Phonotactic knowledge and Universal Learning

James Myers and Jane Tsay
National Chung Cheng University
Lngmyers@ccu.edu.tw Lngtsay@ccu.edu.tw

Workshop on Markedness – Hsinchu, Taiwan – November 2, 2012

Thanks

- Experimental participants and corpus speakers
- Assistants Guo-Ming Hong, Chiung-Wen Hsu, Zi-Ping Hsu, Yu-Leng Lin, Chia-Wen Lo, Chen-Tsung Yang
- National Science Council (Taiwan) grants NSC 97-2410-H-194-067-MY3, NSC98-2410-H-194-086-MY3, NSC100-2410-H-194-109-MY3
- Bits of this were presented at IASCL12 (Montreal, 2011) and SLE 45 (Stockholm, 2012)
- A formal analysis will be presented at TPC4 (Taipei, May, 2013)

Overview

- Universal Grammar vs. Universal Learning
  - UG: Product-oriented
  - UL: Process-oriented
- Predictions for phonotactic judgments
  - UG: Lexicon and universals compete
  - UL: Lexicon and universals cooperate
- Tests on Mandarin and Southern Min
  - They tend to support UL over UG (so far)

Universal Grammar

- UG is product-oriented:
  "It is a common observation that for all their diversity, **languages are made to a great extent of familiar pieces**, much like the wide variety of shapes and objects that can be assembled from a limited array of Lego blocks." (Nevins et al., 2009, p. 359)
- Adult grammar as consisting of bits of UG:
  - Principles & Parameters
  - Distinctive feature theory
  - Optimality Theory constraints

UG and learning

- UG presumes an approach to learning:
  "[Parameterized UG] is thought to dovetail with the results of research into **the interaction of universal and language-particular aspects of grammar** during the actual process of language acquisition."
  (Nevins et al., 2009, p. 359)
- E.g. learning in OT (Tesar & Smolensky, 2000):
  - Complex content: Thousands of innate constraints
  - Simple algorithm: Demote the violated constraints

Universal Learning

- But UL doesn't need UG:
  "By parity of argument, every feature of every language that has ever been spoken must then be part of the language faculty or UG. This seems no more plausible than claiming that, because we can learn to ride a bicycle or read music, these abilities are part of our innate endowment. Rather, **it is the ability to *learn*** bicycle riding by putting together other, more basic abilities **which has to be within our capacities**, not the trick itself."
  (Evans & Levinson, 2009, p. 443)

UL and learning

- UL is process-oriented:
  What's innate are **learning biases**, not bits of grammar
- UL traditionally outside of generative grammar
  - Operating principles (Slobin, 1973)
  - "The myth of language universals" (Evans & Levinson, 2009)
- But not anymore: e.g. Hayes & Wilson (2008)
  - Simple content: No innate constraints
  - Complex algorithm: Constraints built from experience, respecting innate locality biases

How UL explains

- Universals
  Shared learning biases lead to similar outcomes
- Poverty of the stimulus
  Learning biases are corpus analysis algorithms that are shared with the creators of the corpus
- An adult corpus: **A, B, A, B,** …
  Child wonders: What's the next letter?
- If adults are human: **A, B, A, B,** A, B, A, B, …
- If adults are weird: **A, B, A, B,** C, A, B, C, D, …

Arguing for UL

- Universals
  - Shared learning biases lead to *similar* outcomes…
  - …but not necessarily to *identical* outcomes
  - Thus **universals can be fuzzy** (e.g., Mielke, 2008)
- Poverty of the stimulus
  - The innately biased learner finds/creates patterns
  - It doesn't merely use corpus info to fill gaps in UG
  - Thus **"naturalness" should cooperate with lexical experience instead of competing with it**

A schematic phonotactic example

- Naturalness says: A = B > C = D
- But the lexicon says: A > B > C > D
- **UG competition:** If natural, then ignore lexicon

