

Getting to the Bottom of How Language Works

Gilles-Maurice de Schryver

*With Patrick around, it always feels that bit more likely that we shall
get to the bottom of how language works.*

— Adam Kilgarriff¹

1. Genesis of this book

This book originated in 2006, when some friends, colleagues, and admirers of Patrick Hanks got together at the prompting of one of them and agreed to compile a Festschrift for him. It is a pleasure to pay tribute here to the original editors – Gregory Grefenstette, Ramesh Krishnamurthy, Karel Pala, and James Pustejovsky – and to the constant enthusiasm and unwavering support of a few early contributors, in particular Igor Mel'čuk and David Wiggins. Special mention must be made here of Anne Urbschat, who first broached the idea and heroically carried it forward virtually unaided for three years. In the summer of 2009, Anne challenged me to take over the project, knock it into shape, and find a publisher.

When I came to examine the contributions that Anne had already solicited and obtained, I recognized immediately that this was not a mere ragbag of dutiful tributes by a small coterie of colleagues, but a major contribution to understanding the lexicon (an area of linguistics of great current interest), by an unusually wide range of leading lexicographers, lexicologists, philosophers of language, and computational linguists. More specifically, the contributions address issues in the analysis of corpus data and definition writing from both a theoretical and a practical point of view. I solicited some additional contributions and rejected a few that did not seem to fit in well with the theme of the book.

2. Patrick Hanks's contribution to lexicography, corpus linguistics, and lexical theory

Patrick Hanks is a linguistic theorist and empirical corpus analyst, also an onomastician, but above all he is a lexicographer. He has a way with words. And yet, in many of his theoretical writings, he proposes that lexicographers need to 'do away with words' in order to focus on phraseology. In the words of one of his good friends:

He's the ideal lexicographer's lexicographer. I like the way he thinks about words.

— Sue Atkins

Patrick has played a central role in editing no less than four major, highly original dictionaries of the English language – the Hamlyn *Encyclopedic World Dictionary* (1971), the *Collins Dictionary of the English Language* (1979), the *Collins COBUILD English Language Dictionary* (1987), and *The New Oxford Dictionary of English* (1998) – as well as three important dictionaries of personal names (for these, cf. Section 3).²

However, it is for his work as a corpus linguist, building on foundations laid by the late John Sinclair, for which he will probably be best remembered. Having hacked his way through literally hundreds of thousands of corpus 'vines' for three decades now, Patrick has come to answer the question 'Do Word Meanings Exist?' with 'Yes, but ...':

- Yes, but traditional dictionaries give a false impression. What dictionaries contain are (more or less inaccurate) statements of *meaning potentials*, not meanings;
- Yes, but only *in context*;
- Yes, but the meaning potential of a word consists of a cluster of *semantic components*, only *some of which are activated* when the word is used in context. (cf. Hanks 2000c)

A decade on, his work has morphed into the Theory of Norms and Exploitations, 'a principled approach according to which exploitations can be identified as such and set on one side, leaving the distinctive patterns of normal phraseology associated with each word to stand out more clearly' (Hanks forthcoming d; but see also Hanks 2004a).

Certainly, Patrick's most famous publication is 'Word Association Norms, Mutual Information, and Lexicography,' written jointly with Ken Church, which reintroduced statistical methods of lexical analysis into linguistics and which emphasized the importance, for applications such as lexicography, language teaching, and NLP, of measuring the statistical significance of word associations in corpora.³

The Church and Hanks paper had a galvanizing effect on the computational linguistics world in 1989, when it was presented at the 27th *Annual Meeting of the Association for Computational Linguistics* (ACL) in Vancouver. There, it was the only paper to discuss statistical methods in computational linguistics, while at most (if not all) previous meetings of ACL there were none. Nowadays, such papers at ACL are in the majority.

The paper has attracted occasional hostile criticism, but seemingly only by people who feel threatened by it. For example, some people have proposed log-likelihood measures as a means of compensating for the so-called 'sparse data problem.' Computer scientists seem to like log-likelihood very much – it is elegant. But it has not been used in lexicography because it typically produces results that are less useful, practically speaking, than MI score (the statistical measure used by Church and Hanks in their 1989 paper) or t-score, which, as they were to comment in a later paper (Church *et al.* 1994) favours collocating function words, whereas MI favours collocations of pairs of content words.

The argument is that collocations have a large role to play in decoding meaning, and that normal collocations are frequently recurrent in actual usage, so their relative importance can be measured by analysis of a large body of texts. What is more, Church and Hanks found (and published) a methodology for *discovering* the most significant collocates of any selected target word. The importance of this cannot be underestimated. Previous studies measured relations between two pre-selected target words, so they did not give us a *discovery procedure*. Church and Hanks also showed how collocates can be grouped to decide meaning.

When the proceedings of recent corpus linguistics conferences are read, it is surprising and saddening to note that there are many corpus linguists who have still, twenty years on, not yet adjusted their thinking to the most fundamental theoretical implication of this paper, namely that natural languages are analogical systems built around prototypes of many different sorts, and that corpora make it possible to identify these prototypes and measure agreement and variance statistically. If Church and Hanks are right about this (and their implication is hard to refute), it means that all linguistic categorization is a statistical proce-

ture, a point of fundamental importance for lexicography of many different genres, as well as for theoretical and corpus linguistics.

3. Patrick Hanks's career

As an undergraduate at Oxford, studying English Language and Literature, Patrick had dreams of becoming a medievalist, but it was not to be. Like many a young graduate before and since, he drifted around for a while after leaving Oxford and ended up going into publishing as a trainee editor. He has said that he learned more about the practicalities of the English language from Mr Bird and Mr Bell – the chief editor and his deputy – in the ‘morgue’ (the editorial department) of George G. Harrap & Co. Ltd., than from any academic course of study.

He owes his career in lexicography to the millionaire philanthropist publisher Paul Hamlyn, who in 1964 appointed the then 24-year-old Patrick as editor of the *Encyclopedic World Dictionary* (EWD, 1971), an Anglicization of Clarence Barnhart's *The American College Dictionary* (1947). Patrick's brief, as he recalls it, was something like: ‘Minimum of alteration, ... just introduce British spellings and add a few cricket terms.’ However, this ‘on-the-job training’ was to lead immediately to the rather obvious discovery that a dictionary is a ‘collective cultural index’ and that the cultural world of speakers of British English is a world apart from American culture – and of course other English-speaking countries are different again. These differences, obviously, are not restricted to cricket and baseball, nor even to ‘football’ and ‘hockey’ – which mean completely different things in America. A huge number of cultural and vocabulary differences rapidly became apparent as he and the young team appointed to work with him began ploughing their way through the text. These included differences of social structure, law, business practice, leisure activities, language use, and many, many other domains. A team of six lexicographers, all learning ‘on the job,’ together with a dozen expert consultants, was assembled. To take just one example of the need for expert advice: The plant and animal species of Britain and Europe are of course quite different from those of America, while those in Australia, New Zealand, South Africa, and other English-speaking countries are different again. Patrick therefore decided that a full-scale, world-wide review of terms denoting flora and fauna was necessary (from a British standpoint, naturally). This was just one of many special investigations that were undertaken for the Hamlyn dictionary. This dictionary did not make much of a dent in the market supremacy in Britain of Oxford and Chambers, but

it did well abroad, and was in turn to be the foundation for the Australian *Macquarie Dictionary*.

While editing EWD, Patrick got to know the American lexicographer Laurence Urdang (cf. Hanks 2008b) and, through him, the Scottish publisher Jan Collins, who was looking for a British editor for a monolingual dictionary to join the range of new Collins bilingual dictionaries that he had commissioned. Together with Larry Urdang, Alan Isaacs, Paul Procter, and Della Summers, all of whom were to make important contributions to lexicography in English, and a large team of assistants and consultants, Patrick set to work to create a completely new dictionary of English. This was to be the hugely successful *Collins Dictionary of the English Language* (1979), which became the flagship of a range of synchronic monolingual English dictionaries published by Collins. Patrick's introductory essay on 'Meaning and Grammar' in the first edition of this dictionary (Hanks 1979a, unfortunately omitted from subsequent editions), is his first important theoretical statement. It is of interest, not least because it is based on the experience of practical lexicography rather than on theoretical speculation. At the same time, he wrote a fundamental paper on the theory and practice of definition writing (Hanks 1979b).

At this point, Patrick decided that the time had come to put lexicography behind him. After a short period teaching English for Business Purposes in Sweden, he signed up as a PhD student in linguistics at the University of Essex. At this time (1980-82), the Department of Language and Linguistics at the University of Essex was, as he puts it, 'crawling with Chomskyans.' The interaction was one that he describes as one of 'mutual bewilderment.' Patrick's interests in word use and word meaning were satisfied neither by the syntactocentric approach to linguistic theory of generative linguists nor by the logical approaches to meaning ('symbol pushing') of formal semanticists. It was, however, not an entirely unproductive period for him, as he struck up a lifelong friendship with his supervisor, Yorick Wilks, who introduced him to Preference Semantics in the context of Artificial Intelligence and to the work of linguistic philosophers such as Wittgenstein, Putnam, and Grice, all of which were to bear fruit later on. Patrick did not pursue his PhD studies at Essex, however. According to his supervisor:

Patrick turned out not to need a PhD for what he was then doing and decided to leave that until later, like a continental scholar of the age of "habilitation." But I suspect we both gained from our discussions in those times long past – I certainly did.

