Putting pieces together: how logical connectives combine propositions. An fMRI study

Elena Benini[1], Lilia Del Mauro[2], Gianpaolo Basso[3], Giosuè Baggio[4], Carlo Reverberi[2,5]

[1]RWTH Aachen University, Germany, [2]Department of Psychology, University of Milano-Bicocca, 20126, Milan, Italy, [3]Department of Medicine and Surgery, University of Milano-Bicocca, 20900, Monza, Italy, [4]NTNU, Trondheim, Norway, [5]NeuroMi — Milan Center for Neuroscience, 20126 Milan, Italy

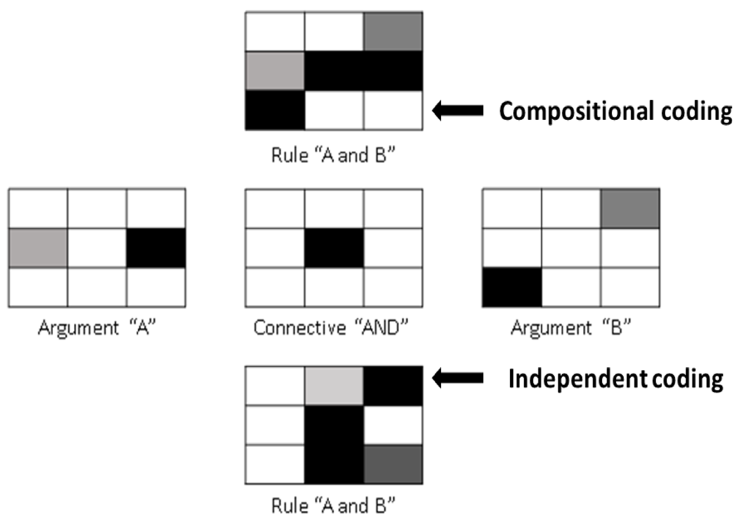
elena.benini@psych.rwth-aachen.de | l.delmauro@campus.unimib.it
carlo.reverberi@unimib.it

Introduction

Rules are widely used in everyday situations to pursue relevant goals and interact strategically with the environment. Simpler **propositions** are often combined into compound sentences ("*if* there is a chicken *or* there is a rabbit don't bring a fox") by an **elementary logical connectives AND, OR, and IF** to form rules. However, where and how our brain represents rules remains unclear.

Rule representation seems to involve a fronto-parietal network (Baggio et al., 2016; Zhang et al., 2013). Moreover, the neural representation of compound rules has been shown to be **compositional** (Reverberi et al., 2012a), with partially segregated brain networks encoding the order of the simple rules composing the compound rule (Reverberi et al., 2012b).

We explored whether the **neural representation of simple rules formed by two propositions and a logical connective** (e.g., "*There is a triangle* and *there is a hexagon*") **is compositional**. To this aim, multivariate pattern analysis (MVPA) was applied to functional magnetic resonance imaging (fMRI) data collected while healthy participants maintained simple rules in working memory. MPVA allows to investigated whether the neural code of a rule contains some features of the neural code of the elements composing the rule



