

Introduction

Patterns, meaningful units and specialized discourses

Rainer Schulze and Ute Römer

University of Hanover / University of Michigan

It would be difficult to imagine someone in Corpus Linguistics who was not interested in multi-word units or patterns in language; someone who, for one reason or another, was never engaged in investigating the real nature of linguistic structures or even the nature of linguistic messages in general. Such a person would have failed to ponder how language users draw on the context of situation when communicating (cf. Malinowski 1923:305) or how linguistic events and the roles performed by the participants contribute to a proper understanding of interpersonal exchanges (cf. Firth 1957). In short, it would be a person who is not concerned with language in performance or with what actually happens in the discourse (and it is only fair to admit that there are still some linguists who renounce performativity of language as a cultural staple in favour of formal and decontextualised representations of language).

Unquestionably, the linguists contributing to this issue all express an intense interest in spoken and written discourse, including language in performance and its relation to the user. Their primary concerns are the very close relationships between vocabulary and grammar, the emphasis on real language or authentic data, and the preference to analyse language on the basis of large computerized data collections.

The importance of the performativity aspect resides in the fact that it reiterates an argument put forth by Austin who pursues the idea that speakers and writers do things with words (1962, 1975), and we are tempted to add that speakers and writers do things by predominantly and unconsciously employing patterns or phraseological items, i.e. strings of words that are highly structured, well-organised and firmly entrenched in the human being's mind. A central question addressed in this context concerns how it is that interactants create and understand meaning through patterns and phraseological items and how these strings of words give shape to our beliefs and values, social and cultural structures (Robinson 2006:8). Taking for granted that the collaborative work of interactants gives patterns their

power to shape reality (see the contributions in Schulze (Ed.) 1998), Corpus Linguistics should be able to show and unravel how and in what ways select strings of words serve specific discursive purposes or how interactants perform implicit or explicit verbal actions in different and specialized contexts of situation.

Despite its centrality in Corpus Linguistics, the scrutiny of patterns or phraseological items in specialist discourse domains is still in its infancy (although some invaluable findings are offered by Biber (2006), Biber, Connor and Upton (Eds.) (2007), Bowker and Pearson (2002), Gavioli (2005), Gledhill (2000), Römer (2008), in the contributions to Connor and Upton (Eds.) (2004), or research projects centring around MICASE, the Michigan Corpus of Academic Spoken English (e.g. Simpson-Vlach & Ellis forthcoming) and has commonly been conducted from a stance that relates specialized discourses to supplementary language skills which are necessary for understanding and producing text in specialized situational settings in order to satisfy particular communicational needs. This approach is firmly associated with aspects of (foreign) language teaching for professional purposes or issues of vocational training, as shown in Bloor and Bloor (1986), Dudley-Evans and St John (1998), Hutchinson and Waters (1987) or Robinson (1991).

The stance adopted in this issue, however, is different: small and large specialized corpora and uses of specialized language becoming manifest in these corpora are put at centre stage. This approach places more weight upon the specialized meanings and their relationships expressed and encoded by patterns or phraseological items, becoming inseparably intertwined with particular domains in which they are being produced. In a nutshell, this view on uses of specialized language is more global, encompassing social as well as interactional and cognitive issues (in the sense of knowledge maintenance and knowledge transfer). This is also to say that the dialectal relationship between specialized language and (discourse) domains is vital: specialized language on the one hand is constitutive of the domain, and the domain on the other hand affects or even shapes specialized language in that it provides knowledge about the specific needs to be covered in such specialized language.

The present issue covers a range of different domains and showcases studies based on text types and corpora such as the *New Scientist* sub-corpus of the Bank of English (Hunston), MICASE, the LIBRA and HEM corpora, covering academic lectures, academic book review articles, and research articles both from the Natural Sciences and the Humanities (Diani), financial texts from an individually assembled collection including *The Financial Times* and *Il Sole 24* (Forchini and Murphy), written political language in the *News Discourse Corpus* (NDC), assembled out of selected articles from the American English-based *The Wall Street Journal* and the British English-based *The Economist* (Milizia and Spinzi), and the language of tourism captured in a corpus of texts culled from British *Farmhouse Holidays* and Italian *Agriturismi* websites (Manca). Concerning different levels of

expertise in these domains, relations between the communicators vary, including expert-to-expert, expert-to-layperson or expert-to-consumer communication. More specifically, the discourses under investigation include language that shows a medium to high degree of abstraction and in places a considerable amount of standardised terminology.

