Language Processing
and language structure

Psycholinguistics

LING/PSYC 27010

Autumn 2016

timleffel@uchicago.edu
lefft.xyz/psycholingAU16

language: structure and ambiguity

human language is unique among communication systems in nature

two properties we'll focus on today:

- language has structure
  - but so do lots of other biological systems
  - and so do lots of other systems in general

- language is ambiguous
  - but so are lots of other kinds of signals
  - and many symbolic systems (but not all)

what (if anything) makes language unique then?

what kind of structure?

> hierarchical structure
  – anything you can diagram as a tree

> recursive structure
  – define a "procedure" in terms of itself

> unbounded structure
  – no upper-limit on size (length, depth)

there is structure in language at every level of representation

structure in sound

- features
- segments
- syllables
- prosodic words
- prosodic phrases

exx. *récord*/*recórd*; *input*/*intact*

structure in "words"

- root
- affixes
- inflection
- compounding
- incorporation
- reduplication
- ...

ex. *unlockable* (cf. *unlikeable*)

structure in phrases and sentences

- "heads"
- complements ("arguments")
- modifiers

ex. *I saw the man with the telescope.*

structure in meaning

- assertion
- presuppositions
- implicatures
- social meaning
- ...

ex. *got a light? ... sorry, it stopped working.*

structure in communication

- conversations
- narratives

ex.
I went to the game last night.
It was great and the Bulls won.
Obama has a pet flamingo.

structure in reasoning about language

- pragmatic "back-and-forth" reasoning
- sarcasm, irony

ex. | video |

ambiguity everywhere!

a potential challenge for the language processing system is the

rampant ambiguity

that pervades every level of linguistic structure

ambiguity everywhere!

- phonological ambiguity
    - > recall *input/intact* (another example later)

- lexical ambiguity
    - > *bat/bat, bank/bank*, *it*, *that*, ...

- structural ambiguity
    - > I saw the man with the telescope.
    - > You can have eggs and ham or bacon.

humans are incredibly good at correctly resolving these

competence versus performance

- grammar is traditionally associated with competence

  > creates (arguably) discrete, idealized representations

- parsing is traditionally associated with performance

  > deals with noisy, continuous input

competence

> **linguistic competence**
> whatever it is we know when we know a language

- knowledge of words
- knowledge of possible words
- knowledge of syntax
- knowledge of conventional mapping between words and meanings

- we know that some sentences are well formed
  (aka "grammatical"), and others aren't

performance

linguistic performance
our ability to use linguistic competence to produce
and comprehend utterances of language

- how we deploy linguistic knowledge to achieve goals
- recruits:
  > sensory-motor system
  > visual system
  > auditory system
- constrained by limited resources (e.g. memory)
- note – parser versus processor

competence and performance

(1) | The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow!

competence and performance

(1)
| The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow! |

(2)
| The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat! |

competence and performance

(1) | The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow!

(2) | The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat!

(3) | The cheese that the mouse that the cat chased that the dog barked bit that is wearing a red hat ate is yellow!

competence and performance

(1)

> The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow!

(2)

> The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat!

(3)

> The cheese that the mouse that the cat chased that the dog barked bit that is wearing a red hat ate is yellow!

(4)

> The cheese that the mouse that the cat that the dog that is wearing a red hat chased bit ate is yellow!

competence and performance

(1)
> The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow!

(2)
> The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat!

(3)
> The cheese that the mouse that the cat chased that the dog barked bit that is wearing a red hat ate is yellow!

(4)
> The cheese that the mouse that the cat that the dog that is wearing a red hat chased bit ate is yellow!

$\implies$ each of these sentences has a different status

competence and performance

(1)
> The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow!   ✓

(2)
> The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat!

(3)
> The cheese that the mouse that the cat chased that the dog barked bit that is wearing a red hat ate is yellow!

(4)
> The cheese that the mouse that the cat that the dog that is wearing a red hat chased bit ate is yellow!