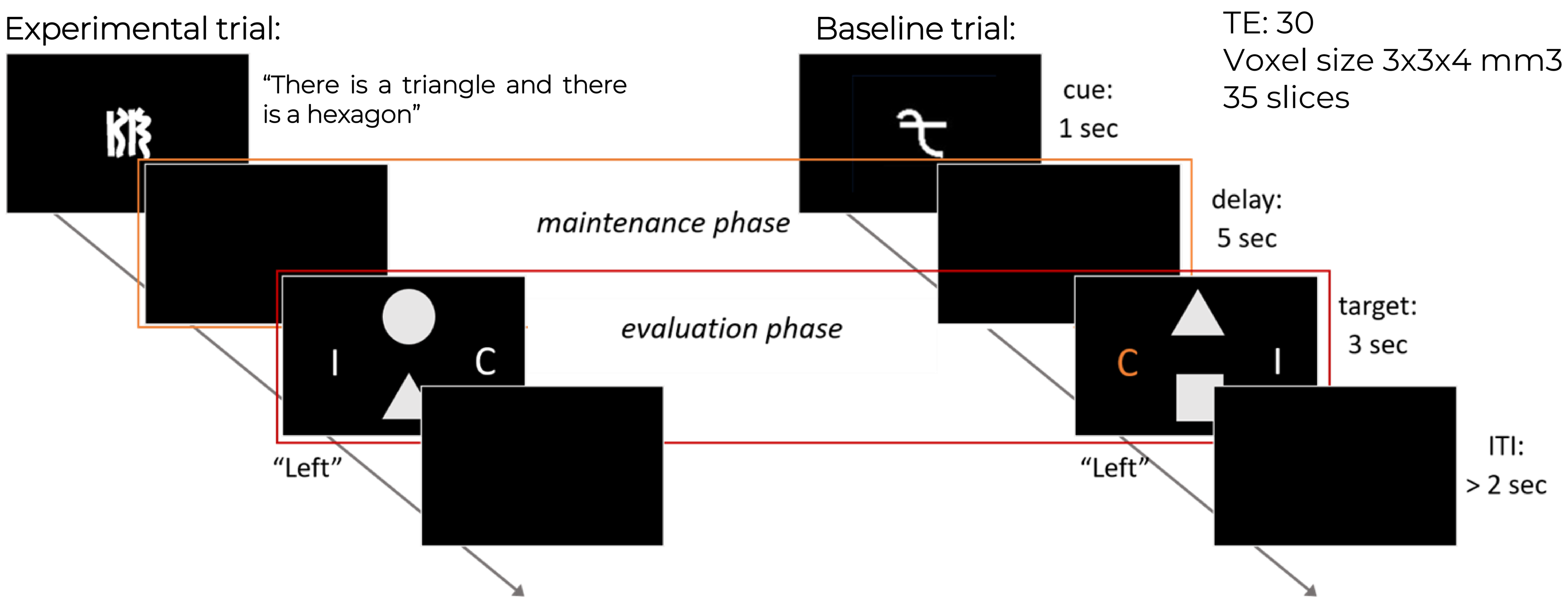
Compositionality
A formulation of the principle of semantic compositionality in linguistic maintains that "at least one meaning of a whole is a function only of the meanings of the parts and of the way they are syntactically combined" (Baggio, 2021, p. 8). Therefore, the neural representation of a rule is compositional if "representations of parts are recoverable from representations of wholes" (Baggio, 2021, p. 12).

Fig1. *Alternative ways of coding a rule. In the middle, examples of activity patterns coding for the single elements composing the rule: proposition A, connective AND, and proposition B. On top and bottom, two alternative coding possibilities for the rule "A and B". A compositional code (top) where the pattern coding for the rule is the combination of the patterns of the single elements composing it. An independent code (bottom) where the pattern for the rule is unrelated to the patterns for its composing elements.*

Methodology

Sample: 9 (planned 34) healthy volunteers

Philips Igenia 3 T
TR: 2000 ms
TE: 30
Voxel size 3x3x4 mm3
35 slices



Experimental timeline. Each trial starts with a cue that prompts participants to retrieve the associated rule. During the *maintenance phase,* participants maintain such a rule in working memory. Then, in the *evaluation phase,* they judge whether a visual scenario composed of two vertically aligned shapes is compatible or incompatible with the rule. In *baseline* trials, participants perform the same motor action (i.e., a keypress) but the rule was "Press the key to the side of the orange letter".

Experimental conditions. 3 logical connectives (AND, OF, IF) * 2 geometric shapes pairs (triangle-hexagon; square-circle) * 2 orders (T-H, S-C and H-T, C-S) + 1 baseline + 1 memory trials + 1 catch trials = 15 conditions.

