

# Explaining Heritage Spanish

Zuzanna Fuchs<sup>1</sup>   Gregory Scontras<sup>2</sup>   Maria Polinsky<sup>3</sup>

<sup>1</sup>Harvard University  
Department of Linguistics

<sup>2</sup>Stanford University  
Department of Psychology

<sup>3</sup>University of Maryland  
Department of Linguistics

Feb. 20, 2016

# Introduction

Heritage speakers perform the costly task of communicating in their non-dominant language, so their grammar might be different, made more economical by principles that make language production and comprehension easier for them.

**Broad question:** In what way can heritage grammars differ from their corresponding native grammars?

# Introduction

At the **phonological** level: differences in pronunciation and intonation/prosody

At the **word/morphology** level: difficulties with using correct gender and number agreement, tendency to eliminate irregular forms

At the **sentence/syntax** level: tendency to impose strict word order, difficulties with non-subject relative clauses

## Questions:

What (if any) differences in the underlying mental representation of language lead to these surface differences?

What principles shape these differences?

## **This talk:**

Heritage speakers often differ from baseline speakers in their ability to produce and perceive correct agreement forms.

How are agreement categories structured in heritage grammars? We use number and gender in Heritage Spanish to address this question.

# Table of Contents

- 1 Background
  - Bundling vs split
  - Agreement attraction
- 2 Experiment
  - Methodology
  - Predictions
  - Baseline Spanish
  - Heritage results
  - Bundling vs split
- 3 Economy in Heritage Spanish
- 4 Conclusions

# Number & Gender

Formal grammars of Spanish tell us gender is most commonly expressed through the feminine word marker *-a* or the masculine word marker *-o*.

(Of course, many exceptions are well known, and these are not the only gender markers in the language.)

Number is expressed through a plural suffix *-s*, or through the lack of a suffix if the word is singular.

# Number & Gender

Nouns are marked for number and gender:

las manzan**as**

los chic**os**

And modifiers (as well as determiners and anaphors) must agree in both number and gender with a head noun:

la**s** manzan**as** roj**as**

lo**s** chic**os** guap**os**



# Number & Gender

Nouns are marked for number and gender:

las manzanas

los chicos

## Questions:

Are number and gender facets of a single category reflected in agreement, or do are they separate from each other?

If they are separate, they receive a **split** representation; if they are facets of the same agreement category, they are **bundled**.

# Split model

Under a **split model** (Picallo 1991; Carminati 2005; Antón-Méndez, et al. 2002; a.o.), number and gender are separate, and agreement in each proceeds independently of the other:

in the syntax: [ *NumP* [ *GenP* ] ]

in practice: l-a-s manzan-a-s roj-a-s

# Bundling model

Under a **bundling model** (Ritter 1993; Carstens 2000, 2003), number and gender are grouped together as one bundle of features (on the Number Projection in the syntax), so agreement in number and gender is a single process.

in the syntax: [ *NumP* [ *Num*<sub>{*Number,gender*}</sub> ] ]

in practice: l-as manzan-as roj-as

## Research question:

Are surface differences between heritage and baseline  
the result of underlying differences?

No

Heritage speakers have  
less online resources.

Yes

A serious instance of grammar divergence.  
What drives these differences?

# Research question

**Task:** Test whether these two categories (number and gender) are split or bundled by putting the categories into conflict with each other, and observe how conflict is perceived in comprehension.

**Phenomenon:** Agreement attraction

# Agreement attraction

## Agreement attraction:

Ungrammatical condition in which a **predicate** incorrectly enters into agreement with a **local noun** rather than the **head noun**.

The **key** to **the cabinets** **were** lost.

**Este libro** sobre **las guerras** **son** muy interesantes.

## The crucial idea:

If number and gender are **bundled**, then if the predicate gets one feature from the wrong noun, it will also get the other feature from that noun.

If number and gender are **split**, this need not be the case.

## Tool: Small clauses

Increased distance between agreeing elements keeps the task from being too easy.

Used small clause structures as in (1) to create conditions with a **head noun**, **local noun**, and **predicate** that allow for increased distance between each part..

