No need to memorize arbitrary exceptions!

Elizabeth Coppock

Department of Linguistics
Stanford University

Center for Language Sciences, Rochester
May 23rd, 2008

Question

How much syntactic information do we store about individual words?

Some answers

- Usage-based models: *A lot*. We record information about how particular words are used in particular syntactic patterns, even if the pattern is productive (Goldberg 2006)
- "Traditional mainstream generative grammar": Not very much.
- Intermediate position: There are some idiosyncrasies to memorize, but there is no need to store arbitrary exceptions to productive patterns.

Baker's Paradox (Pinker 1989)

- **Productivity**: e.g. *text me the address*.
- **Arbitrariness**: **donate the library a book* – just an arbitrary fact about *donate*?
- **No negative evidence**: Well, the idea is, nobody tells you not to use *donate* in the double object construction, or if they do, you ignore them. Controversial. (Braine 1971; Brown and Hanlon 1970; Hirsch-Pasek et al. 1984; Bowerman 1988; Chouinard and Clark 2003)

Why a paradox?

All P's are Q's
There is an x that is a P but not a Q

Possible solutions to the paradox

- Negative evidence: All P's except x are Qs.
  - Implicit: $x \in P$ is expected to show property Q but does not (Braine and Brooks 1995; Brooks et al. 1999; Schütze 1997)
  - Explicit: Through adult reformulations of child speech (Chouinard and Clark 2003)
- Criteria-governed productivity (Pinker 1984): Revise the rule to, "All R's are Q's," where x is not an R.

Domains of claimed arbitrariness

- Prepositions (Culicover 1999)
  - Stranding (*the challenges we persevered* **through/*despite**)
  - Pied-piping (*the door* **through/*out** *which he went*)
- Adjectives (Goldberg 2006)
  - Prenominal (*an* **angry/*aghast** *man*)
  - Predicative (*an imitation that is* **small/*mere**)
- Causatives (Bowerman and Croft 2008)
  - *You're going to* **fill/*overflow** *the spoon*
- Ditransitives (Schütze 1997)
  - *I* **gave/*donated** *the library a book*

Constraints on productivity for ditransitives

- Transfer of possession (Green 1974) $\Rightarrow$ **mend*
- No extra manner specifications (Krifka 1999) $\Rightarrow$ **drag*
- But these don't account for all non-ditransitive verbs: **donate*, **explain*...

Is there a "morphophonological constraint"? (Gropen et al. 1989)

- Verbs of Latinate origin tend to be less acceptable as ditransitives (*donate*, *transfer*, *admit*, *confess*, *exhibit*, *illustrate*, *recommend*...)
- "Presumably children lack a collective racial memory for the history of the language..." (Pinker 1989:46)
- But perhaps there is an observable difference in word shape between Latinate and non-Latinate verbs.

Syllable counting hypothesis (see Green 1974)

- Ditransitives can only have one syllable.
- Counterexamples: ditransitive *promise*, *assign*, *allot*

Prosodic weight hypothesis

- Grimshaw and Prince (1986); Grimshaw (2005): Whether a verb is ditransitive or not is predictable based on prosodic weight.
- More than one metrical foot ⇒ non-ditransitive.
  - Feet contain only one stress ⇒ **(do)(nate)* vs. *(promise)*
  - Feet in English are trochaic ⇒ **(ex)(plain)*
  - Schwas are extra-metrical ⇒ *a(ssign)*, *a(llot)*

Why should feet matter?

- Grimshaw (2005): Multiple lexicons, G(ermanic) and R(omance). Prosodic weight allows learner to classify words into the lexicons.
- Pesetsky (1995): Null affix -*G* which licenses the DO cxn; phonological restrictions on affixation are familiar.
- Anttila (2007): Syntax/metrical phonology interface.

Morphological complexity hypothesis

- Verbs must be monomorphemic in order to be ditransitive (Storm 1977)
- Harley (2006): *ex-* in *exhibit* is an incorporated particle; *exhibit* has the same underlying structure as *show off*

Prosodic weight vs. morphological complexity: Which is right?

- Very correlated in English lexicon
- A nonce word study will help us decide!
  - But first some preliminaries.
  - Hang on until Experiment 2!

Outline

Evidence for a "morphophonological constraint"

THE LEARNABILITY AND ACQUISITION
OF THE DATIVE ALTERNATION IN ENGLISH

JESS GROPEN STEVEN PINKER MICHELLE HOLLANDER
Massachusetts Institute of Technology

RICHARD GOLDBERG
University of Maryland

RONALD WILSON
Digital Equipment Corporation

Gropen et al.'s nonce verbs

Two sets of nonce verbs (8 total)

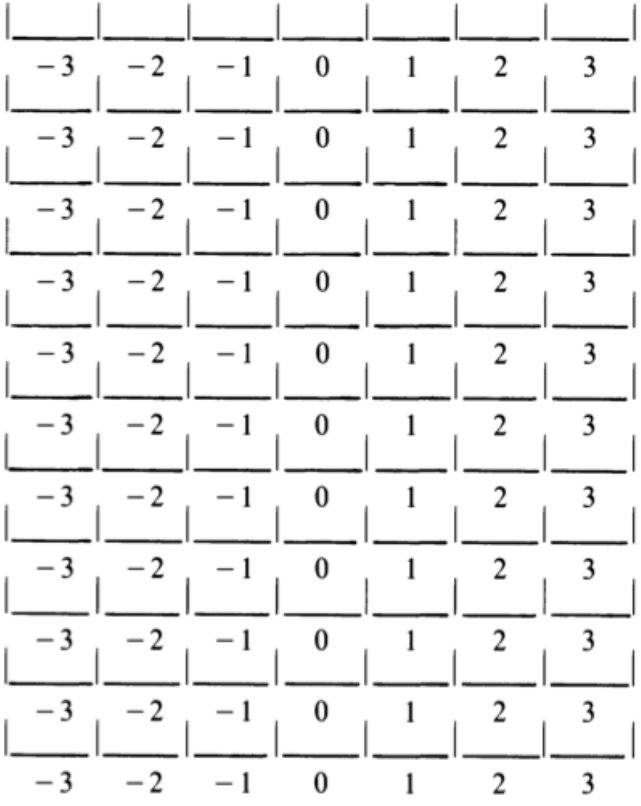
1. Monosyllabic: *norp, pell, moop, tonk*
3. Trisyllabic: *calimode, repetrine, orgulate, dorfinize*

Gropen et al.'s paragraphs (8 total)

John, the star player of the Boston Whalers, was eager to begin the match against the New York Maulers. He knew that he would be facing their champion – Ben – also adept at using the **pell**. And sure enough, at a critical point in the game, John summoned all of his strength and **pelled** the disc to Ben.

	to	*for*
Possessive	2	2
Non-possessive	2	2

Example response form



A ball was pelled to the player by Fred.

Fred pelled a ball to the player.

Fred pelled a ball.

The player was pelled a ball.

A ball was pelled by Fred.

Pell the player a ball!

A ball was pelled.

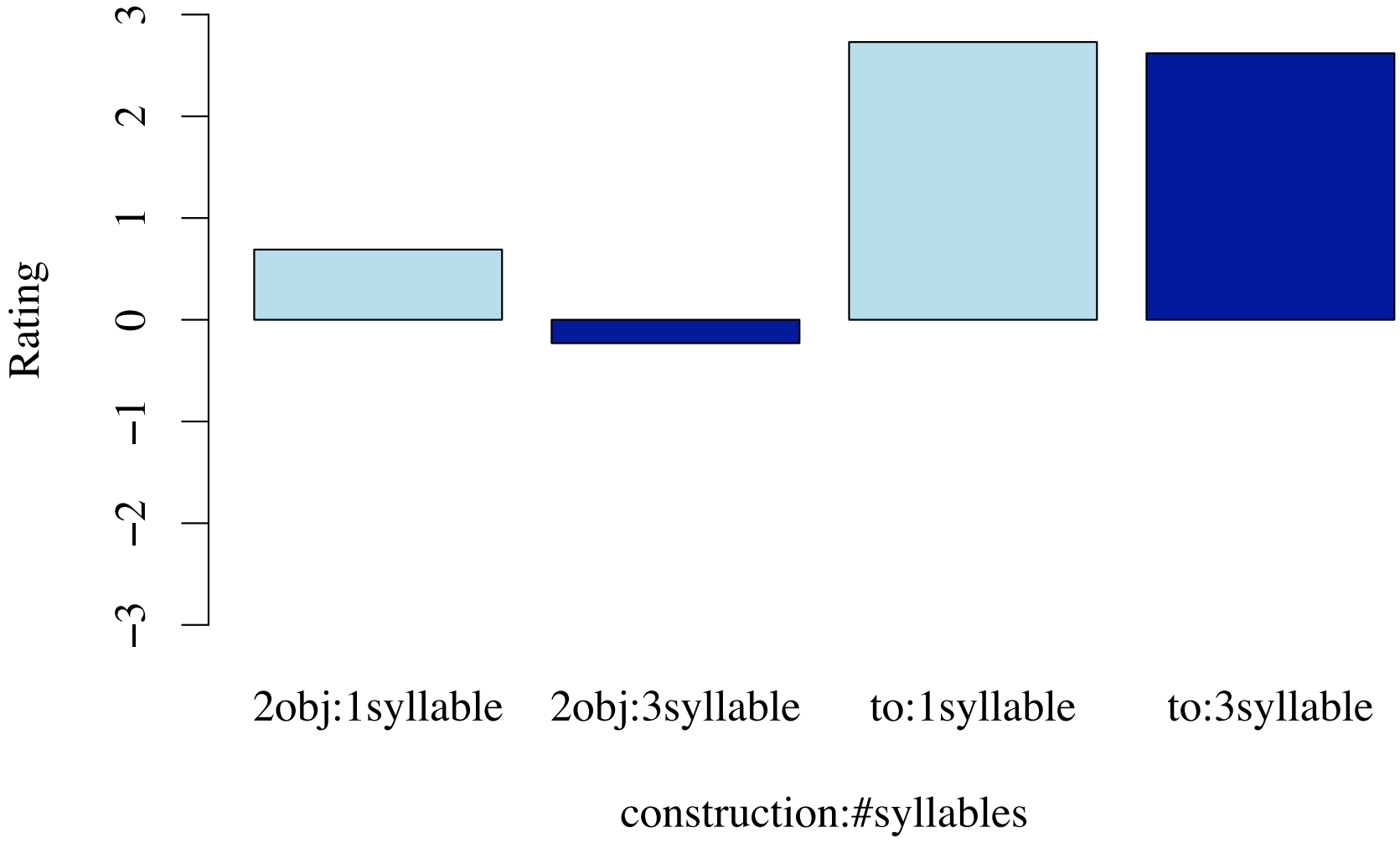
Pell a ball!

