Cognitive Neuroscience

To what extent has Cognitive Neuroscience investigated human natural language understanding?

Goldsmiths University of London

Samora Hunter BSc

33592227

MSc Computational Cognitive Neuroscience

2019

Language is a remarkably broad area of study, it is the focus of several different specialist academic fields and tangentially related to many more. Within the purview cognitive science, cognitive psychology and cognitive neuroscience, machine learning and artificial intelligence we first demarcate our region of interest. The most related field to our interest is natural language understanding. The use of perceptual systems in natural language understanding generally entails the transduction and transmission of sensory information and this information's interactions with memory systems. Similarly, the production of language entails verbal utterances and orthographic writings. Here we make an even narrower distinction by dissociating these more elementary processes (natural language processing) from those of understanding, comprehension and meaning. This essay is concerned with meaning rather than the arguably trivial processes necessary to process raw sensory input. We will however see that this distinction is at times blurry and that the production, representation and transmission of meaning are often intertwined with that of the mere processing. It is however a meaningful one in the discussion of the limitations of current cognitive neurosciences investigation of meaning.

In support of this distinction, it is useful to reference the philosophical demonstration of a qualitative difference between processing and understanding (Searle, 1983). In order to imbue a word with meaning in its pragmatic use, one must understand what that meaning is, not merely follow logical rules. Since we know the brain gives rise to both these cognitive abilities, and the two are distinct in their manifestation, we may logically deduce that at least either their underlying neurobiological process or structure must too be distinct. This would also hold if the ability were heavily reliant on emergent processes. We may also then infer structurally distinct neurobiological cytoarchitecture or neurological information processing pathways. It is here that cognitive neuroscience in its capacities has investigated these structures and proposed information processing pathways as plausible accounts of natural language understanding.

We adopt an approach previously used in the investigation of the neurological basis of another fundamental component of language, “syntax”, in ours of “understanding”. In the illumination of the neuro cognitive basis of syntax Frederici (Frederici, 2017) takes a Marr like level of analysis that organises the overall view of the cognitive function into molecular, neuronal ensemble and a macro functional levels. The same approach is attempted here; we examine what cognitive neuroscience has shown at these levels and discuss how it has informed our knowledge of natural language understanding. We further decompose our phenomena of interest and the point of this essay by asking how the meaning of lexical items is represented in the brain, it is here at this level that the most crucial aspect of meaning is to be found. A complete understanding at this level will answer our question. We see that throughout this essay that cognitive neuroscience struggles to address this question.

We begin by examining the broad approach and findings in the questions of *where* and *when* and finally end by addressing *how* language understanding occurs in the human brain as informed by cognitive neuroscience. Throughout we critically examine the limitations of the methods employed and comment on the extent of the investigations merit with respect to our narrowly defined but philosophically slippery question.

To establish a scale of our understanding of this function we can briefly note current attempts at implementing artificial systems with it. Current implementations of natural language understanding in artificial systems are not capable of human level understanding. It is thought by some that the instantiation of this ability in non-biological substrate is an AI hard problem

(Shapiro & Stuart, 1992). The purposeful construction of such a system would be in the upper tier of such a scale representing our level of understanding. The human being's brain is a known, self-evident example of a system capable of natural language understanding. The algorithms, architecture and environment of the human brain are capable of producing natural language understanding , quod erat demonstrandum. And so, our question is, *"can Cognitive Neuroscience, the study of the biological system underlying this emergent cognitive function, reveal how natural language understanding emerges in the human brain".*