— Yorick Wilks (cf. p. 50)

In 1983, Patrick received ‘an offer he could not refuse.’ He was appointed project manager of John Sinclair’s COBUILD project at the University of Birmingham. COBUILD at this time had been in development for over two years. Patrick describes what he found there as ‘a project bursting with talent and good ideas, but suffering from inadequate management.’ One of the lexicographers on the team takes up the story:

The first thing Patrick said to me was ‘I want to die.’ This was 1983. He had just arrived to bring some sense to the intellectually exciting but organizationally-challenged COBUILD project (on which I was a lowly lexicographer) – and had presumably just been looking at his in-tray. I liked him immediately, and over the years have learned an enormous amount from him, both in conversations and from reading his prolific, original, and insightful output. For my money he is the most important lexicographic thinker since Samuel Johnson.

— Michael Rundell

John Sinclair’s low-key approach – which typically consisted of throwing out a few carefully-chosen, thought-provoking questions and encouraging students to work things out for themselves by analysing data, debate, and other activities – was very well-suited to the context of university teaching and research, but not to the business of dictionary publishing, where something more heavy-handed is called for. Patrick provided the requisite heavy-handedness to the management of the project, as well as more delicate contributions to the content. He restructured the project, negotiated new arrangements with the university bureaucracy and Collins Publishers, and hired staff to be trained as full-time professional lexicographers, rather than part-timers working at home, with only a skeleton staff of full-time professionals. He chaired a committee of the three team leaders (senior lexicographers) and insisted on mutual common sense and practical implementation of collectively agreed editorial policy – i.e. the committee was not just an academic talking shop for airing ideas. He introduced a few adjustments of his own to project policy. Probably the most important of these was insisting that lexicographers should take account of not just one, but three sources of evidence: corpus data, intuitions, and previous dictionaries. That is, he saw a role for respectful evaluation of the English lexicographic tradition and the personal knowledge of lexicographers – in particular the comparison of intuitions of two or more team members – in interpreting corpus data. He was not afraid to make ruthless judgements.

Less than a year before the end of the project, he judged (to everyone's consternation) that the entries that had been drafted so far were riddled with unacceptable vagueness about the relationship between definitions and definienda, and that this would confuse learners. As a result, after some debate, and with the computational support of Jeremy Clear, all the definitions were rewritten during the final editing phase in the now familiar COBUILD style of 'full-sentence definitions,' as a practical implementation of Sinclair's objections to traditional dictionary style. The details are discussed in Hanks (1987, 1988b).

Working on the COBUILD project brought about a 'road to Damascus' conversion for Patrick personally, which was to influence all his subsequent work in corpus linguistics and lexicography. Definitions in his two previous dictionaries had been based on a mixture of comparative introspection (at least two members of the editorial team had to agree that the definitions in an entry correctly represented their mutual beliefs about the word's meaning) and accretion (surveying earlier dictionaries), supported by directed reading programmes, designed to explore the vocabulary of particular domains rather than the language as a whole. COBUILD changed all that. Working systematically through the lexicon, Patrick and his colleagues at COBUILD became aware from a very early stage in the project that pre-corpus dictionaries regularly misreport or distort the basic meanings of words. Dictionaries that claimed to 'put the modern meaning first' often failed to do so, because they did not have enough evidence to know which meaning was the modern meaning. Distortions ranged from the gross to the subtle. An example of a gross distortion was that pre-corpus dictionaries (even those claiming to put the modern meaning first) stated that the most modern meaning of the noun *dope* is 'a kind of varnish,' even though by 1979 the 'drugs' sense had been dominant for at least two decades. Examples of more subtle questions involved deciding how widely to define the scope of *dope*. Is it a slang term used generically for any drug, even aspirins and laxatives? Or does it denote only drugs that have some effect such as impairing athletic performance or imparting a 'high' or a sense of wellbeing? When used of a recreational drug, does it denote primarily cannabis, or is it used more widely to include cocaine, heroin, etc.? Without corpus evidence, lexicographers were powerless to decide such questions by any means other than guesswork. Corpus evidence provided a basis for an attempt to provide answers. The answers were not always correct (subsequent collections of data from larger corpora and from the Internet have since prompted revisions), but COBUILD's definitions were not only radically different in style from those of previous dictionaries; they also represented the first attempt ever to base the definitions of

contemporary words on usage as recorded in a large collection of contemporary texts.

A much debated word in COBUILD at the time was *take*. Should a concrete sense such as ‘remove,’ ‘steal,’ ‘escort,’ etc. be placed first, or should pride of place be given to the much more frequent idiomatic use of the verb in semantically depleted structures such as *take a look, take a step, take a breath, take a walk, take time, take charge, take a chance*?

In Hanks (2000d), Patrick recounts his earliest experiences of studying corpus evidence with John Sinclair. Looking at concordances for the word *lap*, Patrick commented that the several hundred uses of this word in the Birmingham corpus did not contain a single example of ‘going once around a track.’ Sinclair’s response was, ‘I’m much more interested in all those punctuation marks.’ Indeed, the most normal use of *lap* in English is in a prepositional phrase in clause-final position: *in her lap, on his lap*, etc. Obviously, this sense and this phraseology had to be represented first in the COBUILD dictionary, although the other senses had to be recorded as well.

Literally thousands of similar decisions about the meaning and use of English words were supported in COBUILD by corpus evidence, leading its publisher to claim that COBUILD, published in 1987, was the first dictionary designed to help learners with ‘real English.’

For an overview of those revolutionary times in lexicography, see Patrick’s own accounts in Hanks (1990a) – an early take, as actually written in 1984; and Hanks (2009d) – for a perspective 25 years later.

After the publication of the first edition of COBUILD, Patrick spent some time as a visiting scientist working with Ken Church at AT&T Bell Laboratories in New Jersey, while retaining his role as chief editor of English dictionaries for Collins. As pointed out in Section 2, it is then that Ken and Patrick wrote a series of highly-influential papers on statistical analysis of lexical items in corpora. Patrick developed and maintained strong links with the research community, while remaining fully engaged with practical lexicography. Starting around that time, he also became a frequent guest lecturer at various universities and research institutions, first in Britain and America, later around the world. An attendee recalls meeting Patrick:

Patrick was an invited speaker at the Computing Research Laboratory at New Mexico State University in the late 1980’s. He gave a wonderful inspiring talk about the use of corpora in lexicography and I was lucky enough to join the dinner held for him. He entertained all of us with his terrific stories and I have been a fan ever since. His recent work on norms and exploitations

has had a significant influence on my own research, and his insights into the use of words and phrases has led to many helpful discussions for our work on the detection of anomaly in text.

— Louise Guthrie

In 1990 Patrick joined Oxford University Press as manager (later chief editor) of current English dictionaries. Soon after he arrived, he organized a collaboration – known as the Hector Project – between OUP and the Systems Research Laboratory of Digital Equipment Corporation in Palo Alto, California. This was the first systematic attempt ever to link word meaning with word use using corpus evidence. Unfortunately, the results were never adequately reported or publicly evaluated, though the Hector entries were used as a benchmark in Sen-seval.⁴ The best relevant accounts of the project from within were given at successive COMPLEX conferences in Budapest by Atkins (1992) and Hanks (1994). Patrick’s paper, titled ‘Linguistic Norms and Pragmatic Exploitations,’ was the first airing of what later became the corpus-driven Theory of Norms and Exploitations.

Taking account of the findings of the Hector Project, Patrick designed and (with Judy Pearsall as the project editor) edited *The New Oxford Dictionary of English* (1998). This is the only dictionary to be based both on a vast collection of citations (the Oxford Reading Programme, which collects evidence for new, rare, and unusual words and meanings) and a corpus (the British National Corpus). One of its many innovations was to attempt a distinction – not consistently implemented, it must be said – between core meaning and meaning extension or subsense. For example, the core sense of *cocoon* is the most literal one: ‘a silky case spun by the larvae of many insects for protection as pupae.’ Associated with this are two subsenses: a technical one, ‘a covering that prevents the corrosion of metal equipment,’ and a general one, ‘something that envelops or surrounds, especially in a protective or comforting way.’ In everyday English, the last of these is more common than either of the other two. As John Sinclair rightly points out in his paper in this book (p. 38), the core sense is not always the most common one.

In 2000, after ten years in Oxford, Patrick went to Cambridge, Massachusetts to work with James Pustejovsky in a software company (Lexeme, later rechristened LingoMotors) that was developing applications of computational linguistics, including ‘breaking the tyranny of text matching,’ for information retrieval. The company bit the dust along with many other software companies when the bubble burst in 2001-2002, but the ideas continued to be developed at Brandeis University, where James is a professor of computer science. Patrick

became an adjunct professor there. His research activities at Brandeis included exploring ways in which Pustejovsky's Generative Lexicon Theory could be developed and extended using corpus evidence (cf., e.g., Pustejovsky & Hanks 2001, Hanks & Pustejovsky 2005). This work was to lead to the development of a prototype of the Corpus Pattern Analysis project (Hanks 2004a), the aim of which is to map the relationship between word meaning and word use.