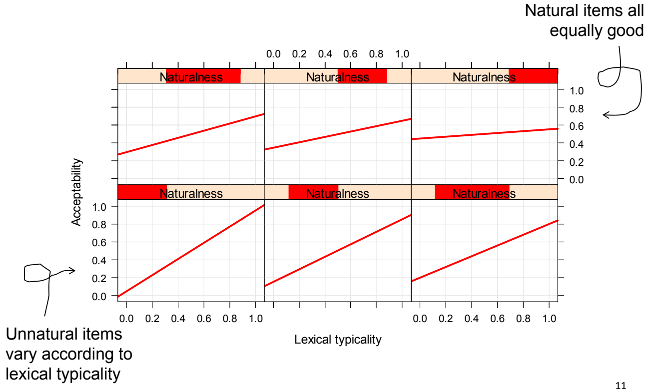
  C > D triggers learning: *D >> Faith >> *C

  So mature grammar says: **A = B > C > D**
- **UL cooperation:** If natural, then learn lexicon

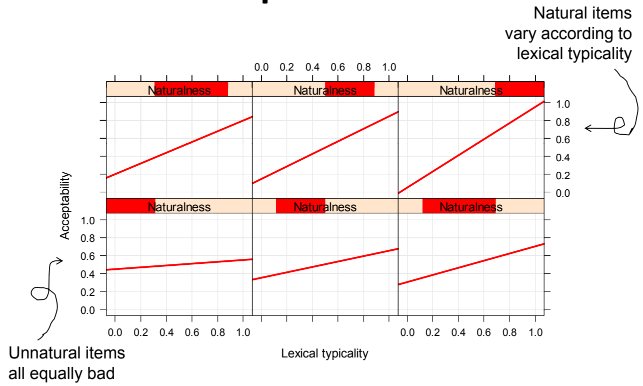
  A & B are *easy to learn from*, so A > B triggers learning

  So mature grammar says: **A > B > C = D**

UG prediction

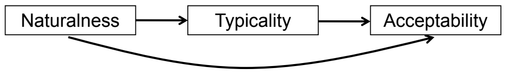


UL prediction

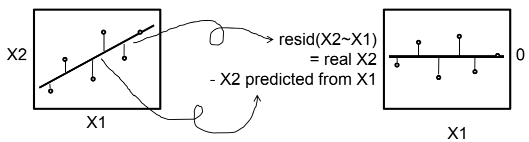


Avoiding statistical artifacts

- Naturalness & lexical typicality are correlated:



- Solution: Remove the correlation by replacing one independent variable with residuals:



Some Mandarin data

- Part of a **mega-experiment** (cf. Balota et al., 2012)

  All 3,274 non-lexical syllables that can be written in BPMF (Taiwan's phonetic orthography)

  Good/bad wordlikeness judgments of BMPF syllables

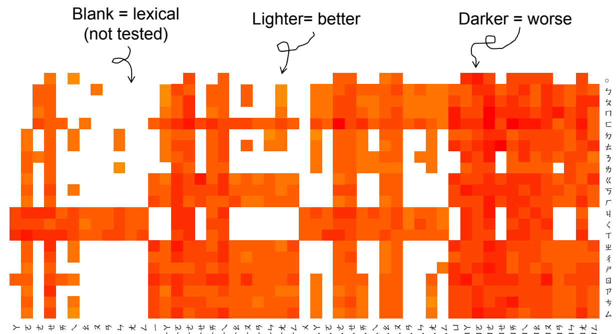
  16 speakers in this pilot (ca. 100 in full experiment)
- Some interesting initial results (details TBA)

  Syllable structure primes across trials, but not onsets

  Slower responses show stronger neighborhood effects

  "Naturalness" and lexical "typicality" interactions

Mandarin nonce syllable judgments



Test 1: Mandarin neighbors

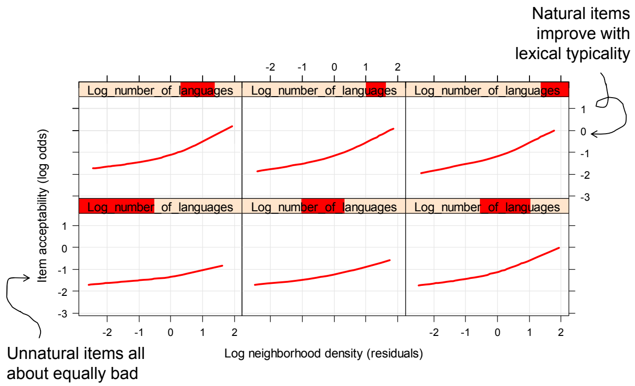
- **Lexical typicality:** Number of lexical neighbors

