Lecture 11: Computational Psycholinguistics

Paula Buttery

Dept of Theoretical & Applied Linguistics, University of Cambridge

What is Psycholinguistics?

Psycholinguistics is concerned with:

- how we acquire, comprehend and produce language;
- understanding how language is stored and processed in the brain.

What is Psycholinguistics?

Psycholinguistics is concerned with:

- how we acquire, comprehend and produce language;
- understanding how language is stored and processed in the brain.

Example research questions in psycholinguistics: Morphology

How are words organised in the brain?

Full listing:

cat	$cat + N + Sing$
cats	$cat + N + PL$
hope	$hope + V$
hopes	$hope + V + 3P + Sing$
fox	$fox + N + Sing$
fox	$fox + V$
foxes	$fox + N + PL$
foxes	$fox + V + 3P + Sing$
foxed	$fox + V + PastPart$

vs.

Minimum redundancy:

cat	N
hope	V
fox	N
fox	V
$\hat{}s$	$+PL$
$\hat{}s$	$+3P + Sing$
$\hat{}ed$	$+PastPart$

Example research questions in psycholinguistics: Morphology

How are words organised in the brain?

Full listing:

cat	$cat + N + Sing$
cats	$cat + N + PL$
hope	$hope + V$
hopes	$hope + V + 3P + Sing$
fox	$fox + N + Sing$
fox	$fox + V$
foxes	$fox + N + PL$
foxes	$fox + V + 3P + Sing$
foxed	$fox + V + PastPart$

vs.

Minimum redundancy:

cat	N
hope	V
fox	N
fox	V
$\hat{}s$	$+PL$
$\hat{}s$	$+3P + Sing$
$\hat{}ed$	$+PastPart$

Example research questions in psycholinguistics: Syntax

Syntactic complexity: What makes a sentence difficult to process?

- The cat the dog licked ran away
- The cat the dog the rat chased licked ran away
- The fact that the employee who the manager hired stole office supplies worried the executive
  Complement clause then relative clause
- The executive who the fact that the employee stole office supplies worried hired the manager
  Relative clause then complement clause
  T. Gibson

Example research questions in psycholinguistics: Syntax

Syntactic complexity: What makes a sentence difficult to process?

- The cat the dog licked ran away
- The cat the dog the rat chased licked ran away
- The fact that the employee who the manager hired stole office supplies worried the executive
  Complement clause then relative clause
- The executive who the fact that the employee stole office supplies worried hired the manager
  Relative clause then complement clause
  T. Gibson

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution .

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was .

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in .

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in the .

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in the back .

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in the back of .

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the   .

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the book.

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the book.
- **Strong garden path effect:**
  The horse raced past the barn

Example research questions in psycholinguistics: Syntax

Parsing: How does the brain perform parse disambiguation?

- **Choosing from multiple parses:**
  He saw the boy with the telescope.
  (He saw (the boy with the telescope))
  vs.
  ((He saw the boy) with the telescope)
- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the book.
- **Strong garden path effect:**
  The horse raced past the barn fell.

Example research questions in Psycholinguistics: Semantics

In what manner is word meaning stored in the brain?

For concepts (nouns):

- categories theory (Lakoff)
  women fire and dangerous things
- decompositional feature based model:
  bird: +feathers +fly +beak
  what shall we use as features?
- prototype theory (canonical examples)
  bird: crow (rather than penguin)
- exemplar theory (multiple good examples)
  bird: {crow, parrot, sparrow}
- Semantic networks (is-a or has-a)
  bird: is-a animal, has-a feathers ...

Example research questions in Psycholinguistics: Semantics

In what manner is word meaning stored in the brain?

For concepts (nouns):

- categories theory (Lakoff)
  women fire and dangerous things
- decompositional feature based model:
  bird: +feathers +fly +beak
  what shall we use as features?
- prototype theory (canonical examples)
  bird: crow (rather than penguin)
- exemplar theory (multiple good examples)
  bird: {crow, parrot, sparrow}
- Semantic networks (is-a or has-a)
  bird: is-a animal, has-a feathers ...

Example research questions in Psycholinguistics: Semantics

In what manner is word meaning stored in the brain?

For concepts (nouns):

- categories theory (Lakoff)
  women fire and dangerous things
- decompositional feature based model:
  bird: +feathers +fly +beak
  what shall we use as features?
- prototype theory (canonical examples)
  bird: crow (rather than penguin)
- exemplar theory (multiple good examples)
  bird: {crow, parrot, sparrow}
- Semantic networks (is-a or has-a)
  bird: is-a animal, has-a feathers ...

Example research questions in Psycholinguistics: Semantics

In what manner is word meaning stored in the brain?

