

WORD VS. RULE FREQUENCIES IN IRREGULAR ACQUISITION

KYLE GORMAN

DAVID FABER

ELIKA BERGELSON

CHARLES YANG

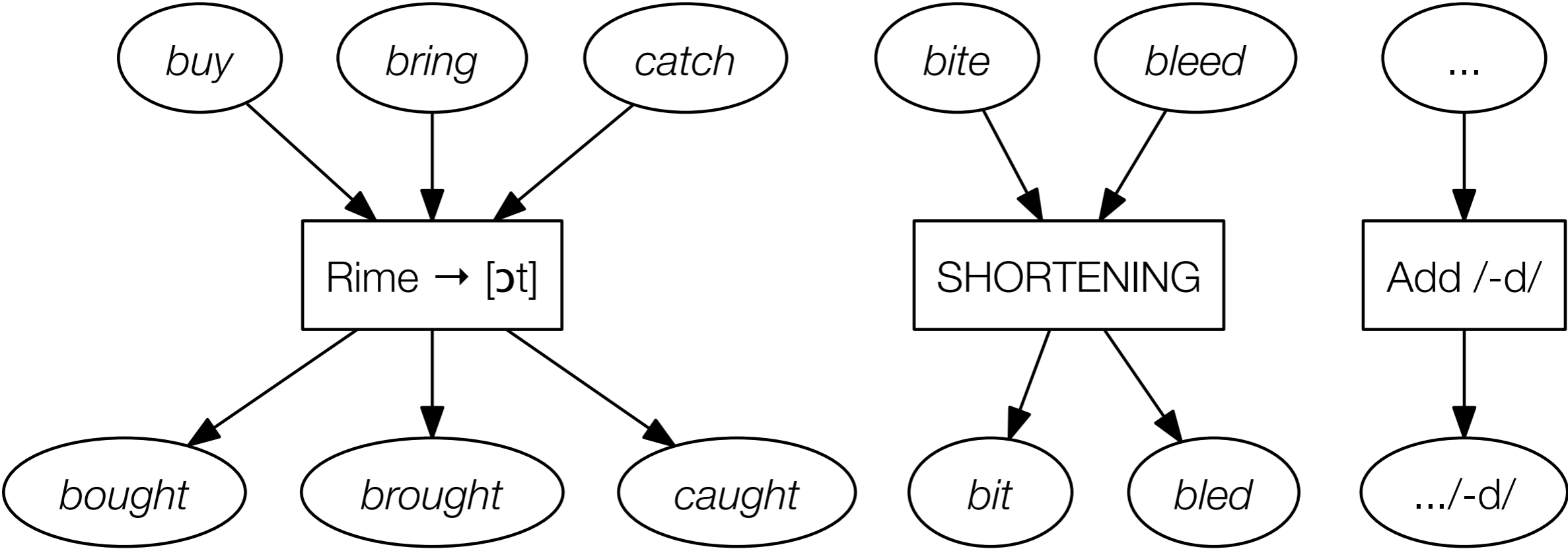
UNIVERSITY OF PENNSYLVANIA

OUTLINE