In 2003, Patrick was invited to spend a year at the Berlin-Brandenburg Academy of Sciences and Humanities, to serve as a consultant to Christiane Fellbaum's research project *Kollokationen im Wörterbuch*, investigating in particular idioms and light verbs. A product of this collaboration was a paper on German light verbs, outlining a new approach to a corpus-based analysis of verbs in dictionary format, while at the same time questioning much of the received wisdom of the Germanistic research community on 'function verbs' and 'support verbs' (cf. Hanks, Urbschat & Gehweiler 2006).

Patrick has for a long time had a happy and fruitful association with scholars in the Czech Republic. In 1995 he had been invited to teach at the seventh Series of Vilém Mathesius Courses at the Charles University in Prague, and since then he has given a number of talks on various aspects of lexicography at meetings of the Prague Linguistic Circle and other Czech Institutions. In 1996 he taught an intensive 'block' course in computational lexicography and corpus analysis at the Faculty of Informatics, Masaryk University, in Brno – and went on to obtain a PhD degree there! The former head of the Information Technology Department at the Faculty of Informatics, explains how this came about:

My close contacts with Patrick Hanks started in 1993, when I had the opportunity to visit Oxford and other British corpus centres (Lancaster and Birmingham), together with colleagues from Prague. During the visit to Oxford it was Patrick who took care of us. He explained how the BNC had been prepared and how it was being used at the time for compiling *The New Oxford Dictionary of English* (NODE, 1998). This was not only interesting for us, but was also a motivating experience for the later development of corpus tools at the Faculty of Informatics, Masaryk University (FI MU), in Brno, and for building the Czech National Corpus in Prague, under the direction of František Čermák at the Charles University.

During one of Patrick's later visits to Brno I was surprised to discover that, despite his achievements in lexicography, he had never actually completed a PhD. I convinced him to go for a PhD degree at FI MU. It took some time, but finally Patrick completed his dissertation at the end of 2001 and in April 2002 he defended it successfully. From that moment our coop-

eration became closer and in 2007 our national grant projects allowed me to employ Patrick at FI MU for two years.

During those two years he put a lot of effort into the development of the Corpus Pattern Analysis (CPA) project,⁵ the output of which is a Pattern Dictionary of English Verbs (PDEV). CPA is a new technique for finding verb meanings based on the analysis of the context in which the individual words (verbs) occur. A lexicographer, i.e. Patrick, starts with a representative sample of the verb tokens in a corpus. Using the CPA tool, developed at the NLP Centre at FI MU by Pavel Rychlý and Adam Rambousek, he classifies the various contexts into groups, which are then characterized as verb patterns capturing verb meanings – or rather, meanings of the whole pattern in which the verb is embedded. In the process information about particular sentence constituents (typically subject, object, adverbial) is added as well.

The editor and browser for CPA take advantage of the DEB platform, also developed at the NLP Centre at FI MU. This software includes the corpus manager Manatee/Bonito, together with an integrated version of the Sketch Engine,⁶ plus editing and browsing functions. Thus the user can build and maintain a database of the context patterns of English verbs (PDEV) and indeed verbs in other languages (presently Italian and Spanish, with the possibility of Czech to follow).

The PDEV project results (close to 700 English verbs so far) are publicly accessible online;⁷ interested users can browse the completed verbs, and see their patterns and distribution, as well as the corresponding concordances for each pattern.

In 2009 Patrick moved to Jan Hajič's Institute of Formal and Applied Linguistics (ÚFAL) at the Charles University in Prague, where one of the projects currently being undertaken is an evaluation of the potential usefulness of PDEV for NLP applications. The software for PDEV is being developed further, with Martin Holub from ÚFAL joining hands with the team at FI MU, and in this way the fruitful cooperation between Patrick and FI MU continues.

— Karel Pala

At the time of writing (June 2010), at an age when most people are enjoying a peaceful retirement, Patrick's appetite for lexicographic work remains undiminished. He has just started not one but two new jobs. As a visiting professor at the University of Wolverhampton, he is co-editing with Ruslan Mitkov the *Oxford Dictionary of Computational Linguistics*. And as a visiting professor at the University of the West of England, in Bristol, he has just started work with

Richard Coates on building a vast database of Family Names of the United Kingdom (FaNUK), recording the origin, meaning, history, and demographics of every name that has any reasonable frequency in Britain. The project will build on Patrick's lifelong interest in names and his three dictionaries of personal names published by Oxford University Press: *A Dictionary of Surnames* (1988) and *A Dictionary of First Names* (1990), both compiled with Flavia Hodges, and the *Dictionary of American Family Names* (2003), which was supported by a team of over 30 specialist consultants. FaNUK has received a large funding grant from the Arts and Humanities Research Council, which will not only enable the research to be carried out on a sound academic basis, but will also train the next generation of onomastic researchers.

4. Patrick Hanks's Theory of Norms and Exploitations

I have quoted Karel Pala at length in the previous section, simply because Patrick's connection with the Czech Republic, and Brno in particular, has been exceptionally important. Here I am not referring to his closing the loop with regard to obtaining a PhD, a process started two decades earlier at Essex, but to the dedicated software tools that have been built for him in Brno. Thanks to these tools, Patrick has been able to put his own theories to the test, producing a new type of data in the process, and refining his main theory – the Theory of Norms and Exploitations – further. Those who have seen drafts of his forthcoming book on the Theory of Norms and Exploitations invariably (and rightly) refer to it as his Magnum Opus.⁸

Patrick's Theory of Norms and Exploitations argues that a natural language is indeed a rule-governed system of linguistic behaviour, but that there are *two systems of rules* – a sort of 'double helix' of rules. One rule system governs normal phraseology and meaning of words in use, while the other allows language users to exploit normal phraseology in all sorts of creative ways. These two rule systems constantly interact, so that it sometimes happens that a creative and original use of a word in one generation of language users becomes established as a new conventional norm in the next generation.

The Corpus Pattern Analysis project applies the Theory of Norms and Exploitations by focusing on the verb, 'the pivot of the clause.' Each verb is associated with a number of patterns based on valencies and collocational preferences; each pattern has a primary implicature (i.e. the meaning of the pattern) and any number of secondary implicatures. Actual meanings in actual texts are created and interpreted by language users relying (in a Gricean fashion) on the meanings associated with each conventional pattern. Thus, an actual meaningful

use of a word in a text is created by a speaker or writer drawing instinctively on his or her individual store of patterns and meaning associated with that word. It is interpreted by hearers and readers in much the same way, by referring to their individual stores of prototypical patterns and meanings. The hearer matches the actual observed use in a text to a stored phraseological prototype, which has a meaning. Speaker's and hearer's stored prototypes are not identical, of course, but any gross and noticeable differences are eliminated by social pressures as the individual grows up and develops in a language community. This general account has many similarities to Michael Hoey's Lexical Priming Theory (2005). Hoey is, of course, another former associate of Sinclair.

According to Patrick, a pattern consists prototypically of a verb, its valency structure, and its collocational preferences in each argument slot. Therefore, a pattern is a prototypical proposition, available for manipulation by language users in a great variety of ways.

It is a lexicographical task to identify, on the basis of corpus evidence, the underlying conventional norms of word use and meaning that speakers rely on when communicating with each other. This involves compiling an inventory of normal uses, which leads in turn to a distinction between norms and exploitations, which implies an interest in figurative language. Literal meaning and metaphorical meaning are complementary concepts – two sides of a single coin. With this we have arrived at yet another research interest of Patrick, namely metaphors and figurative language (cf., e.g., Hanks 2005c, 2006e, 2008a, 2010a).

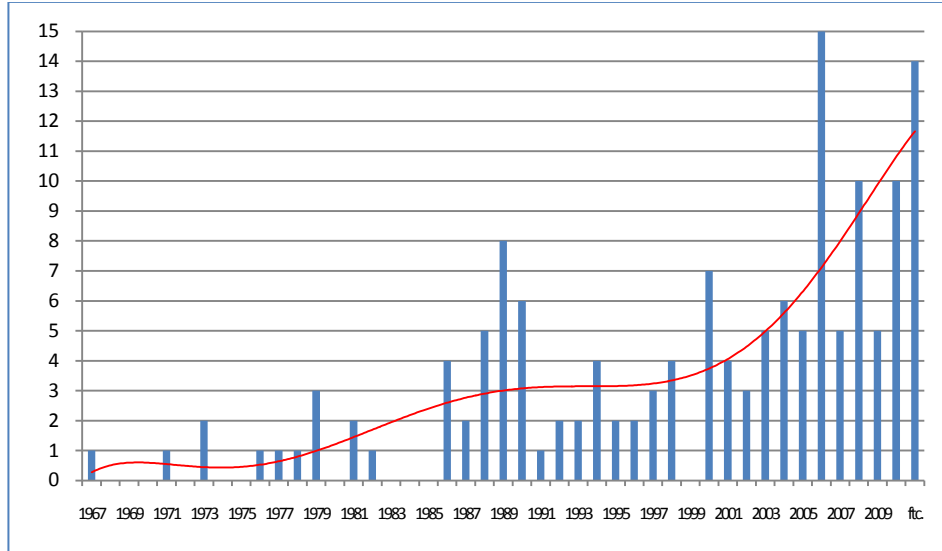
5. Patrick Hanks's publications

As can be deduced from the foregoing, Patrick has been a prolific writer, of all sorts of reference works, as well as all sorts of lexicographic and corpus-linguistic research papers. He has covered the breadth and depth of both those fields, and this while permanently at the forefront of innovation.

