1

MORE QUESTIONS: LSA 347

Rethinking Linguistic Competence

JOAN BRESNAN

2

Question:

Given that there are grammatical rules of a probabilistic character, do there exist rules that maintain a discrete boundary between grammaticality and ungrammaticality?

example:

(1) Who did you discover that they liked?

(2) *Who did you discover the fact that they liked?

3

This example is fraught. It is not, unfortunately, a clear case. (More on unclear cases soon.)

Some believe that Complex NP Constraint violations are not ungrammatical, but only semantically odd. On these accounts both of (1) and (2) are grammatical and contrastive intuitions of difference are attributed to other factors (semantics, information structure). There is an interesting literature on this (Ross, Kuno, Erteschik-Shir, Deane).

`http://lingo.stanford.edu/sag/LSA344/syll/`
`subjacency.html`

4

Another possible example (Labov 1975: 80):[a]

(1) That John told him was a shame.

(2) *John told him was a shame. [meaning (1)]

---
[a]William Labov. 1975. "Empirical Foundations of Linguistic Theory," in *The Scope of American Linguistics*, ed. by R. Austerlitz, 77–113. Lisse: The Peter de Ridder Press.

5

Is this a matter of variety?
Relevant examples courtesy of Liz Coppock:

"It's circumstantial, is the problem," Silvestri grumbled....[a]

"There was a big newspaper scandal and Dad was denounced everywhere... He's upset, is the point."[b]

---

[a] *On, Off* by Colleen McCullough p. 308, 2006]
[b] *Watch Your Mouth* by Daniel Handler, p. 27, 2002

6

David Adger has suggested the following as a sentence which must be ruled out as ungrammatical in English and not improbable:

*Dog the ran.

7

Within a probabilistic framework,

- ungrammaticality can be regarded as asymptotic improbability

- for example, in a stochastic OT framework, the probability of occurrence approaches 0 as the distance between the crucial constraints (measured in standard deviations of the gaussian noise distribution) increases.

- The Gradual Learning Algorithm permits such cases of free fall in ranking when violations of a constraint are never overridden.

8

Nevertheless, *the potential for variation remains as long as there is plasticity in the system.*

Perturbances (for example, contact with Norwegian) could subtly shift the rankings in the stochastic grammars of a community of English speakers to the point where *dog the ran* becomes possible.

9

The plasticity-of-grammar hypothesis:

In both child and the adult individual, grammar is a highly plastic cognitive system sensitively tuned to the frequencies of the environment. At the group level, the social dynamics of linguistic communities change language (in both time and space), selecting from and replicating this inherent variability.

(recall Weinreich, Labov, and Herzog 1968)

10

Question:

Linguist X has made the point that resumptive pronouns are ungrammatical, not because they aren't used, but because they are only used in situations where cognitive problems of speech production and perception—long distance PRO interpretation—are involved.

Examples:

(1) *He's the guy that I like him.

(2) He's a guy that you and I are the only ones that like him.

11

In phonology, the existence of last resorts, repairs, and overrides are ordinarily viewed as part of the grammar, to be explained by the architecture of the system (e.g. global rules or optimization of conflicting constraints).

Hence, an alternative hypothesis: resumptive pronouns are generally dispreferred by a constraint which can be overridden by higher-ranked constraints on the distribution of gaps.

12

	Avoid non-Local gap	Avoid pronoun
the guy that I like him		*!
☞ the guy that I like __		

	Avoid non-Local gap	Avoid pronoun
☞ a guy that you and I are the only ones that like him		*
a guy that you and I are the only ones that like __	*!	

13

The basic point:

limited contexts for grammaticality may arise from constraint interaction. Witness Green's (1971)

He denied it to Kim
*He denied Kim it

where lexical constraints prohibiting a construction type are overridden by the an even stronger constraint.

14

	*V NP Pron	HARMONY	FAITH$_{deny}$(R)
denied the opportunity to Kim		*!	
☞ denied Kim the opportunity			*

	*V NP Pron	HARMONY	FAITH$_{deny}$(R)
☞ denied it to Kim		*	
denied Kim it	*!		*

[using constraints from Bresnan and Nikitina 2007 for convenience]

15

Question: How to decide (i) if data constitute authentic counterevidence to a hypothesis or (ii) if they should be dismissed as the result of error?

The problem is not confined to usage data – it arises also with data from intuitions, as in the previous questions.