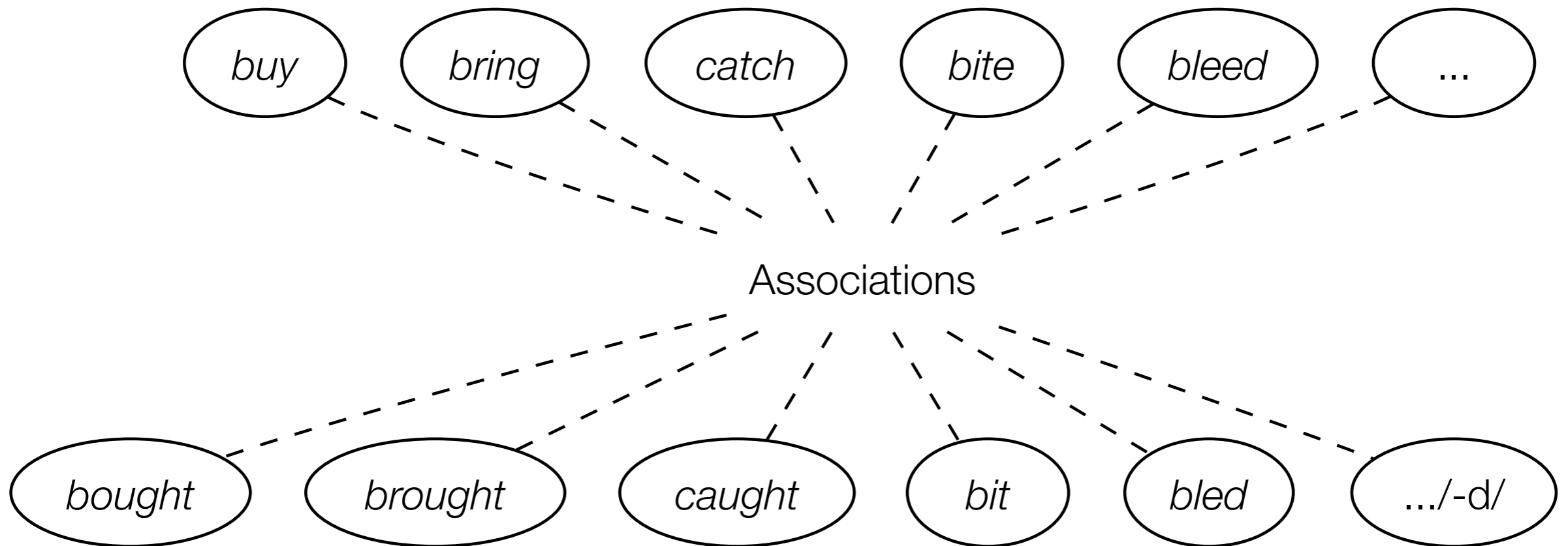
- The state of the past tense debate
- Evidence for morphological rules from overregularization errors
- Evidence against analogy from the absence of overirregularization errors