The Addendum to this chapter lists a first attempt at bringing his published output together. As with all first attempts at completeness, this list is certainly not complete, and may even contain a few 'funnies.' Even so, with about 150 unique entries already, in just three decades, his output can only be considered staggering. The more so that over thirty of the works listed are dictionaries compiled under his editorship, with no fewer than half of those – 16 to be precise – first editions!

If one plots the number of Patrick's publications over time, Figure 1 is obtained.

Figure 1: Number of Patrick Hanks’s publications per year.



While Figure 1 clearly indicates that Patrick moved into a different, higher gear starting around the times of the COBUILD project, the truly astonishing aspect is to notice that he hasn’t merely stepped up his output yet again during the past ten years, but now truly moved into an exponential gear – with no signs whatsoever of slowing down. As a ‘lexicographer’s lexicographer’ Patrick is currently (at seventy) in his ‘prime time.’

If one only considers the books by Patrick which are currently held in the world’s libraries, the statistics are at least as impressive. According to World-Cat,⁹ there are over *ten thousand* of Patrick’s books on the world’s library shelves – 57 different books, totalling 134 editions. The timeline for those library books is as shown in Figure 2.

Figure 2: Patrick Hanks’s publication timeline in the world’s libraries.



Perhaps surprisingly, the top three of his most widely held works are not the four major dictionaries mentioned at the start of Section 2, but the three onomastic reference works mentioned at the end of Section 3: *A Dictionary of Surnames* (held in over 2,500 libraries worldwide), *A Dictionary of First Names* (held in over 2,300 libraries), and the *Dictionary of American Family Names* (held in close to 1,400 libraries).

6. Structure of this book

This book is divided into three parts, very roughly: a theoretical section, a computational section, and a lexicographic section. In each of those sections, the contributions have been placed in an order paralleling Patrick's career.

'Part I: Theoretical Aspects and Background,' starts with the last known paper of John Sinclair, which he was actively working on, for this Festschrift, at the time of his death. This outlines the most radical version of Sinclair's approach to collocational analysis, which is contrasted with the terminological approach of traditional dictionaries. Part I also contains important papers by Wilks (on Preference Semantics), Pustejovsky & Rumshisky (on the Generative Lexicon), Mel'čuk (on the Government Pattern), and Wiggins (on Paradoxes), advancing our theoretical understanding of the nature of word meaning.

Ken Church, quite naturally, opens 'Part II: Computing Lexical Relations,' with some reflections on the nature (size) of (Web) corpora. Grefenstette goes on to use a copy of the Web to predict the number of concepts future (computational) lexicographers will have to describe – the number he predicts is a daunting one! Guthrie & Guthrie, Geyken, and Pala & Rychlý then provide a tour d'horizon, taking us along the places where Patrick has been active over the past few years – the US, Germany and the Czech Republic – and discuss different aspects of computational approaches to the lexicon, in casu of adjectives, nouns and verbs, for English, German and Czech. The last two chapters in Part II revolve around Patrick's PDEV. Cinková, Holub & Smejkalová evaluate the current PDEV, while Jezek & Frontini extend PDEV to now also cover Italian.

Sinclair's chapter at the start of Part I was lightly edited by Rosamund Moon, Sinclair's colleague since the earliest days of the COBUILD project. 'Part III: Lexical Analysis and Dictionary Writing,' includes some thought-provoking reflections on idiom, allusion, and convention by Rosamund herself. John Sinclair's sister, Sue Atkins, then takes over in Part III, with a fascinating comparison of the FrameNet database, and DANTE, the latest lexicographical database that is being populated under Sue's supervision. Kilgarriff & Rychlý then attempt the ultimate dream: from corpus to dictionary, if not automatically, at least semi-automatically. On a lighter note, Bogaards then wonders whether there is any 'theory' in 'lexicography' – the answer, of course, depends on what one understands under these two terms. Next, Baňko provides a lively account of an early adoption of the COBUILD approach, in this case for Polish, while Green delves into the world of (French) argot. Michael Rundell, finally, concludes with a masterfully elegant essay on Patrick's elegance.

Notes

¹ Unless otherwise noted, all quotes in this chapter are ‘personal communication,’ gathered and received during the production of this book.

² The details of all Patrick Hanks’s publications mentioned in this chapter are to be found in the list of his publications included as an Addendum to this chapter.

³ The next four paragraphs are taken from my analysis in De Schryver (2009: 479).

⁴ Cf. <http://www.senseval.org/>.

⁵ Cf. <http://nlp.fi.muni.cz/projects/cpa/> = Hanks (2007a).

⁶ Cf. <http://www.sketchengine.co.uk/>.

⁷ Cf. <http://deb.fi.muni.cz/pdev/> = Hanks (2007c).

⁸ I have been particularly ‘lucky’ in this regard, having worked closely with Patrick for the past three years – in Brno, Witney, Ghent, Barcelona, Cape Town, and Kampala – on the final compilation and editing of this forthcoming book. Much of the contents of this chapter, and especially those on his career, are based on my notes of our conversations.

⁹ Cf. <http://www.worldcat.org/identities/lccn-n83-198535>.

References

[For all publications by Patrick Hanks, see the Addendum to this chapter.]

Atkins, B. T. S. 1992. ‘Tools for Computer-Aided Corpus Lexicography: The Hector Project’ in F. Kiefer, G. Kiss & J. Pajzs (eds.). 1992. *Papers in Computational Lexicography: COMPLEX '92*. Budapest: Linguistics Institute, Hungarian Academy of Sciences, 1–59.

Barnhart, C. L. 1947. *The American College Dictionary*. New York: Random House.

De Schryver, G.-M. 2009. ‘An Analysis of *Practical Lexicography: A Reader* (Ed. Fontenelle 2008)’. *Lexikos* 19: 458–489.

Hoey, M. 2005. *Lexical Priming: A New Theory of Words and Language*. Abingdon: Routledge.

Addendum: Publications by Patrick Hanks

1967

- **Jalovec, Karel.** 1967. *German and Austrian Violin-Makers*. London: Paul Hamlyn Publishers. [Translated from Czech by **George Theiner**, edited by **Patrick Hanks**.]

1971

- **Hanks, Patrick** (ed.) & **Simeon Potter** (editorial consultant). 1971. *Encyclopedic World Dictionary*. London: Paul Hamlyn Publishers. / first edition

1973

- **Guralnik, David B., Martin Joos, Frederic G. Cassidy, James E. Congleton, Patrick Hanks, Marvin Carmony, Bernard van't Hul & Raven I. McDavid, Jr.** 1973. 'Dictionary Treatment of Pronunciation – Regional – Discussion' in R. I. McDavid, Jr. & A. R. Duckert (eds.). 1973. *Lexicography in English* (Annals of the New York Academy of Sciences 211). New York: New York Academy of Sciences, 141–143.
- **Mathiot, Madeleine, Willard Van Orman Quine, Henry A. Gleason, Jr., Patrick Hanks & Sherman M. Kuhn.** 1973. 'Vagaries of Definition – Discussion' in R. I. McDavid, Jr. & A. R. Duckert (eds.). 1973. *Lexicography in English* (Annals of the New York Academy of Sciences 211). New York: New York Academy of Sciences, 251–252.

1976

- **Laurence Urdang Associates** (comp.), **Edwin Riddell** (ed.) & **Patrick Hanks** (managing editor). 1976. *Lives of the Stuart Age, 1603-1714*. New York: Barnes & Noble Books.

1977

- **Hanks, Patrick, Alan Isaacs & John Daintith.** 1977. *The Illustrated Dictionary*. London: Sundial Books.
⇒ Also published as **Hanks, Patrick, Alan Isaacs & John Daintith.** 1980. *Illustrated Children's Dictionary*. London: Octopus Books Ltd.
⇒ Also published as **Hanks, Patrick & Alan Isaacs.** 1987. *Children's Illustrated Dictionary*. London: Bounty Books.

1978

- **Laurence Urdang Associates** (comp.), **William Gould** (ed.) & **Patrick Hanks** (managing editor). 1978. *Lives of the Georgian Age, 1714-1837*. New York: Barnes & Noble Books.

