

Francis Steen  
Communication Studies, UCLA

Alongside linguistic structure, discourse incorporates multimodal dimensions of communication. The integration of language, prosody, gesture, contextual reference, and appeals to recalled and imagined events implies that communication taps into rich multimodal cognitive resources. The sustained deployment of still and moving images in culturally elaborated technologies of communication, from cave paintings to television, is a vivid testament to the power of the underlying cognitive machinery of discourse to adapt to new formats. Visual communication is at once evolutionarily novel, requiring conceptual innovation to become effective and intelligible, and at the same time necessarily grounded in, enabled by, and building on pre-existing structures. It should thus be possible to leverage our existing understanding of discourse into the languages and conventions of visual communication in art, film, television, and the internet. In addition, to the extent that technologies of communication externalize and make tangible some of the backstaged features of cognition, such a project carries in it the potential to illuminate the base case of interpersonal discourse itself.

While cognitive linguists have long appreciated that communication is multimodal, the obstacles to systematic study are formidable. A detailed investigation of the power and practice of multimodal communication requires the development not only of new corpora, but also of new tools and methodologies. The Communication Studies Department at UCLA, in collaboration with the UCLA Library, has initiated the task of assembling a large multimodal corpus of television news for research and education. The news cover important real-world events in a professional manner from multiple perspectives, permitting systematic analyses of effective communicative practices. The ongoing collection is approaching 165,000 programs from the US and other countries, including timestamped video, audio, and text. In addition, we have developed search functions, navigational interfaces for accessing video from transcripts, tools for rapid coding, and an online video analysis tool. Cross-disciplinary research teams from Computer Science, Statistics, Cognitive Science, Linguistics and Communication Studies are already working on low-level text and image parsing as well as higher-level pattern discovery.

As a case study, I will present an analysis of how joint attention (e.g., Tomasello et al. 1993) is deployed in news reporting. I focus on the role of the anchor as a personification of the technologies of television, a role onto which the team-based techniques of visual presentation are tacitly projected. Through an examination of gesture, gaze and head direction, visual transitions, maps, and appeals to possible worlds, I present evidence that joint attention in television news is an emergent and counterintuitive conceptual blend that builds on our naive conceptions, but deploys them in a manner that is evolutionarily novel and innovative.