THE PAST TENSE DEBATE

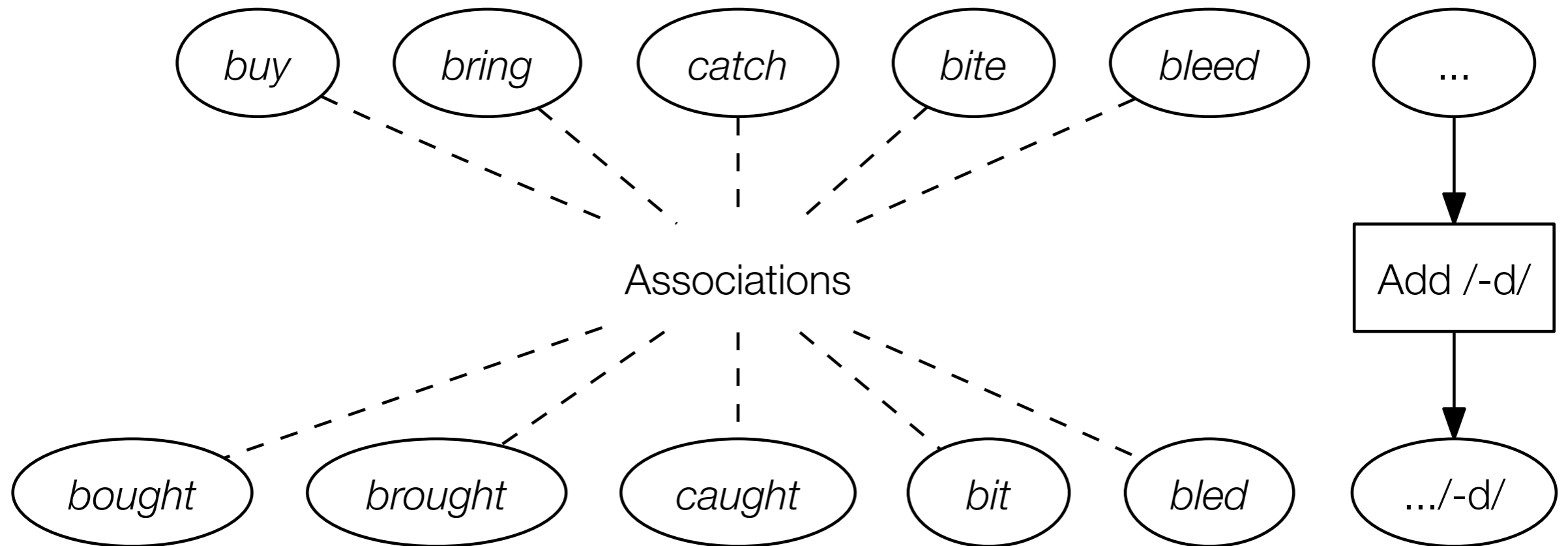
RULES



ANALOGY



RULES & ANALOGY



STORAGE OF IRREGULARS

- “A child may be producing irregular past tense forms as...memorized units rather than as the product of derivational rules.” (Kuczaj 1976: 599, fn. 3)
- “Regular forms are generated by rule, irregular forms are memorized by rote.” (Pinker 1999: 83)

RIME → [ɔt]

buy, bring, catch, seek, teach, think

- This kind of structure is recognized by:
 - Structuralists (e.g., Bloch 1947)
 - Generativists (e.g., SPE)
 - Connectionists (e.g., McClelland & Patterson 2002)

PRODUCTION ERRORS

“My teacher **holded** the baby rabbits...”

“Horton **heared** a Who.”

“I **buyed** a fire dog for a grillion dollars.”

“...the alligator **goed** kerplunk.”

FREQUENCY

- Word frequency: frequencies of past tense form in adult speech to children (Marcus et al. 1992)
- Rule frequency: sum of frequencies of all words following the rule (Yang 2002)

THE FREE- RIDER EFFECT

NO CHANGE

*cut, beat, **hit**, **hurt**, **put**, bet, cost, fit,
let, set, shut, split, spread*

- Massive input frequency variation
- Children no more likely to produce **hitted* for past *hit* than **splitted* for past *split*

THE MEGASTUDY

Study	# children	# tokens
Brown 1973	3	2,626
Kuczaj 1976	1	2,249
Hall et al. 1984	39	3,183
Maslen et al. 2004	1	1,485
Demuth et al. 2006	6	2,192
TOTAL	50	11,735

CODING

VVD*
|
“I eated it”
|
***EATED**

Child	ρ
Adam	0.909
Eve	1.000
Sarah	0.951

*Brill (1995) tagger: <http://gposttl.sourceforge.net/>

RULE COUNTS

No change	13	Rime → [ɔt]	5
V → [oʊ]	10	V → [eɪ]	5
V → [ʌ]	9	V → [ɑ]	3
SHORTENING	8	V → [ɔ]	3
V → [æ]	8	V → [oʊ], [-d]	2
SHORTENING, [-t]	7	V → [ɛ]	2
V → [u:]	7	V → [aʊ]	2
[-t]	5	(Singleton rules)	8

