Corpus analysis of workplace discourse: The case of the construction industry

Michael Handford (Cardiff University, UK)

This paper will discuss how spoken workplace discourse[1] can be analysed using corpus tools, while arguing that corpus software tools alone cannot provide sufficient insight into workplace interactions to answer many of the important questions about meaning in context. It argues that, along with keywords, collocations, concordance lines and discourse prosody, a further 'tool' that needs to be used in corpus analysis of workplace discourse (CAWD) is the 'understanding context' tool (Handford, 2017). This can be achieved through triangulating with other data sources, and building a plausible interpretation of the interaction in question.

The data analysed here is the 'CONIC' Construction Industry Corpus (JSPS Grant No. 25370423), a corpus of over 350,000[2] words of largely professional English as a Lingua Franca spoken interactions recorded in Singapore, Vietnam, Bangladesh, Hong Kong, Japan and the UK, and supplemented with a wealth of other data. It is these other data that facilitate a plausible contextual analysis of the interactions, including at the co-textual, non-verbal, situational and wider sociocultural levels. These data include several hours of onsite and meeting video data transcribed with ELAN, over 20 hours of interviews with more than 25 professional engineers and other expert informants, observation notes, as well as several written documents (e.g. work plans, meeting agendas, presentation slides and written contracts), workplace images (e.g. signs, photographs and diagrams), recordings of focus group discussions with professionals, and training materials. It will be argued that the collection of such data is a requisite of doing CAWD.

The talk draws on Candlin's exhortation that, when doing professional discourse analysis, we should seek to move beyond mere description of the data to a more explanatory position. Analysis thus moves beyond describing what, to making plausible explanations of *how* and *why*: 'Any analysis of text which aspires to some explanatory rather than merely descriptive adequacy presupposes an engagement with social action within the context of the institution in question, and needs to take account of the distinctive perspectives of the involved participants (including the researcher)' (Candlin, 2006: 6). In other words, we need access to scales of context that may not be evident in the co-text if we are to build on textual patterns, and we need to provide emic accounts of communicative action.

The exploration of the relationship between text and context is a crucially important element of the analysis of professional discourse, because without background knowledge it is often impossible to understand what the interaction or document is about (Charles and Charles, 1999). While the

---

[1] Workplace discourse is used here as an umbrella term, capturing business, institutional and professional interactions (Koester, 2010).
[2] As such, the corpus may be the largest corpus of ELF professional interactions to date, and is a relatively substantial corpus of professional spoken discourse.

relationship between text and context is, of course, central in all types of discourse analysis (Gee, 2005; Flowerdew, 2014)[3], it is particularly so in professional and workplace contexts because of the high degree of shared knowledge and practices in workplace communities; linguistically, there is a prevalence of both jargon and highly specific references on the one hand and vague or explicitly deictic language on the other – both of which present challenges for the analyst. This is the reason why specialised corpora are potentially more suitable for research into professional discourse than large corpora.

Extract 1 provides an example of the importance of context in understanding professional or workplace texts, and touches of some of the findings to be discussed in the talk.

Extract 1
Speaker 1: And then and (I go to this here) they can connect + here. . .
Speaker 2: Hmm
Speaker 1: . . .with (an opposite) side. . .
Speaker 2: Hmm
Speaker 1: . . .and then and then we measure there
Speaker 2: Hmm
Speaker 1: and there
(Handford and Matous, 2011: 96).

Without any contextual information it is impossible to make much sense of this interaction, or why it might be worth analysing. In fact, it is taken from a large construction project in Hong Kong, and speaker one is a male Japanese engineer in his early 30s responsible for a part of the project. He is explaining a procedure that needs following to the foreman, a Hong Kongese male in his early 50s, as the foreman needs to communicate this to the local subcontracted labourers. The interaction takes place in the Portacabin on the site. There is a high degree of place deixis present (e.g. 'here' and 'there'), and the analysis of video data shows that speaker 1 is drawing a picture with a pencil to demonstrate the procedure. Interviews with the engineer and foreman shed further light on the motivations for this interaction, in other words an explanatory level of analysis, such as the higher probability of cooperation from the sub-contracted labourers if the foreman issue such directives, rather than the engineer. The engineer, in contrast to the foreman and the labourers, is a permanent employee of the contractors, which has implications for the way he and his actions are perceived. As such, the interaction is interesting in the way power is negotiated in workplace interactions: while the engineer is more senior in terms of status, he is reliant on the foreman for the completion of this task.

In order to explore the issues discussed above, the following questions will be explored in this talk:

---

[3] Discourse analysis is used here as an umbrella term for any type of analysis, including corpus linguistics, that seeks to understand how people create and interpret meanings in actual situations.

1. What lexicogrammatical features are statistically significant in the CONIC corpus, and how do these patterns compare to other professional contexts?
2. How can we interpret, and not merely describe, such patterns, and what do they imply about the construction-industry context?
3. What might some of the implications from the above question be for training in the industry?

The talk will show how deictic features are statistically significant (Scott, 2011) in construction discourse (Emmitt and Gorse, 2003; Handford, 2014) when compared to everyday discourse [4], unlike many other professional contexts. Through analysis of video data, the types of non-verbal communication used in conjunction with such items will be outlined, along with insights from interview data and other sources as offering explanations as to why such patterns are found. It will also be shown how such an approach allows for insights into other discursive practices, as well as the arguably different expectations towards seemingly face-threatening language (Tsuchiya and Handford, 2014), and the considerable communicative, interpersonal and intercultural challenges specific to this industry (see also Tijhuis and Fellows, 2012; Handford 2014).

References

Candlin, C. (2006). Accounting for Interdiscursivity: Challenges to Professional Expertise. In *New Trends in Specialized Discourse Analysis*, Gotti, M & D. Giannoni (Eds). Brussels: Peter Lang.

Charles, M. & Charles, D. (1999). Sales Negotiations: Bargaining through Tactical Summaries." In M. Hewings and C. Nickerson (Eds), *Business English: Research into Practice*, pp. 71 – 99. London: Longman.

Emmitt, S. & Gorse, C. (2003). *Construction Communication*, Oxford: Blackwell.

Gee, J.P. (2005). *An Introduction to Discourse Analysis*. London: Routledge.

Flowerdew J. (Ed) (2014). *Discourse in Context.* London: Continuum

Handford, M. (2014). Construction Industry Discourse. In Bhatia, V. and Bremner, S. (Eds.). *The Routledge Handbook of Professional Discourse*. Abingdon: Routledge.

Handford, M. (2017). Corpus Linguistics. In Vine, B. (ed) *The Routledge Handbook of Language in the Workplace.* Abingdon: Routledge.

Handford, M. & Matous, P. (2011). "Lexicogrammar in the international construction industry: A corpus–based case study of Japanese–Hong-Kongese on-site interactions in English." *English for Specific Purposes*, 30 (2), 87–100.

Koester, A. (2010). *Workplace Discourse*. London: Continuum.

Scott, M. (2011). *Wordsmith Tools*, Version 5. Oxford: Oxford University Press.

---

[4] Using the VOICE corpus of ELF interactions, and also a sub-corpus of the CANCODE corpus as reference corpora.

Tsuchiya, K. & Handford, M. (2014). "A corpus-driven analysis of repair in a professional ELF meeting: not 'letting it pass'." *Journal of Pragmatics*, 64, 117–131.

Tijhuis, W. & Fellows, R. (2012). *Culture in International Construction*. London: Spon Press.