1979

- **Hanks, Patrick.** 1979a. 'Meaning and Grammar' in P. Hanks *et al.* (eds.). 1979. *Collins Dictionary of the English Language*. London: William Collins Sons & Co. Ltd., xxxi-xxxv.
- **Hanks, Patrick.** 1979b. 'To What Extent Does a Dictionary Definition Define?' in R. R. K. Hartmann (ed.). 1979. *Dictionaries and their Users: Papers from the 1978 B.A.A.L. Seminar on Lexicography* (Exeter Linguistic Studies 4 & ITL Review of Applied Linguistics 45-46). Exeter: University of Exeter Press, 32-38.
- **Hanks, Patrick** (ed.), **Thomas H. Long** (managing ed.), **Laurence Urdang** (editorial director), *et al.* 1979. *Collins Dictionary of the English Language*. London: William Collins Sons & Co. Ltd. / first edition

1981

- **Grossmith, George & Weedon Grossmith.** 1981. *The Diary of a Nobody*. London: William Collins Sons & Co. Ltd. [With an introduction by **Patrick Hanks**, illustrated by **Weedon Grossmith**.]
- **Hanks, Patrick.** 1981. 'Book Review: J. Branford. 1980. A Dictionary of South African English: New Enlarged Edition (Oxford University Press)'. *English in Africa* 8.1: 73-85.

1982

- **McLeod, William T.** (managing ed.), **Patrick Hanks** (consultant ed.), *et al.* 1982. *The New Collins Concise Dictionary of the English Language*. London: William Collins Sons & Co. Ltd. / first edition

1986

- **Hanks, Patrick & Jim Corbett.** 1986a. *Business Listening Tasks: Guide for Teachers and Self-Study* (Cambridge Professional English Series). Cambridge: Cambridge University Press.
- **Hanks, Patrick & Jim Corbett.** 1986b. *Business Listening Tasks: Learner's Book* (Cambridge Professional English Series). Cambridge: Cambridge University Press.
- **Hanks, Patrick & Flavia Hodges.** 1986. *The Oxford Minidictionary of First Names*. Oxford: Oxford University Press. / first edition
- **Hanks, Patrick** (ed.), **William T. McLeod** (managing ed.), **Laurence Urdang** (editorial director), *et al.* 1986. *Collins Dictionary of the English Language*. London: William Collins Sons & Co. Ltd. / second edition (first edition in 1979)

1987

- **Hanks, Patrick.** 1987. 'Definitions and Explanations' in J. M. Sinclair (ed.). 1987. *Looking Up: An account of the COBUILD Project in lexical computing and the development of the Collins COBUILD English Language Dictionary*. London: Collins ELT, 116–136.
⇒ Second half (pp. 123–136) reprinted in R. R. K. Hartmann (ed.). 2003. *Lexicography: Critical Concepts*. London: Routledge, Vol. III, 191–206.
- **Sinclair, John M.** (chief ed.), **Patrick Hanks** (managing ed.), *et al.* 1987. *Collins COBUILD English Language Dictionary*. London: William Collins Sons & Co. Ltd. / first edition

1988

- **Hanks, Patrick.** 1988a. 'A New Kind of Dictionary for English Learners: Cobuild'. *Les Cahiers de l'APLIUT* 30-31 (VIII / 1-2): 73–88.
- **Hanks, Patrick.** 1988b. 'Typicality and Meaning Potentials' in M. Snell-Hornby (ed.). 1988. *ZüriLEX '86 Proceedings: Papers read at the EURALEX International Congress, University of Zürich, 9-14 September 1986*. Tübingen: A. Francke Verlag GmbH, 37–47.
⇒ Reprinted in G. Sampson & D. McCarthy (eds.). 2004. *Corpus Linguistics: Readings in a Widening Discipline*. London: Continuum, 58–66.
- **Hanks, Patrick & Flavia Hodges.** 1988. *A Dictionary of Surnames*. Oxford: Oxford University Press. / first edition
- **Hanks, Patrick** (chief ed.), **William T. McLeod, Marian Makins** (managing eds.), *et al.* 1988. *The Collins Concise Dictionary of the English Language*. London: William Collins Sons & Co. Ltd. / second edition (first edition in 1982)
- **Sinclair, John M.** (chief ed.), **Gwyneth Fox, Patrick Hanks** (managing eds.), *et al.* 1988. *Collins COBUILD Essential English Dictionary*. London: William Collins Sons & Co. Ltd. / first edition

1989

- **Church, Kenneth W., William A. Gale, Patrick Hanks & Donald Hindle.** 1989. 'Parsing, Word Associations and Typical Predicate-Argument Relations' in *Proceedings of the International Workshop on Parsing Technologies, 28-31 August 1989*. Pittsburgh: Carnegie Mellon University, 389–398.
⇒ Also published in *Speech and Natural Language: Proceedings of a Workshop Held at Cape Cod, Massachusetts, October 15-18, 1989*.

San Mateo: Morgan Kaufmann Publishers Inc., 75–81.

⇒ Revised version published in M. Tomita (ed.). 1991. *Current Issues in Parsing Technology* (Kluwer International Series in Engineering and Computer Science 126). Norwell: Kluwer Academic Publishers, 103–112.

- **Church, Kenneth W. & Patrick Hanks.** 1989. ‘Word Association Norms, Mutual Information, and Lexicography’ in *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics, 26-29 June 1989*. Vancouver: University of British Columbia, 76–83.
 ⇒ Revised version published in *Computational Linguistics* 16.1 (1990): 22–29.
 ⇒ Reprinted in T. Fontenelle (ed.). 2008. *Practical Lexicography: A Reader*. New York: Oxford University Press, 285–295.
 ⇒ Reprinted in P. Hanks (ed.). 2008g. *Lexicology: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. VI, 150–167.
 - **Hanks, Patrick.** 1989. ‘How common is “common”?’ *The Collins Dictionary Diary* 1989: 2–12.
 - **Hanks, Patrick** (chief ed.), **Marian Makins** (managing ed.), *et al.* 1989a. *Collins Concise Dictionary Plus*. London: William Collins Sons & Co. Ltd. / new edition, combining dictionary with extensive encyclopaedic entries
 - **Hanks, Patrick** (chief ed.), **Marian Makins** (managing ed.), *et al.* 1989b. *Collins Pocket English Dictionary*. London: William Collins Sons & Co. Ltd. / first edition
 - **Hanks, Patrick** (chief ed.), **Marian Makins, Gwyneth Fox** (managing eds.), *et al.* 1989. *The Collins School Dictionary*. London: William Collins Sons & Co. Ltd. / first edition
 - **Sinclair, John M.** (chief ed.), **Patrick Hanks** (editorial director), **Ramesh Krishnamurthy, et al.** 1989. *PONS Cobuild English Learner’s Dictionary*. Stuttgart: Ernst Klett Verlag GmbH. / edition for the German market
 - **Sinclair, John M.** (chief ed.), **Patrick Hanks** (editorial director), **Rosamund E. Moon, et al.** 1989. *Collins COBUILD Dictionary of Phrasal Verbs*. London: William Collins Sons & Co. Ltd. / first edition
- 1990
- **Hanks, Patrick.** 1990a. ‘Evidence and Intuition in Lexicography’ in J. Tomaszczyk & B. Lewandowska-Tomaszczyk (eds.). 1990. *Meaning and Lexicography* (Linguistic and Literary Studies in Eastern Europe 28).

Amsterdam: John Benjamins Publishing Company, 31–41. [Note: This paper was actually written in 1984 but publication was delayed for 6 years by Benjamins.]

- **Hanks, Patrick.** 1990b. ‘Towards a Statistical Dictionary of Modern English: Some Preliminary Reflections’ in T. Magay & J. Zigány (eds.). 1990. *BudaLEX '88 Proceedings: Papers from the 3rd International EU-RALEX Congress, Budapest, 4-9 September 1988*. Budapest: Akadémiai Kiadó, 53–57.
- **Hanks, Patrick & Flavia Hodges.** 1990. *A Dictionary of First Names*. Oxford: Oxford University Press. / first edition
- **Hanks, Patrick, Alan Isaacs & John Daintith.** 1990. *The Hamlyn Illustrated Children's Dictionary*. London: Hamlyn young books.
- **Hanks, Patrick, Marian Makins & Diana Adams.** 1990. *Collins Paperback Thesaurus in A to Z Form*. Toronto: HarperCollins Canada. / second edition (first edition in 1986)
- **Hardie, Ronald G.** 1990. *Collins Gem English Grammar*. London: HarperCollins. [Edited by **Patrick Hanks & Alice Grandison.**]

1991

- **Church, Kenneth W., William A. Gale, Patrick Hanks & Donald Hindle.** 1991. ‘Using Statistics in Lexical Analysis’ in U. Zernik (ed.). 1991. *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*. Hillsdale: Lawrence Erlbaum Associates, 115–164.

1992

- **Hanks, Patrick.** 1992. ‘Computational Analysis and Definitional Structure’. *Lexicographica: International Annual for Lexicography* 8: 100–129.
- **Hanks, Patrick & Flavia Hodges.** 1992. *A Concise Dictionary of First Names*. Oxford: Oxford University Press. / first edition

1992/1993

- **Hanks, Patrick.** 1992/1993a. ‘Lexicography: Theory and Practice’. *Dictionaries: Journal of the Dictionary Society of North America* 14: 97–112.
- **Hanks, Patrick.** 1992/1993b. ‘The Present-Day Distribution of Surnames in the British Isles’. *Nomina* 16: 79–98.

1994

- **Church, Kenneth W., William A. Gale, Patrick Hanks, Donald Hindle & Rosamund E. Moon.** 1994. ‘Lexical Substitutability’ in B. T. S. Atkins & A. Zampolli (eds.). 1994. *Computational Approaches to the Lexicon*. Oxford: Clarendon Press, 153–177.

