# Grammatical categories in the fronto-temporal language network

Mirjana Bozic[1,2], Elisabeth Fonteneau[1,2], Lorraine K Tyler[1], Barry Devereux[1], Paula Buttery[1], William Marslen-Wilson[1,2]

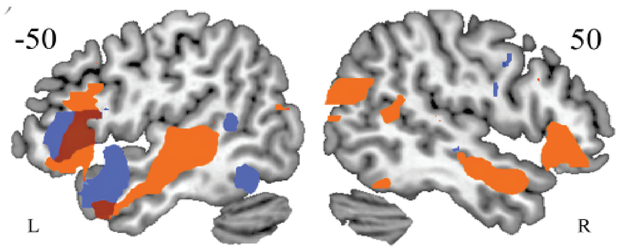
[1] *University of Cambridge, UK;* [2] *MRC Cognition and Brain Sciences Unit, Cambridge, UK*

## Combinatorial processes and grammatical category

Combining smaller elements into larger structures is essential in language, from the combinations of stems and morphemes at the single word level to the combinations of words into phrases and sentences.

At the single word level, some combinations have a purely grammatical function (past tense inflections, e.g. *play+ed*). Other are lexical-semantic (derivations, e.g. *brave+ly*), where the combination process creates a new word in the mental lexicon.

Their processing engages different subsystems of the bilateral fronto-temporal language network: grammatical, morpho-phonological complexity engages the left-lateralised subsystem, while lexical-semantic processing engages the bilateral subsystem [1].




***increase in lexical-semantic complexity***
***increase in grammatical complexity***

Bozic et al (2010) PNAS

Single words are further combined into phrases and sentences, following the rules of syntax. Syntactic processing has been associated with a broad, mainly left-lateralised network of fronto-temporal regions [2].

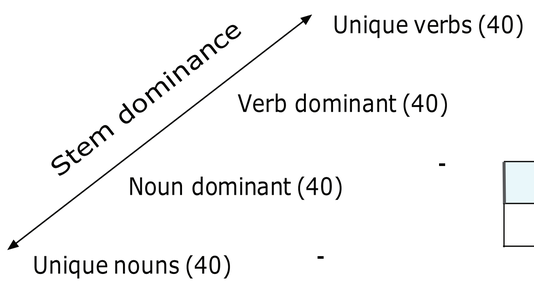
In addition, evidence suggests that combinatorial processing interacts with a word's grammatical category: verbs have rich morpho-syntactic paradigms and engage combinatorial processes more strongly than nouns [3]. Natural language processing (NLP) methods can provide detailed measures of verb paradigm complexity, allowing more specific analyses of their effects on verb processing.

***We investigate whether inflectionally-driven grammatical complexity engages the same processing mechanisms as minimal phrasal complexity, for both verbs and nouns. We employ a multivariate Representational Similarity Analysis (RSA) to assess the patterns of activity evoked by their processing across the regions of the fronto-temporal language network.***

## Design and methods

160 stems were matched on a range of psycholinguistic variables and divided into 4 categories: verb unique, verb dominant, noun dominant, noun unique (based on their frequency of occurrence as a verb or a noun in CELEX).

Two NLP measures were extracted from the VALEX lexicon [4], to express verb paradigm complexity: the average number of subcategorization frames associated with a verb, and the entropy of their distribution.



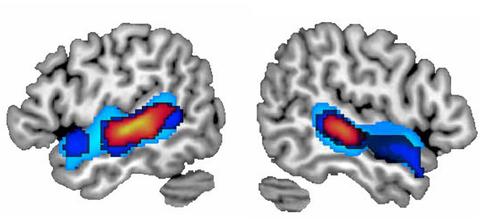

| *Bare stems* | *Inflections* | *Phrases* |
|---|---|---|
| follow / rug | follow**s** / rug**s** | I follow / the rug |

Test words were mixed with 240 acoustic baseline trials (musical rain, MR), and 240 silence trials. Participants listened to them passively and occasionally performed a one-back memory task.