For concepts (nouns):

- categories theory (Lakoff)
  women fire and dangerous things
- decompositional feature based model:
  bird: +feathers +fly +beak
  what shall we use as features?
- prototype theory (canonical examples)
  bird: crow (rather than penguin)
- exemplar theory (multiple good examples)
  bird: {crow, parrot, sparrow}
- Semantic networks (is-a or has-a)
  bird: is-a animal, has-a feathers ...

Example research questions in Psycholinguistics: Semantics

In what manner is word meaning stored in the brain?

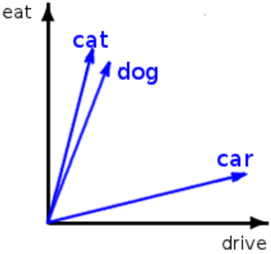
For concepts (nouns):

- categories theory (Lakoff)
  women fire and dangerous things
- decompositional feature based model:
  bird: +feathers +fly +beak
  what shall we use as features?
- prototype theory (canonical examples)
  bird: crow (rather than penguin)
- exemplar theory (multiple good examples)
  bird: {crow, parrot, sparrow}
- Semantic networks (is-a or has-a)
  bird: is-a animal, has-a feathers ...

In what manner is word meaning stored in the brain?

For other words:

- Distributional models (statistical co-occurrence of with other words and grammatical contexts)



Psycholinguists use a range of methodologies

Questionnaires

- Rating experiments
  e.g. how do you rate the grammaticality of this sentence?
- Self evaluations
  e.g. how were you carrying out the task?
- Discovering participant knowledge of the task

Evidence may be obtained from observation

Observations

- Study of speech errors
- Study of the language of aphasics
- Study of language acquisition

Speech Error Data

Evidence may be obtained from observation

Observations

- Study of speech errors
- Study of the language of aphasics
- Study of language acquisition

Speech Error Data

- It's not only us who have screw looses.

Evidence may be obtained from observation

Observations

- Study of speech errors
- Study of the language of aphasics
- Study of language acquisition

Speech Error Data

- It's not only us who have screw looses. (screws loose)

Evidence may be obtained from observation

Observations

- Study of speech errors
- Study of the language of aphasics
- Study of language acquisition

Speech Error Data

- It's not only us who have screw looses. (screws loose)
- He has already trunked two packs.

Evidence may be obtained from observation

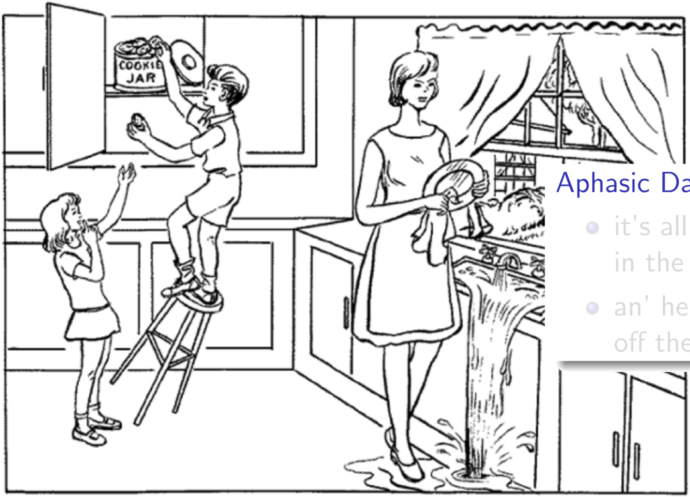
Observations

- Study of speech errors
- Study of the language of aphasics
- Study of language acquisition

Speech Error Data

- It's not only us who have screw looses. (screws loose)
- He has already trunked two packs. (packed two trunks)

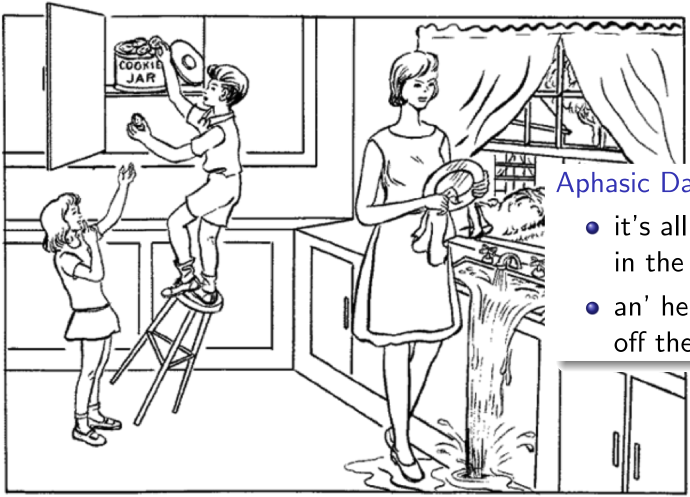
Evidence may be obtained from observation



