Syntactic Processing

Psycholinguistics

LING/PSYC 27010

Autumn 2016

agenda for today (Tues)

1. start thinking about lab topic a bit!

2. reflecting on search engine autocomplete function

3. "Modularity of Mind" and modular systems

4. distinguishing "sentence processing" from "syntactic processing" from "parsing" from...

5. intro to sentence processing
   (w/ focus on syntax today)

6. housekeeping after class or office hours

reflections on autocomplete

question: how does Google autocomplete work?

> **a more productive (two-part) question:**
>
> 1. *what* does Google autocomplete *do*?*****
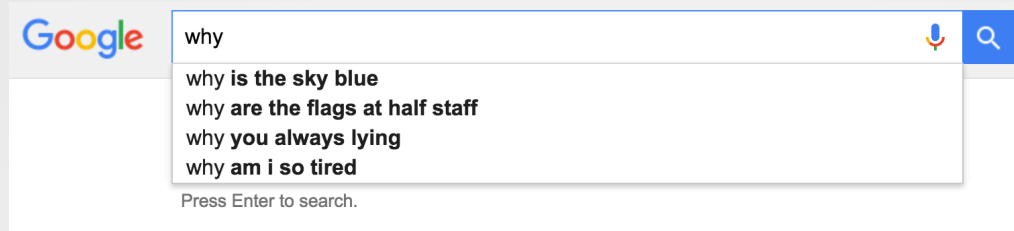> 2. *what kind of data does it produce* in virtue of which we are able to answer 1.?

it makes predictions on the basis of

- incomplete, partial input
- domain-specific "knowledge" about the input
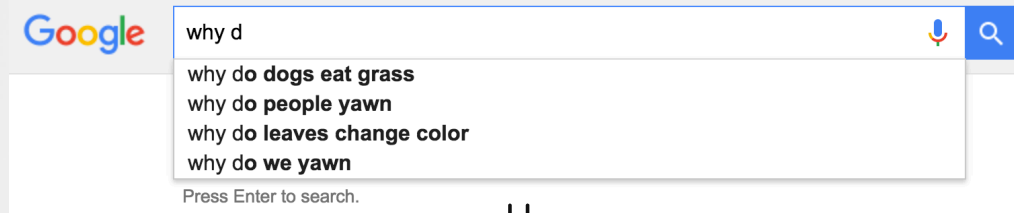- domain-general knowledge that makes some predictions better than others
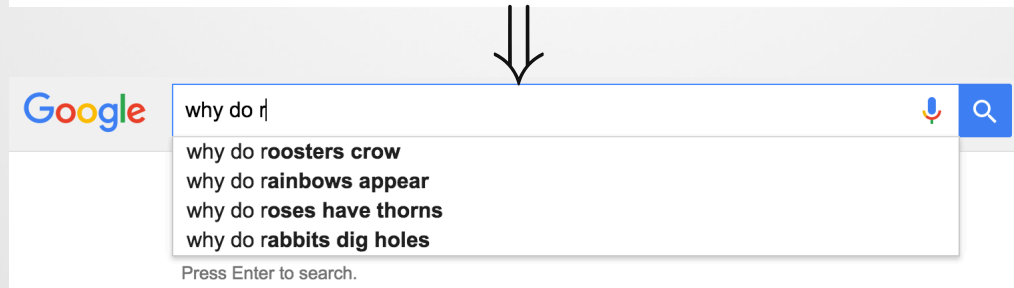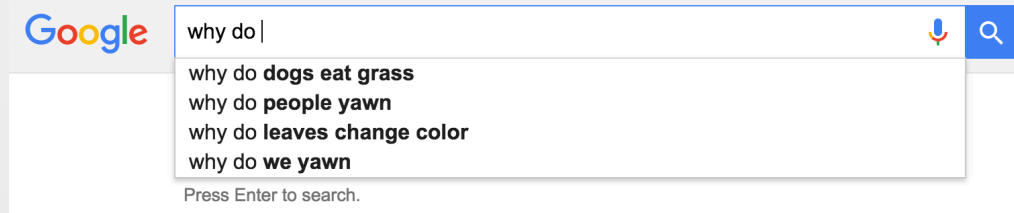

