been shown to exhibit unstable knowledge of Differential Object Marking (DOM) as expressed by ‘a’. DOM is the overt marking of direct objects governed by the lexico-semantic and syntactic features of the object (Aissen, 2003; Bossong, 1991; Leonetti, 2004; Torrego, 1998). Previous research suggests this instability is due to “incomplete acquisition” (Montrul 2004; Montrul & Bowles, 2009) or attrition (Polinsky, 2006) that simplifies the grammar of HS in contact with English. Furthermore, the quality and quantity of input may be reduced when compared to that received by non-heritage speakers during earlier years (Montrul, Sánchez-Walker, 2013; Rothman, 2009; Kupisch & Rothman, 2017). Nevertheless, recent studies suggest that the quality and quantity of input alone are not sufficient explanations for how unstable grammars arise. Instead, the frequency of processing for comprehension and production are crucial factors (Putnam & Sanchez, 2013). Until now, many studies have focused in observing the extent of unstable DOM knowledge among HS (Montrul 2004; Montrul & Bowles, 2009; Montrul & Sanchez-Walker, 2013; Montrul, Bhatt & Girju, 2015). Few studies, however, have investigated in depth the possible factors that may create variability among the speakers. The current study examines: the effect of lexical frequency of the verb on DOM. Specifically it examines whether animate direct objects are more likely to receive DOM if they are complements of high frequency verbs than of low frequency verbs. The main goal is to investigate the connection between verb frequency and the activation of syntactic and semantic features that trigger DOM (Torrego, 1998; Rodriguez-Mondoñedo, 2007; Zdrojweski, 2013) By showing that verb frequency generates the syntactical activation needed for DOM among HS, the study provides further evidence that the quality and quantity of input alone are not the only factors that generate a stable heritage grammar. These findings suggest that frequency of processing for comprehension may be a contributing factor to stable DOM.

*Keywords:* Heritage language, Spanish DOM, lexical frequency.

Word count: X

Verb Frequency and DOM in Heritage Speakers of Spanish

# Methods

There are two main goals of the present section. First, I will provide a description of the criteria that were used to characterize participants as HS and Spanish dominant controls (SDC). The second objective of the present section is to outline the procedure for the present experiment.

## Participants

The subjects are all Spanish-English bilinguals. The experimental group consisted of heritage speakers of Spanish. According to Wiley and Valdés (2000) a Heritage Speaker is an individual “who is raised in home, where a non-English language is spoken, who speaks or at least understands the languages, and who is to some degree bilingual in that language and in English” (p.80). Thus HS of Spanish are Spanish-English bilinguals who speak Spanish at home but live in the United States, where English is the dominant language. In the present study, all HS are native speakers of Spanish who acquired Spanish between 0-3 years, and began learning English in the United States at or before 6.

In order to provide a baseline, an additional group of Spanish Dominant Controls (SDC) was included. The SDC is a group that acquired Spanish as a first language in a naturalistic environment who have spent significant amount of time living in the United States. What differentiates the SDC from the HS is that they were all monolingually raised in Spanish and their age of acquisition of English is at 10 or later. In the present study, all participants acquired English at age 18 or later. The SDC formed a single group (n=15) whereas the HS group (n= 30) was divided into two groups (advanced, n = 20 and intermediate, n = 10) based on their performance on the DELE (see next section).

In order to participate in the study, all HS needed to meet the following criteria: First, they must have grown up speaking Spanish at home. Second, they must have begun learning English in the United States at age 6 or before. Third, they must not speak any native language other than Spanish or English. Finally, they must score 30 or above on a standardized 50 question Spanish proficiency test.

As for the SDC, all participants’ speakers in the study needed to follow the present criteria: First, they must be native speakers of Spanish. Second, they must have lived in a Spanish-speaking country until age 13 or later. Third, Spanish must be their only native language. Finally, they must score 40 or above on a standardized Spanish proficiency test.

## Task and Procedure

In this section I present the procedures for data collection and tasks implemented for the study. Both screening and experimental tasks were used for the study. A total of four tasks (experimental = 1, screening = 2) were used during the experiment as listed below in order that were administered to the participants: 1. DELE, 2. Elicited Production Task, 3. Acceptability Judgment Task, 4. Family Background Questionnaire. In the following section I will outline in detail each procedure of the present experiment.

HS Participants were recruited from Spanish and Latin American studies classes at a large public university in the northeastern United States. SDC participants were recruited from the Spanish Department’s graduate programs. All participants received a $5 gift card after completing the study.

