# Acquisition of Irregular Patterns in Spanish Verbal Morphology

**Laia Mayol** (laia@ling.upenn.edu)
University of Pennsylvania


## 1. Introduction

English past tense acquisition has been the subject of a long controversy between connectionism and generative linguistics. However, English past tense has simple morphology and phonology. Thus, the debate can benefit from looking at a morphologically more complex language.

Goals of this paper:

1. Analyze data from acquisition of irregular patterns in Spanish verbs.
2. Is the data compatible with all models of morphology acquisition?

## 2. Background

### 2.1 Irregular patterns

•Analysis of **lexically arbitrary** verbal alternations.
•They are a minority pattern in the 1st conjugation and more common in the 2nd and 3rd [AAH, 2000]
•All the verbs in the **3rd conjugation** undergo some alternation.

A. Diphthongization in stressed syllables:
[e] → [ie] comenzár vs. com**ié**nzo
[o] → [ue] c**o**ntár vs. c**ué**nto
B. Raising in stressed syllables (only 3rd conj):
[e] → [i] pedír vs. pído
C. Velar insertion:
∅ → [g | k] poner vs. pon**g**o

### 2.2 Morphology acquisition

**Words and Rules** (WR, [PU, 2002]): all regulars are derived by a rule, while irregular forms are individually stored in the lexicon.

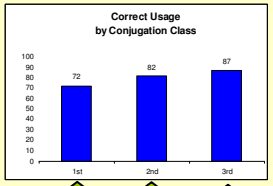
**Rules and Competition** (RC, [YANG, 2003]): there are rules both for the regular forms and for the irregular forms and these rules compete against each other. Acquisition task consists of:

1. Learning whether an irregular rule R applies to a certain subclass of verbs.
2. Learning whether a verb x belongs to a certain irregular class of verbs S

## 3. Data and Results

•Data: transcriptions of six Spanish-speaking children from the CHILDES database [MS, 1985].
• Extraction of:
1. Correct verbs containing one of the irregular patterns: 'quiero'
2. Overregularized verbs: 'quero'
3. Overirregularized verbs: 'cuemo'

•Results: 345 overregularizations found
0 irregularizations found




-1st conj. lower CUR

-2nd conj. larger input

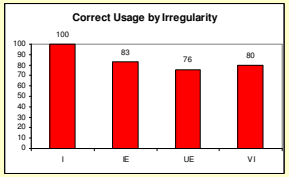
-3rd conj. higher CUR → all 3rd conj. verbs undergo some alternation

| | | | |
|---|---|---|---|
| Child # | 214 | 1456 | 152 |
| Adult # | 740 | 4711 | 213 |

**Figure 1.** Correct Usage Rate (# correct verbs / # total verbs) for each conjugation. Tokens used by children and adults.

- Raising verbs 100% CUR.

-Other three types: the more input from adults the higher the CUR




| | | | | |
|---|---|---|---|---|
| Child tokens | 13 | 1090 | 357 | 362 |
| Adult tokens | 39 | 2373 | 1508 | 1744 |

**Figure 2.** Correct Usage Rate (# correct / # total verbs) for each irregularity. Tokens used by children and adults.

The differences in Correct Usage Rate are statistically significant both for conjugation class and type of irregularity: $p<0.05$ in a $\chi 2$ test for both distributions.

## 4. Analysis

### 4.1 Statistics

- Out of our variables (conjugation class, type of irregularity, adult input), which one gives us the best statistical model?
- Best regression contains 2 variables: (1) input from adults and (2) whether a verb belongs to the 1st conj.
- This model achieves significant results: $F = 6.1$, $p < 0.02$

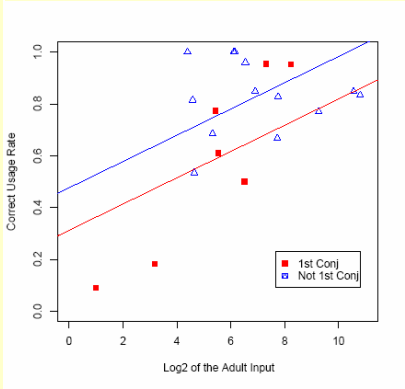



**Figure 3.** Plot of the irregular verbs by their CUR and the adult tokens (transformed by a base 2 logarithm) and of the lines fitted by the statistical model.

• The more input from adults for a particular verb, the easier the verb is for children (as indicated by the slope).
•1st conj. verbs are tougher (the red line is lower than the blue line).

### 4.2 Main findings

1. Many instances of overregularizations were found (345). No instances of overirregularizations were found.
2. The two main predictors of a verb's CUR are:
(a) adult input
(b) membership in the first conjugation
3. There are significant differences in the verb's correct usage according to both conjugation and type of irregularity.
4. Some verbs have a perfect CUR of 1, although there was no adult input in our data

**Not consistent with the WR model!**

## 5. Conclusions

• Can the Words and Rules model account for all the findings?

• NO!
• WR predicts that irregular verbs are individually stored in the lexicon.
• Conjugation or type of irregularity should not have any effect on how well the verb is used, contrary to the facts.
• The gap between the blue and the red line in Figure 3 should not be significant, but it is.

• All the findings are compatible with the Rules and Competition model, which proposes that irregular verbs are organized in subclasses and derived by rules

## References

[AAH] Albright, A., A. Andrade, and B. Hayes (2000). Segmental environments of Spanish diphthongization. UCLA Working Papers in Linguistics.
[MS] MacWhinney, B. and C. Snow (1985). The child language data exchange system. *Journal of Child Language* 12, 271–296.
[PU] Pinker, S. & Ullman, M. (2002) The past and future of the past tense. *Trends in Cognitive Science* 6:456-463.
[YAN] Yang, C. D. (2003) *Knowledge and Learning in Natural Language.* Oxford University Press.

## Acknowledgements

Many thanks to Charles Yang for guiding me in this project and making the automatic data extraction possible. Thanks also to the three anonymous reviewers for their very detailed and useful comments, to Aviad Eilam, Josh Tauberer and Lucas Champollion for suggestions and for proof-reading this poster and to Mercè Pi for help in designing the poster.