

Assignment Extensions

- Extensions given only for unpredictable problems (including illness, but excluding too much work in other classes)
- Direct all requests for extensions to Robert Munro: rmunro@stanford.edu
- Direct any other inquiries to your section leader.

Language and Rationality

SymSys 100, April 23, 2009

Why So Much About Language?

- About 1/3 of the course is on language
- Why devote so much to it?

Why So Much About Language?

- Clear links to other parts of course
 - Much of the discussion of rationality focuses on wording (esp. Tversky & Kahneman)
 - Turing test based on linguistic interaction
- More generally, work on language has played a big role in cognitive science

Why So Much About Language?

- Language is universal to humans
- It is species-specific
- It is the principal mechanism that has allowed human culture to progress across generations

Overview of Language Module

- Today: language & rationality
- Next four sessions relate to question of the extent to which linguistic knowledge is genetically determined
- Final lecture in this module on the relationship between language and thinking

Comments on Today's Reading

- Less connected to lecture than usual
- Wanted a general lead-in to the topic of language
- It's a readable introduction, co-authored by one of our local luminaries

Comments on Today's Reading

- Language is such a pervasive part of our everyday activities that most people never think about it much
- One way to appreciate how complex and multi-faceted it is is to consider what it would take to get a machine to exhibit mastery of language at the level of a 10-year old human
- Jurafsky & Martin survey some of the components of this task

A Puzzle

- It's striking that nobody has succeeded in doing this
- Computers with human-like linguistic abilities have been predicted since the 1940s
- Computers outperform people in some things that seem hard to us (e.g, chess.)
- Yet language technologies still seem quite primitive

One Possible Explanation

- Maybe people have evolved a specialized innate capacity for language
- This idea has been the source of much controversy in cognitive science, and we'll spend a couple of weeks on it.
- But today, I want to focus on some more basic questions about language.

What is Language?

- Conventional answer: A system of connections between sounds and meanings
- What's wrong with that answer?
 - Auditory modality isn't necessary
 - Would cover animal cries

What Distinguishes Human Language?

- Infinity of forms and meanings
- Abstraction from the here and now
- Complexity of the structures employed and meanings conveyed

Are Languages Rational?

- What does that even mean?
- Two possible interpretations:
 - Are the forms employed logical ways of expressing the intended meanings?
 - Do languages have the design features of a rationally designed communication system?

Strange Form-Meaning Pairings

- *bikini* (singular) vs. *bathing trunks* (plural); also *you*
- German *das Mädchen*, ‘the girl’, is grammatically neuter
- *Marysia niczego nie dala Jankowi* (Polish)
Mary nothing not gave John
‘Mary did not give John anything’
- *flammable* is a synonym of *inflammable*
- *Bo took a pee* [not *Bo left a pee*]

So What?

- Languages are the products of evolution (both biological and linguistic)
- So they contain lots of little quirks
- But, on the whole, forms reflect meanings fairly straightforwardly

Is Language Rationally Designed?

- Superficially, the answer seems to be “no”
- Natural languages don't look much like artificial languages, which **are** rationally designed

How Do Natural & Artificial Languages Differ?

- **Syntactic complexity:** Natural languages have things like agreement, grammatical gender, case marking whose function isn't obvious
- **Ambiguity:** An expression in an artificial language has one and only one meaning

Ambiguity in Natural Language

- several varieties
- much more pervasive than most people realize
- seems *a priori* disfunctional

Varieties of Ambiguity

- Lexical: *bank*
- Syntactic: *We saw the man with the telescope*
- Subject vs. Object: *The chicken is ready to eat*
- Scope: *No student solved exactly two problems*
- Pragmatic: *I'm parked on the Oval*

A Riddle

- You're standing on a bridge, and see a boat full of people approaching. It goes under the bridge, and, when it emerges on the other side, there's not a single person on the boat. Nobody climbed onto the bridge or jumped in the water. How is this possible?
- Answer: Everyone on board is married.
- The point: We tend not to notice ambiguities.

Real-life Examples of Ambiguity

- Signs:



- Headlines:

Child's Stool Great for Use in Garden

Squad Helps Dog Bite Victim

Joint Committee Investigates Marijuana Use

Ambiguity in a Contract

Buyer is willing to buy the property at the Purchase Price, net of the full cost of remediating the Hazardous Materials...