Aphasic Data

- it's all happenining in the kitchen.
- an' he's fallenining off the stool

Evidence may be obtained from observation



Aphasic Data

- it's all happenining in the kitchen.
- an' he's fallenining off the stool

Evidence may be obtained from observation

Language Acquisition Data

- Brown's stages:

Stage 2	(2.0–2.5)	-s plurals
Stage 3	(2.5–3.0)	's possessive
Stage 4	(3.0–3.75)	regular past tense
Stage 5	(3.75–4.5)	irregular 3rd person verbs (dos$\rightarrow$ does, haves $\rightarrow$ has)

- show CHILDES example
- The wug test (elicitation)

Jean Berko-Gleason designed the Wug Test

- This is a man who knows how to **SPOW**.
- He is **SPOWING**. He did the same thing yesterday.
- Yesterday he .....

Jean Berko-Gleason designed the Wug Test

- This is a man who knows how to **SPOW**.
- He is **SPOWING**. He did the same thing yesterday.
- Yesterday he .....

Jean Berko-Gleason designed the Wug Test

- This is a man who knows how to **SPOW**.
- He is **SPOWING**. He did the same thing yesterday.
- Yesterday he **.....**

Psycholinguists use a range of methodologies cont.

Questionnaires

- Rating experiments
- Self evaluations

Observations

- Study of speech errors
- Study of the language of aphasics
- Study of language acquisition

Experimental observation as a response to stimulus

- Measurement of brain response
- Measurement of reading times
- Measurement of reaction time to a linguistic task

Psycholinguists use a range of methodologies cont.

Questionnaires

- Rating experiments
- Self evaluations

Observations

- Study of speech errors
- Study of the language of aphasics
- Study of language acquisition

Experimental observation as a response to stimulus

- Measurement of brain response
- Measurement of reading times
- Measurement of reaction time to a linguistic task

Brain responses may be measured by several methods

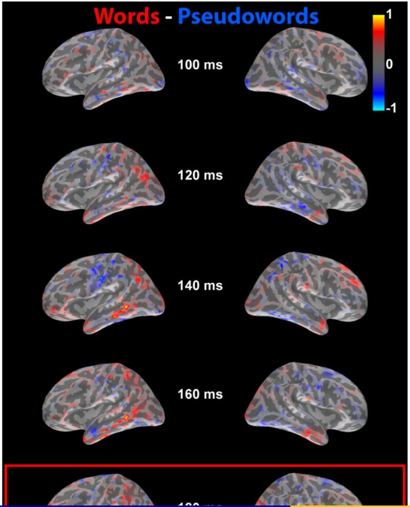
EEG

MEG

- High temporal resolution
- Problematic spatial resolution

Brain responses may be measured by several methods

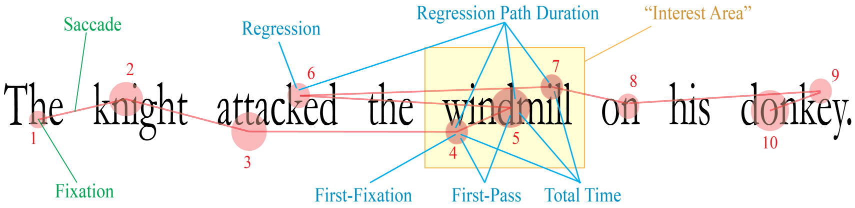
fMRI



- BOLD (Blood-oxygen-level dependent) response—measures the change in magnetization between oxygen-rich and oxygen-poor blood.
- High spatial resolution
- Low temporal resolution

Measuring reading and reacting times

Eye tracking



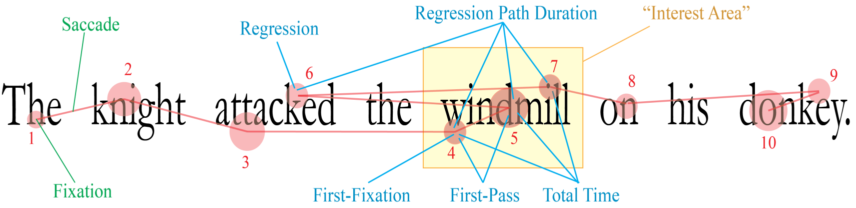
Button Pressing

- Self paced reading
- Completion of a task (e.g. a lexical decision task)

For all reaction time experiments we assume that the time taken to react to a task reflects the 'difficulty' of the cognitive processes involved.

Measuring reading and reacting times

Eye tracking



Button Pressing

- Self paced reading
- Completion of a task (e.g. a lexical decision task)

For all reaction time experiments we assume that the time taken to react to a task reflects the 'difficulty' of the cognitive processes involved.

