

Grammar & Evidence

What are they anyway?

James Myers
National Chung Cheng University
International Workshop on Grammar & Evidence
2007.4.14

Overview

- Grammar
- Evidence

2

But first...

Why should we care?

3

Chomsky (2002:102): “[The development of modern science] had many parts, like the Galilean move towards discarding recalcitrant phenomena if you’re achieving insights by doing so, the post-Newtonian concern for intelligibility of theories rather than of the world, and so on. That’s all part of the methodology of science. It’s not anything that anyone teaches; there’s no course in methodology of physics at MIT. In fact, the only field that has methodology courses, to my knowledge, is psychology. If you take a psychology degree you study methodology courses, but if you take a physics degree or a chemistry degree you don’t do it. The methodology becomes part of your bones or something like that.”

4

Kuhn (1970:12-13): “Each of the corresponding schools derived strength from its relation to some particular metaphysic, and each emphasized, as paradigmatic observations, the particular cluster of ... phenomena that its own theory could do most to explain. Other observations were dealt with by ad hoc elaborations, or they remained as outstanding problems for further research.... Being able to take no common body of belief for granted, each writer ... felt forced to build his field anew from its foundations. In doing so, his choice of supporting observation and experiment was relatively free, for there was no standard set of methods or of phenomena that every ... writer felt forced to employ and explain.”

Funny syntax

Gestern traf sie mich fast.
(yesterday met she me almost)
“Yesterday she almost met me.”

- Meinunger (2001): Grammatical, so....
- Rapp & von Stechow (1999): Not grammatical, so....
- Who’s right? How many Germans do we need to ask...?

6

Funny semantics

Ni xiang zhidao shei mai-le shenme?
(you want know who buy-aspect what)

- Huang (1982) claims that *shenme* (“what”) can have wide scope:
“What was bought, such that you wonder who bought it?”
- Xu (1990) denies this, citing a survey of multiple speakers.

7

Funny phonology

- Leben (1978): Mende tone consistently spreads left to right:

[fɛ̀lámà] “junction” [ndávúlá] “sling”
HLL LHH

- Zoll (2003): Mende tone spread depends on tone quality (based on Leben’s data):

[fɛ̀lámà] “junction” [lèlémá] “mantis”
HLL LLH

8

Funny phonetics

- Halle and Mohanan (1985) devote a stratum in their model to the following:

palatalized /l/: a whale edition
the seal office

velarized /l/: the whale and the shark
the seal offered a doughnut

- Experiments by Sproat and Fujimura (1993) and Hayes (2000) cast doubt.

9

Funny morphology

- Di Sciullo and Williams (1987): Affixes are transparent to semantic role assignment, but roots are not:

a baker of bread
*a bake-man of bread

- Spencer (1991): No, it’s just *verb+*man*:

a baker
*a bake-man

10

So why should we care about methodology?

- Get the study of grammar out of its pre-scientific state
 - Funny data = endless “debates”
- Facilitate communication with psychologists, neurologists, computer scientists, evolutionary biologists
- Improve inferences from *your* data, *now!*

11

What’s grammar?

- It’s what’s covered in grammar books:
 - Pronunciation
 - Morphosyntax
- Grammar is ancient and multicultural:
 - Sumerian morphological analysis
 - Panini’s morphophonology
 - Chinese rhyme books, “empty” & “full” words
- Everybody believes in it, *in some sense*

12

Everybody agrees that ...

- Grammar is systematic
 - Regular, lawful, productive, generative, rule-governed, structured, principled...
 - Not rote memory or ad hoc analogy
 - Coordinates language across speakers
- Grammar is mental knowledge
 - Not just historical residue
 - Not fleeting processes, but “permanent”
 - Not identical to surface linguistic behavior

13

Consequences of mentalism

- Grammar is hidden
 - *Langue* vs. *parole*, competence vs. performance, I-language vs. E-language
 - All psychology is like this, which is why psychologists are super skeptics
- No special “competence data”
 - Grammatical facts are just psycholinguistic / phonetic facts analyzed with different goals

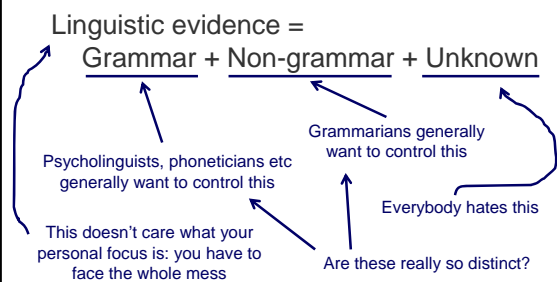
14

More controversial is whether ...

