Linking the verbal and visual: new directions for corpus linguistics

Ronald Carter and Svenja Adolphs

School of English Studies, University of Nottingham

Abstract

This paper discusses an ongoing research project to investigate the compilation of a small corpus and the development of appropriate software tools that enable a more multi-modal approach to language data. The research draws on recent experience developed in the development of spoken corpora to explore alignments of the verbal and the visual and, as a starting point, does so with particular reference to gestures in communication and the role of head nods in particular. Issues of appropriate data capture and description are discussed alongside questions about the nature of language necessarily raised by language research that goes beyond the textual.

1. Introduction

Advances in the field of corpus linguistics over the past two decades have made it possible to develop computerised multi-million word databases of spoken and written language alongside powerful software tools to analyse this data quantitatively and qualitatively, a development that has contributed to pioneering research in many areas of communication studies and language description. However, while the analysis of large-scale text corpora can provide insights into language patterning and can help establish linguistic profiles of particular social contexts, it is limited to the textual dimension of communication. Communication processes are multi-modal in nature and there is now a distinct need for the development of corpora that enable the user to carry out analyses of both the speech and gestures of the participants in a conversation, and of how the verbal and non-verbal complement one another. In other words, corpus linguistics and discourse analysis might begin to be more closely aligned and descriptions made of rich contexts of language use of the kind advocated and illustrated by Michael Stubbs throughout his career.

1.1 Multi-Modal Communication

Recent work in multi-modal communication has seen advances in both theory and practice. The theoretical starting point for much significant work has been systemic-functional linguistics. Systemic linguistics is a theory that focuses on meaning, choice and probability in language and on the significance of language as a social phenomenon, underlining how particular choices of word, grammar and structure encode different meanings in different contexts of language in use.
Foundational work in multi-modal communication such as Kress and van Leeuwen (1996) has illustrated how choices of image can align with verbal choices and this work has been extended in recent years to embrace the multi-modal analyses of word, image and sound within different language varieties, including cartoons, comics, film, information leaflets, maps, advertisements (including TV advertisements), web pages and classroom textbooks (e.g. Baldry and Thibault, 2004, 2006). The emphasis has been on how choices of one image or camera angle or colour tone can cumulatively encode particular meanings. The almost exclusive focus has been on written text.

A particular challenge for current research is therefore to integrate the computer-enabled power of corpus linguistic methods, the theories and practices of multi-modal linguistic research and, with particular reference to the analysis of spoken discourse, the non-verbal signals of human gestures and bodily communication. In other words, one key aim is to provide computerised analyses of patterns of verbal and non-verbal meaning in ways that allow new understandings of textuality to emerge.

1.2 What is a Gesture?

Human communication functions within a variety of direct and indirect 'semiotic channels' (Brown, 1986: 409) which interact with, complement and 'counteract' each other (Maynard, 1987: 590). The occurrence of such channels is affected by modes of communication that differ widely according to their form, function and context-of-use (see foundational work by Argyle, 1969 and Ekman and Friesen 1969, 1976) and more recent studies by Wilcox 2004 and Gu, 2006). However, most studies have been undertaken within a research paradigm of psychology and in experimental rather than naturalistic conditions.

To date, experimental studies of the multi-modal nature of discourse have in general been designed to answer one or both of the following questions (Kendon, 1994: 177):

1 If recipients are offered utterances which include gestures and if they are permitted to see these gestures, do they interpret these utterances differently than when they are not permitted to see them? (examples of such studies include (Dobrogaev, 1929, reported in Kendon, 1980; Rogers, 1978; Riseborough, 1981).

2 If recipients are asked to make judgements about the gestures of others in the absence of speech to which they were related, do they make such judgements in a consistent way, and, if they do, do these judgements show that they have some understanding of the utterance of which they were a part?

Studies of gesture and the multi-modal nature of communication have focused upon gaze, (see Griffin, 2004 and Beattie & Shovelton, 1999, 2002) hand movements (see Rimé & Schiaratura, 1991 and Thompson & Massaro, 1986), head movements and other related gestures. In these studies the focus tends to be on language use in experimental conditions and does not embrace spontaneous,