Patrick Hanks probably no longer needs any introduction to lexicographers and lexicologists, and especially not to readers of the *International Journal of Lexicography*. He has published at least one article or book review every year in IJL since 2004 and his regular contributions to Euralex conferences and to many other journals provide clear evidence that this prolific author cannot be ignored as soon as one discusses modern lexicography. This brief book review will not attempt to describe the multiple facets of Hanks’s contributions to lexicography, lexical semantics, computational and corpus linguistics, and the study of word meaning in general. I encourage the reader to read Gilles-Maurice de Schryver’s excellent introduction to the Festschrift he edited in honour of his friend on the occasion of his 70th birthday. De Schryver rightly points out that ‘Hanks is a linguistic theorist and empirical corpus analyst, also an onomastician, but above all he is a lexicographer’ (p.4). For linguists who, like me, have always taken a lot of pleasure in reading Hanks’s papers on phraseology, on idioms, on metaphors, on word associations, on collocations and collocation extraction, on dictionary definitions, on corpus pattern analysis, on linguistic norms and exploitations, or on proper names, it may be too easy to forget that he is primarily a lexicographer, indeed, and that he has played a pivotal role in several major dictionaries of the English language, including the Collins Dictionary of the English language (1979), the Collins COBUILD English Language Dictionary (1987) and the New Oxford Dictionary of English (1998). Beyond these direct contributions to lexicography, it is safe to claim that if today’s dictionaries, especially learners’ dictionaries, are better at presenting collocational material and phraseological data, it is largely due to the fact that, in 1989, Patrick Hanks showed us how to discover the most significant collocates in a seminal and influential paper he wrote with Ken Church (Church and Hanks 1989).

This Festschrift is divided into three parts: a theoretical section, a computational section, and a lexicographic section. The editor’s introduction sketches Patrick Hanks’s rich career and includes the very long list of his publications. The first part contains five papers, including one by the late John Sinclair (‘Defining the definiendum’), which is published in its unfinished state. Sinclair was writing it when he died unexpectedly in March 2007: he argues that the treatment of multi-word units of meaning should be extended and they should be given the same status as the usual headwords. Yorick Wilks (‘Very large lexical entries and the boundary between linguistic and knowledge structures’) discusses the incorporation of richer semantic structures into the Preference Semantics system. James Pustejovsky and Anna Rumshisky examine some of the mechanisms at work which relate distinct senses of a predicate in
‘Mechanisms of sense extension in verbs’. Igor Mel’čuk then discusses the ‘government pattern in the Explanatory Combinatorial Dictionary (ECD)’, and shows how the semantic actants of a lexeme are implemented on all levels of linguistic representation. The ECD being the core component of Mel’čuk’s Meaning-Text Theory, it is there that information about the syntactic valence of lexical items will be stored. In the last paper of the theoretical section, entitled ‘The paradox of analysis and the paradox of synonymy’, David Wiggins then adopts a more philosophical approach to the analysis of dictionary definitions and the presumed (and analytically uninformative) identity between the definiens and the definiendum (which he calls analysans and analysandum).

In the first paper of the section devoted to ‘computing lexical relations’, Ken Church reminds us that, when many lexicographers study butterflies, Patrick Hanks likes ‘central and typical’ moths (which explains why he and Patrick first applied the calculation of Mutual Information statistics to extract collocations from large corpora). In ‘More is more’, Church argues that it is a good thing that everyone now has more access to lots of (linguistic) data than ever before, and that, for some applications, very large collections of texts may turn out to be more useful than small, balanced corpora. Gregory Grefenstette then tries to use the web as a huge corpus to estimate the number of multi-word concepts used in English. The rough estimate he arrives at (about 200 million two-word combinations in common use on the web) gives an idea of the scale computational lexicographers will have to deal with in the future when they try to model the concepts that these combinations express. In the third paper of the section (‘Identifying adjectives that predict noun classes’), David Guthrie and Louise Guthrie try to make use of adjectives to automatically disambiguate nouns: some adjectives indeed seem to prefer nouns that are, say, animate or inanimate and the authors’ goal is to find sets of adjectives that are predictive of a given semantic category or of a class of named entities. In ‘Statistical variations of German support verb constructions in very large corpora’, Alexander Geyken uses corpus data (a balanced 100-million and an opportunistic 1-billion-token corpus of German) to compare verb-noun expressions (more specifically Nominalisierungsverbgefüge, such as eine Entscheidung treffen, ‘make a decision’) in a German dictionary and identify statistically salient collocations that are missing in the dictionary. In the next paper, Karel Pala and Pavel Rychlý discuss errors that can be found in word sketches produced with Adam Kilgarriff’s Sketch Engine, a powerful lexicographic tool that combines various types of statistical techniques to explore the collocational behaviour of words in large corpora. Given that the performance of the Sketch Engine (and hence the quality of the one-page summaries of collocations it creates) largely depends upon the quality of the underlying grammatical tagging and upon the quality of the formal rules that describe the syntactic relations between sentence constituents, Pala and Rychlý decide to investigate the errors related to
the Czech word *vidět* (‘see’). The overall quality of the Czech version of the tool could be improved by using more training data (to improve the performance of the POS tagger) and by adding heuristic linguistic rules. The last two papers of this section both revolve around Hanks’s Pattern Dictionary of English Verbs (PDEV, Hanks, 2007). In ‘The lexical population of semantic types in Hanks’s PDEV’, Silvie Cinková, Martin Holub and Lenka Smejkalová discuss the assignment of semantic type labels to noun collocates of verbs. In the PDEV, Hanks distinguishes between semantic types, which describe inherent properties of collocates (e.g. Human, Document, Artefact...), and semantic roles, which describe properties assigned to the word in a particular context. The ultimate aim of the PDEV is to map meanings onto patterns of use by means of Corpus Pattern Analysis (CPA). In ‘From pattern dictionary to Patternbank’, Elisabetta Jezek and Francesca Frontini discuss the tension between the semantic types associated with verb arguments and the lexical sets that may fill the different argument positions. They propose an annotation interface to allow the human annotator to link each argument in the pattern to the corresponding portion of text in the corpus instance.