- Grammar is “knowledge that”?
 - Cf. “knowledge how”, “competence to perform” (cf. Derwing & Baker 1978)
- Grammar is categorical?
 - [\pm Grammatical] (cf. Chomsky 1965)
 - Discrete and deterministic (cf. many)
- Grammar is autonomous?
 - Syntax & phonology vs. discourse & psychophysics (e.g. Newmeyer 1998)

15

A framework for controversy



16

Universal grammar

- Don't blame Chomsky
 - Every science strives towards universality
 - Linguists have always assumed universals, at least in their terminology (“nouns”, etc)
- UG : grammars :: a grammar : a language
 - Thus UG need not involve surface universals, anymore than grammar is visible on the surface of a language

17

What's evidence?

- Science = Rationalism + Empiricism
- **Rationalism**: Top-down, deductive logic
 - Formal models of hypotheses and predictions
 - The more elegant and precise, the better
- **Empiricism**: Bottom-up, inductive logic
 - Practical tools for data exploration
 - The more skeptical, the better

18

Rationalists vs. empiricists

- Chomsky prefers rationalism
 - Galileo & Newton favored math over method
- Many others prefer empiricism
 - Boyle (pioneer experimentalist) (Shapin 1996)
 - Psychologists (Miller 1990)
- Example: Views on UG
 - We're all one species, hence UG must exist
 - Only believe in UG if supported by typology

19

Striking a balance

- Rationalists must get their hands dirty
 - Chomsky's history of science is *backwards*: The Scientific Revolution added a radical new empiricism to preexisting rationalism!
 - New ideas are often driven by new tools (telescope; computer; cf. microscope)
- Empiricists need formalism
 - What's wrong with precise hypotheses?
 - Facts do *not* speak for themselves

20

The two kinds of evidence

- Experiments
 - Manipulate input, observe output
 - Can test causality
 - Analysis defined by experimental design
- Corpora
 - Preexisting data: can only observe output
 - Can't test causality, only correlation
 - No design, so no "best" analysis
 - Respectable: Astronomy, archeology, epidemiology

21

Grammarians' preferred evidence

- Syntax: Experiments
 - Native speaker judgments of invented sentences
- Phonology: Corpora
 - Analysis of dictionaries
- Morphology: A little of both
- UG: Corpora
 - Typology of grammars

22

Why judgments in syntax?

- Why aren't corpora preferred?
 - Combinatorics: Key examples can be rare
 - Regular-expression searches miss the complexity of human syntax
- Why judgments?
 - "Offline" tasks tap into long-term memory
 - A learnable, evolvable communication system should involve "obvious" patterns

23

Cautions with judgments

- Real experiments vs. "thought experiments"
- Don't underestimate nuisance variables
 - Schütze (1996) documents the sad history
- Use proper factorial design
 - Cowart (1997) shows how to do it right
- Judges also do intuitive corpus analysis
 - Labov (1996), Luka & Barsalou (2005)
 - Even Chomsky (1970) admits role for analogy

24

Other syntactic evidence

- Corpora
 - May reveal patterns missed by intuitions (Manning 2003)
- Processing experiments
 - Can help show how much grammar is really just parsing (Hawkins 2005)
 - Can help distinguish real grammar from mere parsing (Phillips & Wagers in press)

25

Why corpora in phonology?

- Formal properties
 - Combinatorics limited (in the lexicon)
 - Regular-expression searches work well
- Corpora as models of the lexicon
 - Much phonological knowledge is lexical
- The “best” analysis is definable
 - The grammar acquired by the child, guided by UG, is the one “true” corpus analysis

26

Cautions with phonological corpora

- Don't underestimate nuisance variables
 - Historical relics (Ohala 1986, Blevins 2004)
- Phonetic detail is lost in transcriptions
 - Listeners may “clean up” crucial information (Port & Leary 2005)
- Why stick with second-class evidence?
 - Unlike astronomers and archeologists, phonologists have a choice

27

Other phonological evidence

- Other corpora
 - Natural speech errors (Fromkin 1971)
 - Poetry, loanwords, etc
- Other experiments
 - Judgments (Halle 1962) and other offline tasks (Derwing & de Almeida 2004)
 - Online tasks (Levelt et al. 1999)
 - Phonetics (Kingston & Beckman 1990)

28

Testing Universal Grammar

- Typology as corpus linguistics
 - Hard to get a large, balanced sample
- Child language
 - Look for biases that go beyond adult corpus
 - Competence vs. performance again
- Other evidence?
 - Artificial grammar learning by adults
 - “Subgrammatical” patterns (Hawkins 2005)

29

The best sort of evidence

- As varied as possible
- As informative as possible
 - Not merely conventional or convenient
- Quantitative
 - Face it: All sciences end up mathematized
 - Hypotheses should be maximally precise
 - Fitting hypotheses to data in the face of imperfect knowledge requires statistics

30

Review

- Grammar is real (in some sense)
- Grammar is hard to study
- There's no one right way to do it
- There are even more wrong ways
- We've all got a lot to learn!