  One segment from target (Vitevitch & Luce, 1999)
- **Naturalness:** Number of languages with item's initial consonant (in UPSID; Maddieson, 1984)

  (This and all other analyses reported here ignore tone)

		UPID frequency of onset	
		Lower	Higher
Neighborhood density	Lower	t^hio^{35}	pio^{51}
	Higher	t^hiei^{55}	$piei^{55}$

Test 1: Results



Test 2: S. Min adult judgments

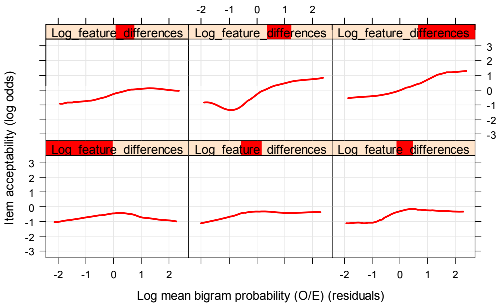
- 255 non-lexical syllables, one per each logically possible bigram of Southern Min phonemes
- 20 native speakers, binary auditory judgments
- **Lexical typicality:** Lexical bigram probability (observed / expected: Frisch & Zawaydeh, 2001)
- **Naturalness:** Bigram feature differences

  (more differences = easier to distinguish perceptually)

  (because task is auditory)

		Mean number of feature changes	
		Lower	Higher
Mean bigram probability (O/E ratio)	Lower	biem1	sot4
	Higher	guoŋ7	piok8

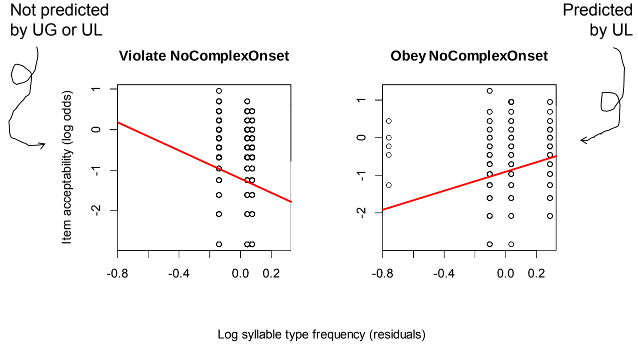
Test 2: Results



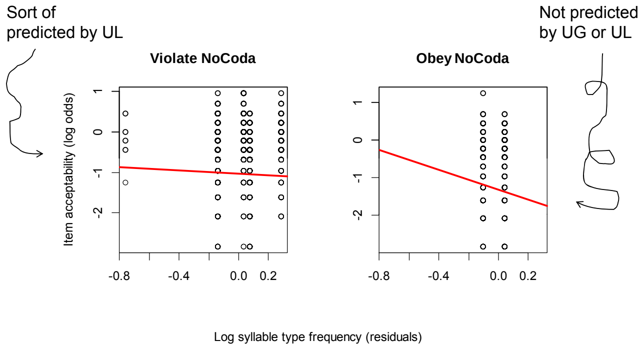
Test 3: Mandarin syllable constraints

- **Lexical typicality:** Syllable type frequency
  - E.g. CV, GV, CGV, CVC, GVC, CGVC, CVG, GVG, CGVG
  - Count number of syllables in syllabary per type
- **Naturalness:** NoComplexOnset, NoCoda
  - Haven't tested Onset yet (no onsetless syllables in this subset of the data)
- Lexical/natural relations positively correlated:
  - NoComplexOnset tends to be obeyed: **C**VX > **CG**VX
  - NoCoda also tends to be obeyed: XV_ > {XV**C**, XV**G**}