Imaging procedure: 18 participants scanned on a 3T Siemens system, using a fast sparse protocol (TR=3.4s, TA=2s). Data were analysed in SPM5, using univariate approaches and multivariate Representational Similarity Analyses [5]

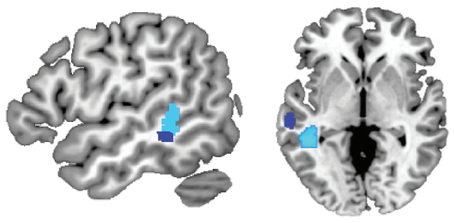
## Univariate results

**Activation for different levels of complexity**



***stems minus acoustic baseline (MR)***
***inflected forms minus acoustic baseline (MR)***
***phrases minus acoustic baseline (MR)***

**Dominance-modulated activation for different levels of complexity**



***modulation for inflected forms (at .001)***
***modulation for phrases (at .01)***

NLP-based measures of verb complexity did not significantly modulate the observed activations.

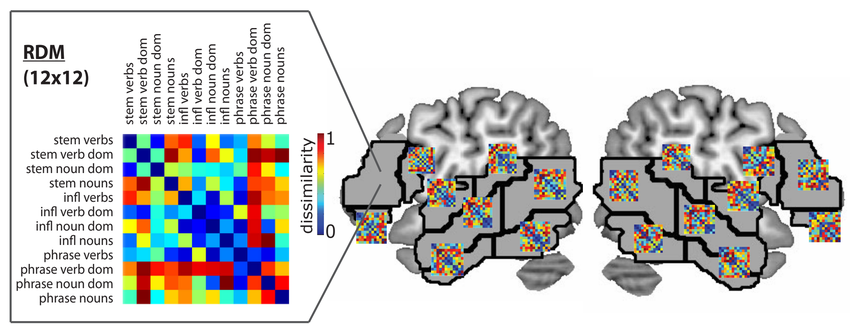
p<.001 voxel , p<.05 cluster, corrected for multiple comparisons

## Multivariate RSA results

RSA is a multivariate pattern analysis method that allows us to assess the information carried by a pattern of activation across multiple voxels.

In RSA, patterns of activation are expressed as Representational Dissimilarity Matrices (RDMs), which show the correlation distance (one minus the correlation value) between activation patterns elicited by pairs of different conditions. Neuroscientific inference is drawn from a second level analysis that compares RDMs to theoretical models, also characterized by RDMs.

**1. RDMs extracted for anatomically defined regions of interest**



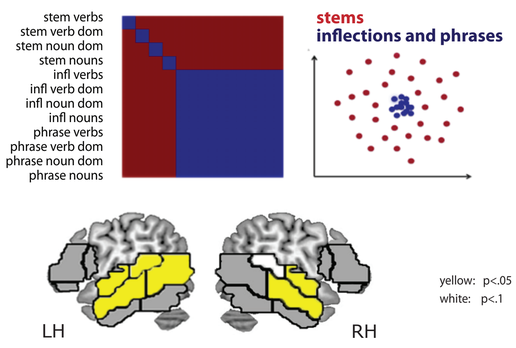

**Regions of Interest (bilaterally)**

Pars opercularis (BA 44)
Pars triangularis (BA 45)
Pars orbitalis (BA 47)
Superior temporal gyrus (BA 22)
Middle temporal gyrus (BA 21)
Inferior temporal gyrus (BA 20)

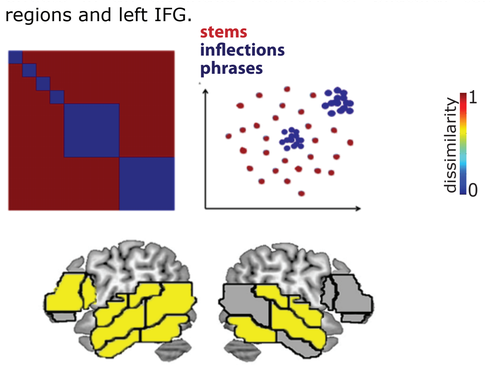
*(all temporal regions further split into anterior and posterior parts)*

