Cornell University

Brain-imaging of *lexical cohesiveness*

Brain-imaging of lexical cohesiveness and sentence-structure building

and sentence-structure building

Murielle Fabre, Shohini Bhattasali, Hazem Al Saied, Mathieu Constant, Christophe Pallier, Wen-Ming Luh and John Hale

Département of Linguistics - Cornell University (NY)

The Department of **Linguistics**

Today's

Question & Approach	**How is sentence comprehension accomplished by the brain** computational models of sentence parsing + fMRI data
Teams	**Trans-Atlantic collaboration**
fMRI Expriment	**Naturalistic story listening of the entire Little Prince**
Results	**Investigate the neural bases of sentences-structure building and lexical cohesiveness**
Next steps	**Comparing French and English and focussing on other computational measures** e.g. modeling dependency-links in the sentence and morpho-syntactic information

The Trans-Atlantic team for this study

Wenming Luh

Shohini Bhattasali

Murielle Fabre

Cornell University

Mathieu Constant

Hazem Al-saied

ATILF UMR 7118
(CNRS/Université de Lorraine)

Christophe Pallier

Inserm

UNICOG
Cognitive Neuroimaging Lab

DeepMind

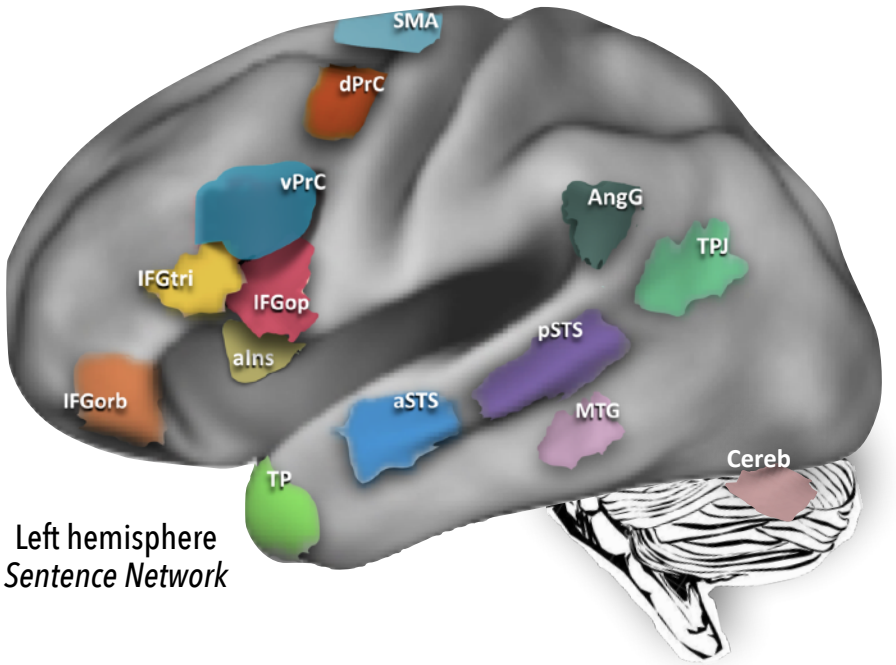
Cornell University

John Hale

Jonathan Brennan

Project & Approach

Bring together computational linguistics and cognitive neuro-imaging to shed light on sentence comprehension and its neural bases



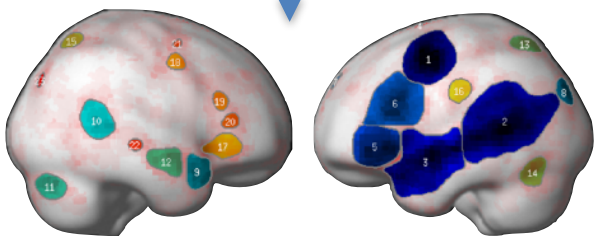
Left hemisphere
Sentence Network

Question

Multi-word expressions

Frequently co-occurring word sequences, known as Multiword Expressions (MWEs)

MWEs are a perfect testing ground to understand how expressions like ***break the ice, boa constrictor, see to it, in spite of*** are processed in the brain.



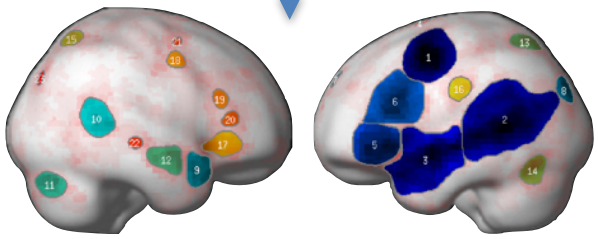
Do MWEs pattern together with sentence-structure building effects ?

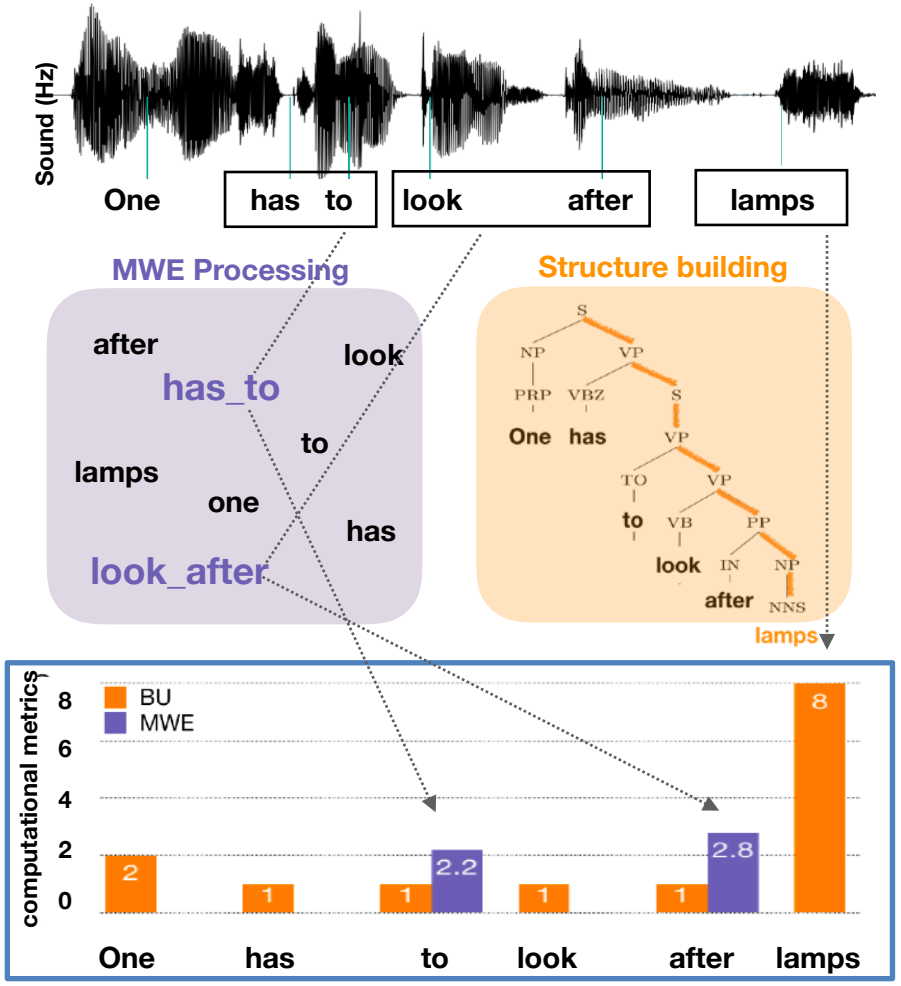
Question

Multi-word expressions

Frequently co-occurring word sequences, known as Multiword Expressions (MWEs)

MWEs are a perfect testing ground to understand how expressions like ***break the ice, boa constrictor, see to it, in spite of*** are processed in the brain.