⟹ each of these sentences has a different status

competence and performance

(1) | The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow! ✓

(2) | The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat! #

(3) | The cheese that the mouse that the cat chased that the dog barked bit that is wearing a red hat ate is yellow!

(4) | The cheese that the mouse that the cat that the dog that is wearing a red hat chased bit ate is yellow!

$\implies$ each of these sentences has a different status

competence and performance

(1)
| The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow!   ✓ |

(2)
| The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat!   # |

(3)
| The cheese that the mouse that the cat chased that the dog barked bit that is wearing a red hat ate is yellow! ✗ |

(4)
| The cheese that the mouse that the cat that the dog that is wearing a red hat chased bit ate is yellow! |

⟹ each of these sentences has a different status
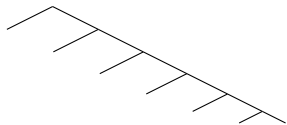
competence and performance

(1) | The cheese that was eaten by the mouse that was bitten by the cat that was chased by the dog that is wearing a red hat is yellow! ✓

(2) | The cheese that chased the cat that bit the mouse that ate the yellow dog is wearing a red hat! #

(3) | The cheese that the mouse that the cat chased that the dog barked bit that is wearing a red hat ate is yellow! ✗

(4) | The cheese that the mouse that the cat that the dog that is wearing a red hat chased bit ate is yellow! ?!?

⟹ each of these sentences has a different status
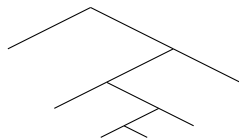
competence and performance

- (4) is definitely well-formed according to the syntax of English

- but it feels very hard to keep 'cheese', 'mouse', 'cat', and 'dog' in your working memory at the same time, and to then link them up to their correct verbs/predicates

- at the same time, (1) contains the same information and has the same syntactic complexity as (4), and yet (1) is much easier to understand

- so wth is going on?!

competence and performance

(3) has a right-branching structure

but (6) has a middle-branching structure

competence and performance

even though the grammar gives sentences (3) and (6) basically the same status, the parser does not

the grammar: determines what structures are and are not part of the language   (part of **competence**)

the parser: the cognitive system that takes raw, unstructured input and converts it into interpretable structures   (part of **performance**)

key properties of the language processor

the language processor is

- temporally incremental (like aging)
- anticipatory and predictive (like attention)
- communicates with other cognitive systems ("integrated")
- automatic (like flinching)
- unconscious (like breathing)
- ...

these properties allow us to cope with ambiguity
(and use it to our advantage)

incremental and anticipatory/predictive

——— —— ——— ——— -— ——— ——— -— ——— ——— ——— ——— —— ——— ———

incremental and anticipatory/predictive

Since ___ _____ ____ _ ____ ___ _ ___ ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay _____ ____ _ ____ ____ _ ____ ____ _____ ____ __ ____ ____

incremental and anticipatory/predictive

<u>Since</u> <u>Jay</u> <u>always</u> ____ _ ____ ____ _ ____ ____ _____ ____ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs _ ____ ___ _ ___ ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a ____ ___ _ ___ ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile ___ _ ___ ___ ____ ___ __ ___ ___

incremental and anticipatory/predictive

Since Jay always runs a mile and _ ___ ___ ____ ___ __ ____ ____

incremental and anticipatory/predictive

<u>Since</u> <u>Jay</u> <u>always</u> <u>runs</u> <u>a</u> <u>mile</u> <u>and</u> <u>a</u> ___ ___ ____ ___ __ ___ ___

incremental and anticipatory/predictive

Since Jay always runs a mile and a half ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

<u>Since</u> <u>Jay</u> <u>always</u> <u>runs</u> <u>a</u> <u>mile</u> <u>and</u> <u>a</u> <u>half</u> <u>this</u> ＿＿＿ ＿＿ ＿＿ ＿＿ ＿＿

incremental and anticipatory/predictive

Since Jay always runs a mile and a half this seems ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half this seems like __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half this seems like an ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half this seems like an easy ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half this seems like an easy task.

