How the Human Brain Solves the Symbol-Grounding Problem

Simone Viganò (simone.vigano@unitn.it)
CIMeC – Center for Mind/Brain Sciences, University of Trento, 31 Corso Bettini
Rovereto, 38068 Italy

Valentina Borghesani (valentina.borghesani@ucsf.edu)
Memory and Aging Center, University of California San Francisco, 675 Nelson Rising Lane
San Francisco, California 94158 USA

Manuela Piazza (manuela.piazza@unitn.it)
CIMeC – Center for Mind/Brain Sciences, University of Trento, 31 Corso Bettini
Rovereto, 38068 Italy

A fundamental issue in cognitive science is the so-called “symbol-grounding problem” (Harnad 1980), related to the question of how symbols acquire meaning. One simple view posits that, for concrete words, our brain solves the problem by creating associations between the neural representations of the surface forms of symbols (spoken or written words) to the one(s) evoked by the object, action, or event classes the symbols refer to (e.g., see Pulvermuller 2013; 2018). Evidence supporting this view comes from the observation that words related to well known concepts such as numerical quantities (Piazza et al. 2007; Eger et al. 2009), colors (e.g. Simmons et al. 2007), manipulable objects (Chao et al. 1999), places (Kumar et al. 2017), or actions (Hauk 2004; 2011), automatically re-activate the same brain regions that are active during the perception/execution of those specific object features/actions. These data, however, are informative on the neural bases of symbol *grounded* representations, but not on those underlying symbol *grounding*: i) they fall short in assessing the role of memory systems implicated in this kind of symbol-to-concept associative learning, and ii) they do not provide a full picture of the effects that symbol grounding has on the brain. Here, to investigate the neural changes generated by this process, we adopted an artificial learning paradigm where 21 adult subjects learned to categorize novel multisensory objects by giving them specific symbolic labels.

Participants studied for 9 days the association between novel multisensory objects and novel words representing their categorical identity (visit http://bit.ly/2MwP92e for audio-visual object exemplars). Objects were created by varying parametrically and crossing orthogonally the size of an abstract shape and the pitch of its associated sound. Before and after learning, participants were presented, during an fMRI scanning session, with pseudorandom sequences of the visual objects that during a small animation produced a sound, and the written words, while performing a one-back task.

Using a set of multivariate pattern activity as well as functional connectivity analyses we found that learning was paired with the development of a bi-directional functional mapping between the neural representations of the objects (in the left angular gyrus) and of the associated symbols (in the left Visual Word Form Area) that, after learning, changed their responses and encoded the newly learnt object-name associations. A key node supporting this bi-directional mapping was the hippocampus, which activity encoded the abstract symbol-to-object mapping rule, and was predictive of inter-individual variations in learning performance. Finally, we also observed concurrent changes in the neural representational geometries of object features in low-level sensory areas, where learning induced both increased sharpening and increased feature selectivity. Taken together, the data indicate that the process of grounding symbol to meaning entails orchestrated changes in the neural activity of both memory and perceptual systems. These findings might provide a fertile ground for inspiring/testing computational algorithms aiming at modelling the acquisition of semantic representations.