SHORTENING

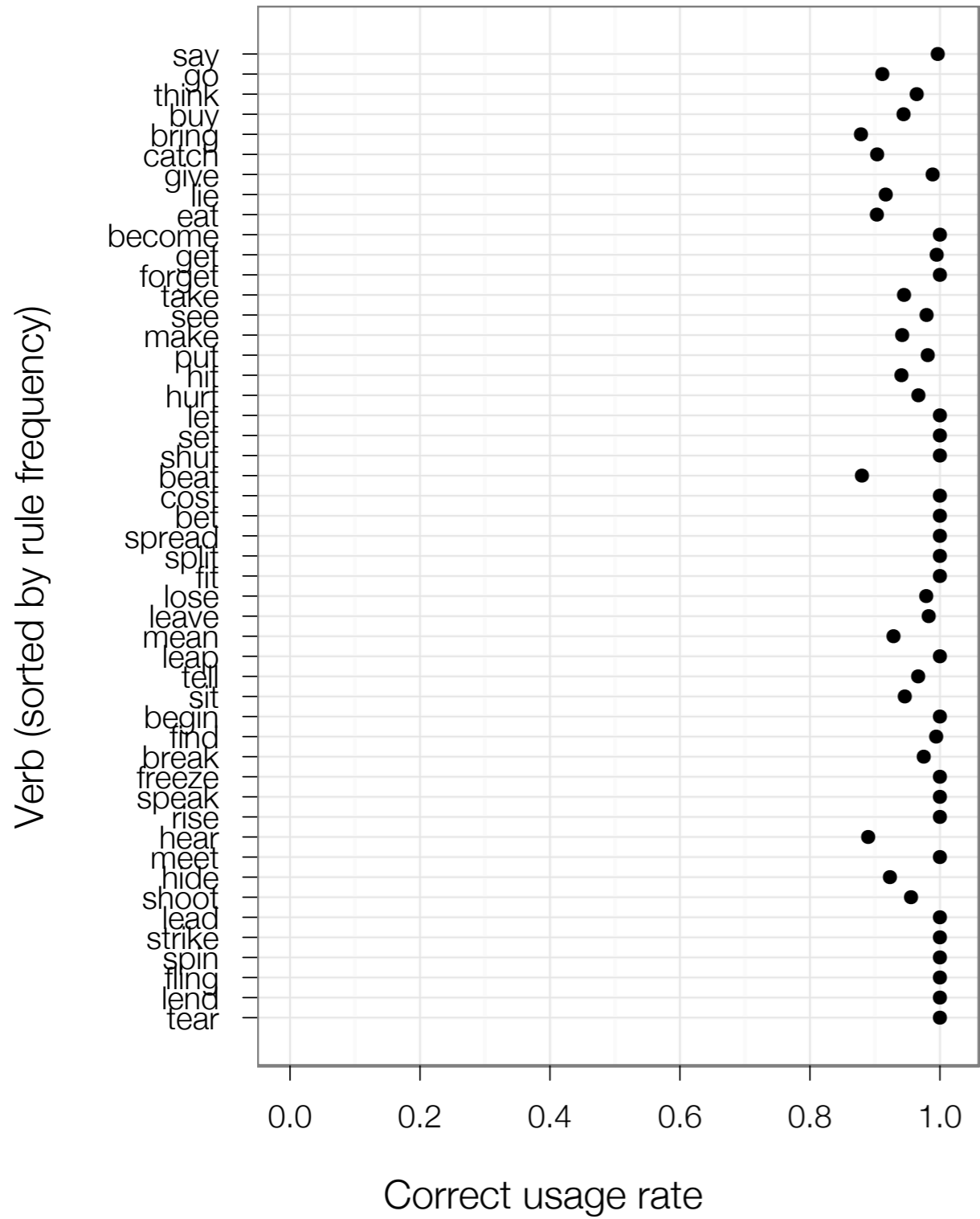
- A classic rule (*SPE*, Jaeger 1984, Halle & Mohanan 1985, Myers 1987, Wang & Derwing 1986, 1994, Halle 1998):

[aɪ]	[ɪ]	[i:]	[ɛ]
<i>divine</i>	<i>divinity</i>	<i>serene</i>	<i>serenity</i>
<i>satire</i>	<i>satiric</i>	<i>athlete</i>	<i>athletic</i>
<i>bite</i>	<i>bit</i>	<i>bleed</i>	<i>bled</i>