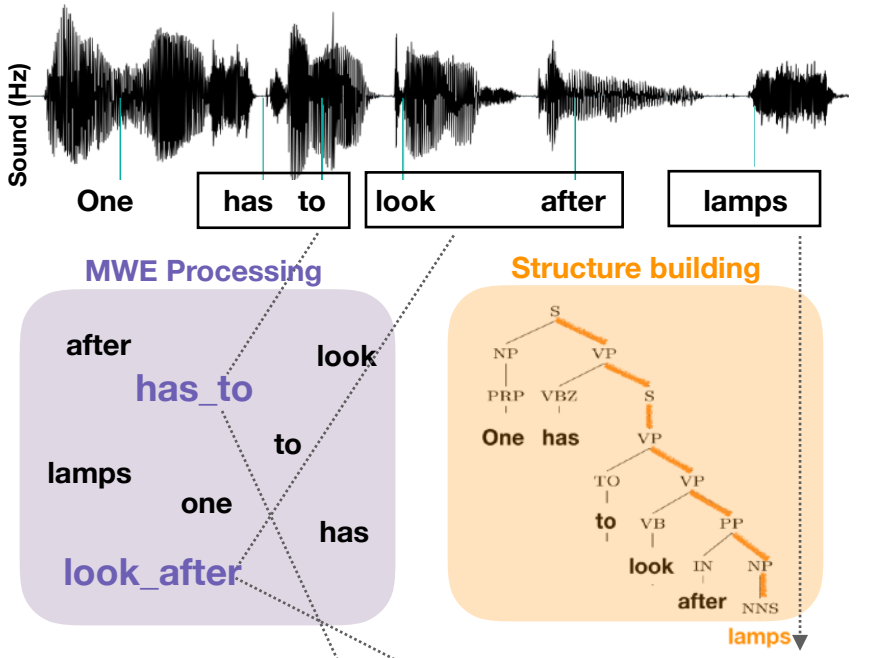


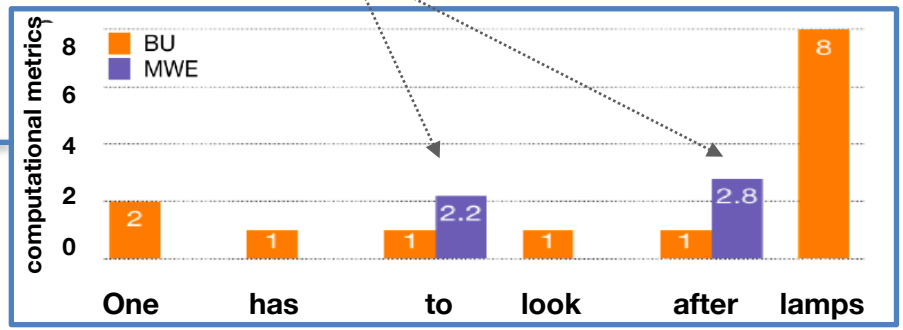
Question

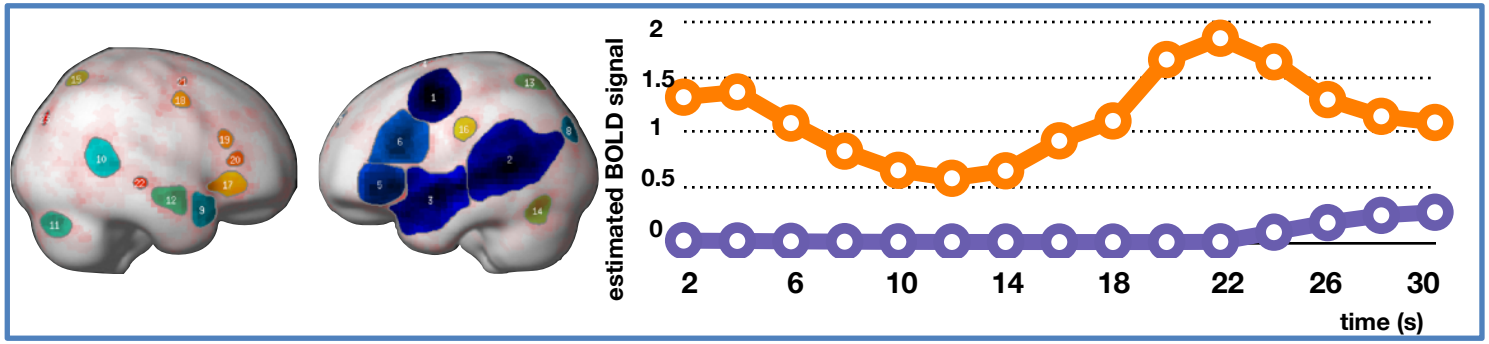
Multi-word expressions

Frequently co-occurring word sequences, known as Multiword Expressions (MWEs)

MWEs are a perfect testing ground to understand how expressions like ***break the ice, boa constrictor, see to it, in spite of*** are processed in the brain.







Naturalistic Corpus The Little Prince in 3 languages

Il y a six ans déjà que mon ami s'en est allé avec son mouton. Si j'essaie ici de le décrire, c'est afin de ne pas l'oublier. C'est triste d'oublier un ami.

French

Six years ago that my friend left wit his sheep. If I try to describe him, it's ignorer not to forget him. It is sad to forget a friend.

English

我的朋友找他的羊已离开六年了。
我在描述他，是了不忘他。
把朋友忘了是一件酸的事。

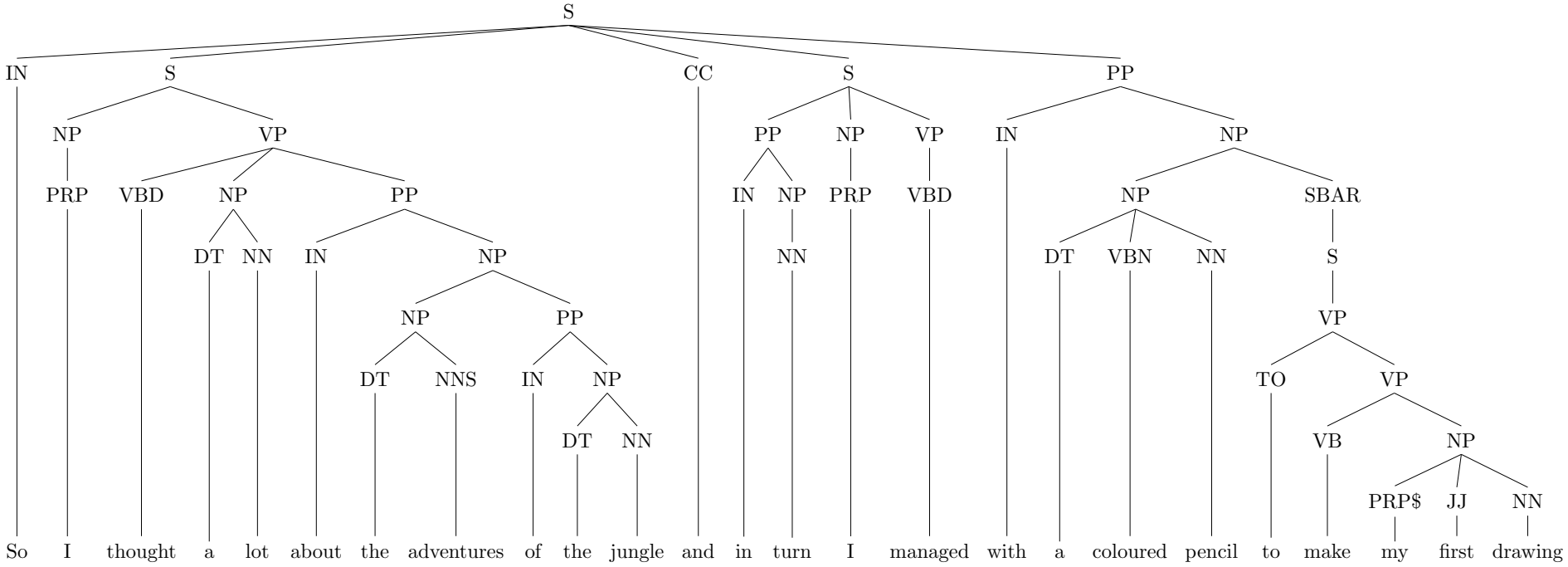
Chinese

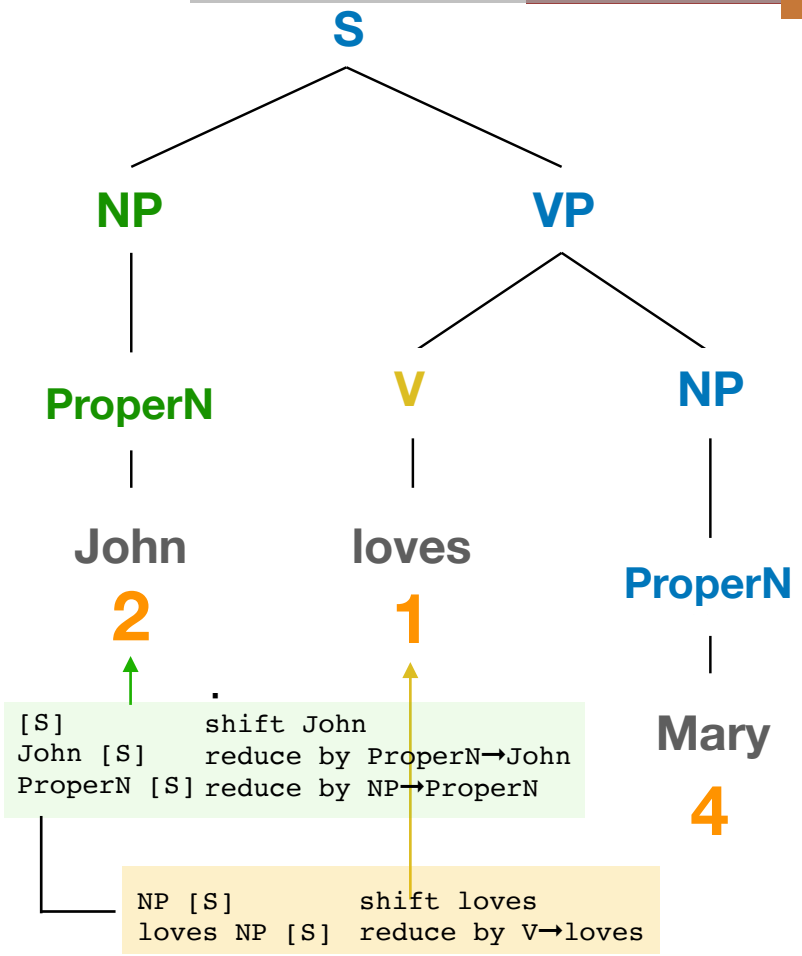
"Everyday listening" conditions

Participants (51):
American native speakers (32 women, 18-37 years old)
Task: Listen to the audiobook The Little Prince (1 h 38 min)+ Comprehension questions after each chapter
Recoding : 3T MRI scanner 32-channel head coil at the Cornell MRI Facility. Muti-echo sequence
Preprocessing : FSL, AFNI + the signal-to-noise ratio, using multi-echo independent components analysis (ME-ICA) (Kundu et al., 2013)
Analysis : General Linear Model (GLM - SPM12)
Control Regressors : pitch (f0), acoustic volume (RMS), Word rate, Word frequency.