FMRI protocol. 6 runs, 64 trials each (4 repetitions of each rule + 4 baseline trials + 6 memory trials + 6 catch trials).

Decoding Analyses

BOLD signal during *maintenance phase* (with a HRF model) or for each 2-seconds time bin (with a FIR model) is predicted for each voxel as a function of the current experimental condition. The obtained maps of beta coefficients are used in a whole-brain searchlight analysis (Eztel et al., 2013), where spheres of 2-voxels radius are fed to a Support Vector Machine (SVM) classifier that learns to classify such maps based on the experimental condition, for example, the logical connective.

Analysis 1. *Decode logical connective across propositions*

Test set	AND rules with triangle and hexagon	IF... THEN rules with triangle and hexagon
Train set	AND rules with square and circle	IF... THEN rules with square and circle

Distinguish AND from IF rules independently from the propositions. The SVM needs to generalize across rules propositions and orders.

Analysis 2. *Decode propositions across logical connective*

Test set	AND and OR rules with triangle and hexagon (both orders)	AND and OR rules with square and circle (both orders)
Train set	IF... THEN rules with triangle and hexagon (both orders)	IF... THEN rules with square and circle (both orders)

Distinguish triangle-hexagon from circle-square rules. The SVM needs to generalize across logical connectives.

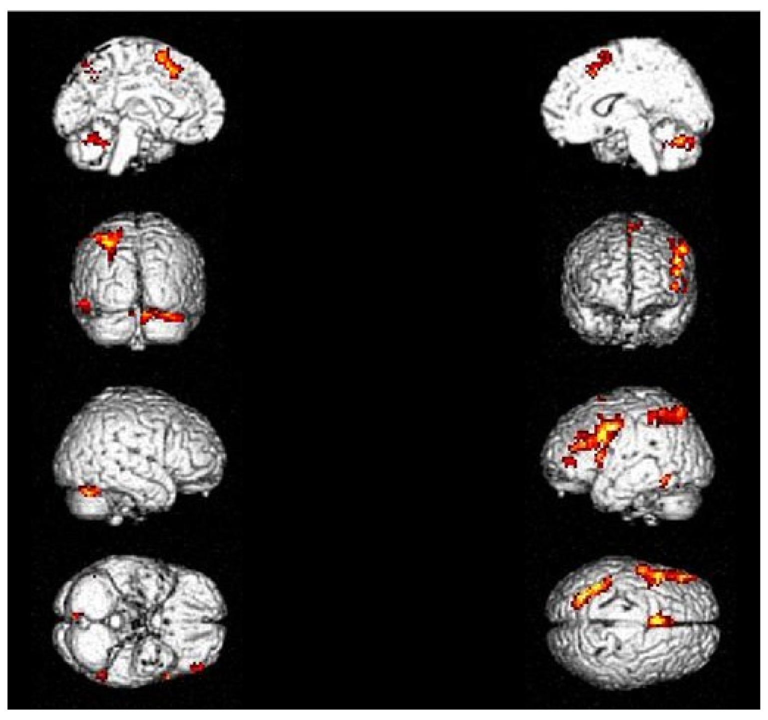
Analysis 3. Decode propositions order

Test set	All rules triangle-hexagon All rules circle-square	All rules hexagon-triangle All rules square-circle
Train set	All rules triangle-hexagon All rules circle-square	All rules hexagon-triangle All rules square-circle

Distinguish rules with the propositions in a certain order from rules with the opposite propositions order? The SVM does not need to generalize.

Results

Univariate Analyses. No mask, no family-wise error correction, alpha < 0.001 at the voxel level, cluster > 30 voxels



Inferior frontal gyrus, supplementary motor area, intraparietal sulcus, superior parietal lobe more active during maintenance of any logical rule than the baseline.
No suprathreshold cluster is found when looking at the difference between the rules (IF > OR, OR > AND nor IF > AND). In these preliminary data, univariate analyses did not detect differences between connectives.