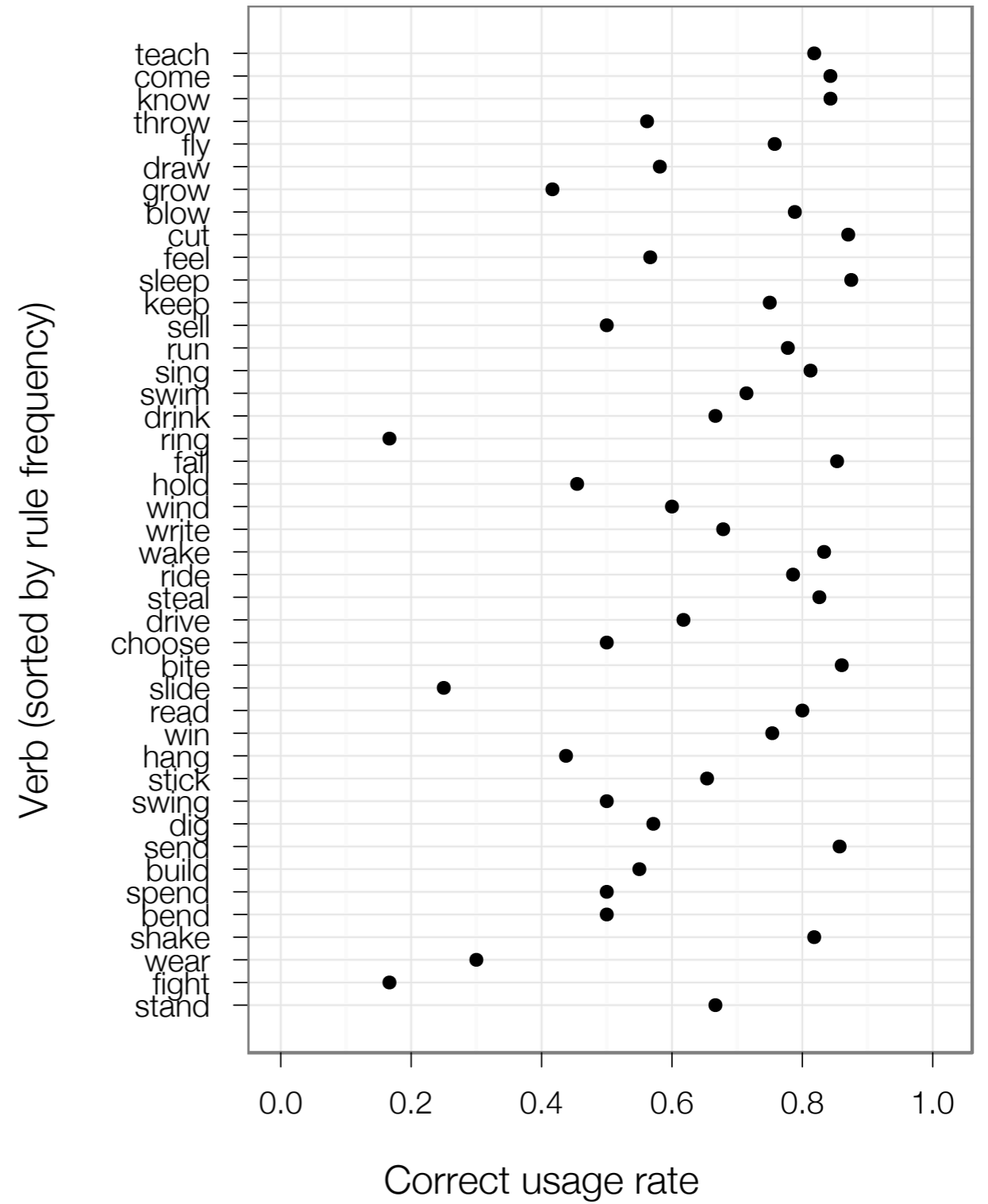
NONPARAMETRICS

	ρ	τ_b	γ
Word frequency	0.128	0.133	0.140
Rule frequency (no SHORTENING)	0.267	0.186	0.190
Rule frequency (w/ SHORTENING)	0.276	0.191	0.202