*s/o to WVO Quine for the wisdom! RIP

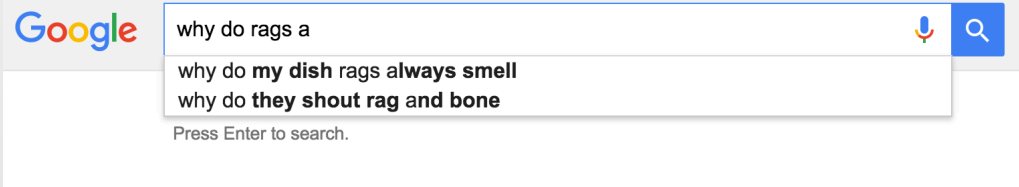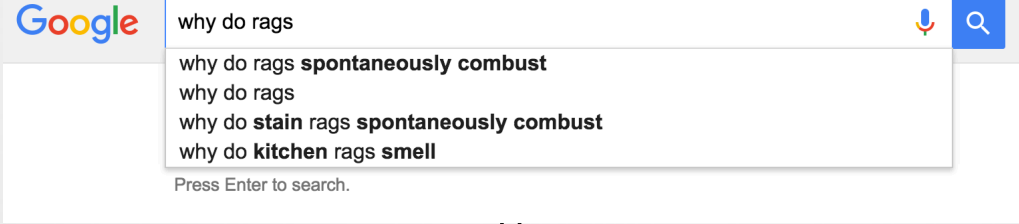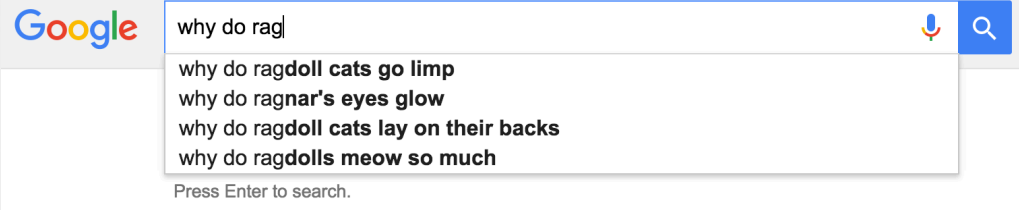# reflections on autocomplete

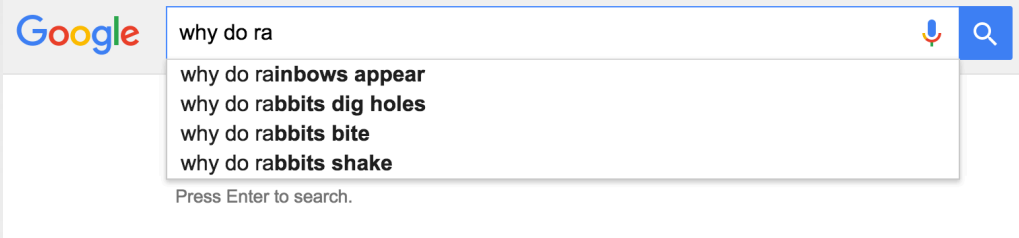committed to verb
to be

wrong! no form of
be starting with '*d*'

reanalyzed to verb
to do

reflections on autocomplete

getting desperate...

reflections on autocomplete
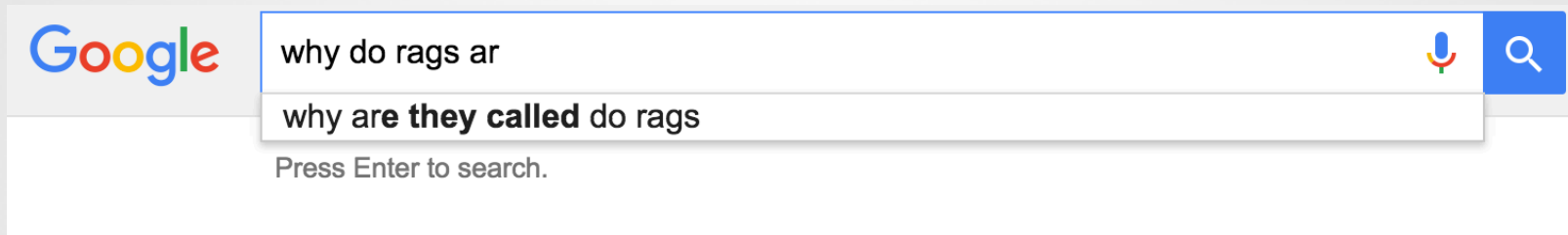
mission...failed? accomplished?!
not obvious!