Fred pelled the player a ball.

The player was pelled a ball by Fred.

Pell a ball to the player!

Gropen et al.'s results (Possessive/*to* condition)



Limitations of Gropen et al.'s study

- Not held constant:
  - Number of syllables
  - Prosodic weight
  - Morphological complexity
    (*calimode*, *repetrine*, *orgul***ate**, *dorfin***ize**)
  - General Latinateness
- Only two relevant paragraphs

Outline

Goals

- Make sure that there really is a difference between 1-foot (prosodically light) and 2-foot (prosodically heavy) verbs, before testing the effect with nonce verbs
- Validate acceptability judgment methodology for nonce word studies

Verbs tested

1 FOOT	2 FEET
allot	entrust
assign	present
award	provide
promise	dictate
render	donate
signal	forfeit

Sentences to judge

- 12 test sentences (Bresnan et al. 2007 factors controlled)
  - 2obj The teacher will **forfeit** [the student] [the trophy] after the game.
  - pp The teacher will **forfeit** [the trophy] [to the student] after the game.
- 48 fillers

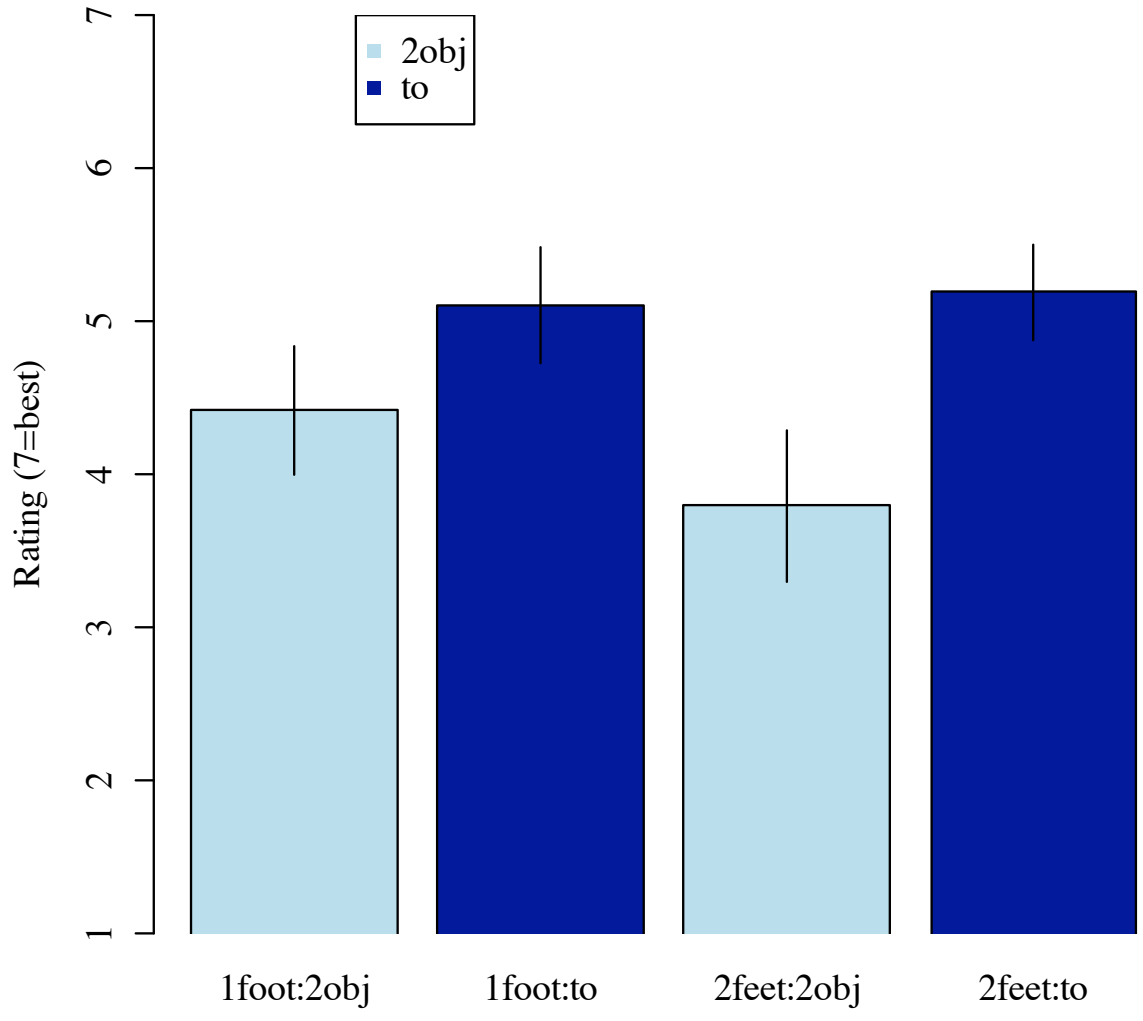
Procedure for each sentence

- The sentence appears on the screen, which the subject reads aloud.
- The subject is asked to rate the sentence on a 1-7 scale (7=best).
- The subject is asked a yes/no comprehension question, and given feedback regarding whether their answer was correct.

Subjects

- 16 subjects
- All Stanford students
- Paid $14

Ratings ~ Metrical feet



Experiment 1: Conclusion

- On the whole, 2-foot verbs were judged less acceptable in the double object construction than prepositional dative construction, relative to 1-foot verbs, as expected.
  - Caveat: this effect was not found for all items: *provide* was judged acceptable in the double object construction.

Now we can move on to nonce words

- There is some evidence for a prosodic weight effect in English.
- Does it govern productivity?

Outline

Goal

Gropen et al. (1989) found:
calimode / repetrine / dorfinize / orgulate < *moop / pell / tonk / norp*

- Would we get the same result with nonce verbs differing only in prosodic weight, holding everything else constant?

Verbs

1 FOOT	2 FEET
feffam	*feffame*
zeepik	*zeepike*
dassud	*dassude*
laundib	*laundibe*
...	...

Paragraphs

Ted, a native of the North, was quite unfamiliar with the customs of the South, where his wife Kate was from. For instance, he had no idea that he was supposed to **feffame** when he proposed. Luckily, the mother of the bride set him straight and graciously explained how to **feffame** the ring to Kate.