incremental and anticipatory/predictive

incremental and anticipatory/predictive

<u>Since</u> ___ _____ ____ _ ____ ____ _ ____ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay _____ ____ _ ____ ____ _ ____ _____ ____ __ ____ ____

incremental and anticipatory/predictive

<u>Since</u> <u>Jay</u> <u>always</u> ____ _ ____ ____ _ ____ _____ ____ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs _ ____ ___ _ ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a ____ ____ _ ____ _____ ____ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile ___ _ ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and _ ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a ___ _____ ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half _____ ___ __ ____ ____

incremental and anticipatory/predictive

<u>Since</u> <u>Jay</u> <u>always</u> <u>runs</u> <u>a</u> <u>mile</u> <u>and</u> <u>a</u> <u>half</u> <u>seems</u> ___ __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half seems like __ ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half seems like an ____ ____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half seems like an easy _____

incremental and anticipatory/predictive

Since Jay always runs a mile and a half seems like an easy task.

led down the garden path

this is (hopefully!) an example of the garden path phenomenon

(which we'll study extensively in the syntactic processing unit)

more potential examples?

global versus temporary ambiguity

temporary ambiguity occurs in probably almost every utterance

yet the language processor is so good that we rarely get consciously
confused to the point that communication completely breaks down

what might garden path sentences suggest about
how we comprehend language?

global versus temporary ambiguity

temporary ambiguity occurs in probably almost every utterance

yet the language processor is so good that we rarely get consciously confused to the point that communication completely breaks down

what might garden path sentences suggest about **how** we comprehend language?

> incremental/anticipatory
> serial or parallel?!

communicates/interacts with other cognitive systems

language perception interacts with other cognitive systems in
certain ways, for example the visual system*

the McGurk Effect! [part1]

the McGurk Effect! [part2]

what do you think this says about how we process language?

*the degree to which language interacts w other cognitive systems is a matter of heated debate – interest??

automatic and unconscious

language comprehension is automatic and unconscious – we can't help but understand language around us, and we sometimes process language without even being aware of it!

the Stroop effect!

masked priming!

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

> spreading of lexical activation ("word association")
  • semantic priming
  • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The

> spreading of lexical activation ("word association")
> - semantic priming
> - form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness

> spreading of lexical activation ("word association")
>   • semantic priming
>   • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined

> spreading of lexical activation ("word association")
  • semantic priming
  • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined by

> spreading of lexical activation ("word association")
  - semantic priming
  - form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined by the

> spreading of lexical activation ("word association")
>   • semantic priming
>   • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined by the lawyer

> spreading of lexical activation ("word association")
>   • semantic priming
>   • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined by the lawyer turned

> spreading of lexical activation ("word association")
  • semantic priming
  • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined by the lawyer turned out

> spreading of lexical activation ("word association")
> * semantic priming
> * form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined by the lawyer turned out to

> spreading of lexical activation ("word association")
  - semantic priming
  - form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

The witness examined by the lawyer turned out to be

> spreading of lexical activation ("word association")
  - semantic priming
  - form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

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

> spreading of lexical activation ("word association")
  • semantic priming
  • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

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

The evidence examined by the lawyer turned out to be unreliable

> spreading of lexical activation ("word association")
  - semantic priming
  - form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

cues from non-linguistic knowledge

inferences about language come from many sources, including...

> encyclopedic knowledge

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

The evidence examined by the lawyer turned out to be unreliable

> spreading of lexical activation ("word association")
  • semantic priming
  • form-based priming
> statistical learning (e.g. co-occurrence frequencies)
> learning by analogy

yxx cxn xndxrstxnd whxt x xm wrxtxng xvxn xf x rxplxcx xll thx
vxwxls wxth xn 'x' – bt t gts hrdr f y dn't vn knw whr th vwls r.

recap

- language is a conventionalized yet fully productive system

- it has hierarchical, recursive structure

- there is ambiguity at every level of linguistic representation

- the language processor is well-suited to cope w ambiguity

next week

language acquisition!

- humans are not born knowing language

- how do humans come to know "Language" and languages?

- is the process of lang. acquisition driven by biology or society?