Evaluating responses

Shadowing

- Participant repeats the stimulus
  e.g. lexical correction, Marslen-Wilson and Welsh
  e.g. missing auxiliaries, Caines

Experimental Strategies: Priming Effect

- Established that exposure to a stimulus influences a response to a later stimulus.
- Caused by spreading activation (the priming stimulus activates part(s) of the brain, then when the second stimulus is encountered less additional activation is needed).
- Priming manifests itself as a measurable change in reaction.
- **Lexical Priming**: e.g. Priming experiments show us that: *lifting* primes *lift*, *burned* primes *burn* but *selective* does not prime *select* (this maybe tells us something about derivational vs. inflectional morphology).
- **Syntactic Priming**: e.g. get candidate to read "the ghoul sold a vacuum cleaner to a witch"; then ask participant to describe picture of a vampire handing a ghost a hat; participant more likely to use the to-construction (i.e. 'the vampire hands a hat to the ghost')
- Also shown for Passive constructions—*Branigan, Pickering and Cleland*; Dialogue modelling—*Pickering and Garrod*

Experimental Strategies: Priming Effect

- Established that exposure to a stimulus influences a response to a later stimulus.
- Caused by spreading activation (the priming stimulus activates part(s) of the brain, then when the second stimulus is encountered less additional activation is needed).
- Priming manifests itself as a measurable change in reaction.
- **Lexical Priming**: e.g. Priming experiments show us that: *lifting* primes *lift*, *burned* primes *burn* but *selective* does not prime *select* (this maybe tells us something about derivational vs. inflectional morphology).
- **Syntactic Priming**: e.g. get candidate to read "the ghoul sold a vacuum cleaner to a witch"; then ask participant to describe picture of a vampire handing a ghost a hat; participant more likely to use the to-construction (i.e. 'the vampire hands a hat to the ghost')
- Also shown for Passive constructions—*Branigan, Pickering and Cleland*; Dialogue modelling—*Pickering and Garrod*

Experimental Strategies: Priming Effect

- Established that exposure to a stimulus influences a response to a later stimulus.
- Caused by spreading activation (the priming stimulus activates part(s) of the brain, then when the second stimulus is encountered less additional activation is needed).
- Priming manifests itself as a measurable change in reaction.
- **Lexical Priming**: e.g. Priming experiments show us that: *lifting* primes *lift*, *burned* primes *burn* but *selective* does not prime *select* (this maybe tells us something about derivational vs. inflectional morphology).
- **Syntactic Priming**: e.g. get candidate to read "the ghoul sold a vacuum cleaner to a witch"; then ask participant to describe picture of a vampire handing a ghost a hat; participant more likely to use the to-construction (i.e. 'the vampire hands a hat to the ghost')
- Also shown for Passive constructions—*Branigan, Pickering and Cleland*; Dialogue modelling—*Pickering and Garrod*

Experimental Strategies: Priming Effect

- Established that exposure to a stimulus influences a response to a later stimulus.
- Caused by spreading activation (the priming stimulus activates part(s) of the brain, then when the second stimulus is encountered less additional activation is needed).
- Priming manifests itself as a measurable change in reaction.
- **Lexical Priming**: e.g. Priming experiments show us that: *lifting* primes *lift*, *burned* primes *burn* but *selective* does not prime *select* (this maybe tells us something about derivational vs. inflectional morphology).
- **Syntactic Priming**: e.g. get candidate to read "the ghoul sold a vacuum cleaner to a witch"; then ask participant to describe picture of a vampire handing a ghost a hat; participant more likely to use the to-construction (i.e. 'the vampire hands a hat to the ghost')
- Also shown for Passive constructions—*Branigan, Pickering and Cleland*; Dialogue modelling—*Pickering and Garrod*

So what is a Computational Psycholinguist?

Computational Psycholinguists generate testable hypotheses by building computational models of language processes and also by drawing on information theory.

Note that Information Theoretic predictions are not always explanatory in terms of processing mechanisms e.g. Uniform Information Density—*Jaeger*

- the girl I saw last Saturday

  vs

- the girl that I saw last Saturday

So what is a Computational Psycholinguist?

Computational Psycholinguists generate testable hypotheses by building computational models of language processes and also by drawing on information theory.

Note that Information Theoretic predictions are not always explanatory in terms of processing mechanisms e.g. Uniform Information Density—*Jaeger*

- the girl I saw last Saturday

  vs

- the girl that I saw last Saturday

How are words organised in the in the brain?