16

the question of "unclear cases"

Chomsky (1957: 14):

"...we may assume for this discussion that certain sequences of phonemes are definitely sentences, and that certain non-sequences are definitely non-sentences. *In many intermediate cases we shall be prepared to let the grammar itself decide,* when the grammar is set up in the simplest way so that it includes the clear sentences and excludes the clear non-sentences."

17

Labov 1975:

Chomsky, who instantiates best practice, nevertheless tends to decide the unclear cases in favor of his own theory and against others'.

This is *the problem of experimenter bias* that perpetually arises with standard practice in syntax.

18

If we adopt this approach, then (and this has really been the practice in much of the field of syntax):

- both of the short and long resumptive pronoun examples are ungrammatical under Linguist X's theory,

- both are completely grammatical under Linguist Y's theory,

- and they are split in grammaticality under the toy OT theory sketched here.

19

The conveniently flexible boundary between competence and performance:

- the competence grammar excludes both types of data, and performance factors make one type sound acceptable;

- the competence grammar includes both types, and performance factors make one type sound unacceptable;

- and the competence grammar includes one type only, perhaps because it is a grammaticalization of performance

20

The result: A proliferation of competing theories each supported by conflicting sets of unclear data!

= The state of the art in syntax?

21

Labov 1975: "The studies of introspective judgments carried out so far show that variation in this field is widespread, uncontrolled and chaotic"

- wholesale rejection of linguists' judgments in controlled experiments,
- instability of claimed idiolectal dialects under replication,
- contradictions between introspections and behavior.

22

The web examples from Bresnan and Nikitina (2003, 2007) are cases of *contradictions between introspections and behavior.*

Pinker, Levin, Krifka, Oehrle, Bruening, Davis et al ⇒ ungrammatical

23

What is to be done?

Labov (1975: 103):

I. THE CONSENSUS PRINCIPLE. If there is no reason to think otherwise, assume that the judgments of any native speaker are characteristic of all speakers of the language.

II. THE EXPERIMENTER PRINCIPLE: If there is any disagreement on introspective judgments, the judgments of those who are familiar with the theoretical issues may not be counted as evidence.

24

III. THE CLEAR CASE PRINCIPLE: disputed judgments should be shown to include at least one consistent pattern in the speech community or be abandoned.

25

In Experiment 2, Bresnan (2006) applies a version of the Clear Case Principle: provides evidence that the usage examples of double object datives with pronominal recipients found by Bresnan and Nikitina are judged by experimental subjects to be significantly better than the types constructed by linguists, and in fact are judged no less natural than theoretically *grammatical* constructions of alternating verbs with lexical NP recipients.

26

Now consider

Linguist Y's "teh" example: we discover thousands of hits of the word *teh* on the web in use in natural contexts. We can even develop an explicit theory of the cognitive processes and psychomotor mechanisms that give rise to the occurrence of these *teh* forms. Must we conclude from their frequency in actual usage and their well-motivated cognitive/psychomotor origins that the rules of English orthography should generate *teh* is an alternating form of *the*?

27

For the sake of the argument we will assume with Linguist Y that all occurrences of *teh* found on the web are typographical errors.

That is, we set aside all instances of *teh* where it is an orthographic rendition of actual variant pronunciations of *the*...

Should we change the rules of English orthography to generate *teh*?

28

"teh" fails all three of Labov's Principles:

I. The Consensus Principle: ask any native speaker who knows how to spell whether *teh* is a word

II. The Experimenter Principle: don't ask Linguist Y

III. The Clear Case Principle: there are no disputed judgments about *teh* – it is (by design) a completely clear case of misspelling.

29

If Linguist Y, from a sincere belief that *teh* is an unclear case, performed an experiment parallel to my experiment 2 (asking subjects to rate the naturalness of *teh* vs. *the* in natural contexts of spontaneous usage), what would he find?

—most likely, very clear evidence with which all speakers who are familiar with English spelling would agree: *teh* is not (currently) a word of English.

The problem of *teh* as an example of the pitfalls of usage data is trivial, easily solved by Labov's 1975 Principles.

30

Are such examples a problem for probabilistic theories of language?

On the contrary, probabilistic theories are preferred in NLP precisely because they are far more *robust* to errors during learning than the categorical kind (see Boersma and Weenink 2005 for examples within the OT framework).

31

Question:

referring to *anymore* and cases where people use forms of expression that they object to, Labov says: *"I argue that intuitions match production in the vast majority of cases, but when they don't, generative grammar has taken the position that the language is what they think they say, not, what they say. Whereas I'd like to say that the language is what people say, not what they think they say. What's your take on this?*

32

Answer: (given in class)