Sentences

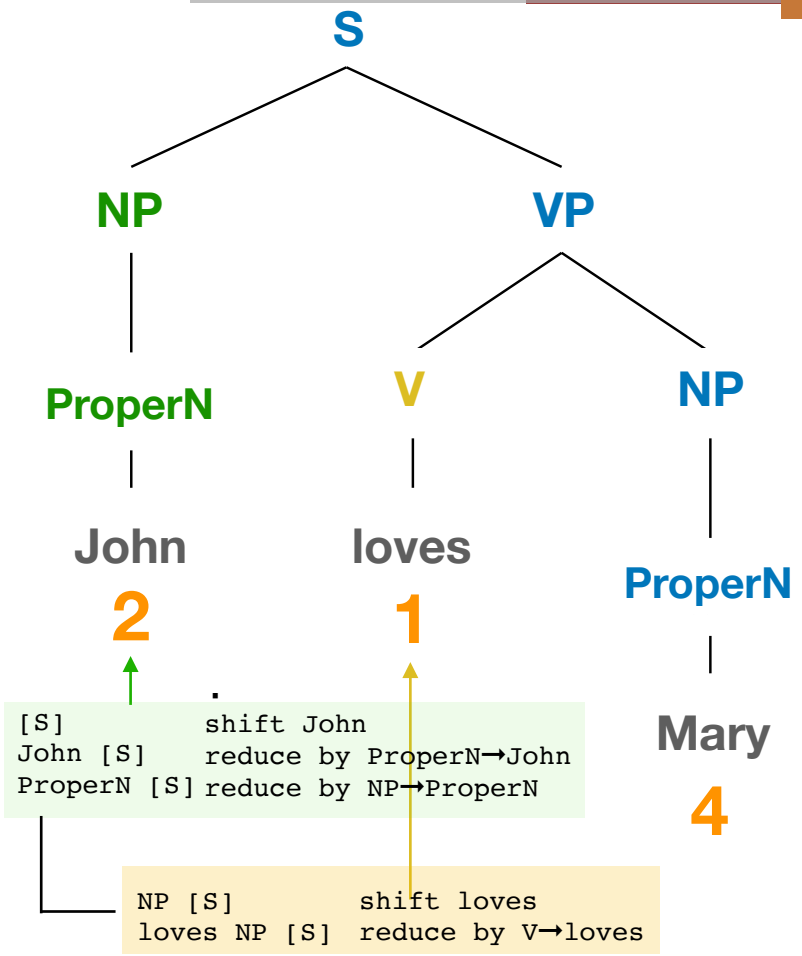
Sentence hierarchical representation





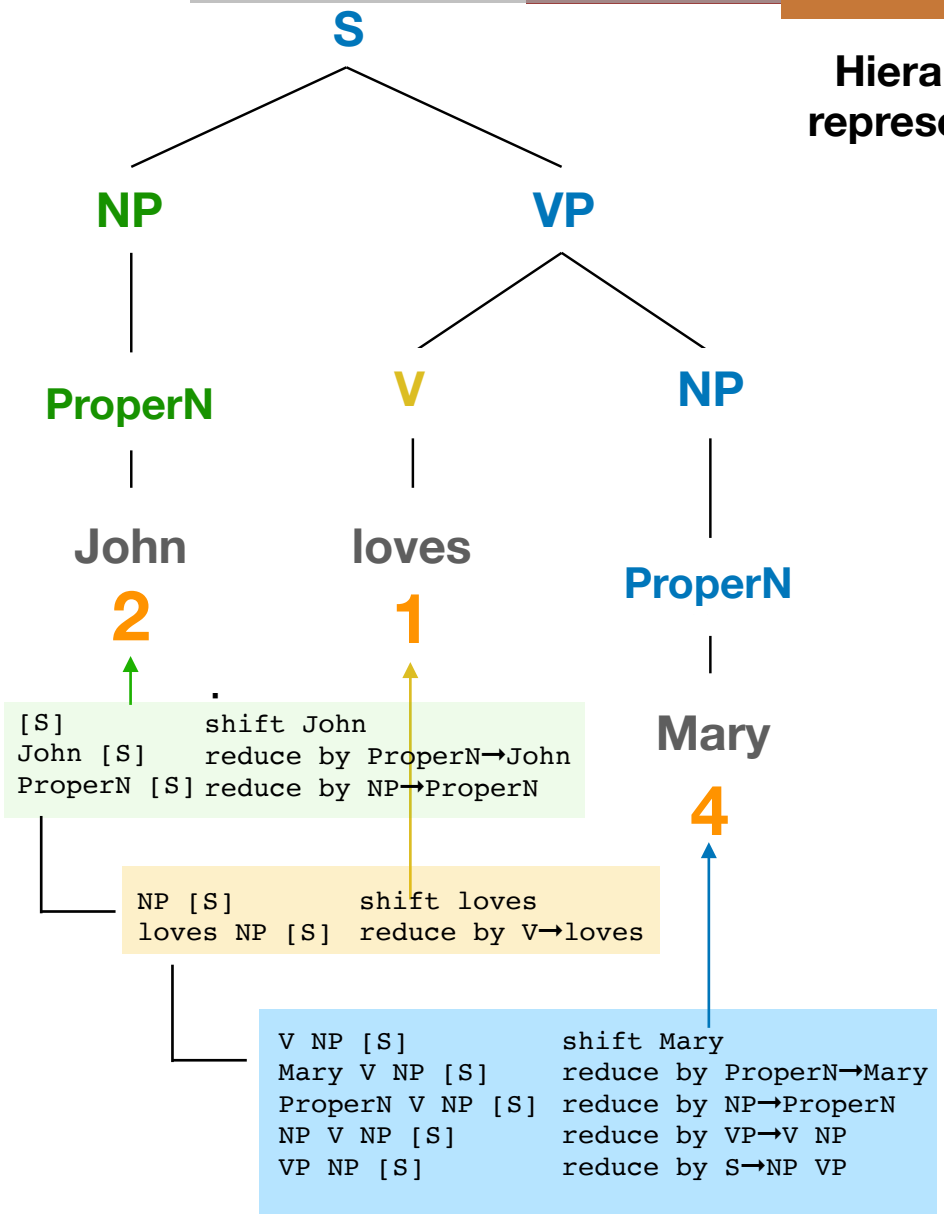
Hierarchical representation

Parser actions Bottom-Up



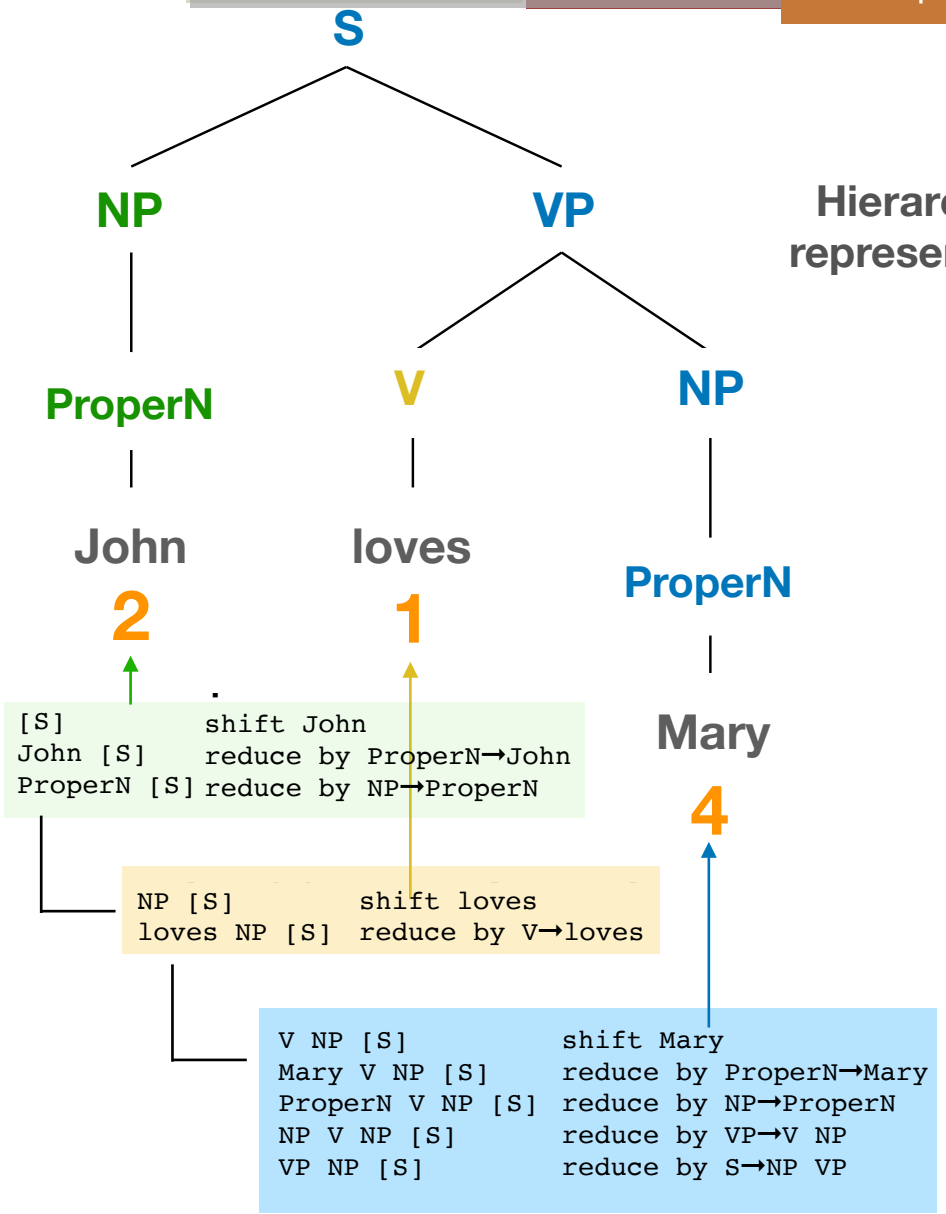
Hierarchical representation

Parser actions Bottom-Up



Hierarchical representation

Parser actions Bottom-Up



Parser actions Bottom-Up

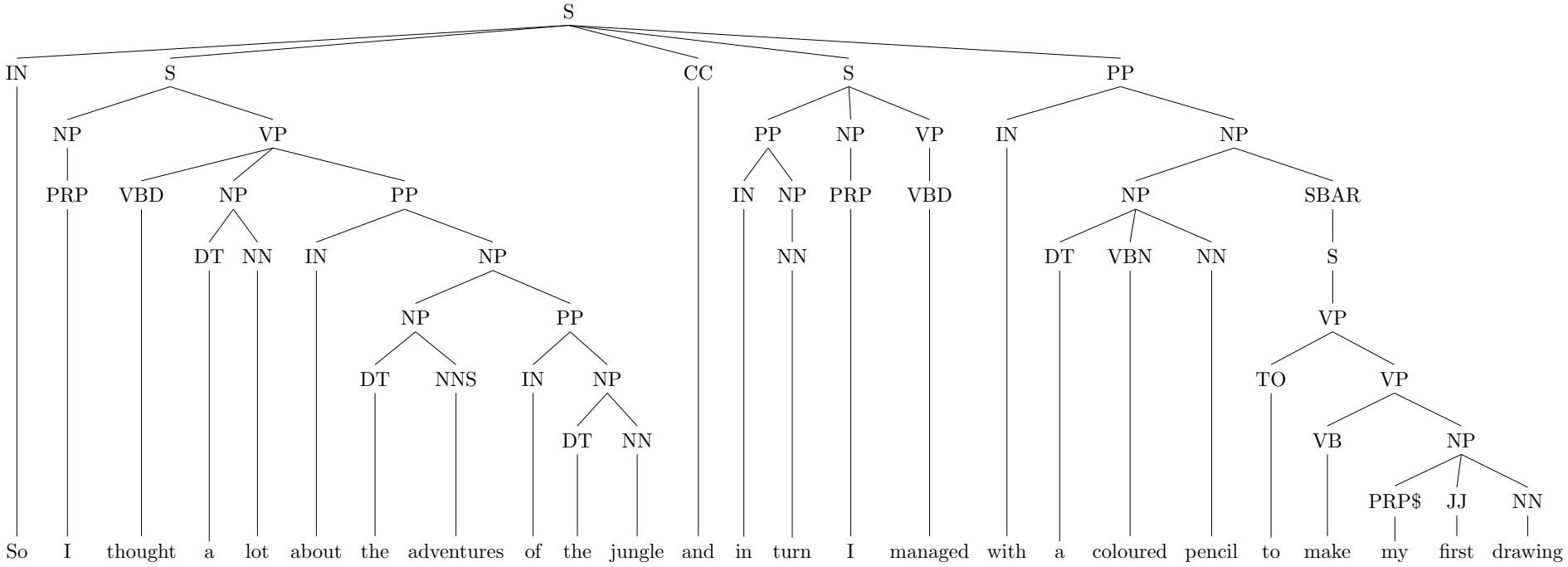
Hierarchical representation → **Parser actions**

Word-by-word measure of syntactic-structure building:

- **Can instantiate constituent-structure building the phrase/sentence.** as it builds and collects sub-parses towards the end of the phrase/sentence.
- **The rules of a grammar are applied at each incoming word**