- **Hanks, Patrick.** 1994. 'Linguistic Norms and Pragmatic Exploitations, or Why Lexicographers need Prototype Theory, and Vice Versa' in F. Kiefer, G. Kiss & J. Pajzs (eds.). 1994. *Papers in Computational Lexicography: COMPLEX '94*. Budapest: Linguistics Institute, Hungarian Academy of Sciences, 89–113.
⇒ Reprinted in P. Hanks (ed.). 2008g. *Lexicology: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. V, 233–255.
- **Hanks, Patrick & Flavia Hodges.** 1994. *Naming Your Baby*. Oxford: Oxford University Press.
- **Hanks, Patrick, Alan Isaacs & John Daintith.** 1994. *The Hamlyn Illustrated Children's Dictionary*. London: Hamlyn young books. / new edition

1995

- **Hanks, Patrick** (chief ed.), *et al.* 1995. *Oxford English Reference Dictionary*. New York: Oxford University Press.
- **Hanks, Patrick & Flavia Hodges.** 1995. *Babies' Names*. Oxford: Oxford University Press. / first edition

1996

- **Hanks, Patrick.** 1996. 'Contextual Dependency and Lexical Sets'. *International Journal of Corpus Linguistics* 1.1: 75–98.
- **Lewandowska-Tomaszczyk, Barbara & Patrick Hanks.** 1996. 'Completive Particles and Verbs of Closing in English' in E. Weigand & F. Hundsnurscher (eds.). 1996. *Lexical Structures and Language Use. Proceedings of the International Conference on Lexicology and Lexical Semantics, Münster, September 13-15, 1994*. Tübingen: Max Niemeyer Verlag, 89–103.

1997

- **Hanks, Patrick.** 1997a. 'Ferocious Empiricism' = 'Book Review: J. A. Foley (ed.). 1996. *J. M. Sinclair on Lexis and Lexicography* (UniPress)'. *International Journal of Corpus Linguistics* 2.2: 289–295.
- **Hanks, Patrick.** 1997b. 'Lexical Sets: Relevance and Probability' in B. Lewandowska-Tomaszczyk & M. Thelen (eds.). 1997. *Translation and Meaning, Part 4. Proceedings of the Łódź Session of the 2nd International Maastricht-Łódź Duo Colloquium on "Translation and Meaning", Held in Łódź, Poland, 20-24 September 1995*. Maastricht: School of Translation and Interpreting, Hogeschool Maastricht, 119–139.
- **Hanks, Patrick & Flavia Hodges.** 1997. *A Concise Dictionary of First Names*. Oxford University Press. / second edition (first edition in 1992)

1998

- **Hanks, Patrick.** 1998a. 'Enthusiasm and Condescension' in T. Fontenelle, P. Hilgsmann, A. Michiels, A. Moulin & S. Theissen (eds.). 1998. *Actes EURALEX'98 Proceedings: Communications soumises à EURALEX'98 (Huitième Congrès International de Lexicographie) à Liège, Belgique / Papers submitted to the Eighth EURALEX International Congress on Lexicography in Liège, Belgium.* Liège: English and Dutch Departments, University of Liège, 151–166.
⇒ Reprinted in W. Teubert & R. Krishnamurthy (eds.). 2007. *Corpus Linguistics: Critical Concepts in Linguistics.* Abingdon: Routledge, Vol. III, 140–153.
- **Hanks, Patrick.** 1998b. 'Problemas e solucións na preparación de dicionarios de *idioms* ingleses' [Problems and Solutions in the Preparation of English Idiom Dictionaries] in X. Ferro Ruibal (ed.). 1998. *Actas do I Coloquio Galego de Fraseoloxía.* Santiago de Compostela: Centro Ramón Piñeiro para a Investigación en Humanidades – Xunta de Galicia, 83–99.
- **Hanks, Patrick & Kate Hardcastle.** 1998. 'The Multiplicitous Origins of American Family Names' in W. F. H. Nicolaisen (ed.). 1998. *Proceedings of the XIXth International Congress of Onomastic Sciences, Aberdeen, August 4-11, 1996.* Aberdeen: Department of English, University of Aberdeen, Vol. III, 164–182.
- **Pearsall, Judy** (ed.), **Patrick Hanks** (chief ed., current English dictionaries), *et al.* 1998. *The New Oxford Dictionary of English.* New York: Oxford University Press. / first edition

2000

- **Hanks, Patrick.** 2000a. 'Contributions of Lexicography and Corpus Linguistics to a Theory of Language Performance' in U. Heid, S. Evert, E. Lehmann & C. Rohrer (eds.). 2000. *Proceedings of the Ninth EURALEX International Congress, EURALEX 2000, Stuttgart, Germany, August 8th - 12th, 2000.* Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 3–13.
- **Hanks, Patrick.** 2000b. 'Dictionaries of Idioms and Phraseology in English' in G. Corpas Pastor (ed.). 2000. *Las lenguas de Europa: Estudios de fraseología, fraseografía y traducción* (Interlingua 12). Granada: Comares, 303–320.
- **Hanks, Patrick.** 2000c. 'Do Word Meanings Exist?' *Computers and the Humanities* 34.1-2: 205–215.

⇒ Reprinted in T. Fontenelle (ed.). 2008. *Practical Lexicography: A Reader*. New York: Oxford University Press, 123–134.

- **Hanks, Patrick**. 2000d. ‘Immediate Context Analysis: Distinguishing Meanings by Studying Usage’ in C. Heffer & H. Sauntson (eds.). 2000. *Words in Context: A Tribute to John Sinclair on his Retirement* (ELR Discourse Analysis Monograph 18). Birmingham: University of Birmingham, 10–30.
- **Hanks, Patrick & D. Kenneth Tucker**. 2000a. ‘A Diagnostic Database of American Personal Names’. *Names: A Journal of Onomastics* 48.1: 59–69.
- **Hanks, Patrick & D. Kenneth Tucker**. 2000b. ‘Two Projects in Onomastic Lexicography’ in U. Heid, S. Evert, E. Lehmann & C. Rohrer (eds.). 2000. *Proceedings of the Ninth EURALEX International Congress, EURALEX 2000, Stuttgart, Germany, August 8th - 12th, 2000*. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 213–221.
- **Hanks, Patrick** (chief ed.), **Maurice Waite, Sara Hawker, et al.** 2000. *The New Oxford Thesaurus of English*. Oxford University Press. / first edition

2001

- **Hanks, Patrick**. 2001a. ‘Los diccionarios fraseológicos en lengua inglesa’ [Dictionaries of Idioms and Phraseology in English] in M. C. Ayala Castro (ed.). 2001. *Diccionarios y enseñanza* (Ensayos y documentos 45). Alcalá de Henares: Universidad de Alcalá, 287–303.
- **Hanks, Patrick**. 2001b. ‘The Probable and the Possible: Lexicography in the Age of the Internet’ in S. Lee (ed.). 2001. *Asialex 2001 Proceedings, Asian Bilingualism and the Dictionary*. Seoul: Center for Linguistic Informatics Development, Yonsei University, 1–15.
⇒ Reprinted in *Studies in Lexicography* 11.1 (2001): 7–36.
- **Hanks, Patrick & Flavia Hodges**. 2001. *A Concise Dictionary of First Names*. Oxford: Oxford University Press. / third edition (first edition in 1992, second edition in 1997)
- **Pustejovsky, James & Patrick Hanks**. 2001. ‘Very Large Lexical Databases’. Tutorial at ACL 2001, Toulouse, 6-11 July 2001. Tutorial notes available online at <http://www.cs.brandeis.edu/~llc/publications/ACL-Toulouse-Slides.pdf>.

2002

- **Hanks, Patrick.** 2002. 'Mapping Meaning onto Use' in M.-H. Corréard (ed.). 2002. *Lexicography and Natural Language Processing: A Festschrift in Honour of B. T. S. Atkins*. S.l.: Euralex, 156–198.
- **Hanks, Patrick, Flavia Hodges, Anthony D. Mills & Adrian Room.** 2002. *The Oxford Names Companion*. Oxford: Oxford University Press.
- **Pustejovsky, James, Luc Belanger, José Castaño, Rob Gaizauskas, Patrick Hanks, Bob Ingria, Graham Katz, Dragomir Radev, Anna Rumshisky, Antonio Sanfilippo, Roser Saurí, Andrea Setzer, Beth Sundheim & Marc Verhagen.** 2002. *NRRC SummerWorkshop on Temporal and Event Recognition for Question Answering Systems*. Online at <http://www.cs.brandeis.edu/~jamesp/arda/time/readings/TERQAS-FINAL-REPORT.pdf>

2003

- **Hanks, Patrick.** 2003a. 'Americanization of European Family Names in the Seventeenth and Eighteenth Century'. *Onoma: Journal of the International Council of Onomastic Sciences* 38: 119–154.
- **Hanks, Patrick.** 2003b. 'Lexicography' in R. Mitkov (ed.). 2003. *The Oxford Handbook of Computational Linguistics*: 48–69. New York: Oxford University Press.
- **Hanks, Patrick.** 2003c. 'Why WordNet Should Not Include Figurative Language, and What Would Be Done Instead' in P. Sojka, K. Pala, P. Smrž, C. Fellbaum & P. Vossen (eds.). 2003. *Proceedings of the Second International WordNet Conference (GWC 2004), Brno, Czech Republic, January 20-23, 2004*. Brno: Faculty of Informatics, Masaryk University, 11.
- **Hanks, Patrick** (chief ed.), **Kate Hardcastle** (managing ed.), *et al.* 2003. *Dictionary of American Family Names* (3 volumes). New York: Oxford University Press. / first edition
- **Pustejovsky, James, Patrick Hanks, Roser Saurí, Andrew See, Robert J. Gaizauskas, Andrea Setzer, Dragomir R. Radev, Beth Sundheim, David Day, Lisa Ferro & Marcia Lazo.** 2003. 'The TIMEBANK Corpus' in D. Archer, P. Rayson, A. Wilson & T. McEnery. 2003. *Proceedings of the Corpus Linguistics 2003 Conference, 28-31 March 2003* (UCREL Technical Paper 16). Lancaster: Lancaster University, 647–656.