→ Who pays for the cleanup?

Multiple Ambiguities in a Sentence

- m ambiguous lexical items and n syntactic ambiguities in a sentence yields 2^{m+n} meanings
- Hence (at least) 32 meanings for:
Old friends and acquaintances remembered Pat's last time in California
- Martin, Church, & Patil reported that their NLP system assigned 455 parses to the sentence:
List sales of the products produced in 1973 with the products produced in 1972

Ambiguity should be hard.

Why?

- A language is a medium for conveying information.
- Unless each signal in the language is identified with unique information, the receiver of the signal has to do extra work to determine the intended interpretation.
- Likewise, ambiguity increases the likelihood of misinterpretation.

Reminder: Where We Are

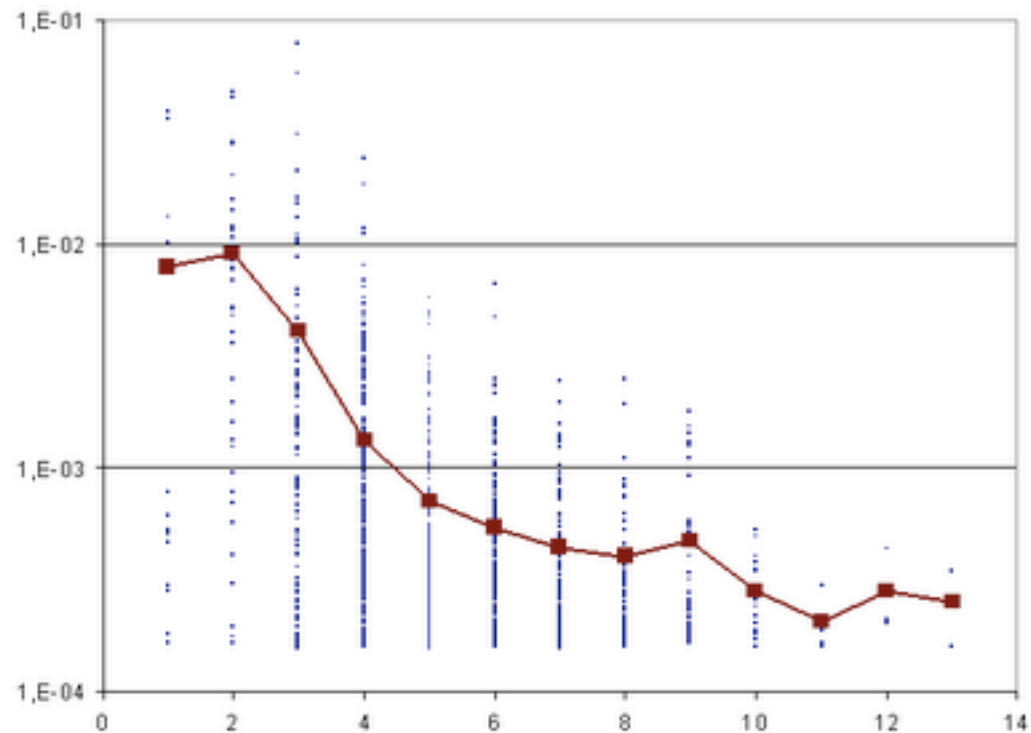
- Considering whether languages are rational -- that is, have the properties a rational designer would give them
- Ambiguity provides what looks like a powerful argument for a negative answer
- But some linguists have argued otherwise

Basic Assumptions

- Primary function of language is communication between people
- Goal: maximize efficiency of information transmission
- Entails a speed/accuracy trade-off
- Primary medium of language is speech, so information transmission constrained by speed of articulation