- 1 (Subject) Verb **NP1** Prep **NP2** ADV **ADJ** ...
- 2 El niño considera **la noticia** en **la revista** terriblemente **aburrida**...



# Agreement attraction predictions: split

(1) El niño considera la notici-a (F-SG) en los periódic-o-s (M-PL) terriblemente aburrid-o-s (M-PL).

(2) El niño considera la notici-a (F-SG) en las revist-a-s (F-PL) terriblemente aburrid-a-s (F-PL).

If number and gender split → listeners will notice that in (1) the **predicate** has incorrectly entered into agreement with the **local noun** in two features (rather than just one, as in (2)).

# Agreement attraction predictions: bundling

(1) El niño considera la notici-a (F-SG) en los periódic-o-s (M-PL) terriblemente aburrid-o-s (M-PL).

(2) El niño considera la notici-a (F-SG) en las revist-a-s (F-PL) terriblemente aburrid-a-s (F-PL).

If number and gender are bundled → both (1) and (2) should be rated equally

- the **predicate** got the whole feature bundle from the **local noun**
- does not matter whether the **predicate** mismatches with the **head noun** in one or both features

## Predictions for bundling vs split

**Prediction 1:** If number and gender are **split**, agreement attraction in one feature **need not** lead to agreement attraction in the other feature.

**Prediction 2:** If number and gender are **bundled**, agreement attraction in one feature **should lead** to agreement attraction in the other feature.

# Table of Contents

- 1 Background
  - Bundling vs split
  - Agreement attraction
- 2 Experiment
  - Methodology
  - Predictions
  - Baseline Spanish
  - Heritage results
  - Bundling vs split
- 3 Economy in Heritage Spanish
- 4 Conclusions

# Goals of the study

**Goal:** To replicate the study in Fuchs et al. 2015 (which tested bundling vs split model in baseline Spanish) with heritage speakers of Spanish, in order to observe how heritage speakers perceive agreement attraction conditions. This in turn allows us to determine how number and gender categories are represented in the heritage grammar.

16 items, each with small clause structure (template in (1), example in (2)). Replicated from Fuchs et al. 2015.

- 1 (Subject) Verb NP1 Prep NP2 ADV ADJ ...
- 2 El niño considera la noticia en la revista terriblemente aburrida...

For each item, different grammatical conditions were formed by systematically manipulating the number and gender of each of NP1, NP2, and ADJ:

NP1 = masculine/feminine x singular/plural = 4 conditions

NP2 = masculine/feminine x singular/plural = 4 conditions

ADJ = masculine/feminine x singular/plural = 4 conditions

**4 x 4 x 4 = 64 conditions**

# Design

Los estudiantes dejaron el cuaderno en el escritorio cuidadosamente cerrado.

Los estudiantes dejaron el cuaderno en el escritorio cuidadosamente cerrada.

Los estudiantes dejaron el cuaderno en la estantería cuidadosamente cerrado.

Los estudiantes dejaron el cuaderno en la estantería cuidadosamente cerrada.

Los estudiantes dejaron la libreta en el escritorio cuidadosamente cerrado.

Los estudiantes dejaron la libreta en el escritorio cuidadosamente cerrada.

Los estudiantes dejaron la libreta en la estantería cuidadosamente cerrado.

Los estudiantes dejaron la libreta en la estantería cuidadosamente cerrada.



Grammatical conditions were recorded by an adult male native speaker of Spanish from Colombia. Ungrammatical conditions were formed by splicing grammatical conditions, in order to avoid phonological cues to grammaticality.

# Design

Participants were presented with each stimulus auditorily and were asked to rate it on a scale from 1 (*completamente inaceptable*) to 5 (*completamente aceptable*).

Pulsa acá para oír una frase. La frase sirve como ejemplo. Cuando la hayas oído, indica si crees que la frase te suena bien. Es decir, si es algo que se puede decir en español.



¿Qué te pareció la frase? ¿Te parece que es algo que se pueda decir en español?  
¿Te parece que es algo tú o alguien que habla español pudiera decir?