Where…

The scientific investigation of the neural basis of the linguistic construct we generally term language has converged on the view that language is broadly lateralised to the left hemisphere in humans. This is greatly evidenced by a meta-analysis of language task related imaging experiments which find the majority exhibiting unilateral hemispheric cellular recruitment (Vigneau et al., 2006). The same meta-analysis and another comprehensive review report strong lateralisation during semantic related tasks (Friederici, 2017). The authors however stress the presence of activity in the right hemisphere under some task conditions. Further isolating the sources of neural activity responsible for language have led to the present perisylvian language network description (Catani et al, 2004). This network spans a region primarily encompassing the areas around the perisylvian fissure. More in line with our question of meaning, Friederici (2017) reports a meta-analysis of the neural basis of language semantic related tasks, identifying the anterior temporal lobe, posterior temporal and middle temporal gyrus. Further probing of this networks functional properties via double dissociations of lesioned clinically distinct aphasics reveal likely candidates for specialised functional integration with extrasylvian areas via association fibers; unicate fasciculus connecting the frontal lobe to the temporal (perisylvian) lobe, occipito frontal fasiculus connecting the occipital lobe to the frontal lobe and the inferior longitudinal fasciculus connecting the occipital lobe to the temporal lobe. Two key regions, Brocas area (Brodmanns area 44 and 45) and Wernikes area (Brodmanns area 22) have also emerged, being strongly related to the production and comprehension of speech respectively. This topographical functional map is however limited in its explanatory power given the implications of leading embodied theories of cognition (Barsalou, 2008). Such theories suggest heavy interconnectivity and interdependency between systems underlying cognitive functions, making localisation particularly challenging for high level functions such as language understanding.

The imaging methods of functional magnetic resonance imaging, event related potentials, positron emission tomography, electroencephalography and diffusion tensor magnetic resonance imaging tractography in conjunction with lesion patient data answer our macro functional level explanation of where meaning with respect to language is understood to some extent. We know now of the region's most active among a general baseline level of activity and noise during these tasks. The explanatory power at this level however limited. Firstly, there exists spontaneous 'noise' activity across the brain (Faisal, Selen & Wolpert, 2008) likely underlying some computations relevant to language understanding of which the nature is unclear and represents a potential confound. Secondly knowledge of increased activity in particular regions gained from fMRI does not tell us much. More generally, the following points surmise the criticism of this knowledge. Information theoretic estimates of the complexity of the human brain in terms of degrees of freedom with respect to causal dynamics and information density put the systems upper bound estimate in the region of the Bekenstein black hole entropy bound on the information content in material systems (Bekenstein, 1981; Bostrom & Sandberg, 2008; Thaggard, 2002; Street S., 2016). Similarly assuming 1,200,000 mm ^{2} volume for an average human brain, we would have about one million voxels per brain. If we take an upper bound estimate for computational capacity of the brain based on individual proteins performing computation, 10^{23} FLOPS, we can suppose this is distributed across the voxels at 10^{18} per voxel, we arrive at quantitative estimate of our ignorance. This is an estimate of the resolution loss in computational terms in attempting inference from the voxel activity to the macro cognitive function with fMRI. It is in those computations that the information underlying meaning is contained. The current resolution seems even useful enough to inform artificial systems architecture, and certainly not the greatly detailed algorithmic steps and operations needed by such systems. It has however shown some promise in neurobiologically grounded computational modelling. For these reasons, at the macro functional level cognitive neuroscience can tell us remarkably little about natural language understanding. However, in certain astronomical problems, as in this aspect of neuroscience we are faced with a vast search space, and with these methods addressing *where*, we are able to orient our telescope to narrow this space for later search, with more advanced technology.