The viewpoint offered in this issue grows out of a general interest in Anthropology (Malinowski), Sociology (Bernstein), and functional or neo-Firthian Linguistics, predominantly of British provenance (e.g. Sinclair, Hunston, Francis, Moon, Hoey, Stubbs, and Wray), offering insights in aspects of more global models of communication in (non-)specialized settings, combining discourse-internal and discourse-external factors in the descriptions and focusing upon linguistic explanations that reject a strict division of labour between lexis and grammar. The lexis-grammar interface in particular is seen as being fundamental and instrumental in the construction and preservation of human (specialized) knowledge.

The contributions

The five articles in this issue provide a varied view on essentially two issues:

1. the degree to which special and specialized information is entrenched in linguistic structures and, hence, found in different domains and
2. the way special and specialized knowledge is encoded, expressed and organised in specialized corpora.

The first paper by Susan Hunston is concerned with the role of select units of meaning in the *New Scientist* corpus that forms part of the Bank of English. The paper presents a new, thought-provoking type of access to corpus research work in that it introduces the notion of ‘semantic sequence’, which, on the surface, is partially synonymous with Sinclair’s (1991) ‘unit of meaning’, but, on closer inspection, reveals that the focus on “sequences of meaning elements rather than [...] formal sequences” (Hunston, this issue) is not a focus on language alone, but also on extra-meanings evoked by strings of words such as *to make sure*, the ‘N that’ pattern or phrases preceded by a preposition that can be found in the *New Scientist* sub-corpus in large numbers. Similar to findings in cognitive semantics, e.g. those made by Talmy (2000), grammatical or grammar words are seen as being essential and thus meaningful in a corpus, contributing to the structure, skeleton or scaffolding for a (non-)linguistic scene, across which the rich “lexical” content of lexical words can be “drawn”. Or, put differently, lexical words contribute to the subject matter of a special discipline, while grammatical or grammar words contribute to the epistemology of specialized discourse.

In a similar vein, Giuliana Diani's contribution supports the significance of micro- and macro-structural analyses of the stance adverbial *really* in different registers and disciplines. Her quantitative results suggest that the frequencies, meanings and uses of *really* vary across spoken and written academic discourse. Evidence for this "variationist" view of select items is found in discourse captured in MICASE, the LIBRA corpus (a corpus of Linguistics Book Review Articles) and the HEM corpus (of History, Economics, and Marketing articles).

The approach taken by Pierfranca Forchini and Amanda Murphy is concerned with a select string of words (the 4-gram *at the end of*), both from a cross text-type (general vs. specialized English) and from a cross-linguistic (English vs. Italian) point of view. Assumptions drawn from the analysis are supported by linguistic evidence from the quality dailies *The Financial Times* and *Il Sole 24*. A revealing and stimulating account of translational and educational implications serves as additional evidence and rounds off an investigation into firmly entrenched constructions, both in English and Italian. Readers are invited to pursue the issue of meaning specification or even grammaticalization further, keeping an eye on the diachronic dimension of n-gram variation with all the concomitant priming implications.

Denise Milizia and Cinzia Spinzi provide a detailed account of the linguistic and rhetorical strategies being employed in public speeches by George W. Bush and Tony Blair. More specifically, they look into particular collocation partners of *terror* in these public addresses, paying careful attention to contiguous and non-contiguous collocations surrounding *terror* (as in *a steadfast ally in the global war on terror* versus *an ally in the war on terror*). The search engine ConcGram has been essential in isolating all the requisite word associations found in BBB (a 10-million-word corpus including speeches of Bush, Blair and Berlusconi). Comparative analyses of selected word association patterns (around *terror* and related items) in corpora compiled from *The Wall Street Journal* and *The Economist* demonstrate that prefabricated structures seem to be more strongly governed by the idiom principle (Sinclair 1987) in spoken than in written political discourse. This observation has a cognitive or mental dimension in that processing time and processing effort seem to play a decisive role in real speech in which the presentation of prefabricated chunks is preferred to non-prefabricated material.

The final paper by Elena Manca considers the co-occurrence of English and Italian adjectives and semantic fields such as 'description of rooms', 'description of surroundings' and 'description of food' in texts being drawn from British farmhouse holidays websites and Italian agriturismo websites. This cross-linguistic analysis which shows the highly phraseological nature of the special language of tourism is complemented by a cross-cultural perspective that utilizes the well-established distinction between high versus low context cultures to demonstrate that the two

languages of tourism are significantly different in coding specialized scenery. A result of Manca's observations is that British English tourism language is more content-oriented and expressed in plain language, with a firm focus on the now, whereas the language of tourism in Italian seems to be much more form-oriented, less plain in style and with a clear focus on the past, not on the now. This is to say that the relationship between language and world-view is bound to an intricate discursive interplay between linguistic cues, cognitive representations and motivated actions. Manca's paper is delivered within the framework of John Sinclair's ideas on the influence of context and register on particular linguistic choices.