1 corresponde a lo siguiente: la frase te parece completamente inaceptable. Es decir, es algo que absolutamente no se puede decir en español, y que ni tú ni nadie que hable español pudiera decir.

5 corresponde a lo siguiente: la frase te parece completamente aceptable. Es decir, es algo que definitivamente se puede decir en español, y que tú o alguien que hable español pudiera decir.

completamente inaceptable  1  2  3  4  5 completamente aceptable

Pulsa "play" para oír otro ejemplo. Una vez que hayas oído la frase, indica si te parece aceptable o inaceptable. Usa la escala de 1 al 5.



¿Qué escogiste para esta frase?

completamente inaceptable  1  2  3  4  5 completamente aceptable

# Participants

- Participants were recruited through Amazon's Mechanical Turk.
- Instructions were given in Spanish.
- Experiment was preceded by a demographics questionnaire:

## Please answer the following questions about your language background

¿Cuál de los siguientes mejor te describe?

- Crecí hablando español y ahora hablo mayormente en inglés
- Crecí hablando español y todavía hablo mayormente en español
- Crecí hablando inglés y aprendí español
- Crecí hablando un idioma que no era español ni inglés, y aprendí español

Si tomó usted clases de español en la escuela, ¿cuántos años de español tomó?

- Nunca he tomado clases de español en la escuela
- 1-2 años
- 3-4 años
- mas de 4 años

¿Tomó usted clases de español después de la escuela secundaria (high school)?

- sí
- no
- no se aplica

Si tomó clases de español en la universidad, ¿cuántos semestres ha tomado?

- no se aplica
- 1 semestre
- 2-4 semestres
- mas de 4 semestres

**¿Ha vivido usted en un país hispanohablante?**

- no, nunca he vivido en un país hispanohablante
- sí, cuando tenía MENOS de 8 años
- sí, cuando tenía MAS de 8 años
- (b) y (c)

**¿Cuántos años vivió en un país hispanohablante?**

- nunca he vivido en un país hispanohablante
- menos de un año
- 1-5 años
- mas de 5 años

**¿Qué idioma habla con su familia?**

- inglés
- español
- español e inglés
- otro idioma

**¿Cómo evaluaría su nivel de proficiencia en español?**

- básico
- intermedio
- con fluidez
- nativo
- prefiero no responder

**¿Cuántos años tiene? (por favor ponga el número)**

# Participants

71 participants were identified as heritage speakers based on satisfying each of the following requirements:

- 1 Grew up speaking Spanish
- 2 Now speak mostly English
- 3 Did not live in a Spanish-speaking country after age 8

# Participants

Mechanical Turk allowed us to gather a substantial sample size, and to identify different proficiency groups based on the demographics questionnaire.

Through one study we were able to gather data on heritage speakers (this project), native speakers (for replicating Fuchs et al.), and L2 speakers (for future work).

# Predictions for bundling vs split

## Predictions for bundling vs split

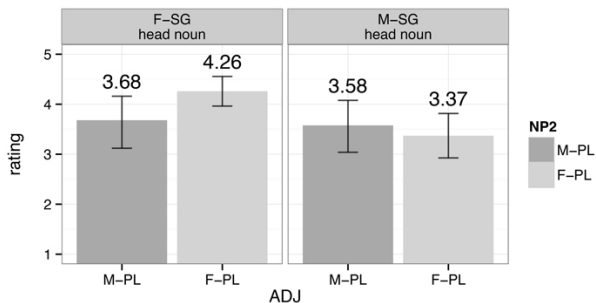
**Prediction 1:** If number and gender are split, agreement attraction in one feature need not lead to agreement attraction in the other feature.

**Prediction 2:** If number and gender are bundled, agreement attraction in one feature should lead to agreement attraction in the other feature.



# Baseline Spanish

Fuchs et al. 2015 tested these predictions on 50 native speakers of Spanish, and found a significant difference between F-SG M-PL M-PL (1) conditions and F-SG F-PL F-PL (2) conditions:



- (1) Los estudiantes dejaron la libreta en los escritorios cuidadosamente cerrados.
- (2) Los estudiantes dejaron la libreta en las estanterías cuidadosamente cerradas.