$\Downarrow$
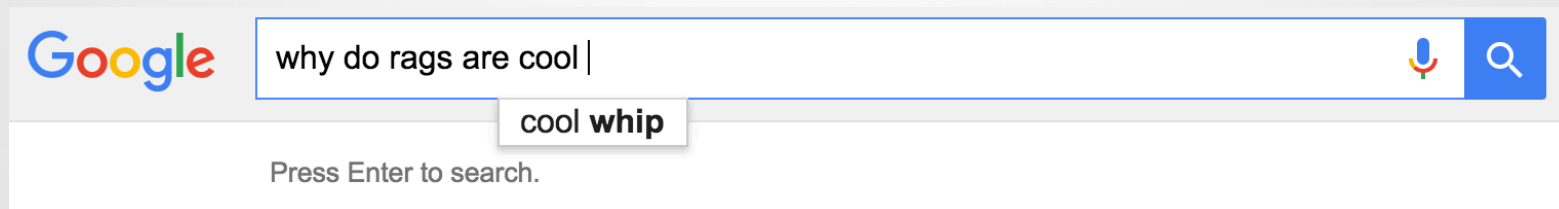
reflections on autocomplete

question: why are we talking about autocomplete?!

>> think of the search bar as an auditory processor

>> think of partial input to the search bar as an unfolding utterance

>> think of the autocomplete algorithm as a sentence processor
(parser + compositional meaning)

>> think of autocomplete results as candidate parses of
(or predictions about) the unfolding utterance

>> think of the searcher as a scientist trying to learn about the underlying
system by trying different queries and seeing how it reacts
(i.e. what data it produces given some input or sequence of inputs)

> there is a non-trivial similarity between the human
> sentence processing system and mechanical prediction-
> generating systems like Google's autocomplete feature

reflections on autocomplete

the language processor...

- is incremental
- is predictive
- is anticipatory
- is constrained by mind architecture (e.g. memory)
- is extremely fast and extremely efficient
- is a biological system of some kind

  question: what are the *key* differences?

the analogy is nice, but sentence processing is a very different kind of animal from search autocomplete (I'd submit, at least)

reflections on autocomplete

with search autocompletion, any information that is relevant and available could (in principle) be used to guide predictions

whether this is also true for human language processing has been debated since the inception of (modern) psycholinguistics

but if a piece of information is relevant and available, *why wouldn't* we use it?

Modularity of Mind

Modularity of Mind

the mind consists (in part) of self-contained, domain-specific, autonomous processing units (*modules*) that are functionally independent and closed off from outside input

>> a natural fit with Chomskyan nativism

>> the idea of a syntactic module was attractive

Mysterian psychologist, Jerry Fodor.
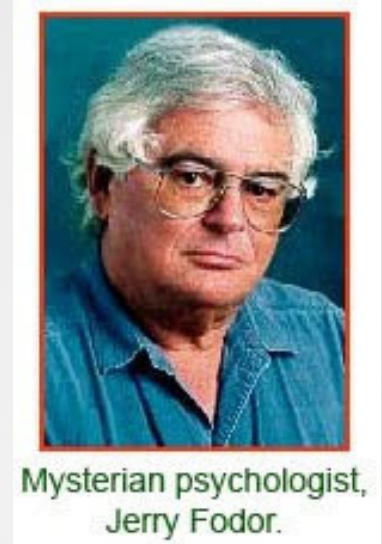
popularized by philosopher Jerry Fodor in the 1980s
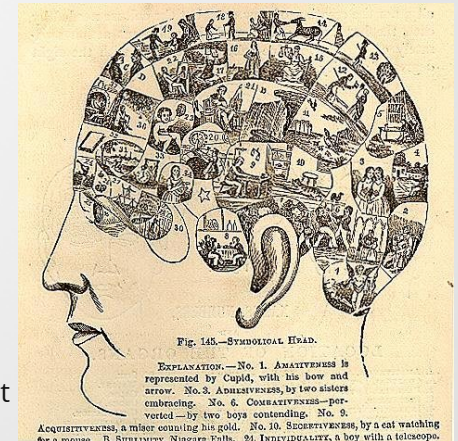
kinda reminiscent of phrenology :/

modular systems

modular systems/mechanisms in human cognition are

- automatic: they operate without deliberation

- domain-specific: they deal with only a single kind of information

- informationally encapsulated: modules do not interact with one another, and are sensitive to only a subset of available information