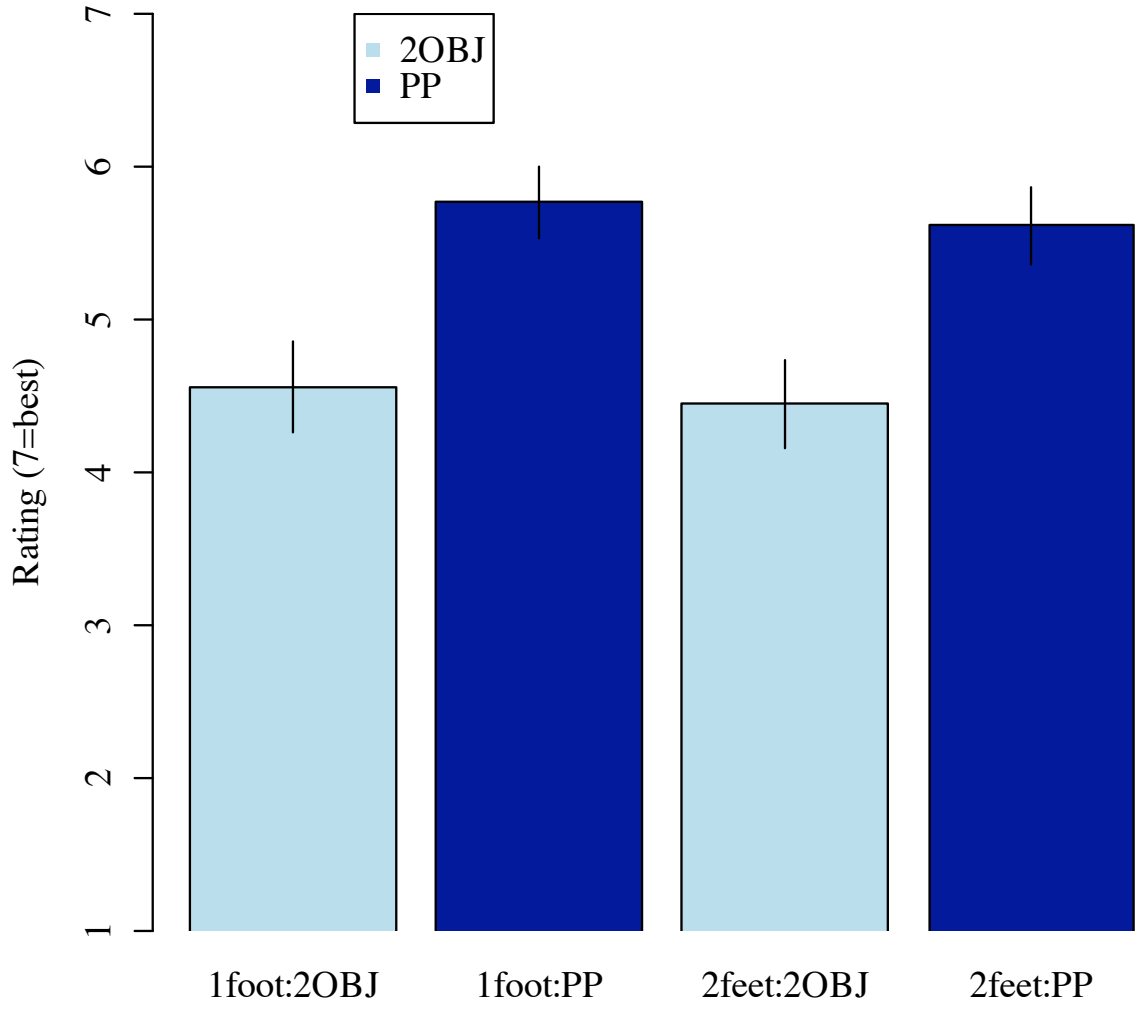
Procedure

- A recording of the paragraph was played over the headphones (in order to ensure that subjects heard the correct pronunciation)
- A series of 4 sentences was presented on the screen. For each sentence, the subject was asked to:
  - Read the sentence out loud (to ensure correct pronunciation)
  - Judge it on a 1-7 scale (7=best).
- For test items, the 4 contained a 2OBJ and a PP.

Subjects

- Same 16 subjects as in English verb experiment
- Completed this experiment first

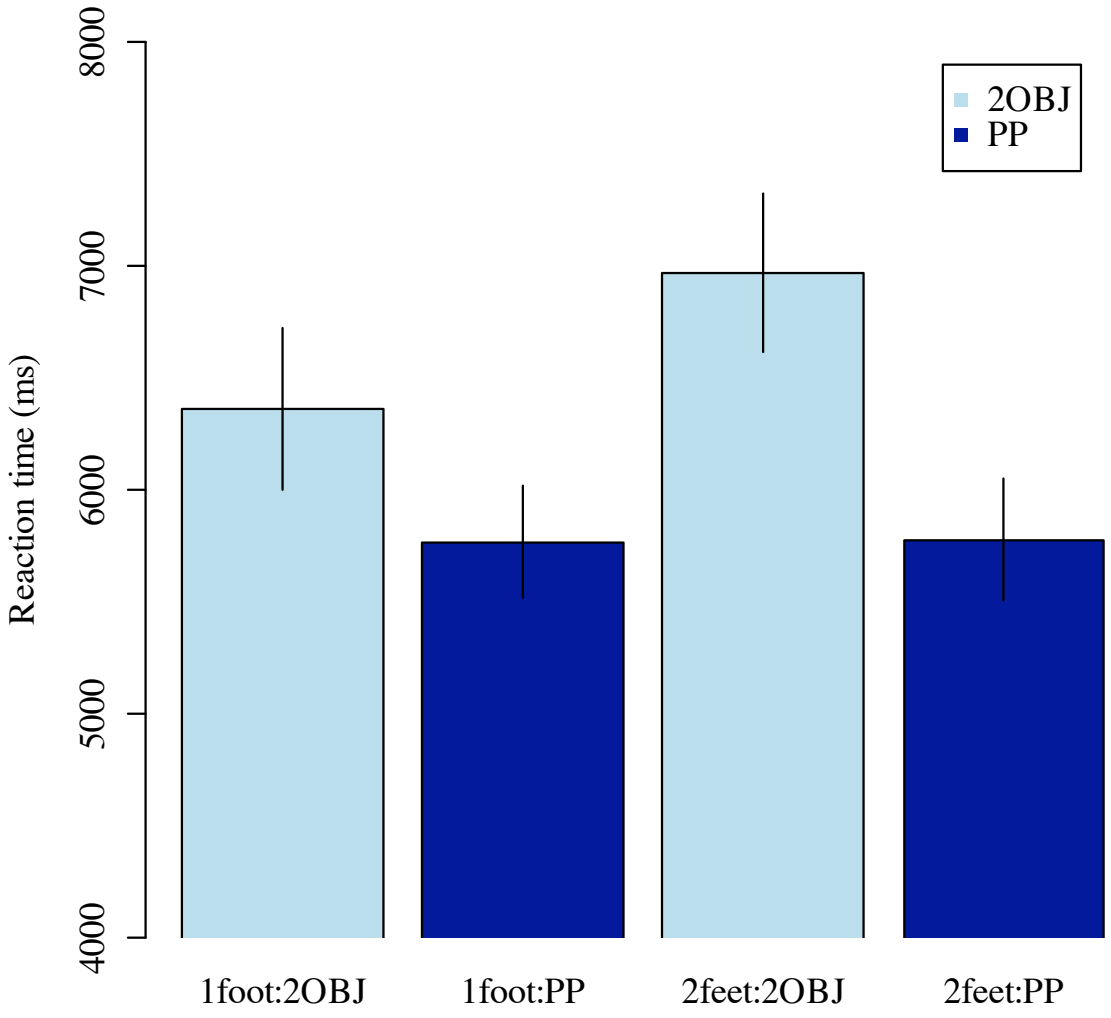
Ratings ~ Metrical feet



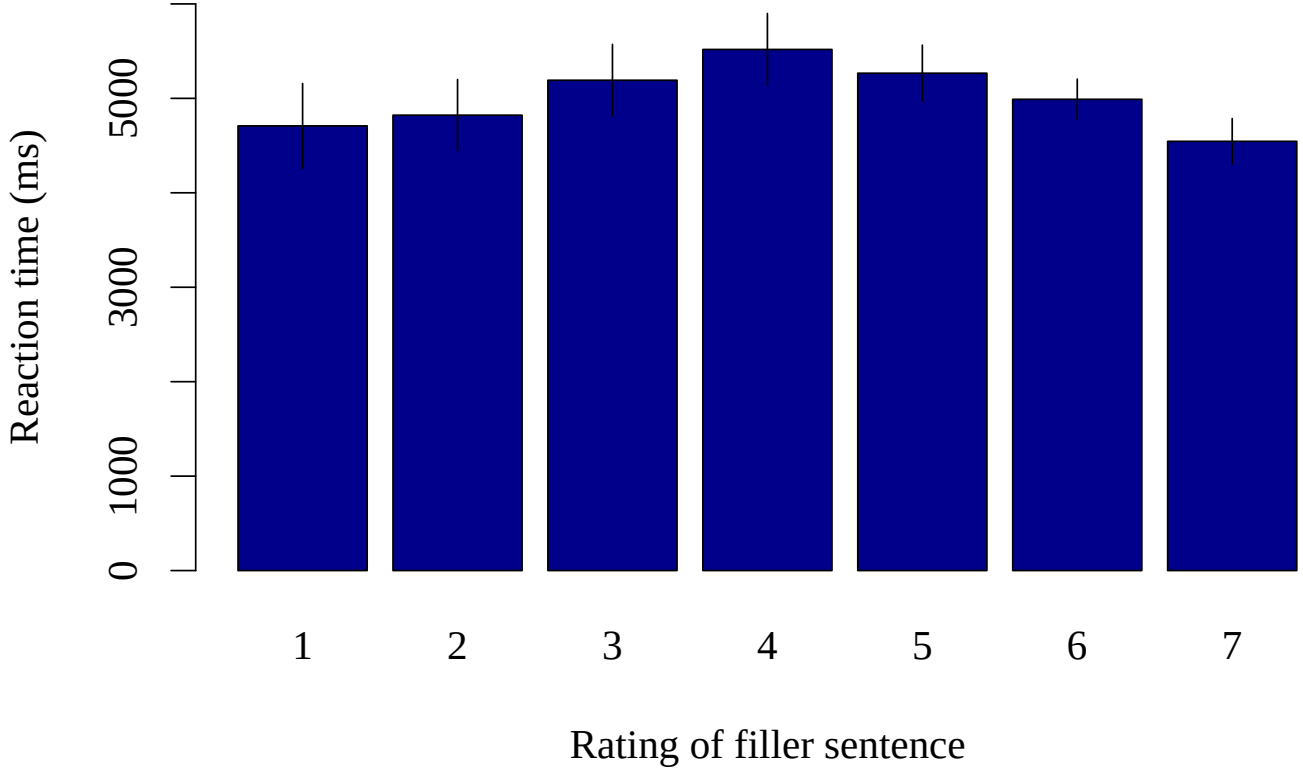
Why no sensitivity to prosodic weight?