## Predictions for bundling vs split

**Prediction 1:** If number and gender are split, agreement attraction in one feature need not lead to agreement attraction in the other feature.

~~**Prediction 2:** If number and gender are bundled, agreement attraction in one feature should lead to agreement attraction in the other feature.~~

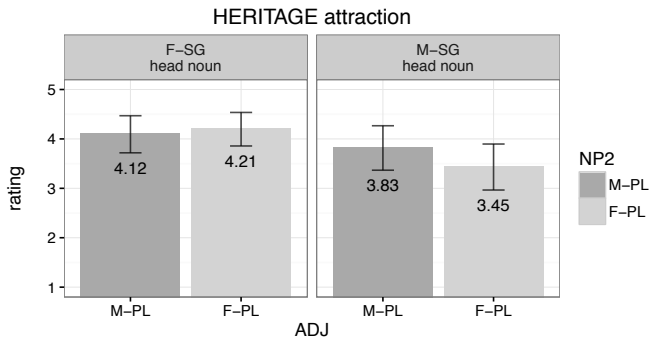
We have evidence that native speakers of Spanish treat number and gender separately, but how do heritage speakers represent these categories?

## Replication

Data from native speaker participants in the study reported here replicated the findings in Fuchs et al. 2015.

# Heritage results

No significant difference between F-SG M-PL M-PL (1) and F-SG F-PL F-PL (2) (or analogous masculine head noun conditions):



(1) Los estudiantes dejaron la libreta en los escritorios cuidadosamente cerrados.

(2) Los estudiantes dejaron la libreta en las estanterías cuidadosamente cerradas.

# Bundling vs split

Conditions with agreement attraction in one feature were rated the same as conditions with agreement attraction in two features, not higher.

## Predictions for bundling vs split

**Prediction 1:** ~~If number and gender are split, agreement attraction in one feature need not lead to agreement attraction in the other feature.~~

**Prediction 2:** If number and gender are bundled, agreement attraction in one feature should lead to agreement attraction in the other feature.

Number and gender are bundled in Heritage Spanish.

# Bundling vs split

Baseline Spanish: Number and gender are separate.

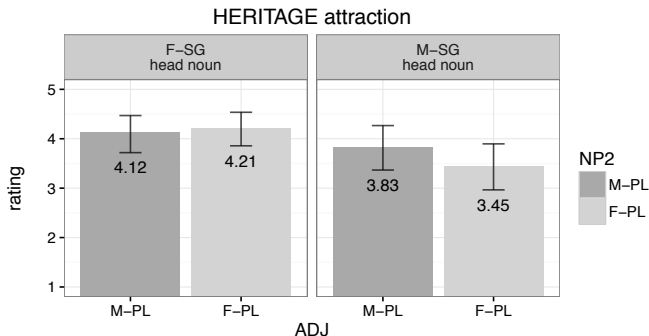
Heritage Spanish: Number and gender are bundled.

→ Heritage Spanish is different.

**A possible counterargument:** What if the heritage results reflect lack of sensitivity to morphology rather than differences in category structuring?

# Bundling vs split

**Response to the counterargument:** If heritage speakers were that insensitive to morphology, there should be no differences between any of the attraction conditions, which is not the case.



# Interim Conclusion

Baseline Spanish: Number and gender are separate.

Heritage Spanish: Number and gender are bundled.

→ Heritage Spanish is different.



# Table of Contents

- 1 Background
  - Bundling vs split
  - Agreement attraction
- 2 Experiment
  - Methodology
  - Predictions
  - Baseline Spanish
  - Heritage results
  - Bundling vs split
- 3 Economy in Heritage Spanish
- 4 Conclusions

# Economy principles

Heritage speakers are performing the costly task of production and comprehension in their non-dominant language. Their grammar may therefore be more sensitive to principles of transparency and economy.