The third part of the Festschrift is entitled ‘Lexical Analysis and Dictionary Writing’. Rosamund Moon carries out a corpus study of the idiom ‘spring to mind’ and discusses the behaviour of this idiom in context. Patterns of usage show that it is a convention for introducing ideas and evaluations, she argues. She also analyses the idiom in terms of the cognitive and Lakoffian models of metaphors (Lakoff and Johnson 1980). In the second paper of this third section, Sue Atkins describes ‘The DANTE database: its contribution to English lexical research, and in particular to complementing the FrameNet data’. The description of this new DANTE lexical database is fascinating and its comparison with the FrameNet database shows both the differences and the similarities between the two projects. The emphasis on the formalization of some types of collocations in DANTE (such as itemisers for units of a substance denoted by a mass noun – *piece/lump/strip/scrap... of metal*) is only one such difference and the project she sketches out aiming at mapping DANTE entries onto the FrameNet data is certainly a most interesting challenge computational lexicographers might be interested in taking on. In ‘Semi-automatic dictionary drafting’, Adam Kilgarriff and Pavel Rychlí try the Sketch Engine to try to cluster corpus instances and apply Hanks’s theory of norms and exploitations to perform word sense disambiguation, a sine qua non if one wants to create ‘intelligent’, disambiguating dictionaries that, for instance, only display the relevant portion and the correct sense of a polysemous dictionary entry in a dictionary look-up perspective. In ‘Lexicography: science without theory’, Paul Bogaards wonders whether there is any such thing as a lexicographic theory. His analysis of the literature shows that there does not seem to be any real theory, but that a good theory of lexicography is longed for in order to improve the practice of dictionary writing. He argues that lexicography should be...
studied as a separate science drawing upon other theories to study the three basic dimensions of dictionaries: data, users and the way users access data. The next paper, by Miroslaw Ban’ko, discusses ‘the Polish COBUILD dictionary and its influence on Polish lexicography’, i.e. a general-purpose dictionary which may not have become a commercial success, but whose lexicographic innovations, based upon the British COBUILD model, exerted some influence on a number of other Polish dictionaries which also adopted the full-sentence definition approach. In ‘ARGOT: The flesh made word’, Jonathon Green, who specializes in the making of slang dictionaries, writes about the development of Francophone slang and notes that argot, criminal slang, remains the primary subject of French ‘slang’ dictionaries. Finally, in the last paper of the Festschrift, ‘Defining elegance’, Michael Rundell tries to clarify the concept of ‘elegance’ when applied to dictionary definitions and shows that this concept indeed has relevance to the lexicographer’s craft. For Rundell, there is a clear connection between elegance and stylishness/good taste, on the one hand, and simplicity, economy, brevity and the avoidance of redundancy, on the other. In addition, elegance is also a quality that is visible when a method is used that does something difficult in a simple and clear way (always think about prioritizing the needs of the users, he reminds us).

‘A Way with Words’ is a compilation of a number of very interesting and thought-provoking essays. Like in any Festschrift of this type, the contributors have been selected because of their special connection with the person who is celebrated. However, what I really liked in this particular instance is that this connection is very visible in the papers themselves. The authors have gone to great lengths to explain how their paper is related to Patrick Hanks’s contribution to lexicology and lexicography (with the exception of the field of proper name dictionaries, which is not represented in the book). Given Patrick’s rich and multi-faceted career, beautifully described by the editor in his introduction, it is not surprising to find so much inspiration in the various papers included in this volume, since, after all, as Michael Rundell points out, ‘we are standing on the shoulders of a giant’.

References


Thierry Fontenelle
Translation Centre for the Bodies of the European Union, Luxembourg
Thierry.Fontenelle@cdt.europa.eu
doi:10.1093/ijl/ecr012
Advance access publication 10 June 2011