Parser actions count

Sentence hierarchical representation + computational complexity metrics

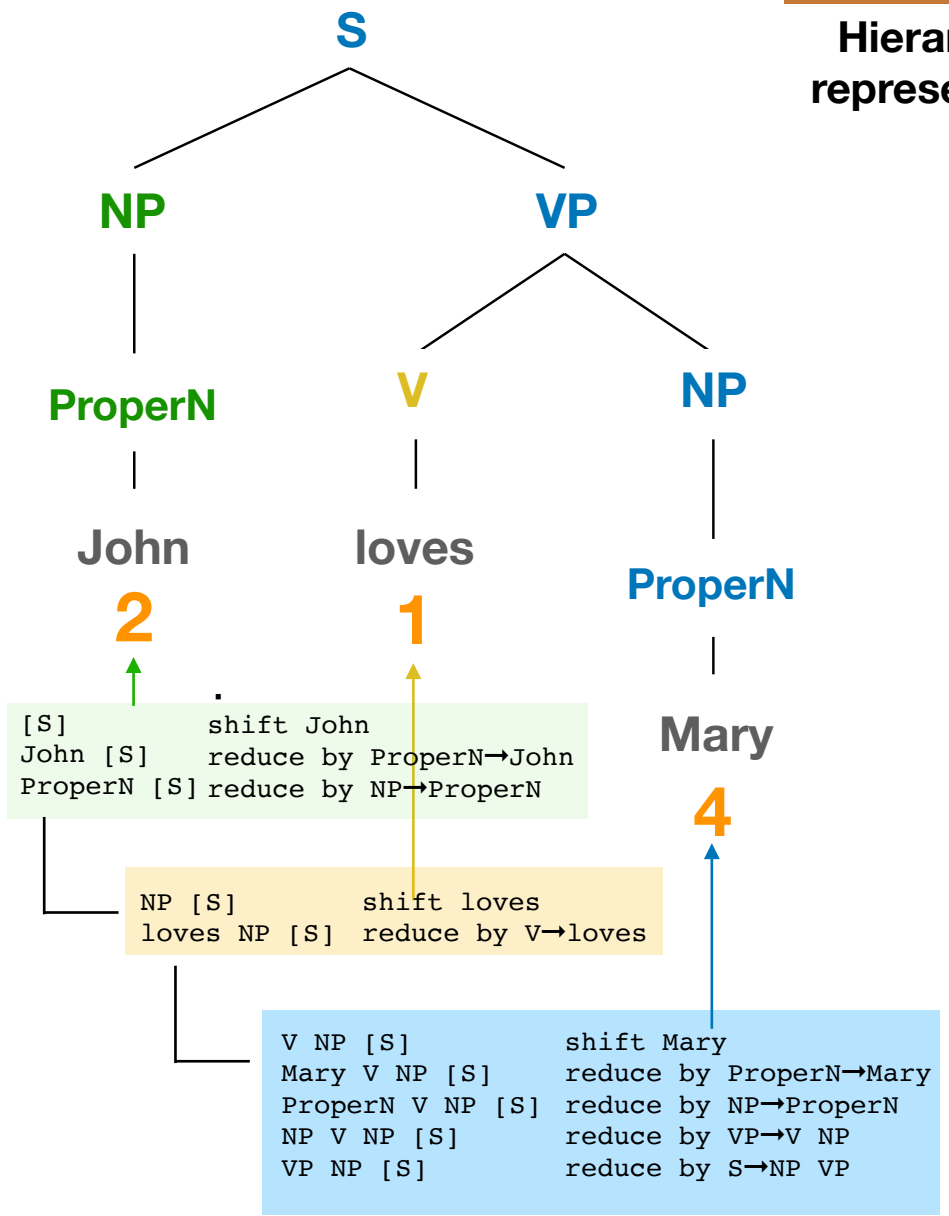


Bottom-up parser action count

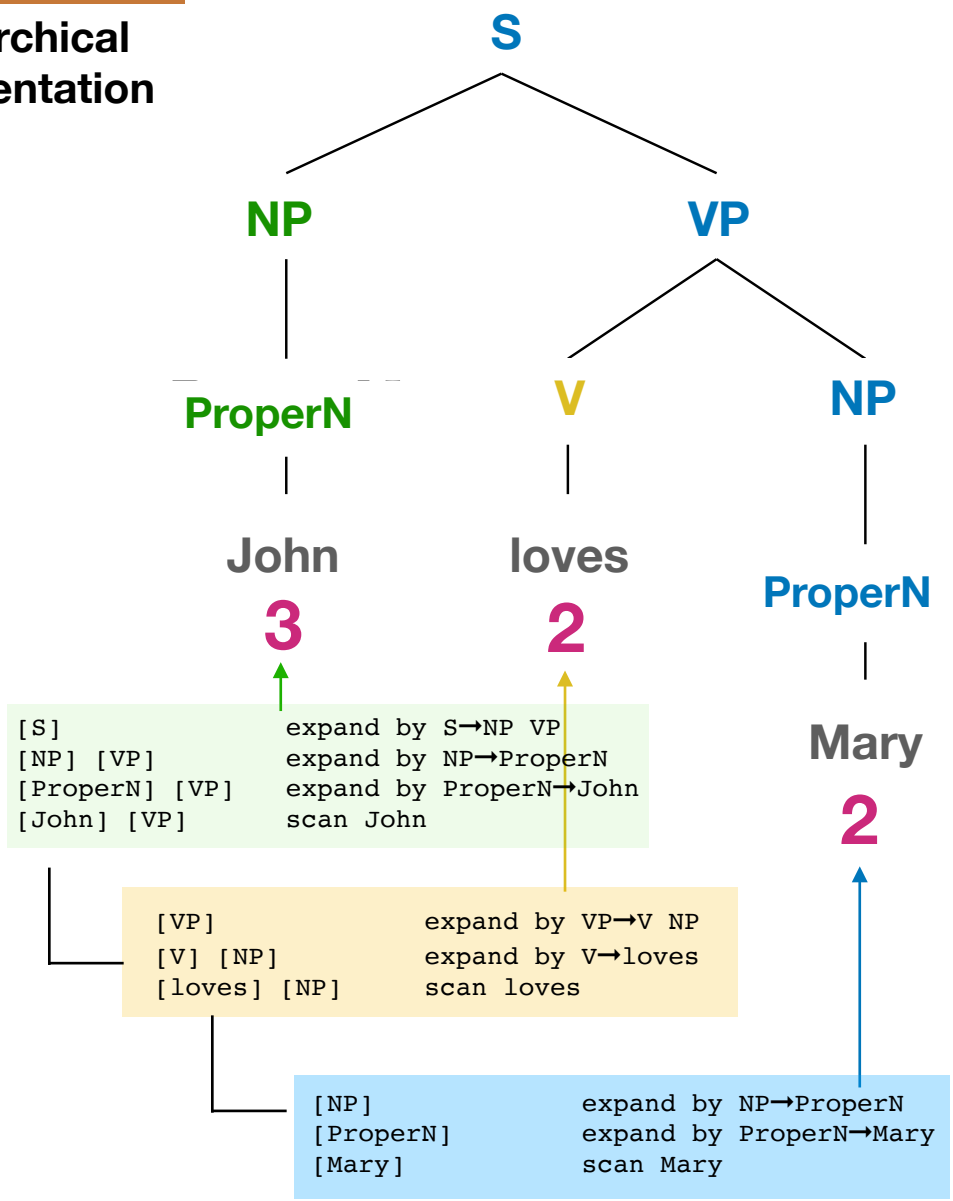
1	2	1	1	2	1	1	2	1	1	7	1	1	3	2	3	1	1	1	2	1	1	1	1	10