- Native English speakers do not show the same sensitivity to prosodic weight when judging the acceptability of nonce verbs as when judging the acceptability of English verbs. Why?
  - Prosodic weight does not govern productivity of 2OBJ cxn?
  - Conscious strategy?

Reaction times



Rating vs. reaction time on fillers



What do reaction time differences mean?

- Slower reaction times → more *intermediate* acceptability levels.
- Thus the reaction time interaction suggests that long verbs are more *intermediate* than short verbs in the double object construction, despite a conscious strategy that masks the effect on acceptability judgments.

Experiment 2: Conclusion

- This is the first evidence that prosodic weight makes an independent contribution to the acceptability of double object constructions.
- But several open questions remain.

How come...

In acceptability ratings of nonce verbs,

$$\textit{feffame} = \textit{feffam} \text{ but } \textit{moop} \neq \textit{dorfinize}$$

- Conscious strategy due to experimental design? Gropen et al. (1989) had more types of filler constructions.
- Speakers are sensitive to number of syllables?
- Speakers are sensitive to morphological complexity?
- Speakers are sensitive to Latinateness?

In fairness...

Grimshaw (2005) emphasizes that prosodic weight is a *cue* telling speakers which lexicon a verb is in. Maybe there are overriding cues of G-lexicon membership here?

Outline

Goals

- Make it much more difficult for subjects to develop a conscious strategy.
- Replicate Gropen et al.'s results.
- Tease apart the effects of number of syllables, number of feet, morphological complexity, and Latinateness.

Procedure

- Subjects read a paragraph containing the word (silently)
- A series of 10 sentences is presented on the screen. For each sentence, the subject is asked to judge how good the sentence is as a use of the word on a 1-7 scale (7=best).
  - 2 sentences were critical: 2OBJ and PP
  - 8 were fillers

Four verb shapes

1 Monosyllabic (*moop*)

3 Trisyllabic (*dorfinize*)

2g Disyllabic with Germanic morphemes (*gelaut*)

2l Disyllabic with Latinate morphemes (*sevolve*)

Nonce verbs (32 total)

1	2G	2L	3
stott	gestosh	submote	orgulate*
tonk*	erskall	obtend	secutize
goam	forhoove	sevolve	volutize
moop*	besloff	expute	sedify
pell*	forqueath	tranject	pugnavate
dorf	gelaut	subfect	crocinize
norp*	erblick	affute	flumenate
blint	begroat	procuse	sulcify

Paragraphs (32 total)

Ron, who had promised Dave that he would provide him with some data for his research, was feeling some regret. It had been a full month since he had last tried to **sevolve**, and he was worried that the information might not go through. He was very relieved after he was able to **sevolve** the crucial statistics to Dave.

Example critical sentences

2obj Ron was happy to **sevolve** Dave the statistics. [declarative]

pp **Sevolve** the statistics to Dave! [imperative]

– *or* –

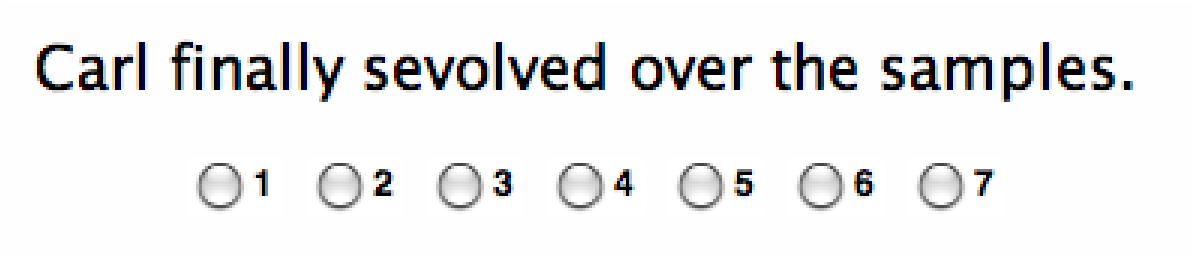
2obj Who **sevolved** Dave the statistics? [*wh* question]

pp Did Ron **sevolve** the statistics to Dave? [*yes/no* question]

Filler sentence types (declarative versions)

- The statistics will be **sevolved** to Dave.
- Dave was **sevolved** the statistics by Ron.
- Ron will **sevolve** soon.
- Ron **sevolved** the statistics perfectly.
- This statistics will **sevolve** well.
- Ron **sevolved** the statistics for Dave.
- Ron finally **sevolved** over the statistics.
- Ron refused to **sevolve** Dave out the statistics.

Dependent variable: Rating (1-7 scale)



Counterbalancing

- Four versions of the experiment, each with a different assignment of words to paragraphs.
  - Each word appeared equally in C and M paragraphs
  - Each paragraph appeared equally with each word shape (1, 3, 2L, 2G)
  - Each construction (2OBJ, PP, etc.) appeared equally with each clause type (IMP, DECL, YN, WHSUBJ)

Subjects

- 32 new subjects, mostly for class credit, some paid $14
- All adult native English speakers

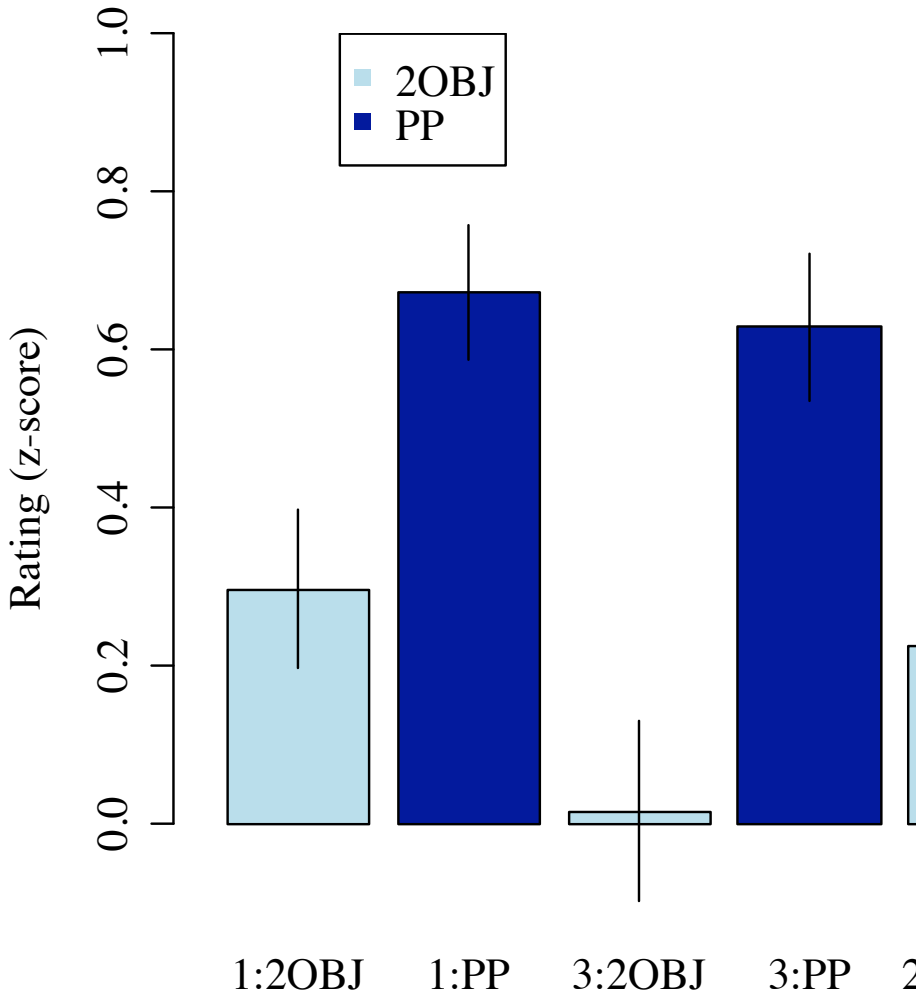
Each subject completed:

- Nonce word rating experiment
- Memory test (not surprise)
- Surprise etymology questionnaire
  (*Old English*, *Ancient Greek*, *French*, or *Latin*)
- Follow-up questionnaire on morphological complexity (online)