Computational Psycholinguists predict reaction times to words based on various organisational models:

- **Morphological family size models** *Baayen*
  Ratio of lexical item to its morphological family size (a larger family co-activates and speeds reaction times).
- **Cohort and Lexical isolation point models** *Marslen-Wilson*
  Fast recognition of high frequency words with low frequency neighbours (recognition point vs. with uniqueness point)
- **Information Residuals** *Moscoso del Prado Martin*
  Showed response latencies in visual lexical decision based on the frequency of the word in a corpus and also the entropy of the morphological paradigm.

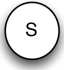
What makes a sentence difficult to process?

Parsing models as predictors for observed patterns in language–*Yngve*

- A sentence is constructed top-down and left-to-right.
- The model consists of a register for the current node being explored and a stack for all the nodes left to explore.
- The size of the stack an approximation to working memory load.
- Yngve predicted that sentences which required many items to be placed on the stack would be difficult to process and also less frequent in the language.
- He also predicted that when multiple parses are possible we should prefer the one with the minimised stack.

Parsing models as predictors for observed patterns in language–*Yngve*

S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased

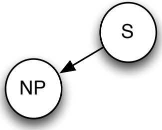


Register	Stack
S	

What makes a sentence difficult to process?

Parsing models as predictors for observed patterns in language–*Yngve*

S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased

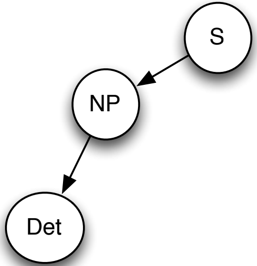


Register	Stack
S	
NP	VP

What makes a sentence difficult to process?

Parsing models as predictors for observed patterns in language–*Yngve*

S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased

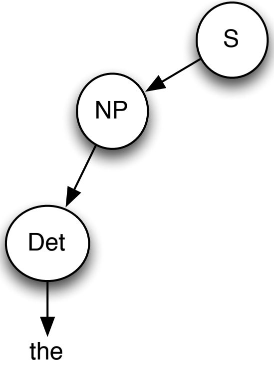


Register	Stack
S	
NP	VP
Det	N VP

What makes a sentence difficult to process?

Parsing models as predictors for observed patterns in language–*Yngve*

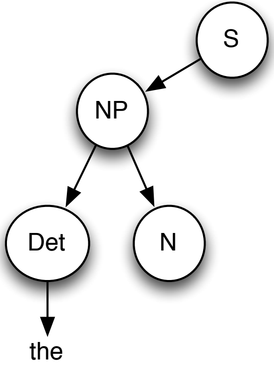
S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased



Register	Stack
S	
NP	VP
Det	N VP
the	N VP

Parsing models as predictors for observed patterns in language–*Yngve*

S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased

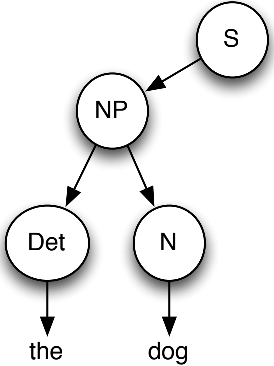


Register	Stack
S	
NP	VP
Det	N VP
the	N VP
N	VP

What makes a sentence difficult to process?

Parsing models as predictors for observed patterns in language–*Yngve*

S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased

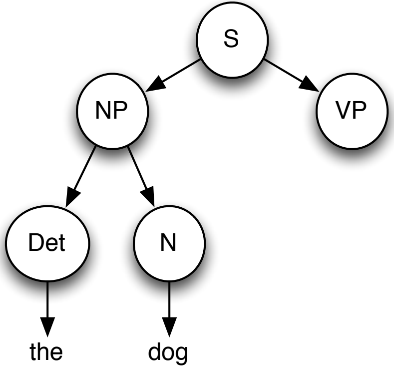


Register	Stack
S	
NP	VP
Det	N VP
the	N VP
N	VP
dog	VP

What makes a sentence difficult to process?

Parsing models as predictors for observed patterns in language–*Yngve*

S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased

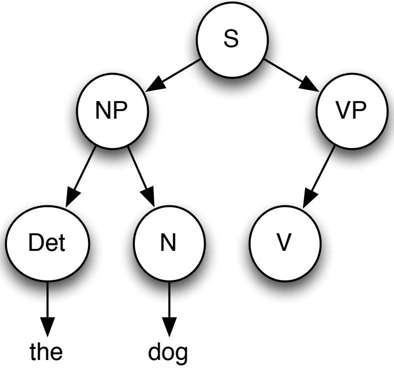


Register	Stack
S	
NP	VP
Det	N VP
the	N VP
N	VP
dog	VP
VP	

What makes a sentence difficult to process?