Number of REDUCE actions taken since last word

Hierarchical representation



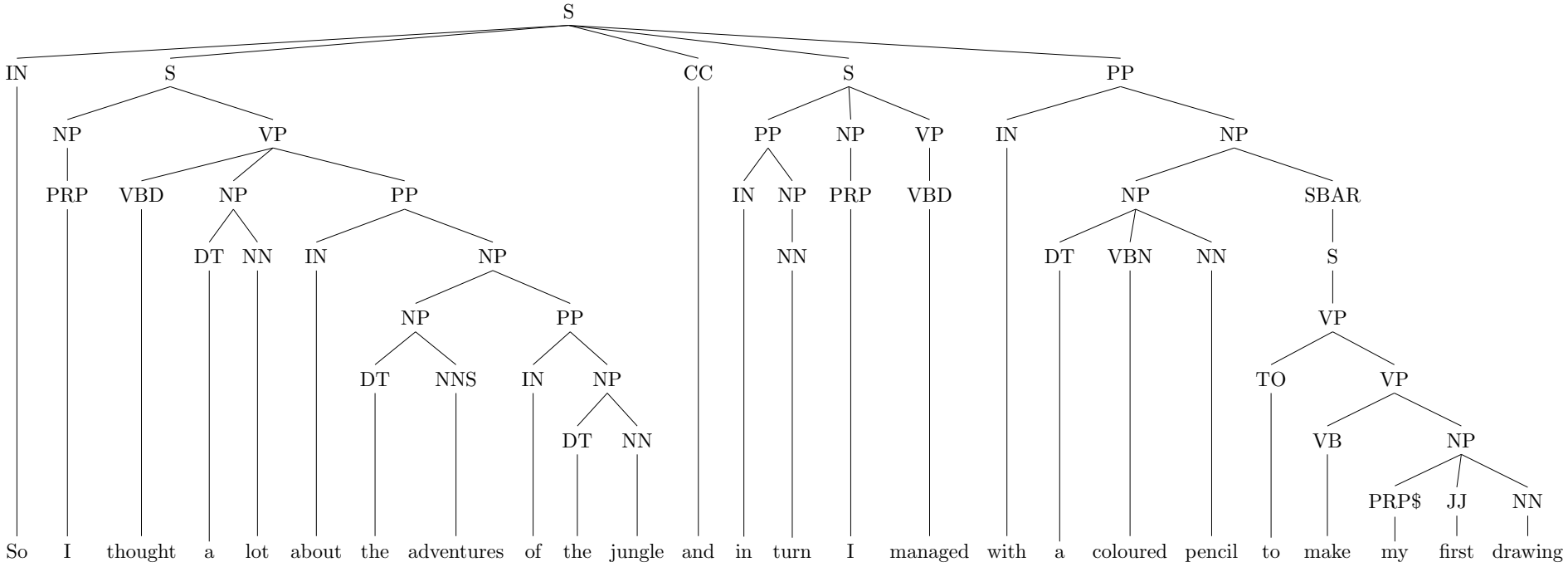
Parser actions Bottom-Up



Parser actions Top-Down

Parser actions count

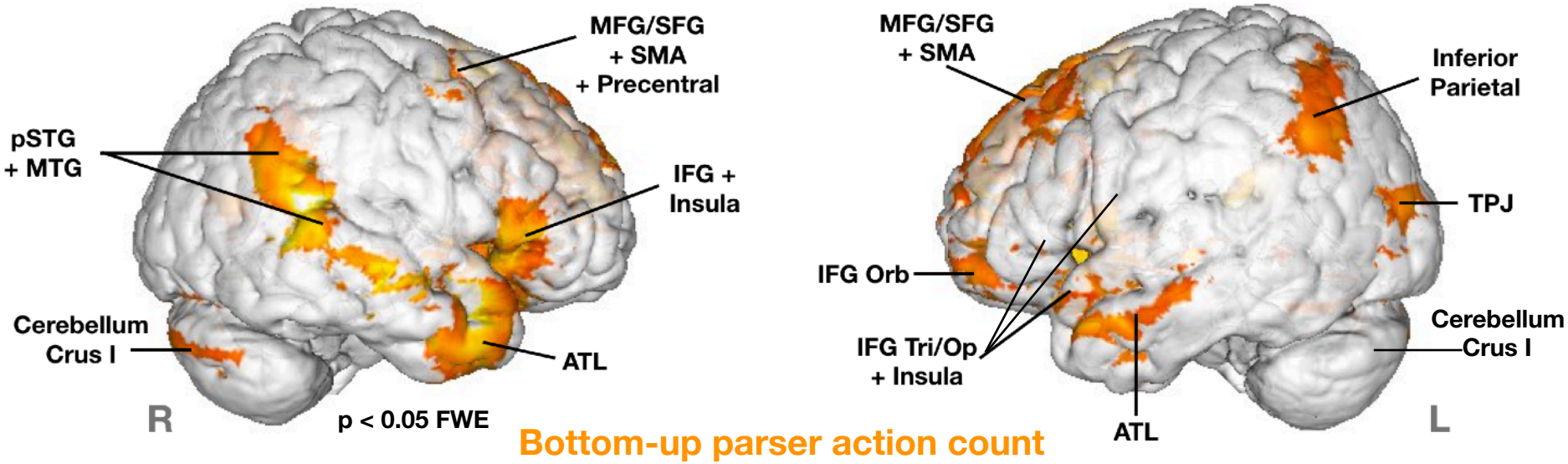
Sentence Hierarchical representation + computational complexity metrics



Bottom-up parser action count

1	2	1	1	2	1	1	2	1	1	7	1	1	3	2	3	1	1	1	2	1	1	1	1	10

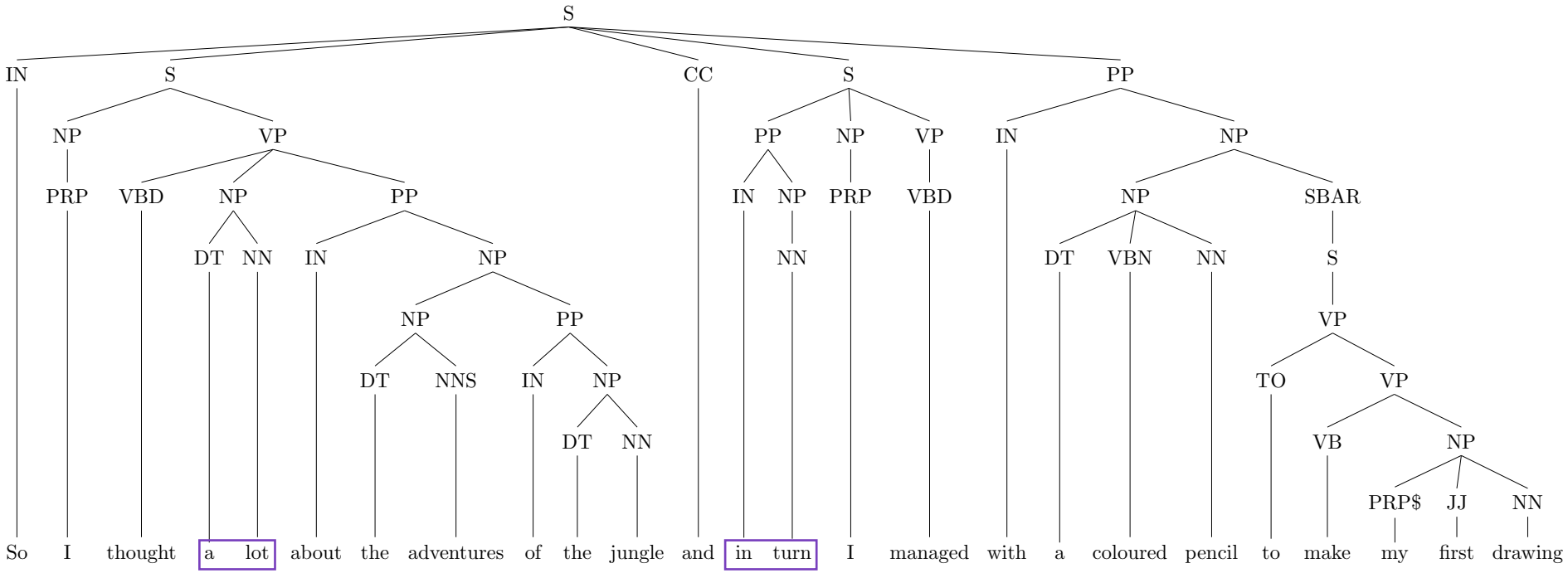
Bottom- up parser action count



Bottom-up parser action count

Bilateral network involving IFG and ATL

What are the processing units and what areas track syntactic-tree building



Bottom-up parser action count

1	2	1	1	2	1	1	2	1	1	7	1	1	3	2	3	1	1	1	2	1	1	1	1	10

One computational graded quantification identifies expressions likely to be processed as units, rather than built-up compositionally, the other tracks tree-building work needed in composed syntactic phrases

Computational toolkit ot identify MWEs

Computational toolkit ot identify MWEs

MWEs were identified using a statistical tagger (Al Saied et al. 2017). trained on Children's Book Test dataset.

You

must

see to it

that

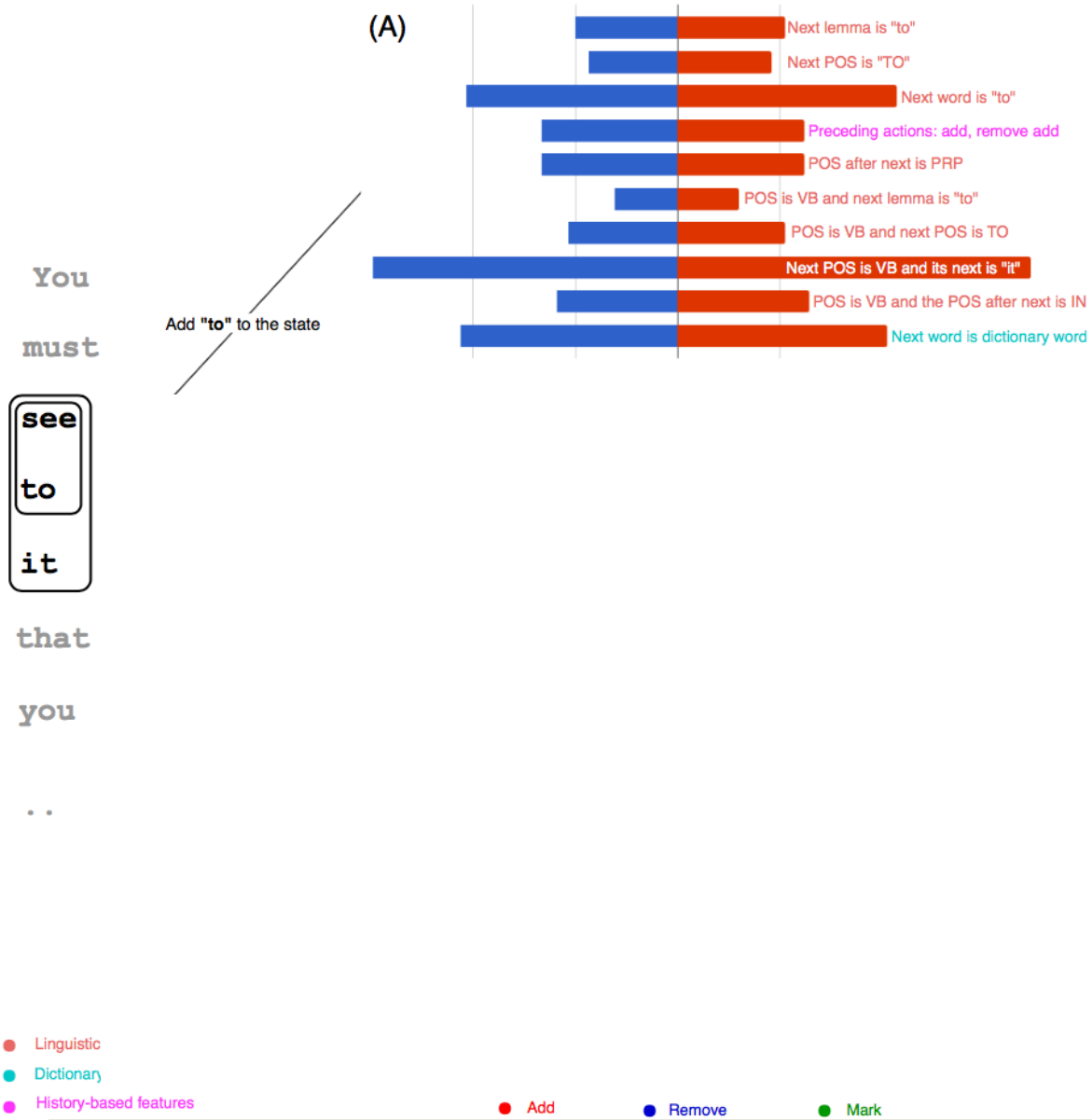
you

..