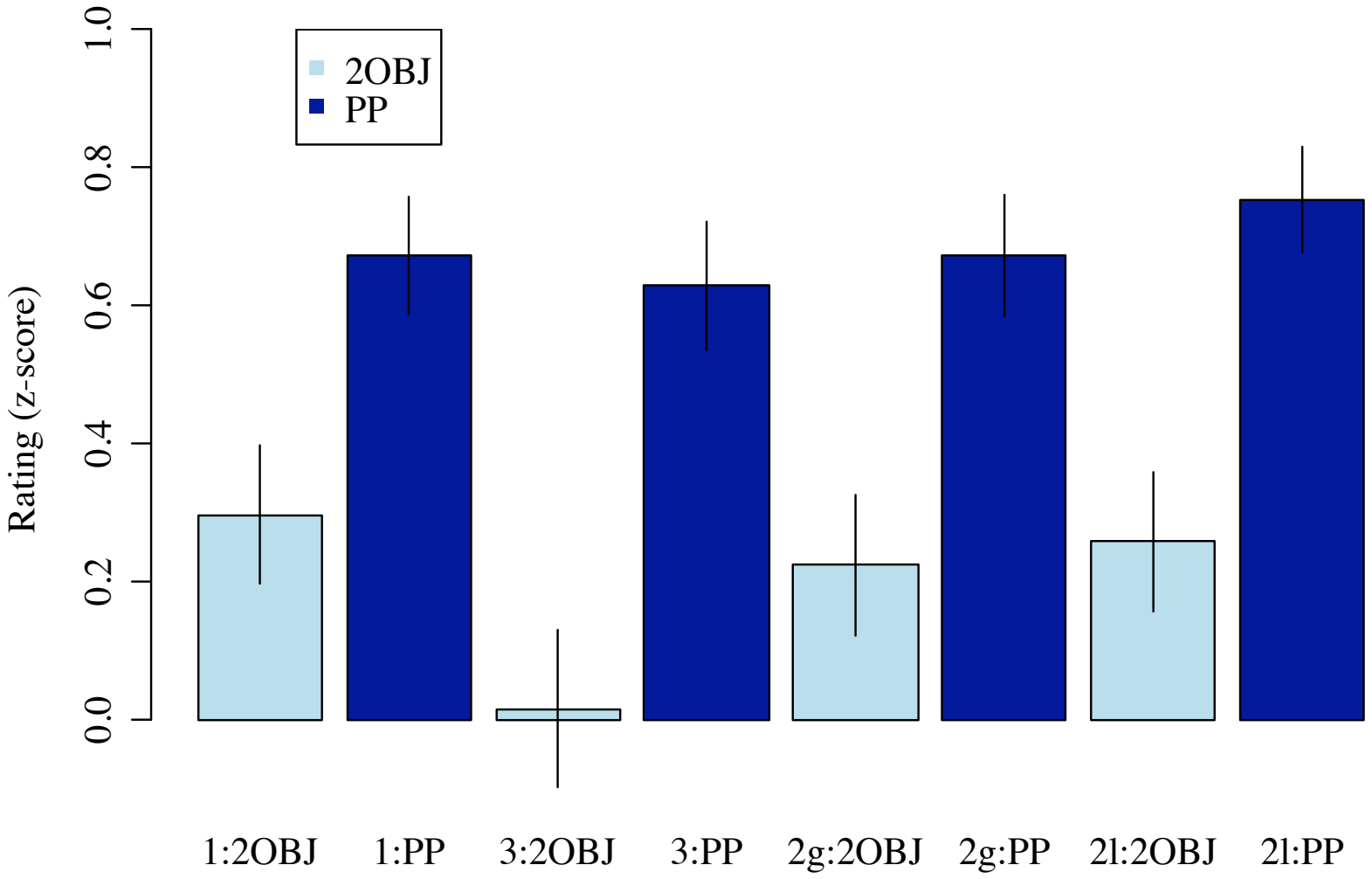
Predictions

- Gropen et al. (1989) should replicate; $1 > 3$
- Strict prosodic weight hypothesis: $1 > 3$, 2_G, 2_L
- Morphological complexity hypothesis: $1 > 3$, 2_G, 2_L
- Two-lexicon hypothesis: 1, $2_G > 2_L$, 3

2OBJ acceptability ~ word shape (replication)



2OBJ acceptability ~ word shape



Prodsodic weight hypothesis?

- Not supported
  - $1 > 3$ in 2OBJ; but $1 \not> 2$G, 2L in 2OBJ
  - Marginal effect of mono- vs. multi-syllable

Morphological complexity hypothesis?

- On surface, predictions are not met; 2G and 2L words were designed as morphologically complex, but...
- Maybe my 2-syllable nonce words aren't really as morphologically complex as the trisyllabic ones

Morphological complexity questionnaire (Hay 2003)

expute

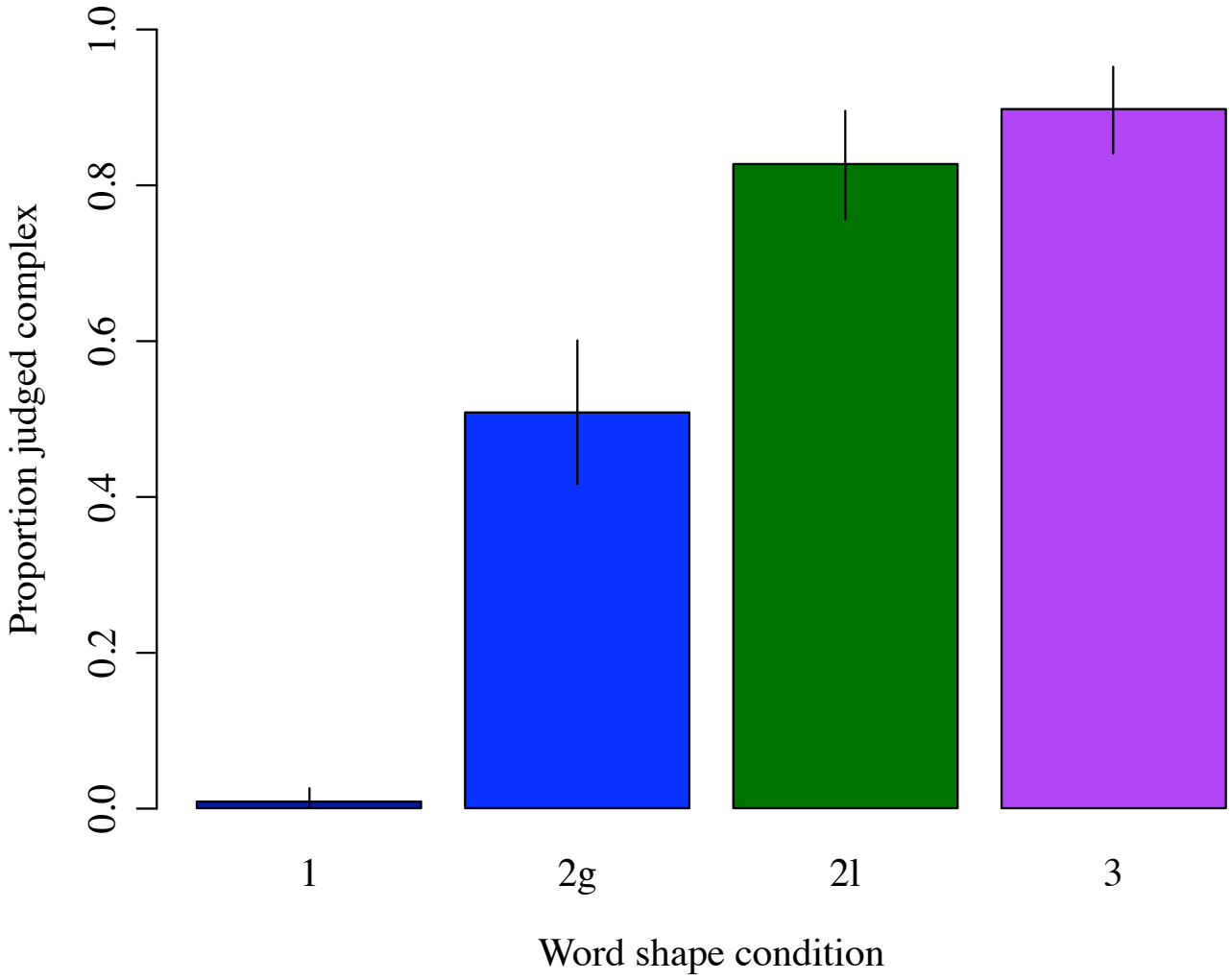
- simple
- complex

How sure are you?