Fig2. *Brain Clusters Differentially Active During Logical-Rules Trials versus Baseline During Maintenance*

Decoding Analyses. No mask, no family-wise error correction, alpha < 0.005 at the voxel level, cluster > 30 voxels

The classifier performed above chance level in decoding (see **Analysis 1** Table) the currently maintained rule in the superior frontal and middle frontal gyri (BA8, BA9, BA45), the inferior temporal gyrus (BA20), and the inferior parietal lobe (BA40). These areas contained information related to the connective alone.

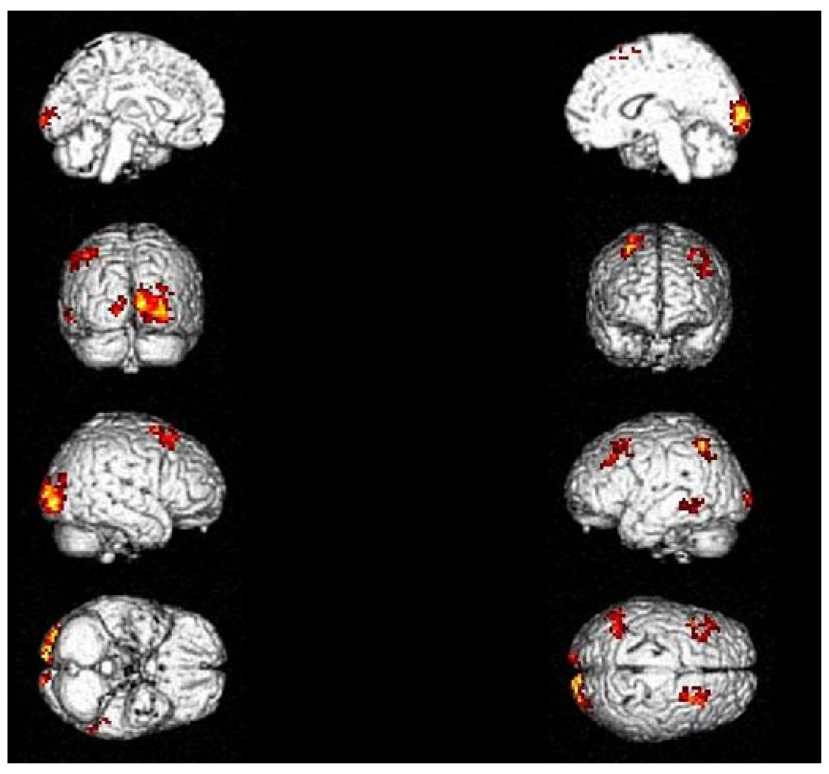
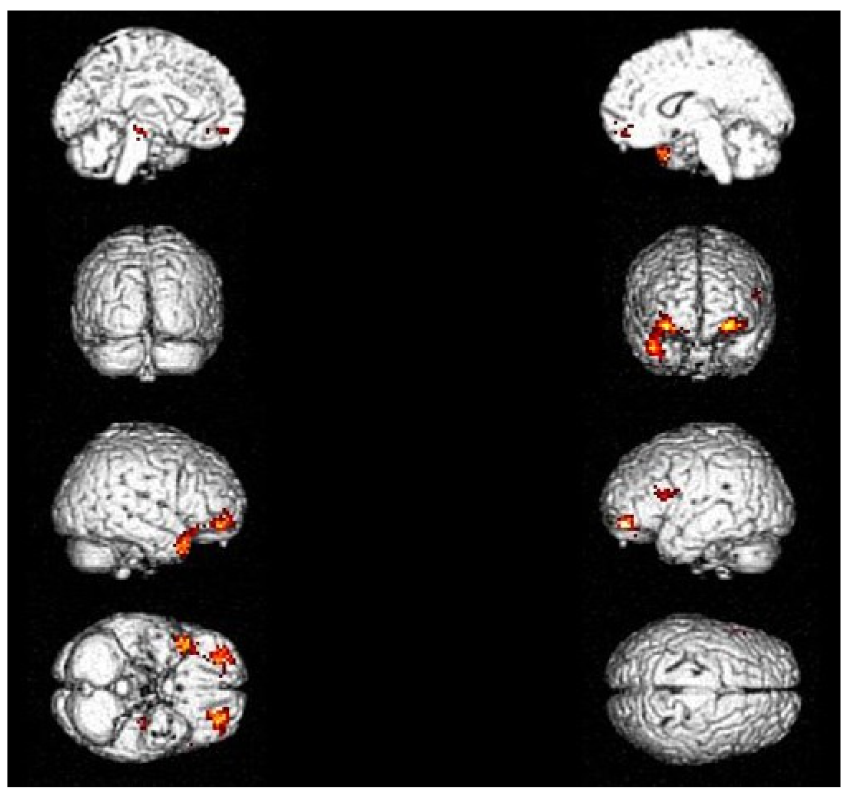