Parsing models as predictors for observed patterns in language–*Yngve*

S→NP VP
NP→Det N
VP→VP NP
Det→the
N→dog
N→cat
V→chased



Register	Stack
S	
NP	VP
Det	N VP
the	N VP
N	VP
dog	VP
VP	
V	NP

What makes a sentence difficult to process?

Yngve's make correct predictions about centre embedding

Consider:

This is the malt that the rat that the cat that the dog worried killed ate.

as opposed to:

This is the malt that was eaten by the rat that was killed by the cat that was worried by the dog.

What makes a sentence difficult to process?

Yngve's make correct predictions about centre embedding

Consider: STACK: N VP VP VP

> *This is the malt that the rat that the cat that the dog worried killed ate.*

as opposed to:

> *This is the malt that was eaten by the rat that was killed by the cat that was worried by the dog.*

Yngve evaluated his predictions by looking at frequencies of constructions in corpus data.

What makes a sentence difficult to process?

Dependency Locality Theory—*Gibson*

- Processing cost of integrating a new word is proportional to the distance between the word and the item with which the word is integrating.
- Distance is measured in words plus new phrases and discourse referents.

DLT will predict that object relative clauses are harder to process because they have two nouns that appear before any verb:

The girl who **likes** *me, went to the party.*
The girl who Peter **likes**, *went to the party.*

What makes a sentence difficult to process?

Dependency Locality Theory—*Gibson*

- Processing cost of integrating a new word is proportional to the distance between the word and the item with which the word is integrating.
- Distance is measured in words plus new phrases and discourse referents.

DLT can also explain:

The fact that the employee who the manager hired stole office supplies worried the executive
vs.
The executive who the fact that the employee stole office supplies worried hired the manager

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution .

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was .

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in .

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the .

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the back .

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the back of .

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the .

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the book.

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the book.
- **Strong garden path effect:**
  The horse raced past the barn

.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the book.
- **Strong garden path effect:**
  The horse raced past the barn fell.

How does the brain perform parse disambiguation?

Surprisal as a measure of lexical and syntactic complexity:

e.g.

$$S(w_i) = \log 1/P(w_i) \tag{1}$$

- **Online parsing ambiguity:**
  The student forgot the solution was in the back of the book.
- **Strong garden path effect:**
  The horse raced past the barn fell.

Surprisal as a predictor of reading times in sentence comprehension—*Levy*.

How does the brain perform parse disambiguation? Hale, 2001

Predictability as a measure of difficulty

- Probabilistic context-free grammars (PCFGs) are a good model of how human sentence comprehension works.
- A probabilistic Earley parser is a good model of online eager sentence comprehension for PCFGs.
- The cognitive effort associated with a word in a sentence can be measured by the word's negative log conditional probability:

$$\log \frac{1}{P(w_i | w_{1 \dots i-1})}$$

Hale, 2001

Example PCFG

S	$\rightarrow$	NP VP	1
NP	$\rightarrow$	N PP	0.2
NP	$\rightarrow$	N	0.8
PP	$\rightarrow$	P NP	1
VP	$\rightarrow$	VP PP	0.1
VP	$\rightarrow$	V VP	0.2
VP	$\rightarrow$	V NP	0.4
VP	$\rightarrow$	V	0.3
N	$\rightarrow$	{it, fish, rivers, December, they}	0.2
P	$\rightarrow$	{in}	1
V	$\rightarrow$	{can, fish}	0.5

How does the brain perform parse disambiguation? Hale, 2001

Probabilistic Early Parsing

- Prefix Probability: makes use of the fact that all the trees for the grammar must sum to 1.

How does the brain perform parse disambiguation? Hale, 2001

$edge_n$	DOTTED RULE	[start at, wait at]	HISTORY	word n
e_0	S → • NP VP	[0,0]		**word 0**
e_1	NP → • N	[0,0]		**word 1**
e_2	NP → • N PP	[0,0]		
e_3	N → they •	[0,1]		
e_4	NP → N •	[0,1]	(e_3)	
e_5	NP → N • PP	[0,1]	(e_3)	
e_6	S → NP • VP	[0,1]	(e_4)	
e_7	PP → • P NP	[1,1]		**word 2**
e_8	VP → • V	[1,1]		
e_9	VP → • V NP	[1,1]		
e_{10}	VP → • V VP	[1,1]		
e_{11}	VP → • VP PP	[1,1]		
e_{12}	V → can •	[1,2]		
e_{13}	VP → V •	[1,2]	(e_{12})	
e_{14}	VP → V • NP	[1,2]	(e_{12})	
e_{15}	VP → V • VP	[1,2]	(e_{12})	
e_{16}	S → NP VP •	[0,2]	(e_4, e_{13})	
e_{17}	VP → VP • PP	[1,2]	(e_{13})	
e_{18}	NP → • N	[2,2]		**word 3**
e_{19}	NP → • N PP	[2,2]		
e_{20}	VP → • V	[2,2]		
e_{21}	VP → • V NP	[2,2]		
e_{22}	VP → • V VP	[2,2]		
e_{23}	VP → • VP PP	[2,2]		
e_{24}	PP → • P NP	[2,2]		
e_{25}	N → fish •	[2,3]		
e_{26}	V → fish •	[2,3]		