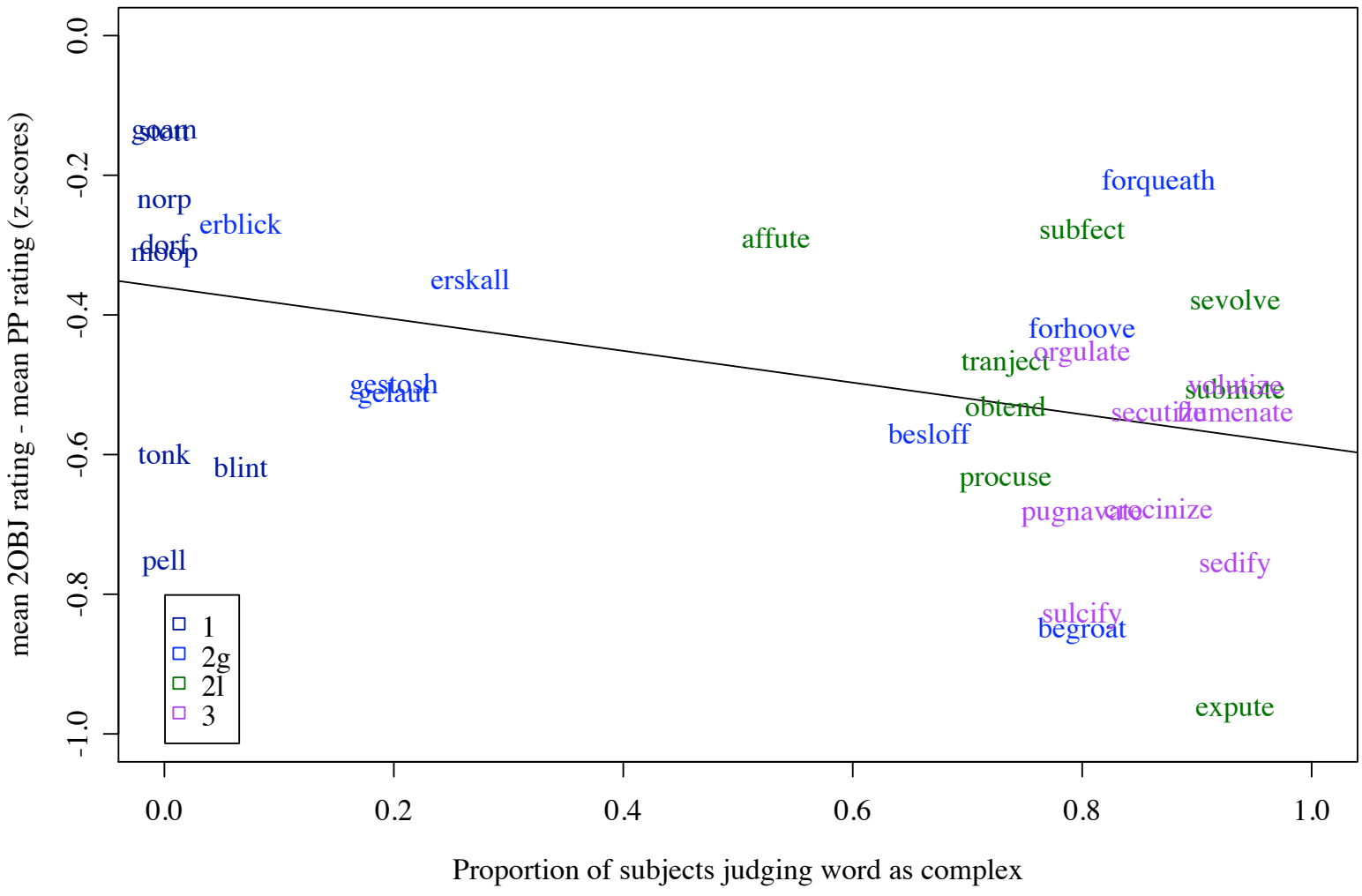
- 1 (very unsure)
- 2
- 3
- 4 (very sure)

Next

Complexity rating ~ Word shape



2OBJ acceptability ~ Perceived complexity



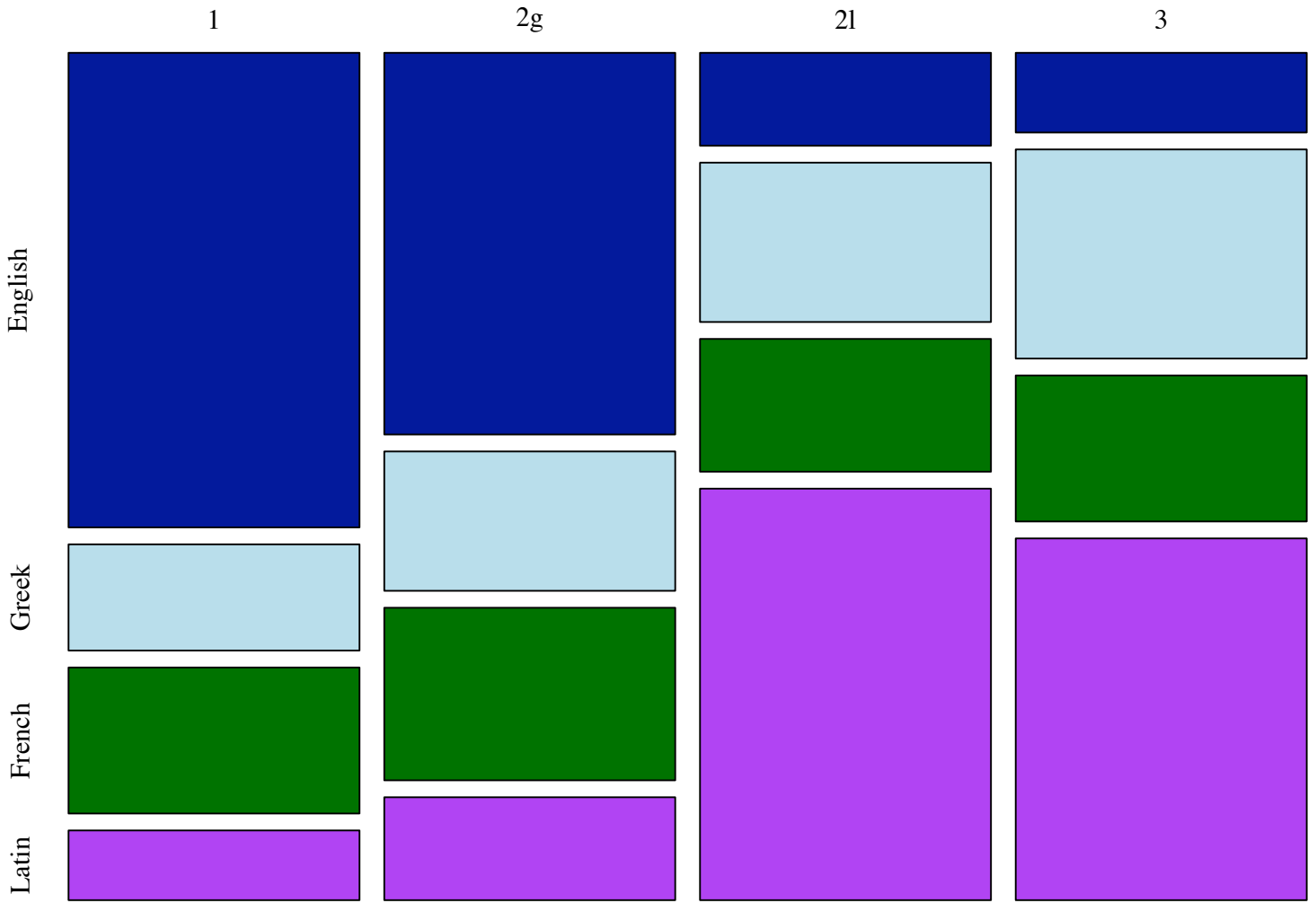
But could it just be number of syllables?

- Morphological complexity and number of syllables are highly correlated.
- Restricting the data to just 2-syllable words, morphological complexity is no longer a significant factor.

Two-lexicon hypothesis?

- On surface, predictions are not met; no difference between 2G and 2L
- But again, maybe what matters is how the words are *perceived*.

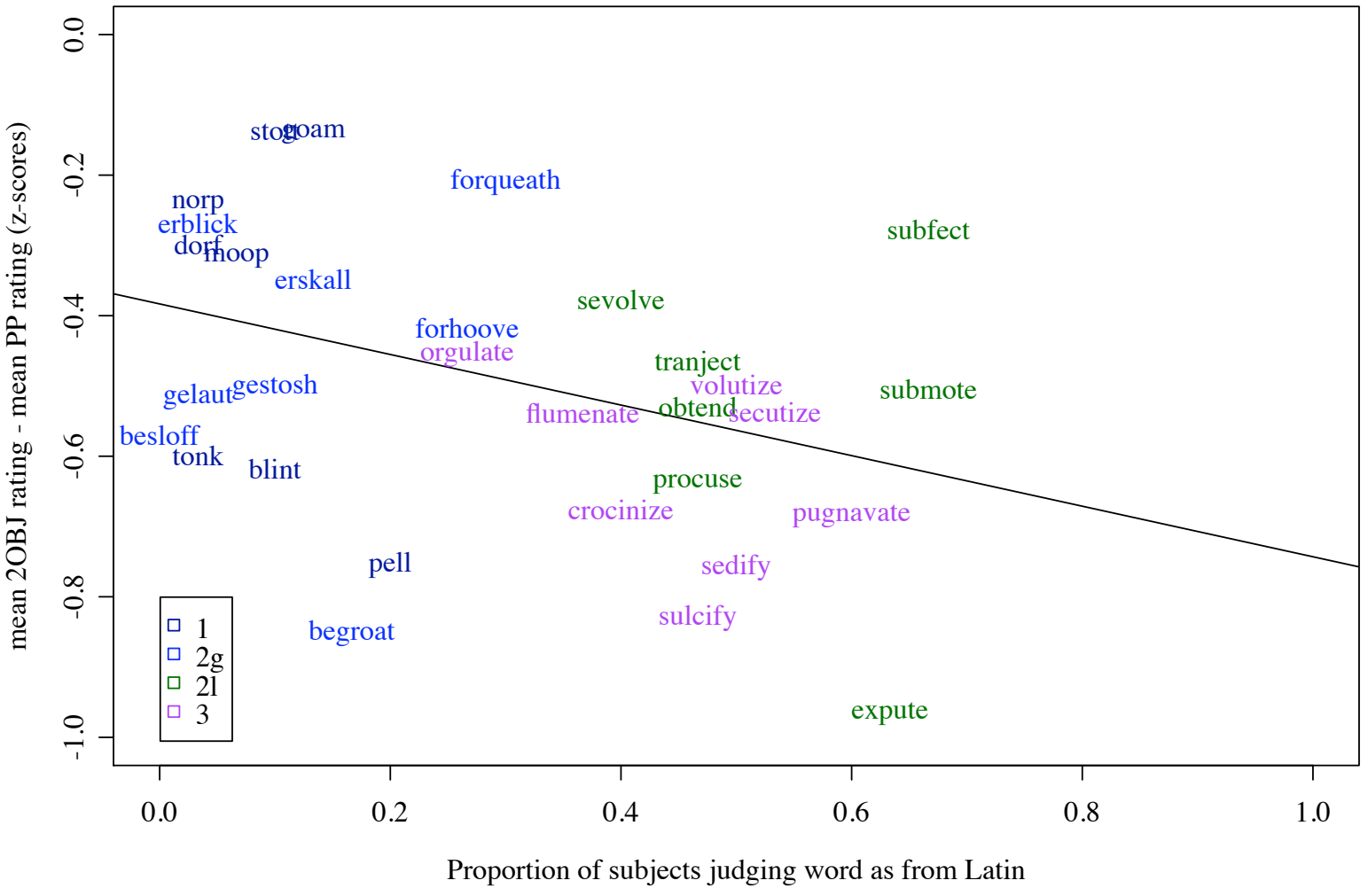
Etymology responses by word shape condition