**2. RDMs compared to a hierarchy of theoretical models**

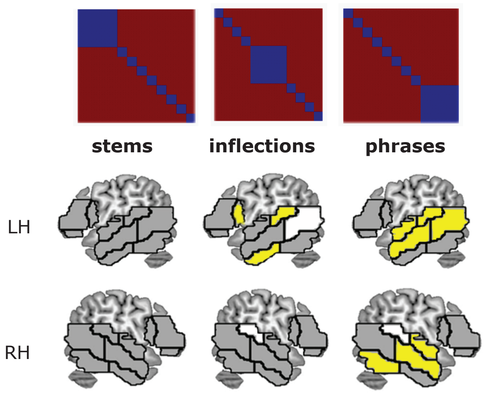
**Complexity model** tests for the sensitivity to complexity processing, regardless of its type. Specifically, this model assumes that any complex item creates an activation pattern similar to the pattern triggered by other complex items, but dissimilar to that triggered by simple words. Blue indicates high correlation between activation patterns, red indicates the absence of correlation. The results show significant effects in STG and MTG bilaterally.



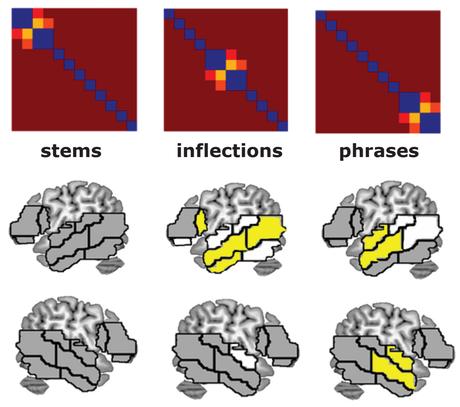

**Complexity type model** tests for the sensitivity to complexity processing, but also differentiates between inflectional and phrasal complexity. It assumes that inflected forms and phrases create different activation patterns, which are also dissimilar to the activation triggered by simple words. The results show significant effects in an extended network of bilateral temporal regions and left IFG.




**Detector models** code for the processing of stems, inflected forms and phrases separately. They show that inflectional complexity correlates with a strongly left-lateralised set of regions, while phrasal complexity correlates with a distinctly bilateral network. There was no significant correlation in any region with the processing of simple stems.




**Dominance-modulated models** test whether verb dominance modulates any of the correlations observed in the 'detector' models. They show a significant correlation increase for inflections in the L perisylvian network, as well as an apparent decrease for phrases in bilateral temporal regions. There were no effects for stems.




**Models modulated by NLP-based measures of verb complexity** failed to explain any additional variance over and above the variance captured by the detector models.

## Summary and conclusions

***Grammatical complexity driven by inflectional and by phrasal combinations engage overlapping, yet distinct processing mechanisms:***

***1. Inflectionally-driven grammatical complexity engages left-lateralised fronto-temporal regions.***

***2. Processing of minimal phrases engages bilateral temporal regions.***

***Grammatical category (verb dominance) modulates the processing of inflected forms and phrases, but not the processing of bare stems.***

***The more detailed characterizations of verb complexity provided by NLP methods (the number of subcategorization frames associated with a verb, and the entropy of their distribution) did not add explanatory power to our models.***

***Multivariate analyses allow more detailed dissociations of linguistic processing dimensions than conventional univariate approaches.***

## References

[1] Bozic M, Tyler LK, Ives DT, Randall B & Marslen-Wilson WD (2010) PNAS, 107, 17439-44; [2] Kaan E & Swaab T (2002) TICS, 8, 350-356; [3] Tyler LK, Randall B & Stamatakis EA (2008) J Cogn Neurosci, 20, 1381-1389; [4] Korhonen A, Krymolowski Y & Briscoe T (2006) Proceedings of the 5th international conference on Language Resources and Evaluation; [5] Kriegeskorte N, Mur M & Bandettini PA (2008) Front Syst Neurosci, 2, 1-27.

## Contact

mb383@cam.ac.uk