All in all, this issue attempts to bring together some recent work in the area of Corpus Linguistics (putting some new-generation software tools and corpora to use) with a clear focus both on specialized domains and discourse in specialized settings. The approaches and the data in the following papers, more than anything else, suggest a wealth of innovative approaches to the study of recurrent features in specialized languages. A specific intention addressed by all the papers in this issue is to focus on spoken and written texts as a collective memory bank and the central object of study, assembled and taken from a range of different corpora. There is overwhelming consensus that the traditional disjunction between lexis and grammar is convincingly and vehemently rejected in favour of a phraseological viewpoint according to which the linguistic organisation, representation and handling of specialized information and knowledge is inseparably connected with the needs of the interactants, irrespective of different communicative situations in which sender and receiver or addresser and addressee have different levels of knowledge concerning the specialized domain of communication.

Acknowledgment

We are very grateful to a number of colleagues who helped us in the peer reviewing process for this special issue and provided insightful comments on earlier versions of the selected articles. We would also like to thank Susanne Röltgen for her diligent assistance in copy-editing the texts for this issue.

References

- Austin, J. L. (1962, 1975). *How to Do Things with Words*. Oxford: Oxford University Press.
- Biber, D. (2006). *University Language: A Corpus-Based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.
- Biber, D., Connor, U. & Upton, T. A. (2007). *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam: John Benjamins.

- Bloor, M. & Bloor, T. (1986). *Languages for Specific Purposes: Practice and Theory*. Dublin: Trinity College, Centre for Language and Communication Studies.
- Bowker, L. & Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. New York and London: Routledge.
- Connor, U. & Upton, T. A. (Eds.). (2004). *Discourse in the Professions: Perspectives from Corpus Linguistics*. Amsterdam: John Benjamins.
- Dudley-Evans, T. & St John, M. J. (1998). *Developments in English for Specific Purposes*. Cambridge: Cambridge University Press.
- Firth, J. R. (1957). *Papers in Linguistics 1934–1951*. London: Oxford University Press.
- Gavioli, L. (2005). *Exploring Corpora for ESP Learning*. Amsterdam: John Benjamins.
- Gledhill, C. (2000). *Collocations in Science Writing*. Tübingen: Gunter Narr.
- Hutchinson, T. & Waters, A. (1987). *English for Specific Purposes*. Cambridge: Cambridge University Press.
- Malinowski, B. (1923). The problem of meaning in primitive languages. In C. K. Ogden & I. A. Richards (Eds.), *The Meaning of Meaning* (pp. 296–346). New York: Routledge & Kegan Paul.
- Robinson, D. (2006). *Introducing Performative Pragmatics*. New York and London: Routledge.
- Robinson, P. (1991). *ESP Today: A Practitioners' Guide*. New York: Prentice-Hall.
- Römer, U. (2008). Identification impossible? A corpus approach to realisations of evaluative meaning in academic writing. *Functions of Language*, 15 (1), 115–130.
- Schulze, R. (Ed.). (1998). *Making Meaningful Choices in English: On Dimensions, Perspectives, Methodology and Evidence*. Tübingen: Gunter Narr.
- Simpson-Vlach, R. & Ellis, N. C. (Forthcoming). An academic formulas list (AFL). *Applied Linguistics*.
- Sinclair, J. McH. (1991). *Corpus Concordance Collocation*. Oxford: Oxford University Press.
- Sinclair, J. McH. (1987). The nature of the evidence. In J. McH. Sinclair (Ed.), *Looking Up: An Account of the COBUILD Project in Lexical Computing* (pp. 150–159). London: HarperCollins.
- Talmy, L. (2000). *Toward a Cognitive Semantics*. 2 vols. Cambridge, MA: MIT Press.

Editors' addresses

Rainer Schulze
 Department of English
 Leibniz University of Hanover
 Königsworther Platz 1
 30167 Hanover
 Germany
 raischulze@gmx.de

Ute Römer
 English Language Institute
 University of Michigan
 500 E. Washington St.
 Ann Arbor, MI 48104–2028
 U. S. A.
 uroemer@umich.edu