2003/2004

- **Saurí Colomer, Roser & Patrick Hanks.** 2003/2004. 'Iberian Names in North America: The Case of Asturian'. *Revista de Filoloxía Asturiana* 3-4: 89–114.

2004

- **Hanks, Patrick.** 2004a. 'Corpus Pattern Analysis' in G. Williams & S. Vessier (eds.). 2004. *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004, Lorient, France, July 6-10, 2004*. Lorient: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud, 87–97.
- **Hanks, Patrick.** 2004b. 'The Syntagmatics of Metaphor and Idiom'. *International Journal of Lexicography* 17.3: 245–274.
- **Hanks, Patrick.** 2004c. 'WordNet: What is to be Done?' *Panel Presentation at the Second International WordNet Conference (GWC 2004), Brno, Czech Republic, January 20-23, 2004*. Brno: Faculty of Informatics, Masaryk University, 1–6. Online at <http://www.fi.muni.cz/gwc2004/presentation/Hanks/hanks-panel.pdf>.
- **Hanks, Patrick & James Pustejovsky.** 2004. 'Common Sense about Word Meaning: Sense in Context' in P. Sojka, I. Kopeček & K. Pala. (eds.). 2004. *Text, Speech and Dialogue. Proceedings of the 7th International Conference (TSD 2004), Brno, Czech Republic, September 8-11, 2004* (Lecture Notes in Artificial Intelligence 3206). Berlin: Springer-Verlag GmbH, 15–17.
- **Pustejovsky, James, Patrick Hanks & Anna Rumshisky.** 2004. 'Automated Induction of Sense in Context' in *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004), 23-27 August 2004*. Geneva: University of Geneva, 924–930.

2005

- **Hanks, Patrick.** 2005a. 'Johnson and Modern Lexicography'. *International Journal of Lexicography* 18.2: 243–266.
- **Hanks, Patrick.** 2005b. 'Metaphors and Meanings: A Lexicographical Approach' in F. Kiefer, G. Kiss & J. Pajzs (eds.). 2005. *Papers in Computational Lexicography, COMPLEX 2005*. Budapest: Linguistics Institute, Hungarian Academy of Sciences, 81–106.
- **Hanks, Patrick.** 2005c. 'Similes and Sets: The English Preposition *like*' in R. Blatná & V. Petkevič (eds.). 2005. *Jazyky a jazykověda: Sborník k 65. narozeninám prof. Františka Čermáka* [Languages and Linguistics: A

Festschrift in Honour of Professor František Čermák's 65th Birthday]. Prague: Philosophy Faculty, Charles University in Prague.

⇒ Reprinted in P. Hanks & R. Giora (eds.). (to appear in 2010). *Metaphor and Figurative Language: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. VI.

- **Hanks, Patrick & James Pustejovsky**. 2005. 'A Pattern Dictionary for Natural Language Processing'. *Revue française de linguistique appliquée* 10.2: 63–82.
- **Nicholas, Nick & Patrick Hanks**. 2005. 'Robert Ingria, 1953-2003'. *Journal of Greek Linguistics* 6.1: 243–244.

2006

- **Hanks, Patrick**. 2006a. 'Conventions and Metaphors (Norms and Exploitations): Why Should Cognitive Linguists bother with Corpus Evidence?' in *Meeting of the German Cognitive Linguistics Association, Theme Session "Cognitive-Linguistic Approaches: What can we Gain by Computational Treatment of Data?"*, 7 October 2006, Munich, Germany, 9–12.
- **Hanks, Patrick**. 2006b. 'Definition' in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier, Vol. III, 399–402.
- **Hanks, Patrick**. 2006c. 'English Lexicography' in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier, Vol. IV, 184–194.
- **Hanks, Patrick**. 2006d. 'Lexicography: Overview' in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier, Vol. VII, 113–128.
- **Hanks, Patrick**. 2006e. 'Metaphoricity is Gradable' in A. Stefanowitsch & S. T. Gries (eds.). 2006. *Corpus-based Approaches to Metaphor and Metonymy* (Trends in Linguistics: Studies and Monographs 171). Berlin: Mouton de Gruyter, 17–35.
⇒ Reprinted in P. Hanks & R. Giora (eds.). (to appear in 2010). *Metaphor and Figurative Language: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. VI.
- **Hanks, Patrick**. 2006f. 'Nicknames' in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier, Vol. VIII, 624–626.

- **Hanks, Patrick.** 2006g. ‘Personal Names’ in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier, Vol. IX, 299–311.
 - **Hanks, Patrick.** 2006h. ‘Proper Names: Linguistic Status’ in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier, Vol. X, 134–137.
 - **Hanks, Patrick.** 2006i. ‘The English Language: An International Medium of Communication’ in J. Wu *et al.* (eds.). 2006. *Proceedings of the International Conference on EU-Fujian, China: Cross-Cultural Dialogue*. Xiamen: Xiamen University.
 - **Hanks, Patrick.** 2006j. ‘The Organization of the Lexicon: Semantic Types and Lexical Sets’ in E. Corino, C. Marello & C. Onesti (eds.). 2006. *Atti del XII Congresso Internazionale di Lessicografia, Torino, 6-9 settembre 2006 / Proceedings XII Euralex International Congress, Torino, Italia, September 6th-9th, 2006*. Alessandria: Edizioni dell’Orso, 1165–1168.
 - **Hanks, Patrick** (section ed.). 2006k. Section: “Lexicography” in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier.
 - **Hanks, Patrick, Kate Hardcastle & Flavia Hodges.** 2006. *Oxford Dictionary of First Names*. Oxford: Oxford University Press. / second edition (first edition in 1990)
 - **Hanks, Patrick, Anne Urbschat & Elke Gehweiler.** 2006. ‘German Light Verb Constructions in Corpora and Dictionaries’. *International Journal of Lexicography* 19.4: 439–457.
 - **Rumshisky, Anna, Patrick Hanks, Catherine Havasi & James Pustejovsky.** 2006. ‘Constructing a Corpus-based Ontology Using Model Bias’ in G. C. J. Sutcliffe & R. G. Goebel (eds.). 2006. *Proceedings of the Nineteenth International Florida Artificial Intelligence Research Society Conference, Melbourne Beach, May 11-13, 2006*. Menlo Park: AAAI Press, 327–332.
 - **Sheidlower, Jesse & Patrick Hanks.** 2006. ‘American Lexicography’ in K. Brown (chief ed.), *et al.* 2006. *Encyclopedia of Language and Linguistics, Second Edition*. Oxford: Elsevier, Vol. I, 184–193.
- 2007
- **Hanks, Patrick.** 2007a. *Corpus Pattern Analysis (CPA) – Project Page*. Online at <http://nlp.fi.muni.cz/projects/cpa/>.

- **Hanks, Patrick.** 2007b. ‘John Sinclair (1933-2007)’. *International Journal of Lexicography* 20.2, Euralex Newsletter Summer 2007: 209–215.
- **Hanks, Patrick.** 2007c. *Pattern Dictionary of English Verbs (PDEV) – Project Page*. Online at <http://deb.fi.muni.cz/pdev/>.
- **Hanks, Patrick.** 2007d. ‘Preference Syntagmatics’ in K. Ahmad, C. Brewster & M. Stevenson (eds.). 2007. *Words and Intelligence II: Essays in Honor of Yorick Wilks* (Text, Speech and Language Technology 36). Berlin: Springer-Verlag GmbH, 119–135.
- **Hanks, Patrick, Karel Pala & Pavel Rychlý.** 2007. ‘Towards an Empirically Well-founded Semantic Ontology for NLP’ in *Proceedings of the 4th International Workshop on Generative Approaches to the Lexicon, Paris, May 10-11, 2007*.