Mathieu Constant

Hazem Al-saied

Computational toolkit ot identify MWEs

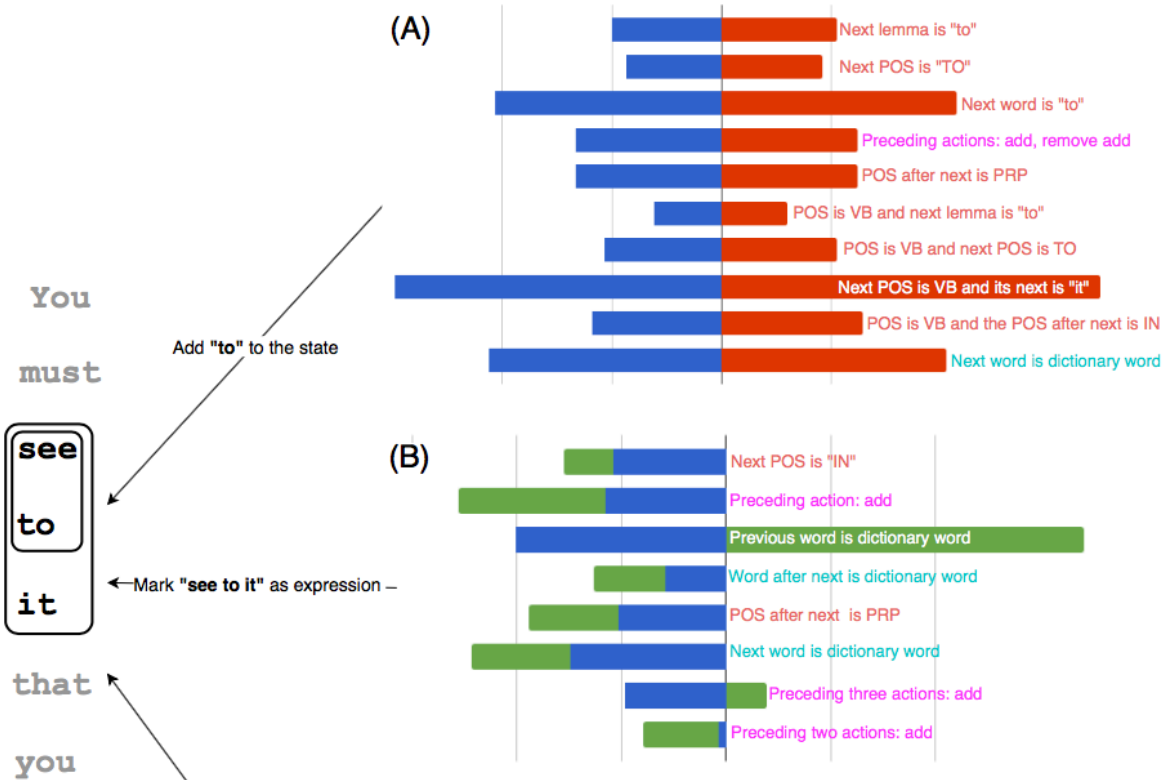


MWEs were identified using a statistical tagger (Al Saied et al. 2017). trained on Children's Book Test dataset.

Mathieu Constant

Hazem Al-saied

Computational toolkit ot identify MWEs

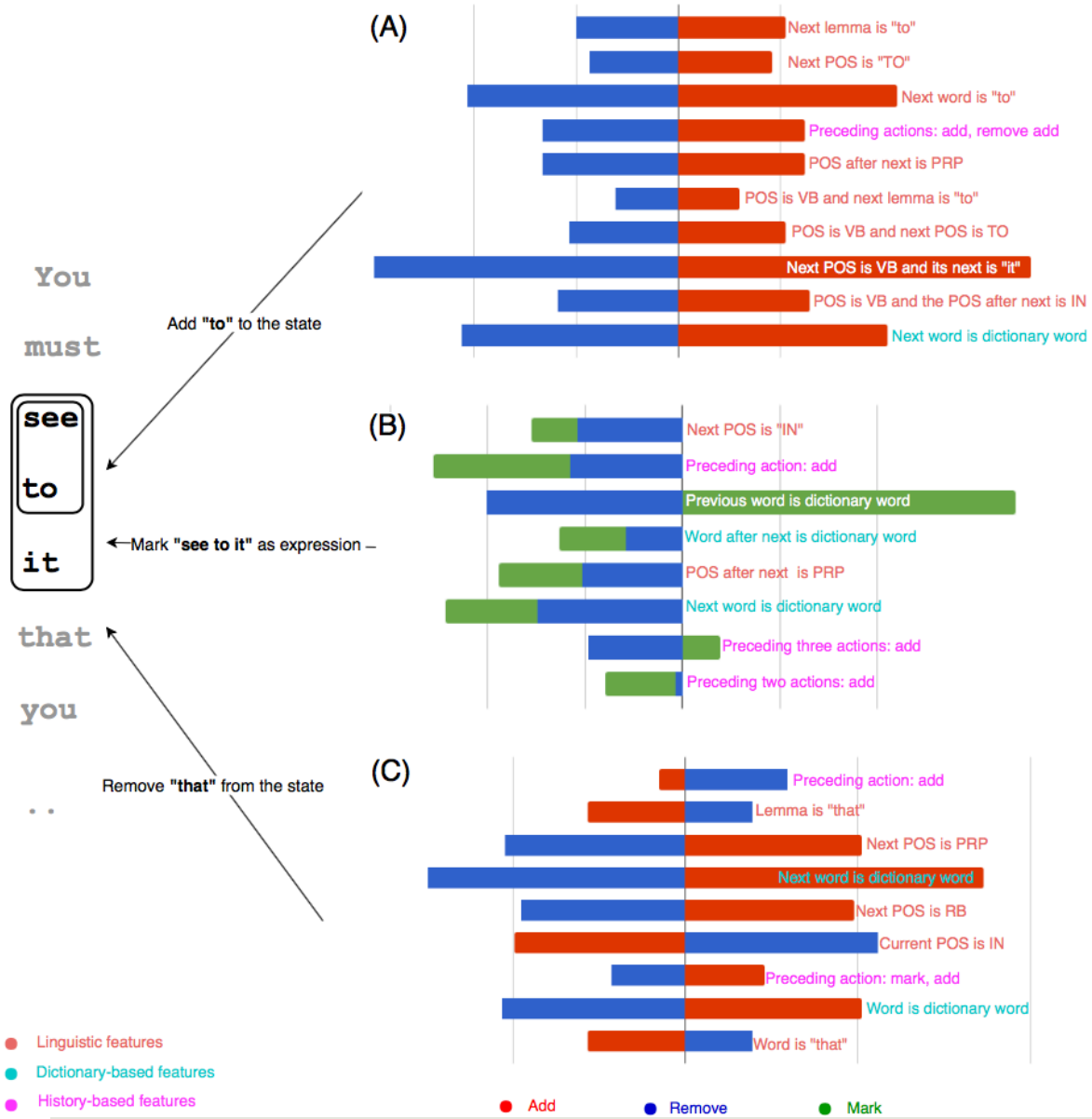


MWEs were identified using a statistical tagger (Al Saied et al. 2017). trained on Children's Book Test dataset.

Mathieu Constant

Hazem Al-saied

Computational toolkit ot identify MWEs

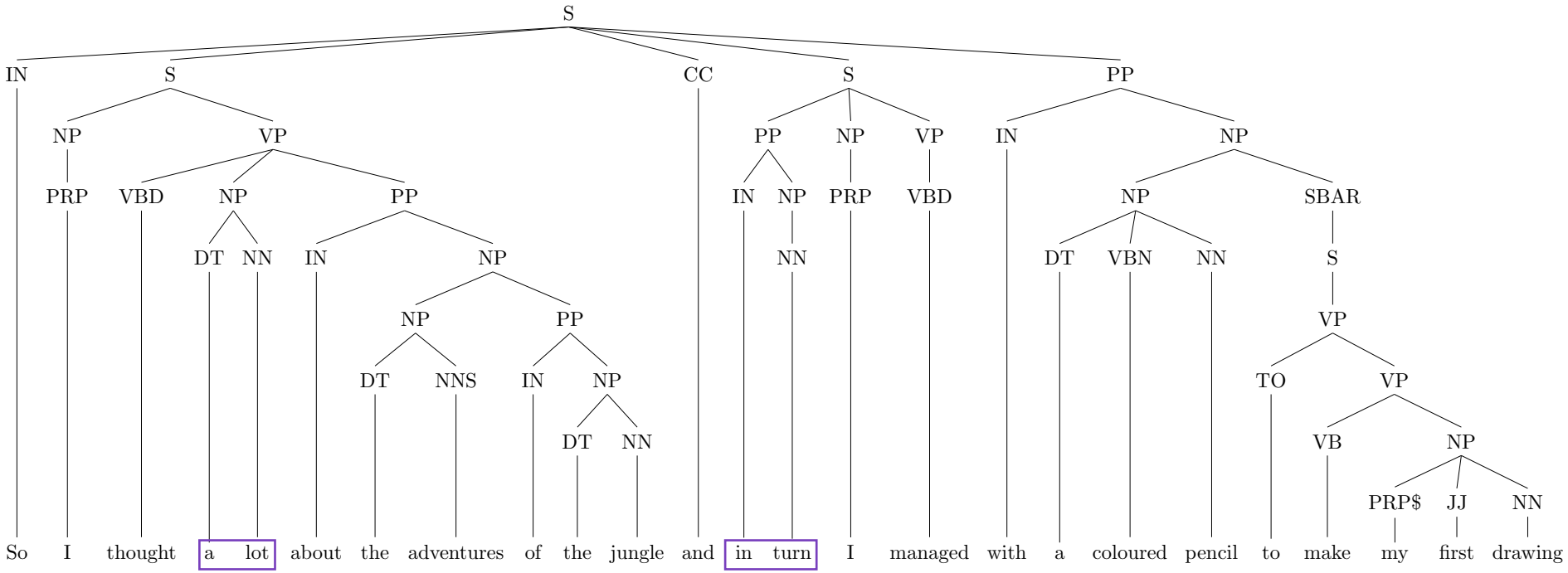


MWEs were identified using a statistical tagger (Al Saied et al. 2017). trained on Children's Book Test dataset.