31

References (I)

- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Chomsky, N. (1970). Remarks on nominalization. Reprinted 1972 in *Studies on semantics in generative grammar* (pp. 11-61). Mouton.
- Chomsky, N. (2002). *On nature and language*. Cambridge University Press.
- Cowart, W. (1997). *Experimental syntax: Applying objective methods to sentence judgments*. Sage Publications.
- Derwing, B. L., & Baker, W. J. (1978). On the re-integration of linguistics and psychology. In R. N. Campbell & P. T. Smith (Eds.), *Recent advances in the psychology of language: Formal and experimental approaches*, pp. 193-218. New York: Plenum Press.
- Derwing, B. L., & de Almeida, R. G. (2004). Non-chronometric experiments in linguistics. University of Alberta and Concordia University ms. In D. Eddington (Ed.), *Experimental and quantitative linguistics*. <http://alcor.concordia.ca/~almeida/Derwing-deAlmeida-2004Sub.pdf>
- Di Sciullo, A. M., & Williams, E. (1987). *On the definition of word*. MIT Press.
- Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27-52.
- Halle, M. (1962). Phonology in generative grammar. *Word*, 18, 54-72.
- Halle, M., & Mohanan, K. P. (1985). Segmental phonology of Modern English. *Linguistic Inquiry*, 16, 57-116.
- Hawkins, J. A. (2005). *Efficiency and complexity in grammars*. Oxford University Press.

32

References (II)

- Hayes, B. (2000). Gradient well-formedness in Optimality Theory. In J. Dekkers, F. van der Leeuw, & J. van de Weijer (Eds.), *Optimality Theory: Phonology, syntax, and acquisition* (pp. 88-120). Oxford University Press.
- Huang, J. (1982). *Logical relations in Chinese and the theory of grammar*. MIT Ph.D. thesis.
- Kingston, J. & M. E. Beckman (Eds.) (1990). *Papers in laboratory phonology 1: Between the grammar and physics of speech*. Cambridge: Cambridge University Press.
- Kuhn, Thomas S. (1970). *The structure of scientific revolutions*. (Second edition, enlarged). University of Chicago Press.
- Labov, W. (1996). When intuitions fail. L. McNair (Ed.) *CLS 32: Papers from the Parasession on Theory and Data in Linguistics* (pp. 77-105). University of Chicago.
- Leben, W. (1978). The representation of tone. In V. A. Fromkin (ed.), *Tone: A linguistic survey* (pp. 177-219). Academic Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75.
- Luka, B. J., & Barsalou, L. W. (2005). Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language*, 52, 436-459.
- Manning, C. D. (2003). Probabilistic syntax. In R. Bod, J. Hay, & S. Jannedy (Eds.) *Probabilistic linguistics* (pp. 289-341). MIT Press.

33

References (III)

- Meinunger, A. (2001). Restrictions on verb raising. *Linguistic Inquiry*, 32, 732-740.
- Miller, G. A. (1990). Linguists, psychologists and the cognitive sciences. *Language*, 66, 317-322.
- Newmeyer, F. J. (1998). *Language form and language function*. MIT Press.
- Ohala, J. J. (1986). Consumer's guide to evidence in phonology. *Phonology Yearbook*, 3, 3-26.
- Phillips, Colin, & Matthew Wagers. 2006. Relating structure and time in linguistics and psycholinguistics. In G. Gaskell (Ed.) *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press.
- Port, R., & Leary, A. P. (2005). Against formal phonology. *Language*, 81 (4), 927-964.
- Rapp, I., & von Stechow, A. (1999). Fast. Handout from talk presented at ZAS/Berlin. [Cited in Meinunger 2001]
- Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
- Shapin, S. (1996). *The scientific revolution*. University of Chicago Press.
- Spencer, A. (1991). *Morphological theory: An introduction to word structure in generative grammar*. Basil Blackwell, Oxford.
- Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, 21, 291-311.
- Xu, L. (1990). Remarks on LF Movement in Chinese questions. *Linguistics*, 28, 355-382.
- Zoll, C. (2003). Optimal tone mapping. *Linguistic Inquiry*, 34 (2), 225-268.

34