One Immediate Consequence

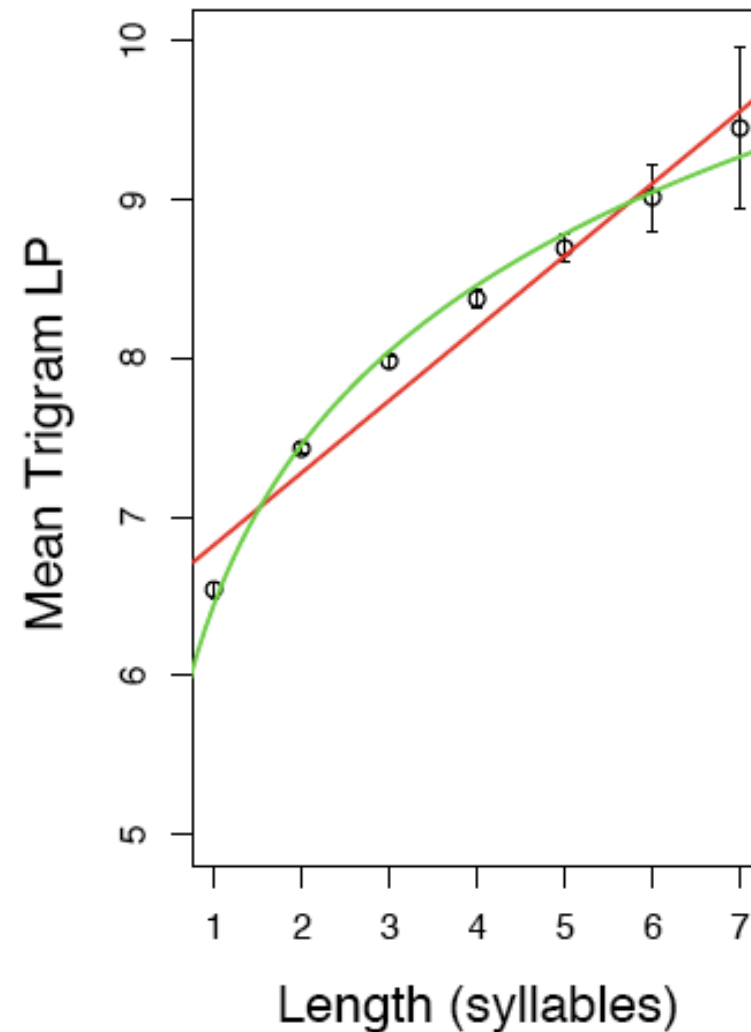
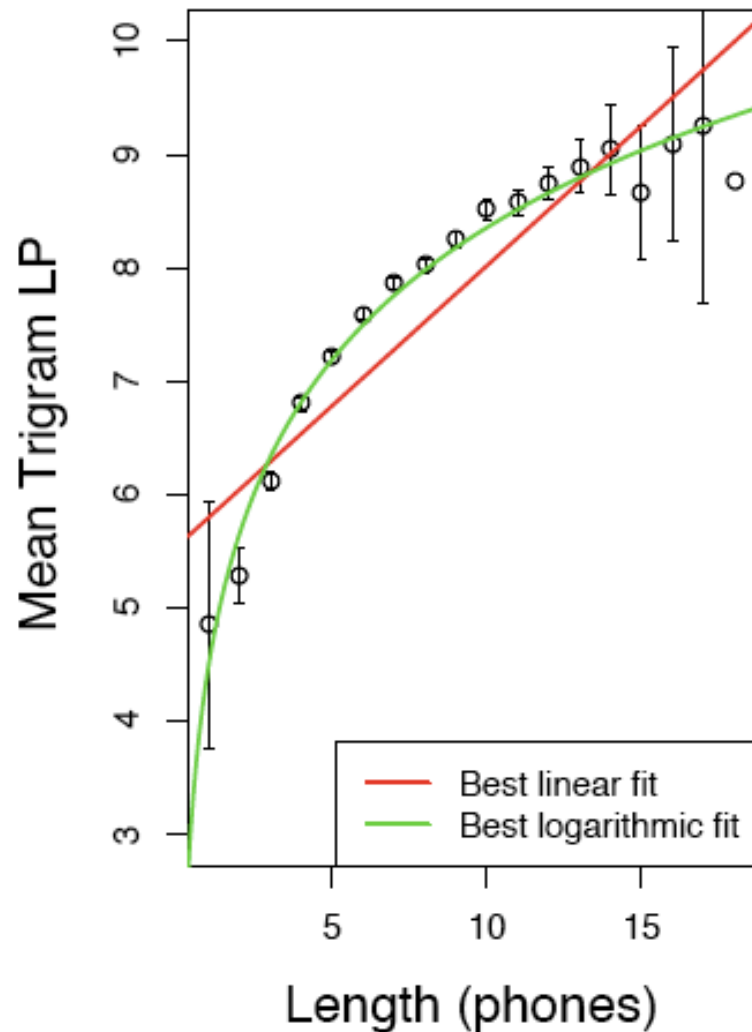
- Frequent words should be short (to minimize articulation time and effort)
- This holds, at least for words more than two letters long:



Another Consequence

- To facilitate comprehensibility, words that are less predictable in context should be longer
- Grad students Steven Piantadosi (MIT) & Hal Tily (Stanford) tested this, as follows:
 - Predictability of a word in context measured (crudely) in terms of its conditional probability in a large corpus, given the two preceding words
 - Length measured in both phones (speech sounds) and syllables.

Piantadosi & Tily's Results



Yes, but....

- These are fairly low-level facts about languages
- The arguments against the rationality of language structure was based on high-level things like grammar and ambiguity
- The challenge is to show that all these things make sense from a functional perspective

Can Ambiguity Be Explained Functionally?

- People have tried
- Any suggestions?

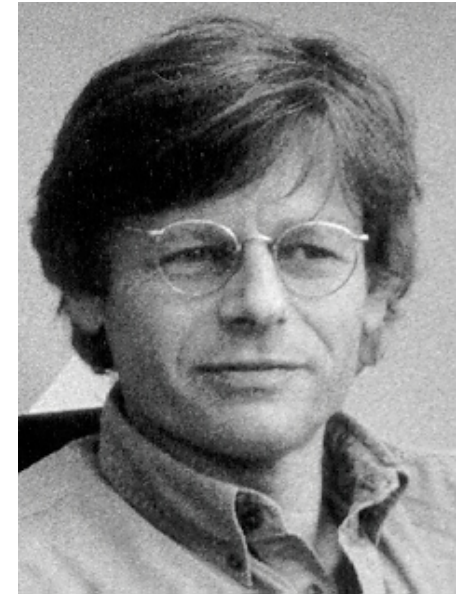
Zipf's Explanation for Lexical Ambiguity

- Principle of Least Effort: “a person...will strive to solve his [sic] problems in such a way as to minimize the total work that he must expend in solving both his immediate problems and his probable future problems”
- Speaker's work is minimized if one word expresses all meanings (maximal ambiguity)
- Listener's work is minimized if there is one word for each meaning (no ambiguity)
- The solution is a compromise, in which the number of words is between one and the number of meanings.
- Hence, “we may expect that at least some words must have multiple meanings”



Levinson's Explanation for Ambiguity

- Articulation is a bottleneck for communication
- People are very good at contextual inference (cf Jay's lectures)
- We can keep utterances short by relying on listeners' inferential abilities to disambiguate
- The cost of occasional miscommunication is offset by the gain in speed of information transmission



Other Possible Functions of Ambiguity: Diplomacy and Deception

- Sometimes a speaker wants to be understood one way by one listener and another way by another listener:

Nothing is better than Pat's cooking.

- Ambiguous language facilitates plausible deniability, e.g., *lifetime guarantee*
 - Most people take it to mean 'guaranteed for the lifetime of the consumer'
 - Manufacturers use it to mean 'guaranteed for the lifetime of the product'

From Roger Cohen's column in today's *New York Times*

“Language is lethal. The Bush administration’s legal memos opening the way to torture are a reminder of the intimate link between a bureaucrat’s lawyerly subordinate clause and a man’s near drowning.

...

“Don’t say what you mean when you mean to do the unspeakable. That’s an old rule. It was perfected in the 20th century from Moscow to Buenos Aires.

“Opacity is the refuge of the faceless tormentor. The constitutions of totalitarian states are always unreadable, impenetrable — and very long. In a thicket of words lies plausible deniability when the time for horror’s accounting arrives.”



So...Is Language Rationally Designed?

- Question remains open
- Trying to explain as much about language as possible in terms of its functions is a good research strategy