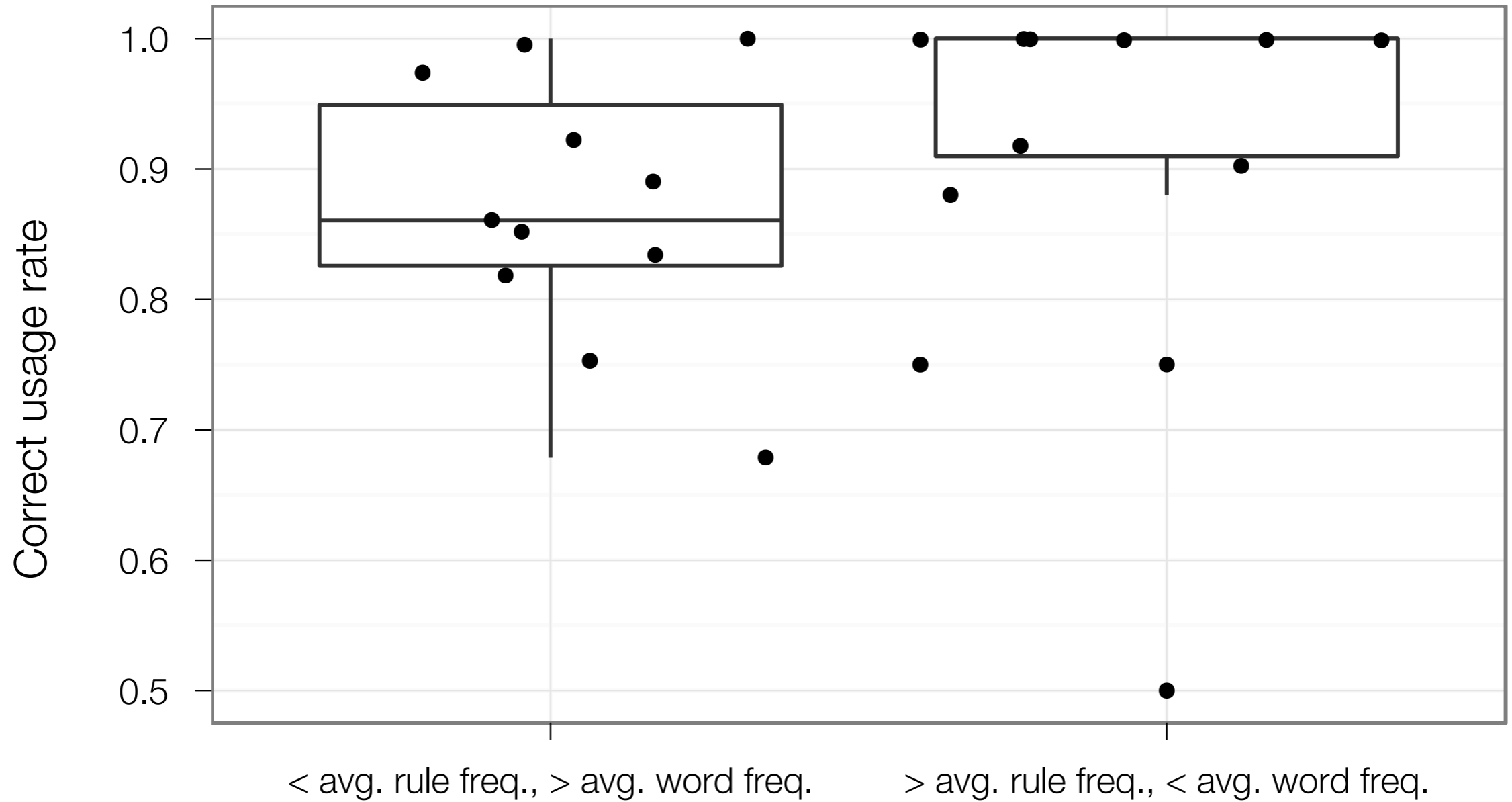
CORRECT USAGE RATE PER VERB



CORRECT USAGE RATE PER VERB



FREQUENCY DISSOCIATION



Two-tailed Mann-Whitney $W = 193$, $p = 0.039$

LOGISTIC REGRESSION

- Differences between subjects and words require a full random effects structure
- Word & rule frequency multicollinearity ($R = 0.639$) mandates residualization*

*For residualization technique, see Gorman 2010

LOGISTIC REGRESSION

	beta	std. err.	<i>p</i> -value
(Intercept)	2.563	0.333	1.5E-14
Rule frequency	2.501	0.514	1.5E-06
Residual word frequency	2.134	0.748	0.004

ANALYSIS

- Effect of rule frequency indexes strength of hypothesized rule
- Effect of word frequency indexes strength of word-rule mapping

(IR)REGULARS IN ADULT LANGUAGE

- Emmorey 1989, Allen & Badecker 2002, Stockall & Marantz 2006 (*pace* Stanners et al. 1979, Marslen-Wilson et al. 1993)
- Lignos & Gorman submitted (*pace* Alegre & Gordon 1999, Baayen et al. 2007, etc.)

ANALOGY?

ERROR RATES

- Overregularization is common (7.1% here)
- Overirregularization is very rare (0.23% in Xu & Pinker 1995)

XU & PINKER 1995

- Dialectical: *bring/brang*
- (td)-deletion (Guy & Boyd 1990): *sleep/slep*
- Wrong structural description: *bite/bat, bite/bet, beat/bait, fight/food, fit/feet, lift/left, trick/truck, say/set, sit/sought*
- Double marking: *crush/crooshed, jump/janged, bring/brunged*
- Overextension of *-en*: *dranken, aten, sawn, shotten, shooten, closen, trippen*
- **Overirregularization:** *fling/flang*

BERKO 1958

bing, bod, gling, mot, rick, spow

- Past tense forms of 6 nonce verbs elicited from 6-year-olds and adults