When…

Outside of single cell recordings that are grossly impractical in the investigation of a cognitive function that likely recruits a great number of functional neural units, electroencephalography typically offers us the greatest temporal resolution. The seminal paper with regard to the processing of meaning across the time dimension in language is undoubtedly that by Kutas and Hillyard (1980) whom report an event related potential finding of what has since been termed the N400 component. This negative wave robustly emerges in ERP signals analysis of EEG recordings during tasks in which a subject is presented with a semantic incongruity in a stimuli sentences terminus word. Two degrees of semantic incongruity were originally examined, the examples given were “He took a sip from the waterfall" as a moderate condition and "He took a sip from the transmitter" as a strong condition. Interestingly the linguistically more incongruent trials produced greater amplitude in the signal. The authors propose that the N400 could represent a semantic reprocessing for the information. More than one thousand follow up studies support this original finding and have expanded the paradigm to a vast range of modalities with similar results. Similarly, many document the components dissociation from syntactically incongruent sequences, from semantically congruent sequences and from mere sequential pattern sequence violations such as those possible to simulate in music (Kutas & Federmeier, 2011). The aggregate of these investigations suggests this component represents some kind of semantic information processing, more specific theorizing about the nature of its representation can be found in Kutas and Federmeier 2011 and Federici, 2017. In constraining our view of these findings to the sphere of relevance to our question regarding natural language understanding we again find critical issues. On the one hand these findings can and have informed cognitive models of the role of semantic processing in signal flow type diagram explanations (Federici, 2017). However whatever meaningful that is happening at the neural level that produces this signal in time is not captured by this component, the representation of the information necessary for understanding is not captured either. These range of studies again address the macro functional level, this relatively superficial level is insufficient to tell us something meaningful about meaning itself. Even assuming the thoughts of Kutas and Hillyard about the nature of the processing marked at the N400 being a reprocessing of semantic information, we are still left with the question of how this processing is occurring, what does it mean to process meaning and how does the biological substrate of the human brain achieve this.

How…

The question of *how* brings us to our conclusion. Cognitive neuroscience has but scratched the surface of the representation, understanding and comprehension of meaning in language. It is obvious that the most valuable details lie at the bottom of our level of analysis, between the molecular and cell assembly level, beyond the reach of current approaches. This is most striking when we imagine what it would take to endow an artificial system with this cognitive ability using just the knowledge gained from these approaches. The methods such as functional imaging are however useful in mapping the outlines of the neural basis of language understanding. N.B. semantic memory is likely relevant to the current essay though has been omitted as its investigation through cognitive neuroscience suffers from similar critical faults.

References

Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature reviews*. Neuroscience, 9(4), 292–303. doi:10.1038/nrn2258

Barsalou, L.W. (2008) Grounded Cognition. *Annual Review of Psychology*, 59, 617-645. http://dx.doi.org/10.1146/annurev.psych.59.103006.093639

Catani, M., Jones, D. K., & Ffytche, D. H. (2004). Perisylvian language networks of the human brain. *Annals of Neurology*, 57(1), 8-16. doi:10.1002/ana.20319

Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. Cambridge, MA: The MIT Press.

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP*). Annual Review of Psychology*, 62(1), 621-647. doi:10.1146/annurev.psych.093008.131123

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843), 150-157. https://doi.org/10.1038/35084005

Sandberg, A. and Bostrom, N. (2008) Whole Brain Emulation: A Roadmap. Future of Humanity Institute, *Oxford University, Technical Report* #2008-3.

Searle, John (1983), "Can Computers Think?", in Chalmers, David (ed.), Philosophy of Mind: Classical and Contemporary Readings, Oxford: *Oxford University Press,* pp. 669–675, ISBN 978-0-19-514581-6

Shapiro, Stuart C. (1992). Artificial Intelligence In Stuart C. Shapiro (Ed.), *Encyclopedia of Artificial Intelligence* (Second Edition, pp. 54–57). New York: John Wiley. (Section 4 is on "AI-Complete Tasks".)

Street S. (2016). Neurobiology as Information Physics. Frontiers in systems neuroscience, 10, 90. doi:10.3389/fnsys.2016.00090

Vigneau M , Beaucousin V, Hervé PY, Duffau H , Crivello F, Houdé O, Mazoyer B Tzourio-Mazoyer N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*, 30(4):1414-1432 DOI: 10.1016/j.neuroimage.2005.11.002

Wang, Y., Liu, D. & Wang, Y. *Brain and Mind* (2003) 4: 189. https://doi.org/10.1023/A:1025405628479