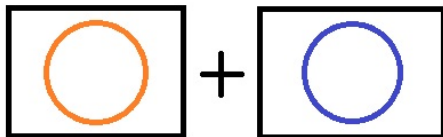
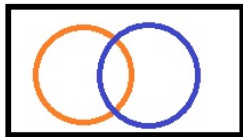
The question is, which type of economy guides the restructuring of the grammar? **Representational economy** or **processing economy**?

# Representational economy

## Representational economy:

Storing a single form in memory that may be used in multiple contexts is more economical than storing several distinct forms.

→ **bundling will be preferred**, because the feature bundle serves as a single unit stored in memory, rather than two separate features

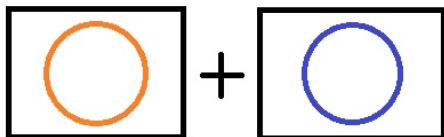
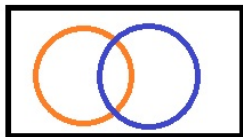


# Processing economy

## Processing economy:

A single and invariant form-meaning mapping is more economical than resolving ambiguity in context.

→ **split model will be preferred**, because one-to-one mappings are easier to check at a glance; disentangling a complex category to check a particular feature is taxing



# Economy principles

Representational economy → bundling

Processing economy → split

We normally assume that heritage speakers are always trying to facilitate processing ...

... but these results show divergence in heritage grammars is evident not just in processing but also in the structure of the grammar.

# Table of Contents

- 1 Background
  - Bundling vs split
  - Agreement attraction
- 2 Experiment
  - Methodology
  - Predictions
  - Baseline Spanish
  - Heritage results
  - Bundling vs split
- 3 Economy in Heritage Spanish
- 4 Conclusions

# Conclusions

- Presented results of a comprehension experiment in Spanish that employed agreement attraction.
- **Main question:** Whether or not baseline and heritage speakers handle structuring of number and gender in the same way.
- **Result:** Heritage speakers have bundled number and gender, unlike baseline...
- ... but not because they don't see morphology
- **Implications:** Differences in heritage grammar from baseline can't all be boiled down to processing, they reflect differences in underlying structure

## Going further

It seems that representational economy has more of a role in shaping the structure of number and gender, but might there be instances in the heritage grammar in which economy of processing will outweigh representational economy?

Does the structure of number and gender in the L2 grammar resemble that of the heritage grammar or the baseline? What does this suggest about the principles that shape mental representations in L2 grammars?



## Selected References

- Antón-Méndez**, Inés, Janet L. Nicol & Merrill F. Garrett. 2002. The relation between gender and number agreement processing. *Syntax* 5. 1–25.
- Carminati**, Maria Nella. 2005. Processing reflexes of the Feature Hierarchy (person > number > gender) and implications for linguistic theory. *Lingua* 115. 259–285.
- Carstens**, Vicki. 2000. Concord in minimalist theory. *Linguistic Inquiry* 31. 319–355.
- Carstens**, Vicki. 2003. Rethinking complementizer agreement: Agree with a case-checked goal. *Linguistic Inquiry* 34. 393–412.
- Fuchs**, Zuzanna, Maria Polinsky & Gregory Scontras. 2015. The Differential Representation of Number and Gender in Spanish. *The Linguistic Review* 32. 703 – 737.
- Picallo**, María Carmen. 1991. Nominals and nominalization in Catalan. *Probus* 3. 279–316.
- Ritter**, Elizabeth. 1993. Where's gender? *Linguistic Inquiry* 24. 795–803.

# Acknowledgments

Work on this project was supported in part by grants from the Center for Advanced Study of Language (CASL) at the University of Maryland, from the NSF (SMA-1429961) to Maria Polinsky, and from the Observatory of the Spanish Language in the U.S. at Harvard University to Maria Polinsky and Zuzanna Fuchs. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors.

We would like to thank Giuli Dussias, Ruth Kramer, Brooke Larson, Nicholas Longenbaugh, Francisco Moreno-Fernández, Nina Radkevich, Rodrigo Ranero, Allison Ray, Andrés Sanín, and the audiences at GLOW 37 and The Polinsky Language Sciences Lab for helpful comments on this project.