2008

- **Gehweiler, Elke & Patrick Hanks.** 2008. ‘The Linguistic Field: An Investigation’ in P. Hanks (ed.). 2008g. *Lexicology: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. II, 22–44. [Translation of **Jost Trier**’s (1934) ‘Das sprachliche Feld: Eine Auseinandersetzung’ in *Neue Jahrbücher für Wissenschaft und Jugendbildung* 10: 428–449.]
- **Hanks, Patrick.** 2008a. ‘How to Say New Things: An Essay on Linguistic Creativity’. *Brno Studies in English* 34: 39–50.
- **Hanks, Patrick.** 2008b. ‘Laurence Urdang (1927-2008)’. *International Journal of Lexicography* 21.4, Euralex Newsletter Winter 2008: 467–471.
- **Hanks, Patrick.** 2008c. ‘Lexical Patterns: From Hornby to Hunston and Beyond’ in E. Bernal & J. DeCesaris (eds.). 2008. *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)* (Sèrie Activitats 20). Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, 89–129.
- **Hanks, Patrick.** 2008d. ‘Lexicology: General Introduction’ in P. Hanks (ed.). 2008g. *Lexicology: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. I, 1–35.
- **Hanks, Patrick.** 2008e. ‘The Lexicographical Legacy of John Sinclair’. *International Journal of Lexicography* 21.3: 219–229.
- **Hanks, Patrick.** 2008f. ‘Towards a Diachronic Structural Semantics’ in P. Hanks (ed.). 2008g. *Lexicology: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. II, 140–193. [Translation of **Eugenio Coşeriu**’s (1964) ‘Pour une sémantique diachronique structurale’ in *Travaux de linguistique et de littérature* 2.1: 139–186.]

- **Hanks, Patrick** (ed.). 2008g. *Lexicology: Critical Concepts in Linguistics* (6 volumes). Abingdon: Routledge.
- **Hanks, Patrick** (guest ed.). 2008h. Special Issue: “The Legacy of John Sinclair”. *International Journal of Lexicography* 21.3.
- **Hanks, Patrick & Elisabetta Jezek**. 2008. ‘Shimmering Lexical Sets’ in E. Bernal & J. DeCesaris (eds.). 2008. *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)* (Sèrie Activitats 20). Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, 391–402.

2009

- **Hanks, Patrick**. 2009a. ‘Common Sense Blossoms in Springfield, MA’ = ‘Book Review: S. J. Perrault (ed.). 2008. Merriam-Webster’s Advanced Learner’s English Dictionary (Merriam Webster Inc.)’. *International Journal of Lexicography* 22.3: 301–315.
- **Hanks, Patrick**. 2009b. ‘Dictionaries of Personal Names’ in A. P. Cowie (ed.). 2009. *The Oxford History of English Lexicography*. New York: Oxford University Press, Vol. II, 122–148.
- **Hanks, Patrick**. 2009c. ‘Sestavljanje enojezičnega slovarja za domače govorce’ [Compiling a Monolingual Dictionary for Native Speakers]. *Jezik in slovstvo* 54.3-4: 7–24.
- **Hanks, Patrick**. 2009d. ‘The Impact of Corpora on Dictionaries’ in P. Baker (ed.). 2009. *Contemporary Corpus Linguistics* (Series: Contemporary Studies in Linguistics). London: Continuum, 214–236.
- **Hanks, Patrick**. 2009e. ‘The Linguistic Double Helix: Norms and Exploitations’ in D. Hlaváčková, A. Horák, K. Osolsobě & P. Rychlý (eds.). 2009. *After Half a Century of Slavonic Natural Language Processing (Festschrift for Karel Pala)*. Brno: Masaryk University, 63–80.

2010

- **Hanks, Patrick**. 2010a. ‘Nine Issues in Metaphor Theory and Analysis’ = ‘Book Review: A. Deignan. 2005. Metaphor and Corpus Linguistics (John Benjamins Publishing Company)’. *International Journal of Corpus Linguistics* 15.1: 133–150.
- **Hanks, Patrick & Kate Hardcastle**. 2010. *Babies’ Names*. New York: Oxford University Press. / second edition (first edition in 1995)
- **Jezek, Elisabetta & Patrick Hanks**. 2010. ‘What Lexical Sets Tell Us about Conceptual Categories’. *Lexis: E-Journal in English Lexicology* 4 (Corpus Linguistics and the Lexicon / La linguistique de corpus et le

lexique), 7–22. Online at <http://screcherche.univ-lyon3.fr/lexis/spip.php?article136>.

To appear in 2010

- **Hanks, Patrick.** (2010b). ‘Compiling a Monolingual Dictionary for Native Speakers’. *Lexikos* 20.
- **Hanks, Patrick.** (2010c). ‘Elliptical Arguments’ in S. Granger & M. Paquot (eds.). *Proceedings of ELEX 2009 – eLexicography in the 21st century: New challenges, new applications* (Cahiers du Cental). Louvain-la-Neuve: Presses universitaires de Louvain.
- **Hanks, Patrick.** (2010d). ‘Lexicography, Printing Technology, and the Spread of Renaissance Culture’ in A. Dykstra (ed.). *Proceedings of the 14th EURALEX International Congress, 6-10 July 2010, Leeuwarden/Ljouwert, The Netherlands*.
- **Hanks, Patrick.** (2010e). ‘Making Meanings’ in B. Sharp & M. Zock (eds.). *Proceedings of the 7th International Workshop on Natural Language Processing and Cognitive Science (NLPCS 2010), 8-9 June 2010, Funchal, Madeira, Portugal*.
- **Hanks, Patrick.** (2010f). ‘Marcus Fabius Quintilianus, *Institutio Oratoria* (Institutes of Oratory)’ in P. Hanks & R. Giora (eds.). *Metaphor and Figurative Language: Critical Concepts in Linguistics*. Abingdon: Routledge, Vol. I. [Translation of **Quintilian**’s ‘*Institutio Oratoria*’ (AD 95), Bk. 8, Ch. 6.]
- **Hanks, Patrick.** (2010g). ‘Terminology, Phraseology, and Lexicography’ in A. Dykstra (ed.). *Proceedings of the 14th EURALEX International Congress, 6-10 July 2010, Leeuwarden/Ljouwert, The Netherlands*.
- **Hanks, Patrick & Rachel Giora** (eds.). (2010). *Metaphor and Figurative Language: Critical Concepts in Linguistics* (6 volumes, with 6 introductory articles). Abingdon: Routledge.

In preparation

- **DeCesaris, Janet & Patrick Hanks.** (forthcoming). ‘The Persistence of Technology in Lexis’.
- **Hanks, Patrick.** (forthcoming a). ‘Book Review: D. Geeraerts. 2010. *Theories of Lexical Semantics* (Oxford University Press)’. *International Journal of Lexicography*.
- **Hanks, Patrick.** (forthcoming b). ‘Coding Semantic Properties of Words in Computational Dictionaries’ in R. H. Gouws, U. Heid, W. Schweickard & H. E. Wiegand (eds.). *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Develop-*

ments with Special Focus on Computational Lexicography (Handbooks of Linguistics and Communication Science). Berlin: Walter de Gruyter.

- **Hanks, Patrick.** (forthcoming c). ‘English and American II: Synchronic Lexicography’ in R. H. Gouws, U. Heid, W. Schweickard & H. E. Wiegand (eds.). *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Special Focus on Computational Lexicography* (Handbooks of Linguistics and Communication Science). Berlin: Walter de Gruyter.
- **Hanks, Patrick.** (forthcoming d). *Lexical Analysis: Norms and Exploitations*. Cambridge, MA: MIT Press.
- **Hanks, Patrick.** (forthcoming e). ‘Lexicography’ in R. Mitkov (ed.). *Oxford Handbook of Computational Linguistics, Second Edition*. New York: Oxford University Press.
- **Hanks, Patrick.** (forthcoming f). ‘Lexicography from Earliest Times to the Present’ in K. Brown & K. Allan (eds.). *Oxford Handbook of the History of Linguistics*. New York: Oxford University Press.
- **Hanks, Patrick.** (forthcoming g). ‘Linguistic Creativity’ in T. Veale (ed.). *The Agile Mind*.
- **Hanks, Patrick.** (forthcoming h). ‘Monolingual Lexicography’ in C. Chapelle (chief ed.), et al. *Encyclopedia of Applied Linguistics*. Hoboken, NJ: Wiley-Blackwell.
- **Hanks, Patrick.** (forthcoming i). ‘Representing the Unrepresentable: Dictionaries, Documents, and Meaning’ in P. M. Bertinetti (ed.). *Atti del XLII Congresso Internazionale di Studi, Pisa, Scuola Normale Superiore, 25-27 settembre 2008*.
- **Hanks, Patrick.** (forthcoming j). ‘The Lexicon’ in R. Mitkov (ed.). *Oxford Handbook of Computational Linguistics, Second Edition*. New York: Oxford University Press.
- **Hanks, Patrick.** (forthcoming k). ‘Wie aus Wörtern Bedeutung entsteht: Semantische Typen treffen auf syntaktische Abhängigkeiten’ in *Sprachliches Wissen zwischen Lexikon und Grammatik, 46. Jahrestagung des Instituts für Deutsche Sprache (IDS), 9. - 11. März 2010*.
- **Hanks, Patrick & Peter McClure.** (forthcoming). ‘Some Methodological Considerations in Approaching the Study of Family Names’. *Nomina: Journal of the Society for Name Studies in Britain and Ireland*.
- **Mitkov, Ruslan & Patrick Hanks** (eds.). (forthcoming). *Oxford Dictionary of Computational Linguistics*. New York: Oxford University Press.