Test 3a: NoComplexOnset results



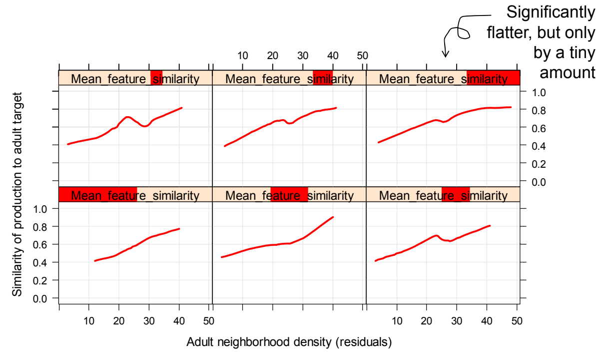
Test 3b: NoCoda results



Test 4: What about kids?

- Seven children from the longitudinal Taiwanese Child Language Corpus (Tsay, 2007)
  - 88,280 syllable tokens with at least one bigram
  - (For other findings, see Myers & Tsay, 2011)
- Production accuracy (similarity to adult targets)
- **Lexical typicality:** Neighborhood density
  - (based on the adult lexicon)
- **Naturalness:** Bigram feature similarities
  - (more similar = easier to produce)

Test 4: S. Min child production



Results summary

- Results that point clearly towards UL:
  - Mandarin adults: Neighbors x Onset typological freq
  - S. Min adults: Bigram frequency x Feature diffs
- Mixed results:
  - Mandarin adults: Syllable constraints (partly pro-UL)
  - S. Min kids: Neighbors x Feature sims (tiny pro-UG)
- Results that point clearly towards UG:
  - None (so far)

The future

- More tests are needed
  - Other measures of lexical typicality and naturalness
  - Other languages, tasks, populations (kids vs. adults)
- UL needs to be formalized
  - The interaction of lexical typicality and naturalness should depend on the details
    - E.g. neighborhood density and phonotactic probability involve different processes (Pylkkänen & Marantz, 2003)
- Logic beyond phonotactics: Test other "interactions of universal and language-particular aspects of grammar"

Phonotactic knowledge and Universal Learning

James Myers and Jane Tsay
National Chung Cheng University
Lngmyers@ccu.edu.tw Lngtsay@ccu.edu.tw

Workshop on Markedness – Hsinchu, Taiwan – November 2, 2012

http://upload.wikimedia.org/wikipedia/commons/2/20/Copper_Beech_Fagus_sylvatica_f._purpurea_Autumn_Leaves_Closeup_3008px.jpg

References (1/4)

Balota, D. A., Yap, M. J., Hutchison, K.A., & Cortese, M. J. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed). *Visual word recognition, Vol. 1*. Psychology Press.

Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429-492.

Frisch, S. A., & Zawaydeh, B. A. (2001). The psychological reality of OCP-Place in Arabic. *Language, 77*, 1:91-106.

References (2/4)

Hayes, B., & Wilson, C. (2008). A Maximum Entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry, 39* (3), 379-440.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press.

Mielke, J. (2008). *The emergence of distinctive features*. Oxford University Press.

Myers, J., & Tsay, J. (2011, July). Phonetic, phonotactic, and neighborhood effects on syllable production in child Southern Min. Poster presented at the 12th International Congress for the Study of Child Language, Montreal. <http://www.ccunix.ccu.edu.tw/~lngproc/MyersTsay_IASCL2011_HO.pdf>

References (3/4)

Nevins, A., Pesetsky, D., & Rodrigues, C. (2009). Pirahã exceptionality: A reassessment. *Language, 85* (2), 355-404.

Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Science, 7* (5), 187-189.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.) *Studies of child language development* (pp. 175-208). Hold, Rinehart, and Winston.

References (4/4)

Tsay, J. S. (2007). Construction and automatization of a Minnan child speech corpus with some research findings. *Computational Linguistics and Chinese Language Processing, 12* (4), 411-442.

Tesar, B., & Smolensky, P. (2000). *Learnability in Optimality Theory*. MIT Press.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*, 374-408.