Participants came to an office in the Spanish department to complete the experiment. The study was explained in Spanish where the participants had time to read the consent form and ask any questions that they may have about the study. Participants completed three tasks in total. First, they completed the DELE in order to place them appropriately into the two experimental groups (intermediate or advanced). Then, they completed two experimental tasks, an Elicited Production Task (EPT). Finally, they completed the Family Background Questionnaire. The majority of the participants took approximately 50-70 minutes to complete the whole experiment. In the following paragraphs, I will provide specific descriptions of each of these tasks and their relevance for the main questions of the present study.

## Data analysis

Data from the categorization task were analyzed using a generalized linear mixed effects (GLMM) model with a binomial linking function (as implemented in the lme4 package1.1–10 in R 3.2.2). The model included DOM response as the criterion, and frequency (frequent and infrequent) and group (SDC, Adv. HS, Int.HS) as predictors.

# Results

I will start the results by presenting data from the Elicited Production Task and then presenting the data from the Acceptability Judgment Task. The first research question asked whether verb frequency would play a role in the knowledge of DOM in animacy contexts among the HS. The second research question asked whether proficiency played a role in the productive knowledge of the DOM.

## Results from the Elicited Production Task

source("../scripts/eptdata.R")

ept\_fig1

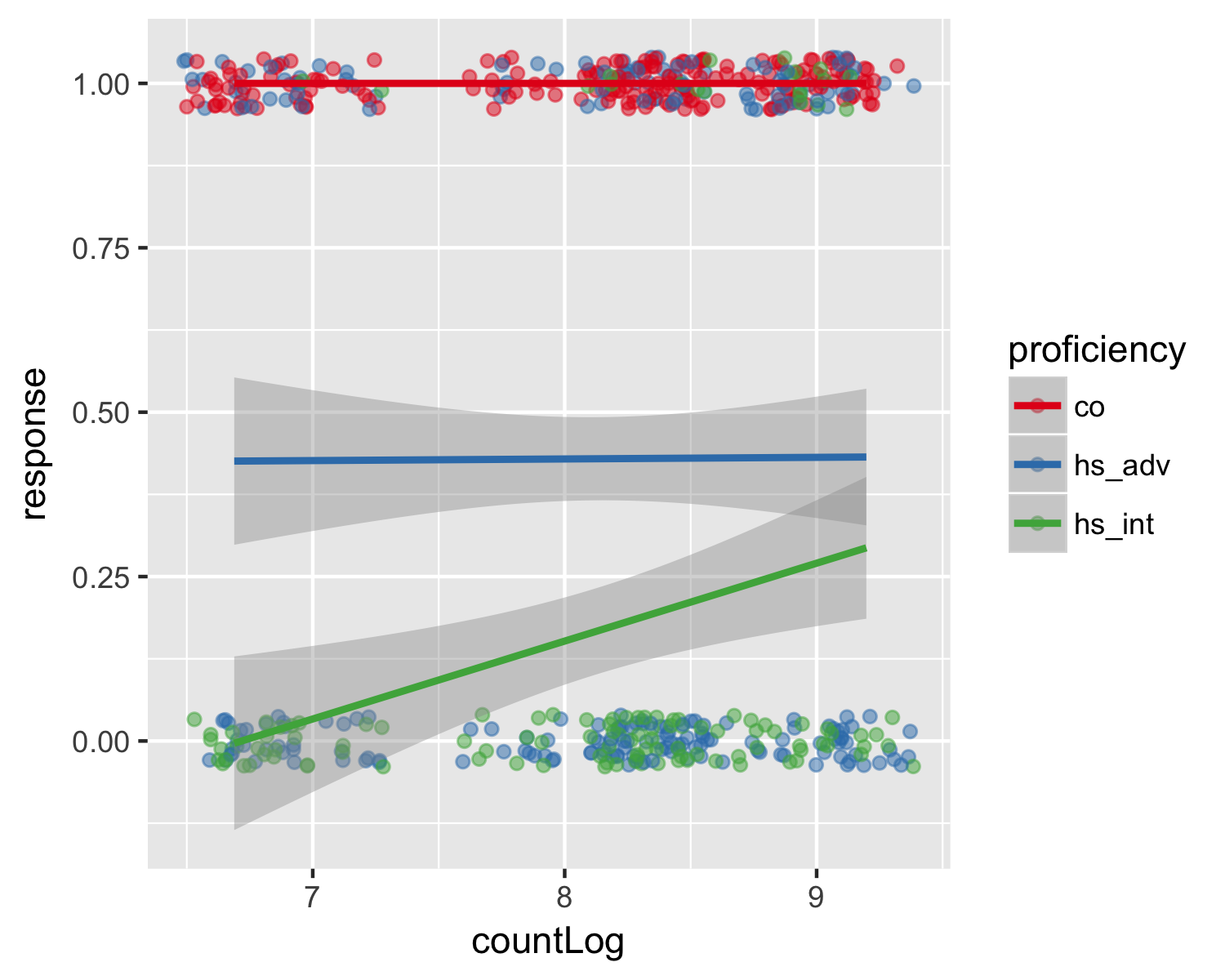


Figure 1 shows the results of the DOM production as a function of group (SDC, Adv. HS and Int. HS) and frequency (frequent and infrequent). The model included DOM response as the dependent variable, and group (SDG, Adv. HS, Int. HS), and frequency (frequent, infrequent) as predictors. DOM production was coded as “1” and no DOM production was coded as “0”. The model provides main effects of group (F (2) = 206.6 , p < 0.001), but no main effect of frequency. However, there were group by frequency interaction (F (1) = 3.48, p < 0.1).

anova(mod\_nul, mod\_cat, mod\_con, mod\_int)

## Analysis of Variance Table  
##   
## Model 1: eLog ~ 1  
## Model 2: eLog ~ countLog  
## Model 3: eLog ~ countLog + proficiency  
## Model 4: eLog ~ countLog \* proficiency  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 539 199.16   
## 2 538 198.75 1 0.411 1.9703 0.16100   
## 3 536 112.60 2 86.145 206.6365 < 2e-16 \*\*\*  
## 4 534 111.31 2 1.294 3.1045 0.04566 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The model that contained the interaction provided the best fit data (R2 = 0.44). Overall, DOM production increased as a function of proficiency. In addition, there was an interaction between the frequency and proficiency type where only the intermediate group showed effects of frequency in their production of DOM t(6.54) = 2.21, p < 0.05).

summary(mod\_int)