Mathieu Constant

Hazem Al-saied

What are the processing units and what areas track syntactic-tree building



Bottom-up parser action count

1	2	1	1	2	1	1	2	1	1	7	1	1	3	2	3	1	1	1	2	1	1	1	1	10

One computational graded quantification identifies expressions likely to be processed as units, rather than built-up compositionally, the other tracks tree-building work needed in composed syntactic phrases

Association measure : Point-wise Mutual Information

PMI : A computational measure to link the degree of cohesiveness of MWEs

$$\text{PMI} = log_2\left(\frac{O}{E}\right)$$

$$O = \frac{count(\text{whole expression})}{\text{corpus size}}$$

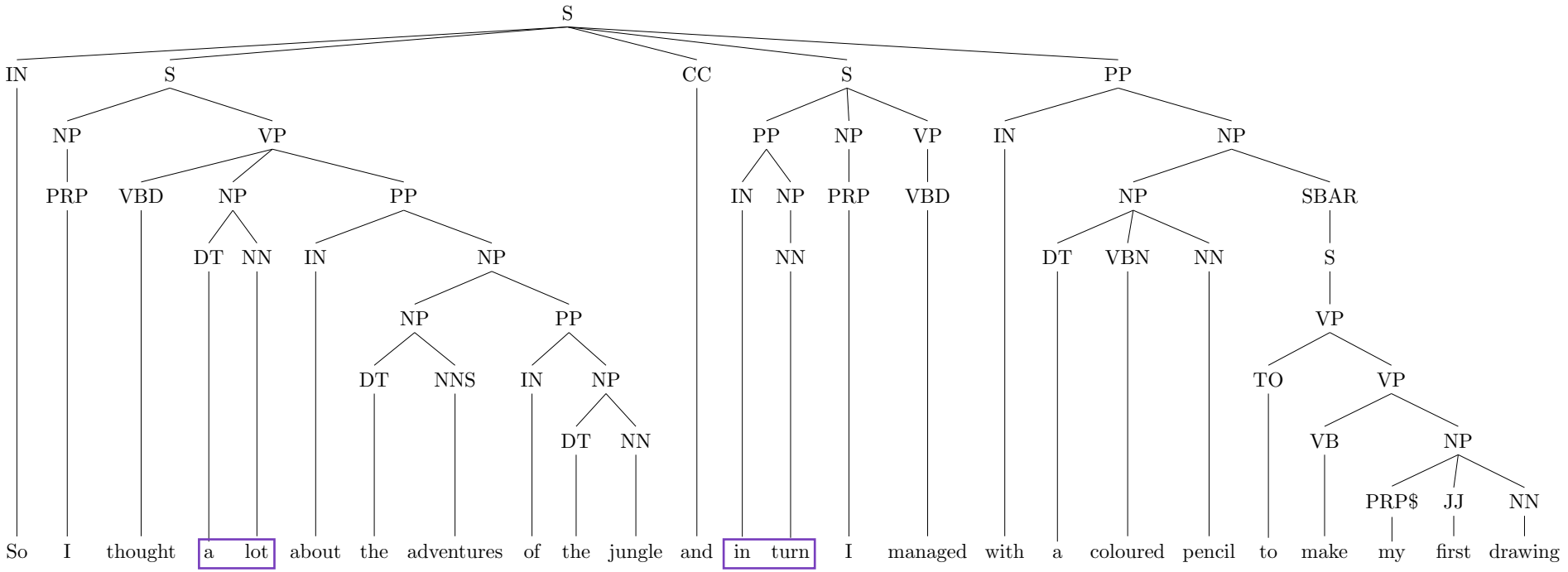
$$E = \frac{\text{count}(w_1) * \text{count}(w_2) * \cdots * \text{count}(w_n)}{\text{corpus size}^n}$$

PMI	multiword expression receiving this score
26.59474426	heart skipped a beat
23.79983038	have nothing to do with
21.25998782	forehead with a handkerchief
21.17721316	burst into tear
20.17480668	once upon a time
20.15121667	boa constrictor
18.85209561	peal of laughter
-2.336733827	be order
-2.493268369	do calculation
-2.721901963	be object
-2.982215241	be hundred
-3.152845604	a well
-3.501675488	drink anything
-3.635409951	have plan

This study investigates whether MWEs with different levels of cohesiveness evoke different patterns of activation in the brain using fMRI.

We ask if less cohesive MWEs pattern together with sentence-structure building effects observed by correlating brain activity with Bottom-up parser's action count.

What are the processing units and what areas track syntactic-tree building



Bottom-up parser action count

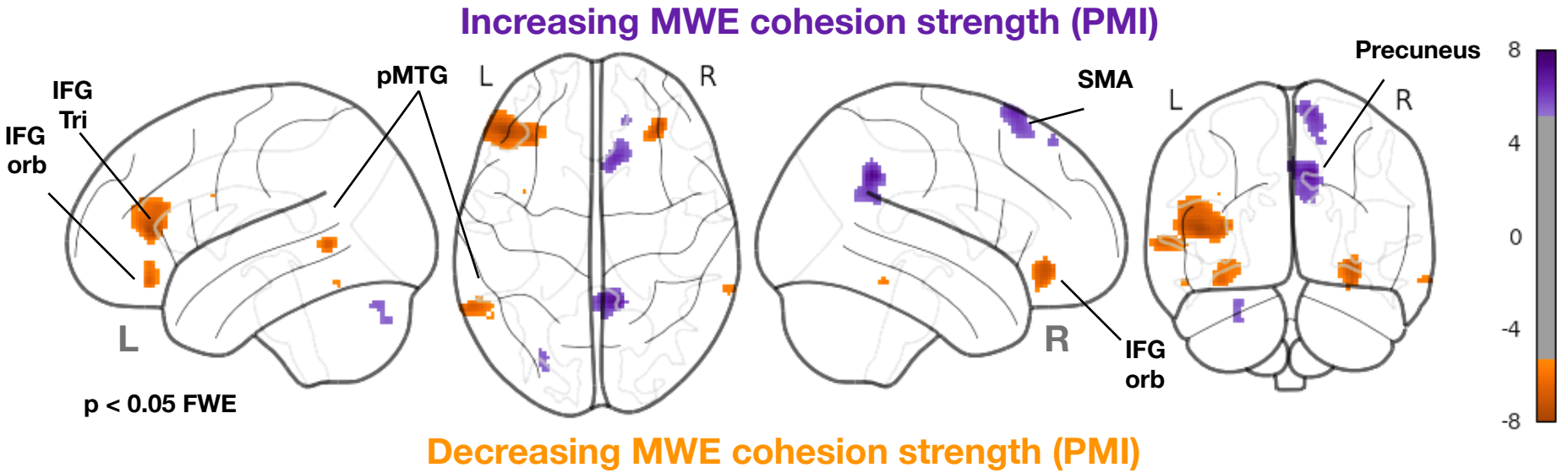
1	2	1	1	2	1	1	2	1	1	7	1	1	3	2	3	1	1	1	2	1	1	1	1	10

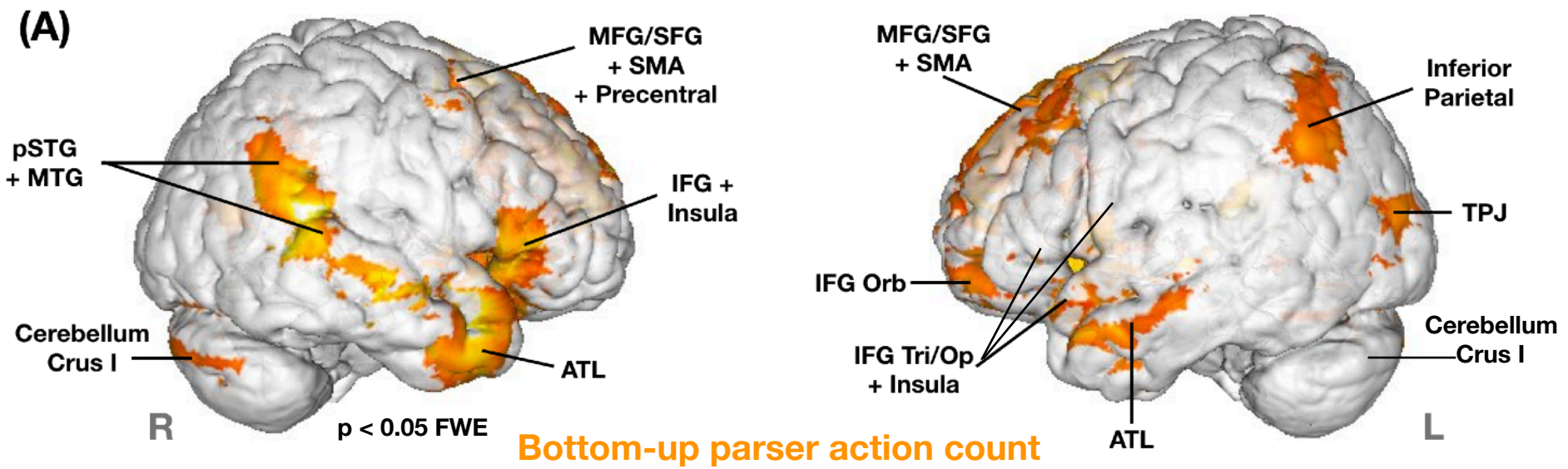
Cohesion strength MWEs (PMI)

0	0	0	0	2.62	0	0	0	0	0	0	0	0	4.083	0	0	0	0	0	0	0	0	0	0	0	0	0

One computational graded quantification identifies expressions likely to be processed as units, rather than built-up compositionally, the other tracks tree-building work needed in composed syntactic phrases

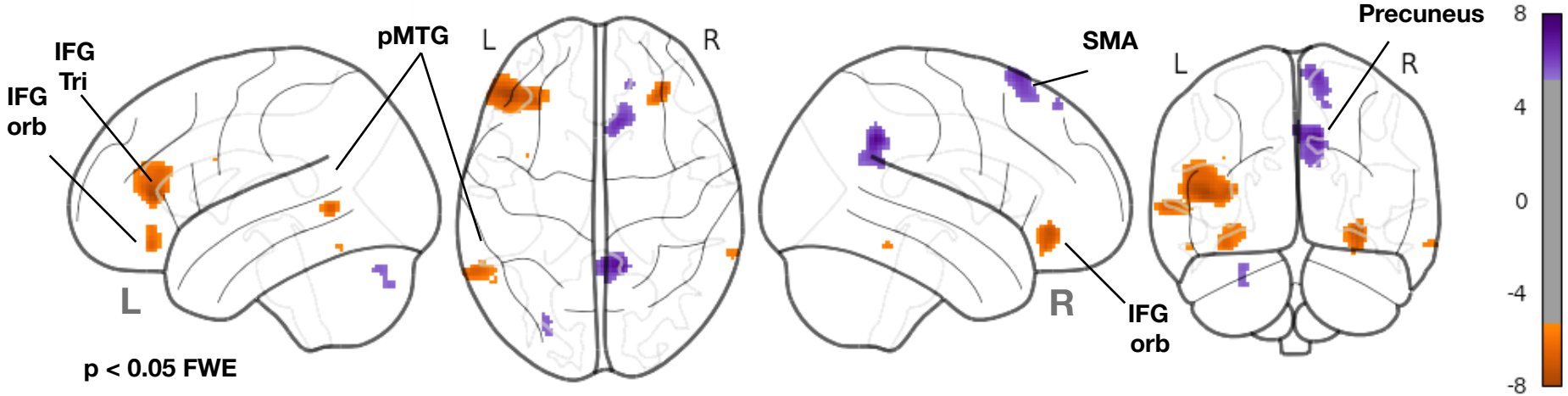
Positive and negative correlation with lexical cohesiveness





Bottom-up parser action count

Decreasing MWE cohesion strength (PMI)



Increasing MWE cohesion strength (PMI)

Summary

1 - PMI : word-by-word computational measure of lexical cohesiveness

—> cognitively plausible computational measure of the balance between compositionality and cohesiveness in MWEs

Our study is showing that this association measure in MWEs produces a neurocognitive effect during naturalistic story listening.

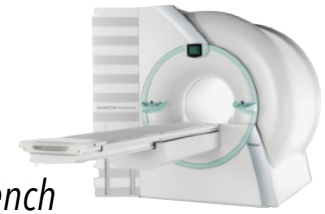
2 - The effect of less cohesive MWEs vs. phase-structure building effect (BU)

We observe an overlap between the significant effect for decreasing MWE cohesiveness and Bottom-up parser action count in left IFG and posterior temporal lobe.