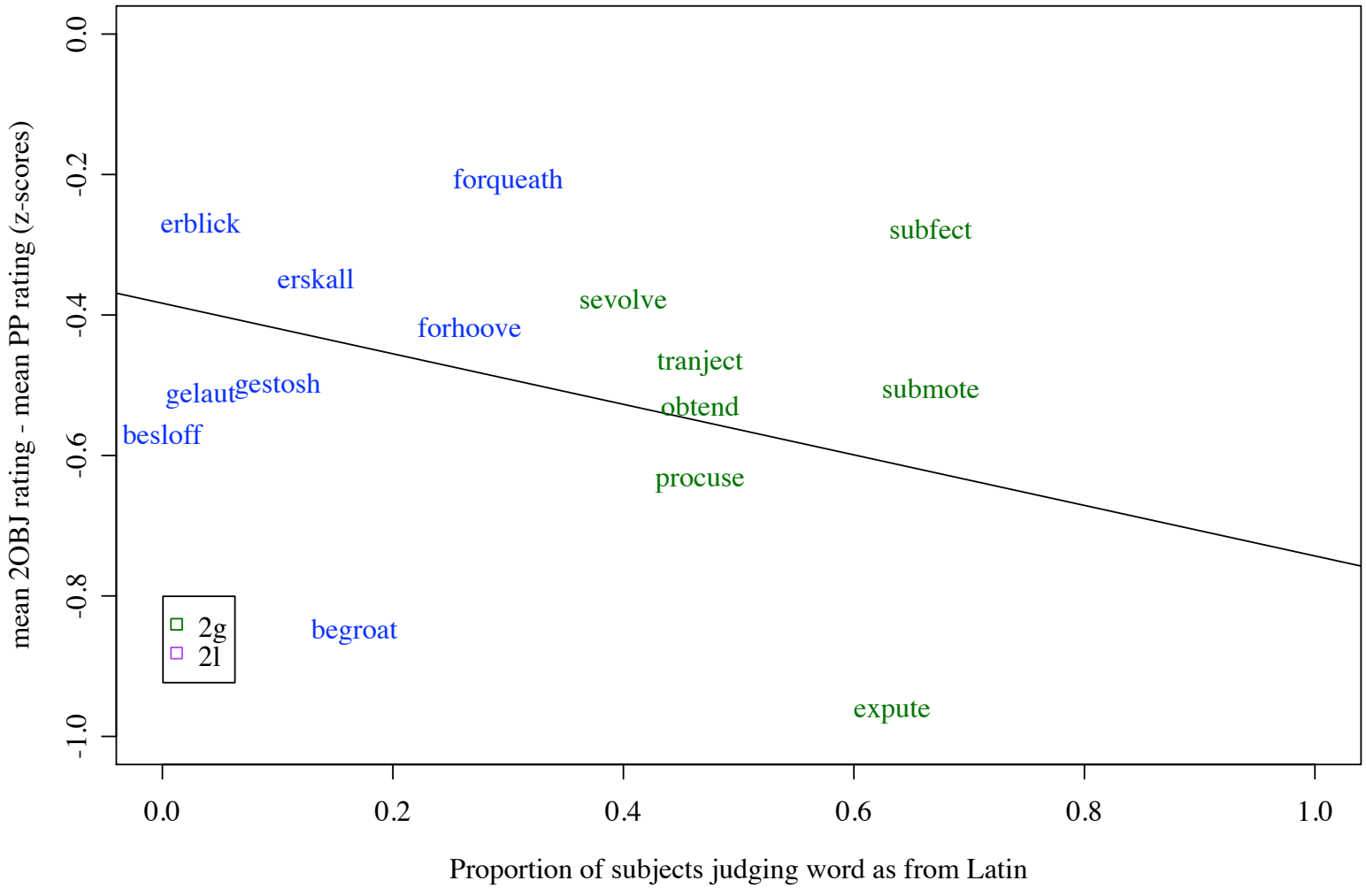
Perceived etymology: recoding

- Subjects' responses on etymology section recoded as: *Latin* vs. *Other*

2OBJ rating ~ perceived Latinateness



Among only 2-syllable words, Latinateness survives



Experiment 3: Summary

- Gropen et al. (1989) replicated
- Strict prosodic weight hypothesis not supported
- Morphological complexity hypothesis somewhat supported but further study is needed
- Two-lexicon hypothesis *was supported*
  - Perceived rather than actual Latinateness matters

Back to Baker's paradox

- We now have evidence that speakers are sensitive to perceived Latinateness, controlling for number of syllables.
- Since this generalization applies to words like *donate*, it is not necessary to store a particular constraint against any individual word.

What type of constraint is "Latinateness"?

1. A grammatical constraint, a criterion governing the productivity of the ditransitive construction
2. A stylistic constraint, perhaps related to formality
   - Support: *to*-datives more frequent in written corpora (Bresnan et al. 2007)
   - Also: Gropen et al.'s (1989) morphophonological effect was weaker than the effect of using non-possessive semantics
   - Implication: words like *donate* are not really "exceptions" in the grammatical sense

