

Learning English irregular plurals – why irregulars are like regulars.

Inbal Arnon & Eve V. Clark,
Linguistics, Stanford University, inbalar@stanford.edu



ABSTRACT. How irregular forms are acquired has been a source of debate between single-route and dual-route models of lexical representation. Both models can accommodate over-regularization errors, but while the dual-route model predicts that all irregular items should induce an equal number of errors (Marcus et al. 1992), the single-route one predicts that error rate will be related to token frequency, as found in an earlier corpus study (Maslen et al., 2004). The current study provides support for a single-route model by showing that over-regularization errors are affected by token frequency and modality type (recall vs. recognition) both of which also affect lexical access for regular forms. The study also tests the effect of the situational context (naming vs. teaching) on children's errors to add a social dimension to what affects the activation level of lexical items.

Introduction

Why do children produce forms like 'mouses' and 'foots'?

- Dual-route models: Regulars and irregulars are stored and learned in different ways.
 - Errors reflect application of a regular rule to the irregular form

Prediction: Error rate should be equal for all irregular items

Support: Constant 4% error rate in a corpus analysis (Marcus 1992: 1996, 2000)

- Single-route models: Regulars and irregulars are stored in the same way.
 - Errors reflect competition from the greater type and token frequency of the regular plural marker

Prediction: Error rate is related to token frequency – fewer errors on high token frequency irregulars.

Support: Higher token frequency yields fewer errors in a corpus (Maratsos, 2000; Maslen et al: 2004)

The current study – eliciting English irregular plurals from children

Prediction: The same factors affect lexical access for regular forms

- Frequency effects: fewer errors for irregulars with high token frequency
- Modality effects: fewer errors for recognition compared to recall
- Situational frequency: fewer errors in more familiar social settings

Method

Participants. 22 children, 10 four-year-olds (mean 4;0, range 3;6-4;2) and 12 five-year-olds (mean 5;0, range 4;9-5;2)

Material: 9 irregular nouns (feet, snowmen, mice, deer, geese, sheep, teeth, fish, children). Each noun was given a frequency estimate (token frequency / total irregular token frequency) calculated using speech from a 4 million word subset of the English corpora in CHILDES (MacWhinney, 2000).

Item	Raw Count	Freq Estimate
Snowmen	9	.002
Geese	59	.01
Deer	102	.02
Mice	206	.04
Sheep	412	.07
Children	643	.12
Teeth	967	.18
Feet	1215	.22
Fish	1813	.34
Total	5426	1

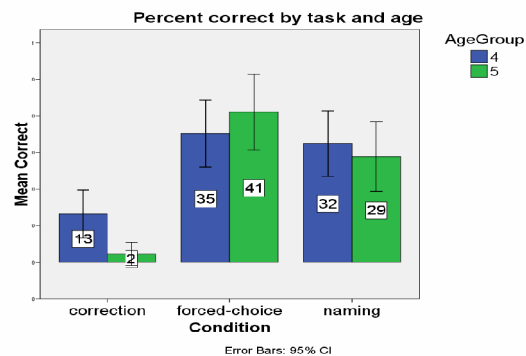
Tasks: Each child completed 3 tasks: picture naming task, correction, and forced choice. Each task had colour-naming fillers.

- Picture-naming task: Children saw pictures and had to name them.
- Correction task: Children helped a puppet learn English by correcting its mistakes when naming pictures. The puppet always produced over-regularized forms.
- Forced-choice task: children had to choose between the correct and the over-regularized forms presented by the puppet (e.g. do you say foos or feet?)

Results

1. High error rate overall

- 38% correct in forced-choice, 31% in naming and 8% in correction.



2. What affects children's error rate?

- Lexical frequency: fewer errors when token frequency was higher ($B = 5.69, p < .001$)
 - Controlling for phonological similarity between singular and plural (e.g. deer, sheep, fish).
- Situational frequency: more errors in the correction task ($B = -1.87, p < .001$)
- Task type: marginally fewer errors ($B = .39, p = .09$) for recognition
- No effect of age. Five-year olds still made as many errors as four-year-olds, $p = .6$

1. The analyses were done using mixed effect logistic regression models, with subject and item as random effects.

3. What kinds of errors do children make?

- For naming errors, children produced the over-regularized form 80% of the time (e.g.foots, toots, fishes), the singular form 18% of the time, and a 'double' plural 2% of the time (e.g. feets, teeths)
- For forced-choice errors, children chose the over-regularized form 64% of the time, produced a singular 9% of the time and accepted both forms 27% of the time
 - Some interchangeable use of the correct and over-regularized form

Discussion

The results show an effect of token frequency on over-regularization errors. As predicted by single-route models, items with higher corpus frequency elicited fewer errors. This result held when controlling for the phonological similarity between the singular and plural forms. Children's performance was influenced by the task. They chose the correct form more often than they produced it. This mirrors differences between recall and recognition in lexical access. Children accepted both forms in the forced-choice task 17% of the time, demonstrating some interchangeable use (also observed in spontaneous speech). Children's performance was worst in the correction task (only 8% correct), suggesting that it is not only token frequency but also situational frequency that influences performance. This resembles the finding that lexical access is influenced by affective and situational factors (Goldinger, 1998; Nygaard & Lunders, 2002). The results reveal effects of token frequency, some interchangeable use of correct and incorrect forms, and effects of task type, all results that are easier to accommodate within a single-route model.

References

- Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425-455
- Goldinger, S.D (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251-279
- Maratsos, M. (2000). More overregularizations after all. *Journal of Child Language*, 28, 32-54.
- Marcus, G. F. (1996) Why do children say 'breaaked'? *Current Directions in Psychological Science*, 5, 81-85
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., and Xu, F. (1992). Overregularization in Language Acquisition. *Monographs of the Society for Research in Child Development*. 57 (4, Serial No. 228)
- Maslen, R., Theakston, A., Lieven, E., & Tomasello, M. (2004). A dense corpus study of past tense and plural overgeneralizations in English. *J. of Speech, Language, & Hearing Research*
- Nygaard, L.C. & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Journal of Memory and Cognition*, 30, 583-593
- Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of parallel distributed processing models of language acquisition. *Cognition*, 28, 73-193.