S	→	NP VP
NP	→	N PP
NP	→	N
PP	→	P NP
VP	→	VP PP
VP	→	V VP
VP	→	V NP
VP	→	V
N	→	{they, fish, rivers, ...}
P	→	{in}
V	→	{can, fish}

$$\mathcal{N} = \{NP, VP, PP, N, V, P\}$$
$$\mathcal{N}_{PofS} = \{N, V, P\} \subset \mathcal{N}$$

• they • can • fish •

0 1 2 3

Hale, 2001

$edge_n$	DOTTED RULE	[S, W]	HIST	**Prob**	**MaxProb**
e_0	S → ● NP VP	[0,0]			P(S → NP VP)=1
e_1	NP → ● N	[0,0]			P(e_0)P(NP → N)=1*0.8=0.8
e_2	NP → ● N PP	[0,0]			P(e_0)P(NP → N PP)=1*0.2=0.2
e_3	N → they ●	[0,1]		P(N → they)=0.2	
e_4	NP → N ●	[0,1]	(e_3)	P(e_3)P(NP → N) =0.2*0.8 =0.16	
e_5	NP → N ● PP	[0,1]	(e_3)		
e_6	S → NP ● VP	[0,1]	(e_4)		
e_7	PP → ● P NP	[1,1]			P(N → they)P(e_2)P(PP → P NP) =0.2*1*0.2*1=0.04
e_8	VP → ● V	[1,1]			P(N → they)P(e_1)P(VP → V) =0.2*1*0.8*0.3=0.048
e_9	VP → ● V NP	[1,1]			P(N → they)P(e_1)P(VP → V NP) =0.2*1*0.8*0.4=0.048
e_{10}	VP → ● V VP	[1,1]			P(N → they)P(e_1)P(VP → V VP) =0.2*1*0.8*0.2=0.032
e_{11}	VP → ● VP PP	[1,1]			P(N → they)P(e_1)P(VP → VP PP) =0.2*1*0.8*0.1=0.0016
e_{12}	V → can ●	[1,2]		P(V → can)=0.5	
e_{13}	VP → V ●	[1,2]	(e_{12})	P(e_{12})P(VP → V) =0.5*0.3 =0.15	
e_{14}	VP → V ● NP	[1,2]	(e_{12})		
e_{15}	VP → V ● VP	[1,2]	(e_{12})		
e_{16}	S → NP VP ●	[0,2]	(e_4,e_{13})	P(e_4)P(e_{13})P(S → NP VP) =0.2*0.8*0.5*0.3*1 =0.024	
e_{17}	VP → VP ● PP	[1,2]	(e_{13})		

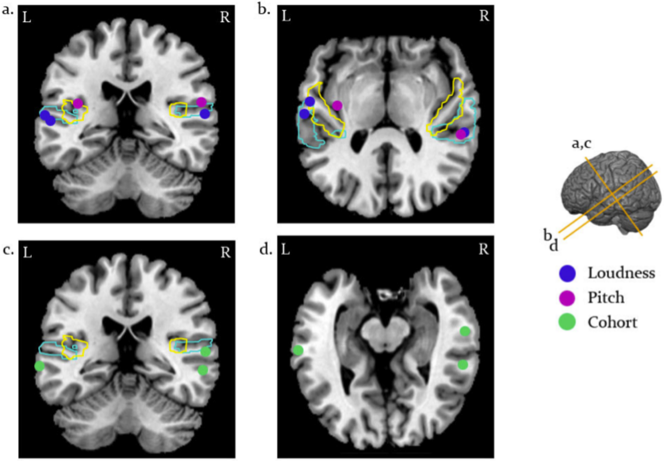
What makes a sentence difficult to process?