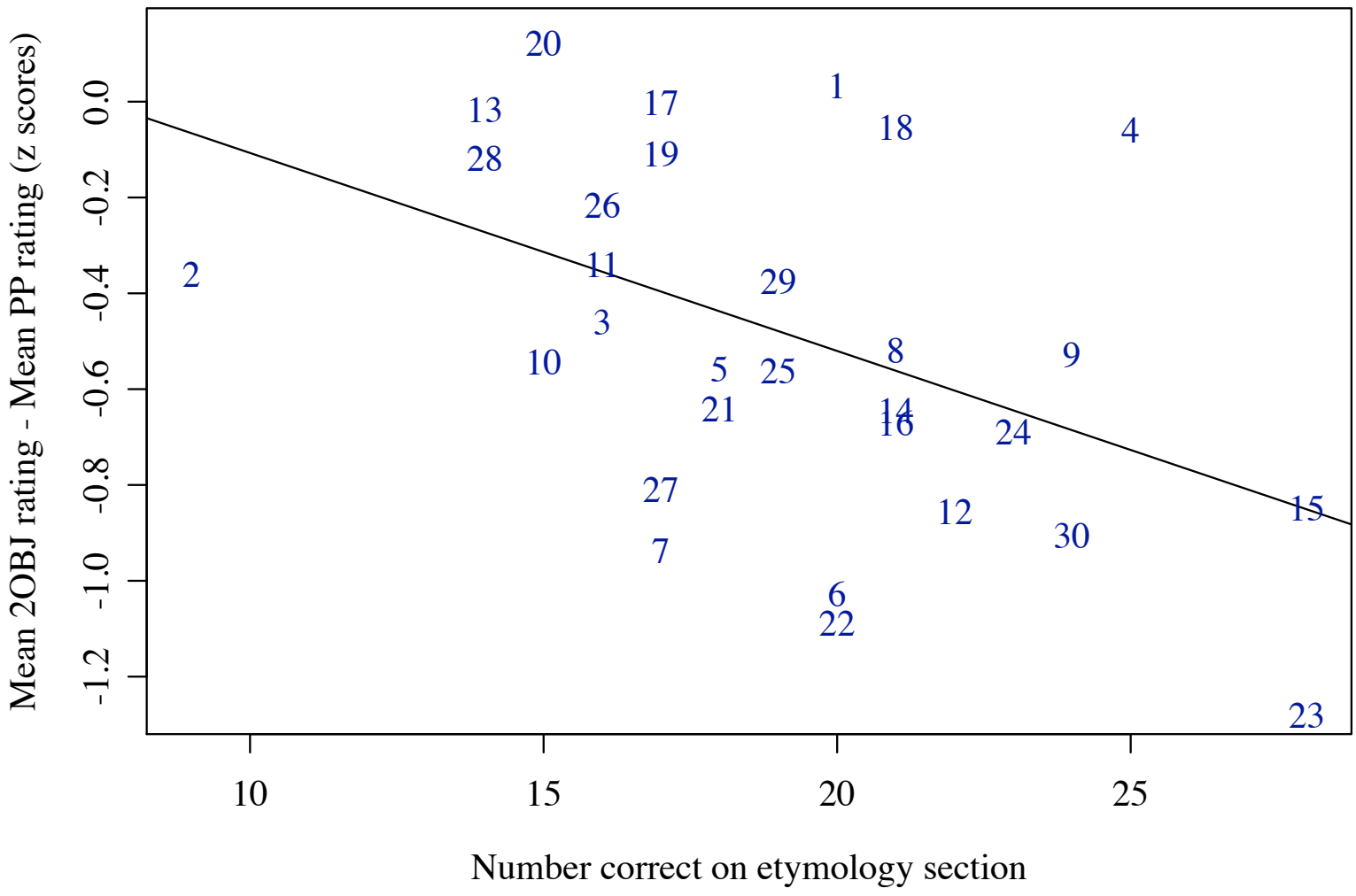
Larger claim about lexical storage

- There are certain lexical idiosyncrasies:
  - e.g., *I would rather that ...*
  - But this is not arbitrary negative exception to a productive pattern; this is an unexpected positive ability
- There are even some arbitrary negative exceptions to productive patterns
  - **goed* (preempted by *went*)
  - **to farmer* 'to be a farmer' (preempted by *farm*; Clark 1987)
- But in the absence of a preempting form, there is no need to memorize arbitrary exceptions to productive patterns.

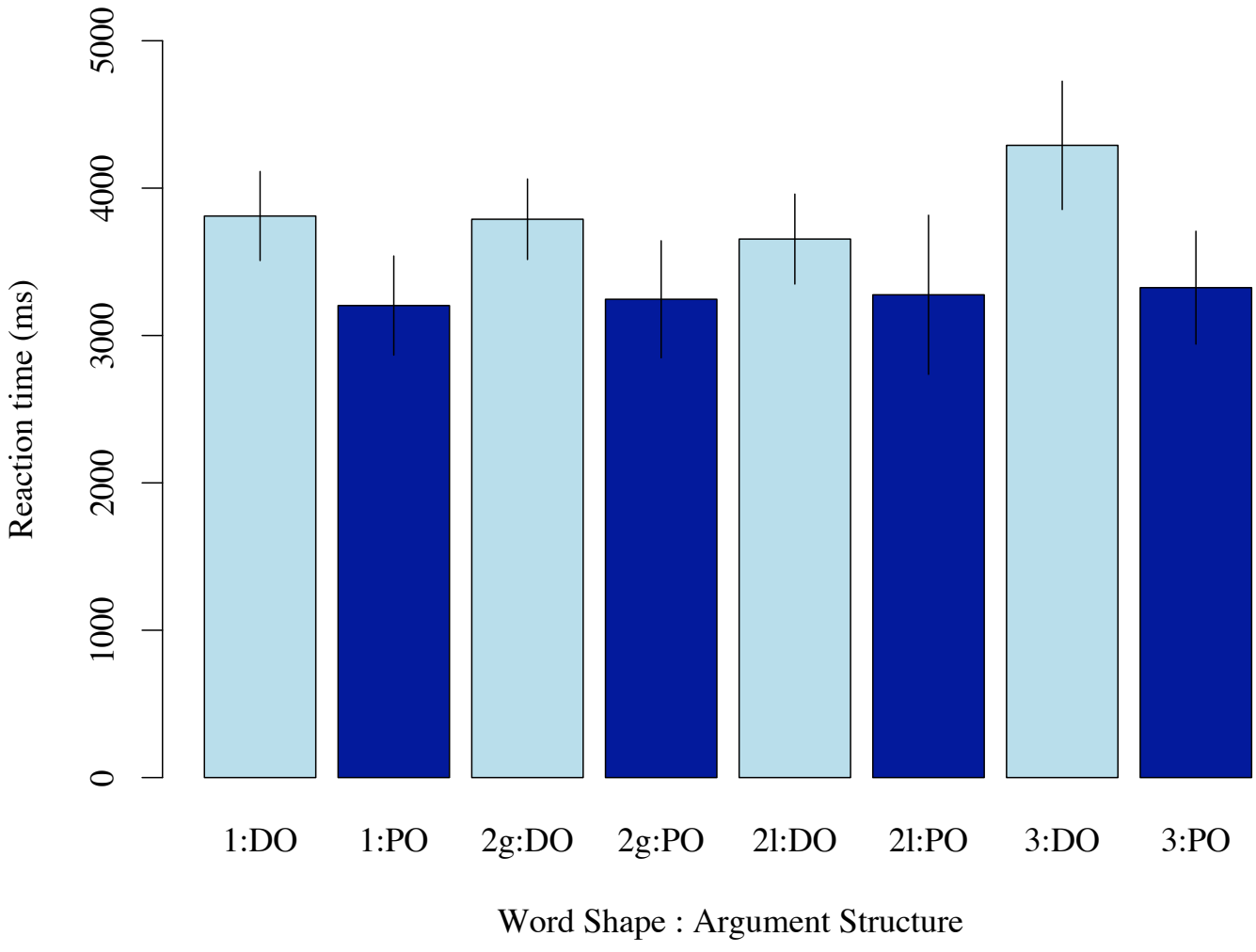
Thank you!

And thanks to Ivan Sag, Beth Levin, Dan Jurafsky, Eve Clark, Arnold Zwicky, Meghan Sumner, Charles Fillmore, *inter alii*.

Appendix A: 2OBJ rating ~ Etymology score



Appendix B: reaction times in Experiment 3



Anttila, A. (2007). Phonological constraints on constituent ordering. *Proceedings of WCCFL*.

Baayen, H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. To appear in *Journal of Memory and Language, special issue on Emerging Data Analysis Techniques*.

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry*, 10:533–581.

Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In Hawkins, J. A., editor, *Explaining language universals*, pages 73–101. Basil Blackwell, Oxford.

Bowerman, M. and Croft, W. (2008). The English lexical causative alternation. In Bowerman, M. and Brown, P., editors, *Crosslinguistic perspectives on argument structure: implications for learnability*. Lawrence Erlbaum Associates, Mahwah, New Jersey.

Braine, M. D. S. (1971). On two types of models of the internalization of grammars. In Slobin, D. I., editor, *The ontogenesis of grammar*, pages 153–186. Academic Press, New York.

Braine, M. D. S. and Brooks, P. J. (1995). Verb argument structure and the problem of an overgeneral grammar. In Tomasello, M. and Merriman, W. E., editors, *Beyond names for things: young children's acquisition of verbs*, pages 353–376. Lawrence Erlbaum, Hillsdale, NJ.

Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R. H. (2007). Predicting the dative alternation. In *Proceedings of the Royal Netherlands Academy of Science Workshop on the Foundations of Interpration*.

Brooks, P. J., Tomasello, M., Dodson, K., and Lewis, L. B. (1999). Children's overgeneralizations with fixed transitivity verbs. *Child Development*, 70(6):1325–1337.

Brown, R. and Hanlon, C. (1970). Derivational complexity and the order of acquisition in child speech. In Hayes, J. R., editor, *Cognition and the development of language*, pages 11–53. Wiley, New York.

Chouinard, M. M. and Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30:637–669.

Clark, E. V. (1987). The principle of contrast. In MacWhinney, B., editor, *Mechanisms of Language Acquisition*, pages 1–33. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ.

Culicover, P. W. (1999). *Syntactic Nuts: Hard cases, Syntactic theory, and Language acquisition*. Oxford University Press, Oxford.

Goldberg, A. (2006). *Constructions at Work*. Oxford University Press.

Green, G. M. (1974). *Semantics and syntactic regularity*. Indiana University Press, Bloomington.

Grimshaw, J. (2005). *Words and Structure*, chapter 4. CSLI Publications, Stanford.

Grimshaw, J. and Prince, A. (1986). A prosodic account of the *to*-dative alternation. Ms., Rutgers University.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., and Wilson, R. (1989). The learnability and acquisition of the dative alternation in English. *Language*, 65(2):203–257.

Harley, H. (2006). The bipartite structure of verbs cross-linguistically (or Why Mary can't exhibit John her paintings). Ms., University of Arizona.

Hay, J. (2003). *Causes and Consequences of Word Structure*. Routledge, New York.

Hirsch-Pasek, K., Treiman, R. A., and Schneiderman, M. (1984). Brown and Hanlon revisited: mothers' sensitivity to ungrammatical forms. *Journal of Child Language*, 11:81–88.

Krifka, M. (1999). Manner in dative alternation. In *WCCFL 19*, pages 260–271.

Pesetsky, D. (1995). *Zero Syntax: Experiencers and Cascades*. MIT Press, Cambridge, MA.

Pinker, S. (1984). *Language learnability and language development*. Harvard University Press, Cambridge, MA.

Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. MIT Press, Cambridge, MA.

Schütze, H. (1997). *Ambiguity Resolution in Language Learning*. CSLI Publications.

Storm, P. (1977). Predicting the applicability of dative movement. In *The book of squibs*. Chicago Linguistics Society, University of Chicago, Chicago.