- Only 1 subject out of 86 produced *bing/bang, gling/glang*

WUG TEST AS TURING TEST

- Computational models of inflection overpredict irregular *wug* responses:
 - Minimum generalization learner (Albright & Hayes 2003)
 - TiMBL (Daelemans et al. 2009)
 - Fragment grammars (O'Donnell 2011)

PRODUCTIVITY

- Irregular rules are lexically conditioned:
 - No overirregularization of similar regulars
- Regular rules are Elsewhere Rules
(over)applying in the absence of positive evidence or in the case of memory failure:
 - They must first reach critical mass
(Marchman & Bates 1994, Yang 2005)

THANKS

- The CHILDES contributors
- Gary Marcus for the Brown data

**OUR DATA WILL BE
AVAILABLE ON
CHILDES SHORTLY**

BIBLIOGRAPHY

- A. Albright & B. Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90(2): 119-161.
- M. Alegre & P. Gordon. 1999. Is there a dual system for regular inflections? *Brain and Cognition* 68(1-2): 212-217.
- M. Allen & W. Badecker. 2002. Inflectional regularity: Probing the nature of lexical representation in a cross-modal priming task. *Journal of Memory and Language* 46(4): 705-722.
- H. Baayen, L. Wurm, J. Aycocock. 2007. Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon* 2(3): 419-463.
- J. Berko. The child's learning of English morphology. *Word* 14: 150-177.
- B. Bloch. 1947. English verb inflection. *Language* 23(4): 399-418.
- E. Brill. Transformation-based error-driven learning and natural language processing: A case study in part of speech tagging. *Computational Linguistics* 21(4): 543-565.