##   
## Call:  
## lm(formula = eLog ~ countLog \* proficiency, data = ept\_elog2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.5284 -0.5209 0.0000 0.0000 1.1701   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.609e+00 3.381e-01 -4.760 2.49e-06 \*\*\*  
## countLog -1.270e-15 4.140e-02 0.000 1.0000   
## proficiencyhs\_adv -7.231e-01 4.473e-01 -1.617 0.1065   
## proficiencyhs\_int -2.198e+00 5.346e-01 -4.111 4.56e-05 \*\*\*  
## countLog:proficiencyhs\_adv 3.015e-03 5.477e-02 0.055 0.9561   
## countLog:proficiencyhs\_int 1.450e-01 6.546e-02 2.214 0.0272 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4566 on 534 degrees of freedom  
## Multiple R-squared: 0.4411, Adjusted R-squared: 0.4359   
## F-statistic: 84.29 on 5 and 534 DF, p-value: < 2.2e-16

# Discussion

In the present study, I have been able to show that verb frequency plays a role in the production of the DOM, but the effects are restricted to the proficiency of the HS. Only the Intermediate HSs showed frequency effects in their production. Following Putnam and Sanchez (2013) the HS are more accurate with DOM when they activate Spanish more frequently for production. However, for those who activate less, their DOM production highly relies on the frequency of activation.

# References

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. Natural Language & Linguistic Theory, 21(3), 435-483.

Bossong, G. (1991). Differential object marking in Romance and beyond. New analyses in Romance linguistics, 143-170.

Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. International Journal of Bilingualism.

Leonetti, M. (2004). Specificity and differential object marking in Spanish. Catalan Journal of Linguistics, 3(1), 75-114.

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, & Bowles. M. (2009). Back to basics: Differential object marking under incomplete acquisition in Spanish heritage speakers. Bilingualism: Language and Cognition 12(3). 363–383.

Montrul, S., Bhatt, R., & Girju, R. (2015). Differential object marking in Spanish, Hindi, and Romanian as heritage languages. Language, 91(3), 564-610.

Polinsky, M. (2006). Incomplete acquisition: American Russian. Journal of Slavic Linguistics 14. 191–262.

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

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

Torrego, E. (1998). The dependencies of objects (Vol. 34). Mit Press.

Zdrojewski, P. (2013). Spanish DOM as a case of lacking case. Unpublished Ms., Universidad Nacional de General Sarmiento/Universidad de Buenos Aires.