The results show an overlap between the significant effect for decreasing MWE cohesiveness and Bottom-up parser action count in left IFG and posterior temporal lobe. Highly cohesive MWEs implicate the Precuneus and the SMA, suggesting that only truly lexicalized linguistic expressions rely on these areas rather than traditional frontal and temporal nodes of the language network.

meanwhile … on the other side of the ocean

18 french subjects have been scanned while listening to *The Little Prince in French*



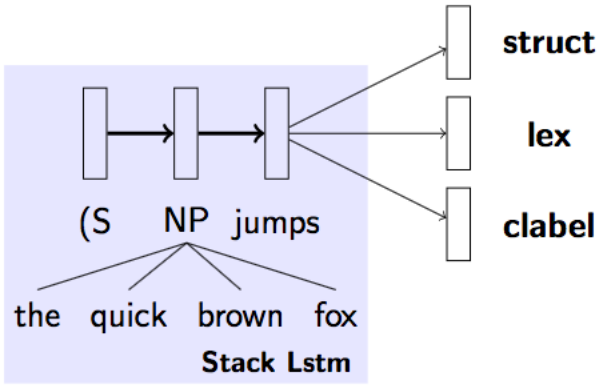
fMRI

UNICOG Cognitive Neuroimaging Lab

Inserm

The computer science French team proceeded by creating computer models of the comprehension process. By integrating parsing algorithms with information-theoretical complexity metrics which have been applied at the behavioral level in psycholinguistics.

Eric de la Clergerie



Benoit Crabbé

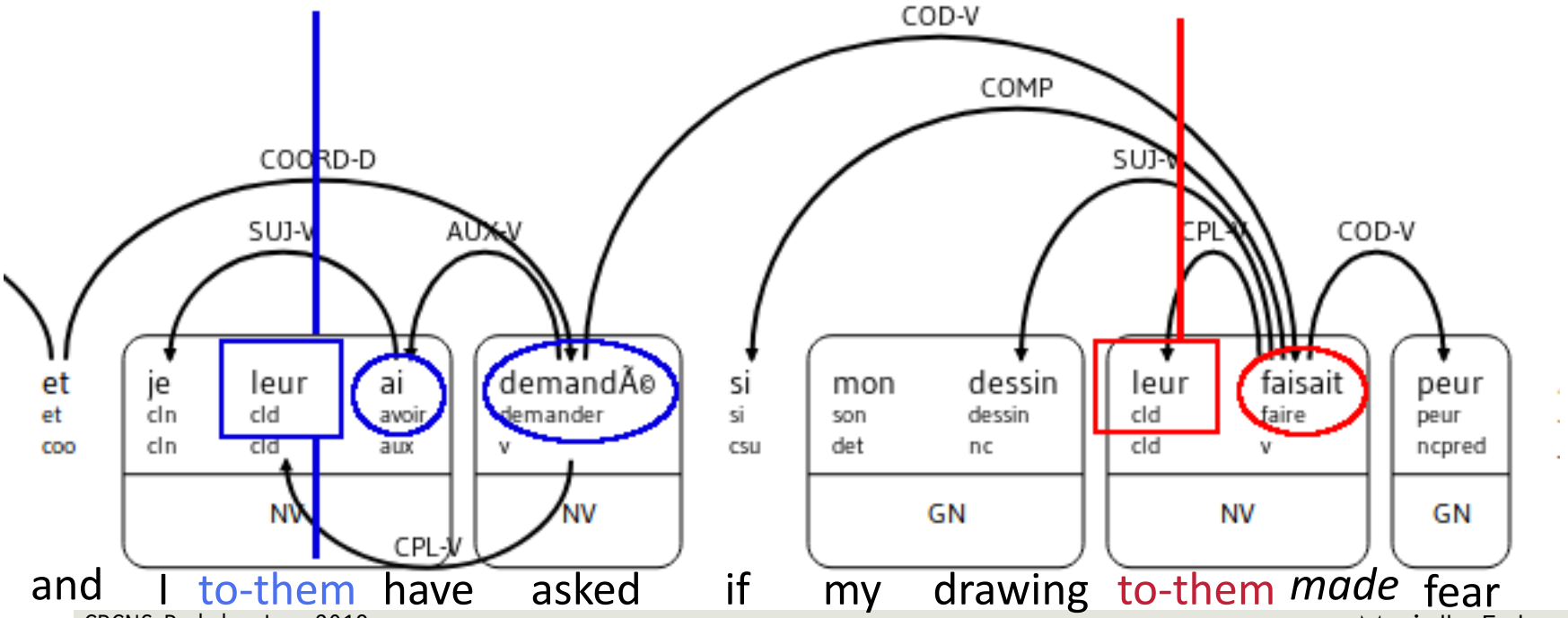
Dependency-links in the sentence

Word by word measure of the relationships between words, quantifying for example how many words are governed the present word.

Eric de la Clergerie

Pauline Brunet

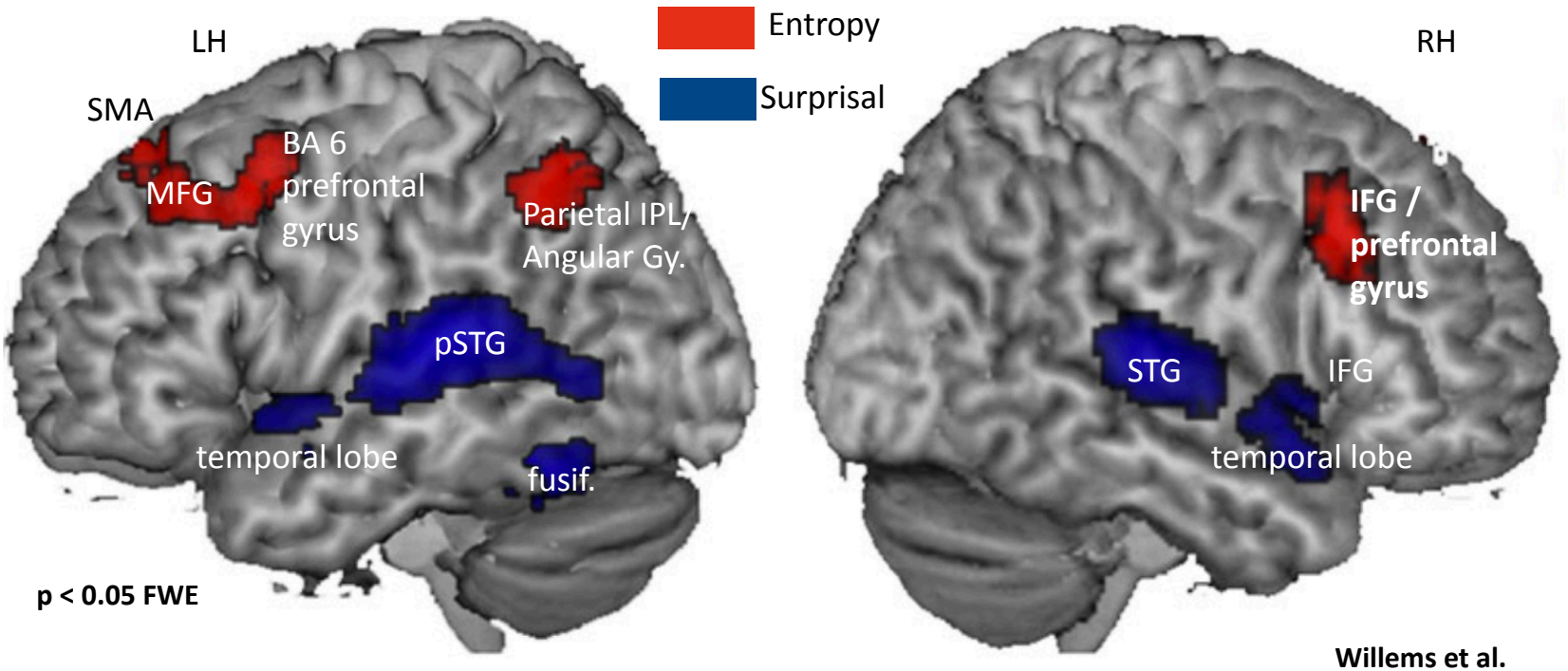
Deriving word-by-word processing difficulty predictions linked to the number of dependency relations that end, start or cross a given word



Thanks for your attention

Bonus Slides

word-by-word entropy (trigrams)



Willems et al. 2016

fMRI results of the whole-brain analysis. Brain areas that were significantly activated in response to the regressors modeling entropy and surprisal.

Entropy results for whole-brain analysis

Region	MNI coordinates (X Y Z)	Cluster extent (voxels)	Maximum t-value
Right inferior frontal gyrus	54 18 40 48 18 48	180	–3.99
Left middle frontal gyrus/ventral precentral sulcus	–46 12 50 –42 8 56 –36 30 46	353	–3.53
Left SMA	–4 42 44		–3.14
Left inferior parietal lobule	–44 –56 50 –52 –56 48 –46 –48 48	235	–3.50

Entropy expresses the strength of expectations about what the next coming word will be.

Surprisal expresses how unexpected was a word

word-by-word high surprisal vs. Low surprisal ()

Fig. 2. Left hemisphere anterior temporal lobe activation in the contrast of high- versus low-surprisal conditions (whole-brain analysis, group maps thresholded at voxelwise $p < 0.05$ and corrected for multiple comparisons by removing clusters with below-threshold size to achieve a mapwise error corrected $p < 0.05$). Numbers represent the Talairach x-coordinate of the slice.

fMRI results of the whole-brain analysis (z-scores). Brain areas that were significantly activated in response to the contrast high surprisal versus low surprisal.

High surprisal versus low surprisal.

Henderson et al. 2016

Talairach coordinates, volume of the cluster (μl), maximum z-score, and the label of anatomical structure for the high-surprisal vs. low-surprisal analysis, L = left hemisphere, R = right hemisphere.

Volume	Max	x	y	Z	Anatomical structure
High surprisal > low surprisal					
2322	3.711	31	−1	5	R putamen
	3.553	25	13	8	R putamen
	3.178	37	−16	2	R insula, R inferior frontal gyrus (pars opercularis)
2268	4.323	−31	1	0	L insula, L inferior frontal gyrus (pars opercularis), L putamen
837	3.368	28	−43	−9	R fusiform gyrus, R lingual gyrus
810	3.324	−31	−46	−15	L fusiform gyrus
756	3.758	10	−10	−3	R ventral diencephalon
Low surprisal > high surprisal					
729	−4.174	43	19	38	R middle frontal gyrus

Correlation matrix of the regressors in the GLM model

FRANÇAIS > 0