Predicting brain response from a probabilistic parser:

> *We used functional Magnetic Resonance Imaging (fMRI) to monitor brain activation while subjects passively listen to short narratives. The texts were written so as to introduce various syntactic complexities (relative clauses, embedded questions, etc.) not usually found (in such density) in actual corpora. With the use of a computationally implemented probabilistic parser (taken to represent an ideal listener) we have calculated a number of temporally dense (one per word) parametric measures reflecting different aspects of the incremental processing of each sentence. We used the resulting measures to model the observed brain activity (BOLD). We were able to identify different brain networks that support incremental linguistic processing and characterize their particular function.*

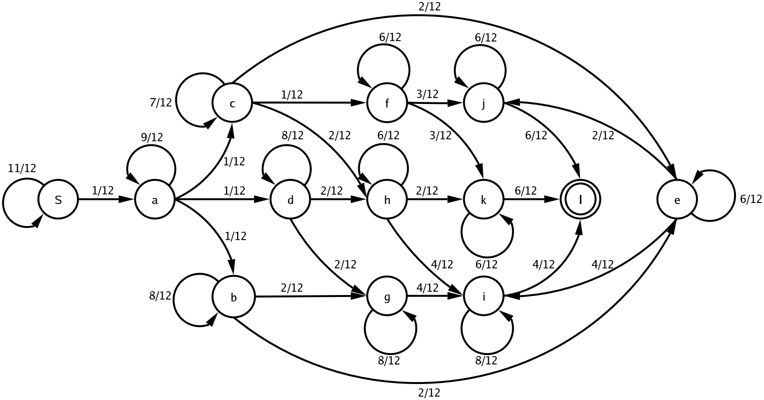
Asaf Bachrach

What makes a sentence difficult to process?



'Identifying computable functions and their spatiotemporal distribution in the human brain'—*Andrew Thwaites et al.*

Modelling acquisition of syntax as a process



state S-$\{\}$
state a-$\{s\backslash np\}$
state b-$\{s\backslash np, s/s\}$
state c-$\{s\backslash np, (s\backslash np)/np\}$

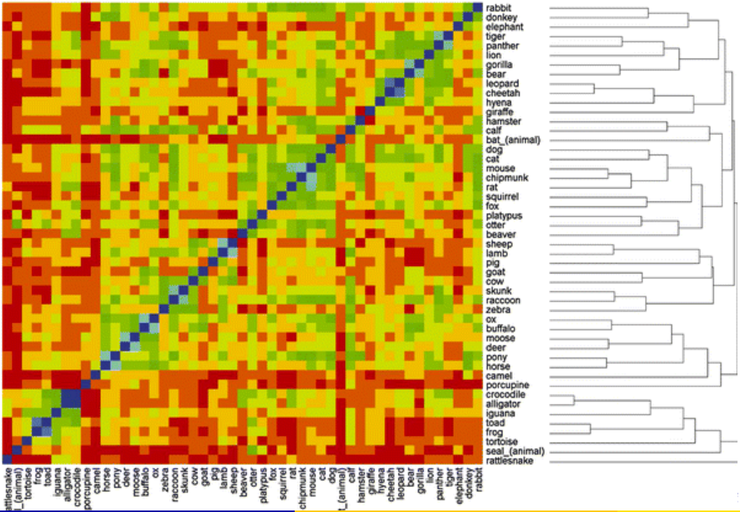
How is meaning represented in the brain?

Vector space models (VSMs) and semantic priming—*Pado and Lapata*

- take word pairs from the psychological literature
- compute vector representations for target words and related and unrelated prime words
- distance between related prime and target should be smaller than distance between unrelated prime and target.

How is meaning represented in the brain?

Towards Unrestricted, Large-Scale Acquisition of Feature-Based Conceptual Representations from Corpus Data—*Devereux et al.*



- Psycholinguistics is concerned with understanding how language is stored and processed in the brain.
- Computational Psycholinguistics contributes to the field by making predictions using information theory or computational models of language.
- These predictions are tested through observations or various experimental measurements.

To find out more:

- **Harley, T. (2001) The psychology of language from data to theory**.
- Berko, J. (1958) The Child's Learning of English Morphology.
- Marslen-Wilson, W. & Welsh, A. (1978) Processing interactions and lexical access during word recognition in continuous speech.
- Pickering, M.J., & Garrod, S. (2004). The interactive-alignment model
- Jaeger, T. (2010). Redundancy and Reduction: Speakers Manage Syntactic Information Density.
- de Jong, N., Schreuder, R. & Baayen, H. (2000) The morphological family size effect and morphology.
- Moscoso del Prado Martn, F., Kostic, A. & Baayen, H. (2004) Putting the bits together: An information theoretical perspective on morphological processing.

- Gibson, E. (2000) The Dependency Locality theory: A Distance-based theory of linguistic complexity.
- Levy, R (2008) Expectation-Based Syntactic Comprehension.
- Bachrach, A. (2008) Imaging Neural Correlates of Syntactic Complexity in Naturalistic Context.
- Pado, S. & Lapata, M. (2007) Dependency-based Construction of Semantic Space Models.
- Devereux, B. et al (2013) The Centre for Speech, Language and the Brain (CSLB) concept property norms.