Sandy Ideas and Coloured Days: Some Computational Implications of Embodiment

RONAN G. REILLY

Department of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland, Email: rreilly@nova.ucd.ie

Abstract. This paper is an exploration of the relationship between language and vision from the perspective of language evolution on the one hand, and metaphor on the other. Recent research has suggested that the origins of human language capacity can be traced to the evolution of a region in the brain that permits the interaction of information from sensory and motor cortices. In light of this, it is hypothesised that the computational mechanisms of language are derived from those of the sensory-motor domain, and that the pervasiveness of metaphor is one manifestation of language's computational antecedents. A variety of cognitive and computational implications are drawn from these hypotheses.

Key words: language evolution, metaphor, synesthesia, cortical computation

1. INTRODUCTION

When I was an undergraduate I remember reading a book by R. L. Gregory (1974) which contained a chapter entitled "The Grammar of Vision." It discussed the possibility that language might have its roots in visual perception, and described parallels that existed between vision and grammar. Visual perception, Gregory maintained, involved "parsing" the sensory input using a lexicon of features into meaningful "sentences". The process was driven by a set of "grammatical" rules which could be violated in the case of so-called impossible objects, such as those devised by the artist M. C. Escher. Gregory suggested that this rule-based parsing aspect of vision may have been exploited during the evolution of language. I found Gregory's paper very exciting, having previously been puzzled by the evolutionary recency of language and not persuaded by those who argued that it had arisen from a serendipitous mutation. Here was a proposal that the computational machinery underlying language could also be found in a more ancient perceptual system. Nonetheless, as I will argue later, by seeking a connection between language and vision, and indeed between language and other sense modalities, we are forced to look at the computational mechanisms underlying language in a rather different light than that initially envisaged by Gregory. His view of this relationship was very much language-centred, heavily influenced by the work of Chomsky and particularly by the latter's distinction between
deep and surface structure (Chomsky 1967). I hope to show in this paper, however, that an exploration of the possible sensory and motor roots of language allows us to see its underlying mechanisms in a way that is fundamentally at odds with the view of language articulated by Chomsky.

2. EVOLUTION OF LANGUAGE

If Gregory's hypothesis is at least broadly correct, then how and why did the machinery of sensation and perception suddenly become utilised in this radically new way? Wilkins and Wakefield (in press) argue that the neural preconditions for language evolved slowly and that language emerged abruptly only when the brain achieved the appropriate internal configuration. They argue that language did not evolve directly from communicative-based precursors. They disagree with the contention of, for example, Lieberman (1992), that language arose as a result of selective pressures for improved speech capacity. In fact, Wilkins and Wakefield maintain that language capacity emerged 2 million years before the development of a modern vocal articulatory system. Instead, they propose that the relevant internal configuration giving rise to language was the emergence of regions of cortex in which information could be represented amodally, independent of any particular sense modality. Such regions came about, they argue, because of the need to coordinate information from different sense modalities in, for example, the performance of complex eye-hand coordination tasks. In particular, they arose as a result of the hominid shift to bipedalism and the resultant potential for the development of novel manipulative abilities. Wilkins and Wakefield identify one region in particular, the parieto-occipito-temporal (POT) junction, which Geschwind (1965) has called the "association area of association areas," as being implicated in the emergence of language. A central argument in Wilkins and Wakefield's thesis is that the POT region did not evolve to serve as the processing base for language, but rather was reappropriated for language, having evolved for visual-somasthetic-motor coordination. It will be argued in the rest of the paper that a better way of conceptualising the role of the POT is as a region of multi-modal rather than amodal interaction, and that this multi-modality is a core computational mechanism for language processing. The amodal/multi-modal distinction may seem a little artificial, since for modalities to interact there needs to be some form of common basis for interaction. What I am trying to emphasise here is that the modalities interact without sacrificing their essential sensory nature. Sensory and motor areas project to the POT, but there is no explicit translation of the sensory-motor information into some kind of neural *lingua franca*. In fact, the very form that this information takes, neuronal firing patterns, is itself a common cortical "language."