

**Do grammars do best what speakers did most: the Social Cognition Parallax Interview
Corpus (SCOPIC) cross-linguistic corpus on social cognition in grammar**

Nicholas Evans
Australian National University

‘Languages do best what speakers do most’ is a famous functionalist motto, attributed to Jack Du Bois. Few would disagree that it is a key insight in understanding how language structures evolve. But this elegant formulation conceals two deep questions. Firstly, by being formulated in the present it conceals what may be a very slow time-scale for the evolution of structures. Does the fact that English and Japanese have a grammaticalised past tense, while Indonesian does not, reflect something about what is going on in the speech of people today, or in the past? Hence the ‘did most’ of my title. Secondly, are ‘grammars’ and ‘speakers’ intended as a single homogenous group across the world’s cultures, or do different effects get played in different cultures (as has been argued since Vico and Herder, and more recently in studies of ‘Ethnosyntax’ by Enfield and others).

These are fundamental questions for our understanding of why languages differ so much from each other. To answer them, for a targeted domain of language (social cognition), a team of us have created the Social Cognition Parallax Interview Corpus (SCOPIC). This grows out of a project involving Alan Rumsey, Barb Kelly, Andrea Schalley, Lila San Roque and others, which uses a series of picture cards and a narrative problem-solving task (the Family Problems Picture Task) to elicit rich talk about social cognition – in descriptive, conversational and narrative modes – in a way that avoids privileging the categories of any one language in the way that can happen with questionnaires or parallel translation tasks. In other words, speakers of each language are left free to highlight whatever it is about the situations they consider worthy of emphasis.

SCOPIC is now being built as a large interrogable corpus by Danielle Barth and myself. SCOPIC is an open-ended, accessible corpus that balances the need for language-specific annotation with typologically-calibrated markup. It provides richly annotated data, focusing on functional categories relevant to social cognition, the social and psychological facts that place people and others within an interconnected social context and allow people to interact with one another. By ‘parallax corpus’ we mean ‘broadly comparable formulations resulting from a comparable task’, to avoid the implications of ‘parallel corpus’ that there will be exact semantic equivalence across languages.

In this talk I will sketch out the most interesting features of social cognition as a domain for semantic typology, illustrate how the Family Problems Picture Task works, describe the technical structure of the corpus, and show an example of a typological analysis, using recursive partitioning, a modern statistical technique.