BIBLIOGRAPHY

- R. Brown. 1973. *A first language: The early stages*. Cambridge: Harvard University Press.
- N. Chomsky & M. Halle. 1968. *The sound pattern of English*. Cambridge: MIT Press.
- W. Daelemans, J. Zavrel, K. van der Sloot, & A. van den Bosch. 2009. TiMBL: Tilburg Memory-Based Learner. Technical report ILK 09-01, Induction of Linguistic Knowledge Research Group, Tilburg University.
- K. Demuth, J. Culbertson, & J. Alter. 2006. Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language & Speech* 49(2): 137-173.
- K. Emmorey. Auditory morphological priming in the lexicon. *Language & Cognitive Processes* 4(2): 73-92.
- K. Gorman. 2010. The consequences of multicollinearity among socioeconomic predictors of negative concord in Philadelphia. In M. Lerner (ed.), *Penn Working Papers in Linguistics* 16.2, 66-75.

BIBLIOGRAPHY

- G. Guy & S. Boyd. 1990. The development of a morphological class. *Language Variation & Change* 2(1): 1-18.
- W. Hall, W. Nagy, & R. Linn. 1984. *Spoken words: Effects of situation and social group on oral word usage and frequency*. Hillsdale, NJ: Erlbaum.
- M. Halle & K. Mohanan. 1985. Segmental phonology of Modern English. *Linguistic Inquiry* 16(1): 57-116.
- M. Halle. 1998. The stress of English words 1968-1998. *Linguistic Inquiry* 29(4): 539-568.
- J. Jaeger. 1984. Assessing the psychological status of the Vowel Shift Rule. *Journal of Psycholinguistic Research* 13(1): 13-36.
- S. Kuczaj. 1976. *-ing, -s, and -ed: A study of the acquisition of certain verb inflections*. Doctoral dissertation, University of Minnesota.
- S. Kuczaj. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* 16(5): 589-600.

BIBLIOGRAPHY

- V. Marchman & E. Bates. 1994. Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language* 21(2): 339-366.
- G. Marcus, S. Pinker, M. Ullman, M. Hollander, J. Rosen, & F. Xu. 1992. *Overregularization in language acquisition*. Chicago: University of Chicago Press.
- W. Marslen-Wilson, M. Hare, & L. Older. 1993. Inflectional morphology and phonological regularity in the English mental lexicon. In *Proceedings of the 15th Annual Meeting of the Cognitive Science Society*, 693-698.
- R. Maslen, A. Theakson, E. Lieven, & M. Tomasello. 2004. A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language and Hearing Research* 47(6): 1319-1333.
- J. McClelland & K. Patterson. 2002. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Science* 6(11): 465-472.

BIBLIOGRAPHY

- S. Myers. 1987. Vowel shortening in English. *Natural Language and Linguistic Theory* 5(4): 485-518.
- T. O'Donnell. 2011. Productivity and reuse in language. Doctoral dissertation, Harvard University.
- S. Pinker. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.
- R. Stanners, J. Neiser, W. Herson, & R. Hall. 1979. Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18(4): 399-412.
- L. Stockall & A. Marantz. 2006. A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon* 1(1): 85-123.
- F. Xu & S. Pinker. 1995. Weird past tense forms. *Journal of Child Language* 22(3): 531-556.

BIBLIOGRAPHY

- H. Wang and B. Derwing. 1986. More on English vowel shift: The back vowel question. *Phonology Yearbook* 3: 99-116.
- H. Wang and B. Derwing. 1994. Some vowel schemas in three English morphological classes: Experimental evidence. In M. Chen & O. Tzeng (ed.), *In honor of Professor William S-Y Wang: Interdisciplinary studies on language and language change*, 561-575. Taipei: Pyramid.
- C. Yang. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.
- C. Yang. 2005. On productivity. *Linguistic Variation Yearbook* 5: 333-370.