- cognitively impenetrable: not open to introspection, not affected by existing beliefs, etc.

ex: perception of visual depth, distance, etc.

modularity is a nice theme that we'll
encounter periodically, but back to
sentence processing for now!

back to sentence processing

terminology:

- sentence processing
- syntactic processing
- parsing

segmentation and lexical access

⬇

syntactic processing { parsing and structure-building

⬇

compositional interpretation

⬇

pragmatic reasoning

language processing doesn't (always) happen in a neat
sequence of disjoint stages, but as a rough
approximation it looks something like this:

ambiguity everywhere!

a recurring theme:

there is pervasive ambiguity and uncertainty
at every level of linguistic structure

this is probably easiest to see at the sentence level

The lion was shot by the dangerous poacher with a scar.
The lion was shot by the dangerous poacher with a rifle.

I told you I'm going to do it tomorrow!
I told you I'm going to do it yesterday!

The witness examined by the lawyer turned out to be unreliable.
The evidence examined by the lawyer turned out to be unreliable.

intuition: language-independent knowledge seems to affect amount of uncertainty in parsing

ambiguity everywhere!

these strings are "ambiguous" in two
different senses of the word

I've seen cuter dogs than Fido.

I've seen cuter dogs than Fido...

You can have eggs and ham or bacon.

You can have eggs and ham...

Teens who drink piss off their parents.

Teens who drink piss...

globally ambiguous

temporarily ambiguous

ambiguity everywhere!

the study of temporary ambiguity is king in
sentence processing research, esp. syntax

While mom was dressing the baby wouldn't stop crying!

> backtracking due to incorrect interpretation
> of *the baby* as object of *dress* is an example of
> **the garden path effect**

nb. in an extended sense of the term, almost all sentences
have the potential to be garden path sentences...why/how?!

garden path sentences

crucial observation: there are systematic regularities in the kinds of sentences that garden path people

The woman who sews dresses well.

A dog that I had really loved bones.

Since Jay always jogs a mile and a half is nothing to him.

Mary gave the child a dog bit a bandaid.

The old man the boat.

Mary knew the answer was "B".

...

questions:

why do these sentences cause garden paths?

where are the disambiguating regions?

what do the structures have in common?

garden path sentences

and of course, the famously incomprehensible

The horse raced past the barn fell.

other problematic sentences

If you win, then I lose.

If if you win, then I lose, then the game is fair.

If if if you win, then I lose, then the game is fair, then it's probably no fun.

Either you win or I win.

Either either you win or I win or we draw.

Either either either you win or I win or we draw or we didn't even play.

That grass is green is obvious.

That that grass is green is obvious doesn't matter.

That that that grass is green is obvious doesn't matter is no concern of mine.

the cheese is yellow

the cheese the mouse ate is yellow

the cheese the mouse the cat chased ate is yellow

the cheese the mouse the cat the dog bit chased ate is yellow

other problematic sentences(?)

I thought the Cavs won.

Bill knew I thought the Cavs won.

You told Mary Bill knew I thought the Cavs won.

I explained that you told Mary Bill knew I thought the Cavs won.

My girlfriend is cool.

My girlfriend's roommate is cool.

My girlfriend's roommate's aunt is cool

My girlfriend's roommate's aunt's yard is cool.

My girlfriend's roommate's aunt's yard's secret fort is cool.

these don't seem as difficult...why?

goals for remainder of syntax

1. common experimental paradigms and methodologies (throughout)

2. computational modeling, example and quick demo (Thurs)

- left-, right, and middle-branching structures
- top-down, bottom-up, and left-corner parsing algorithms

3. theoretical models of parsing/ambiguity resolution (next wk)

- The Garden Path Theory
  (the specific theory, not the phenomenon)
- constraint-based models
- the "close enough" theory

4. theoretical debates in sentence processing (next wk)

- serial versus parallel structure-building;
- modular versus non-modular syntactic processing;
- single- versus multi-stage models; and
- heuristics and biases

5. important experimental results interleaved throughout!