Fig3. *Brain Regions Containing Information about Rules Independently from the Propositions and Their Order During Maintenance*



The classifier performed above chance level in decoding (see **Analysis 3** Table) the order of the propositions (e.g., triangle-hexagon vs. hexagon-triangle) in the currently maintained rule in the inferior frontal gyrus (BA 44, 45, 47), orbitofrontal cortex (BA11), temporal pole (BA38), and parahippocampus (BA30).

Fig4. *Brain Regions Containing Information about Propositions Order during Maintenance*

In these preliminary data, we found no suprathreshold cluster containing information about propositions (see **Analysis 2** Table).

Discussion

Univariate analyses revealed a prefrontal network, mostly left, which is preferably active during the maintenance of logical rules than baseline. The network is consistent with previous findings (e.g., Baggio et al., 2016; Zhang et al. 2013).

The possibility to decode the connective suggests that these logical rules were represented in a compositional fashion in the brain since information related to the connective can be detected independently from information related to the propositions and their order.

Interestingly, propositions order is represented in a different network as the connective. This suggests that also the order of the propositions might be represented compositionally.

Our preliminary results suggest that, at least at certain stages, neural representations of simple logical rules is compositional. Therefore, we found evidence for a compositional processing path, from sentences to their meaning (Baggio, 2021).

Literature Selection

Baggio, G., Cherubini, P., Pischedda, D., Blumenthal, A., Haynes, J. D., & Reverberi, C. (2016). Multiple neural representations of elementary logical connectives. *NeuroImage*, 135, 300–310. https://doi.org/10.1016/j.neuroimage.2016.04.061

Etzel, J. A., Zacks, J. M., & Braver, T. S. (2013). Searchlight analysis: Promise, pitfalls, and potential. *NeuroImage*, *78*, 261–269. https://doi.org/10.1016/j.neuroimage.2013.03.041

Reverberi, C., Görgen, K., & Haynes, J.-D. (2012a). Compositionality of Rule Representations in Human Prefrontal Cortex. Cerebral Cortex, 22(6), 1237–1246. https://doi.org/:10.1093/cercor/bhr200

Reverberi, C., Görgen, K., & Haynes, J.-D. (2012b). Distributed Representations of Rule Identity and Rule Order in Human Frontal Cortex and Striatum. *The Journal of Neuroscience*, 32(48), 17420–17430. https://doi.org/10.1523/JNEUROSCI.2344-12.2012

Zhang, J., Kriegeskorte, N., Carlin, J. D., & Rowe, J. B. (2013).Choosing the rules: Distinct and overlapping frontoparietal representations of task rules for perceptual decisions. *The Journal of Neuroscience*, 33(29), 11852–11862. https://doi.org