Hanks, Patrick. *Lexical Analysis: Norms and Exploitations.* 2013, 462 pp. ISBN 978-0-262-01857-9. Cambridge, Mass./London, England: The MIT Press.

Patrick Hanks is well known as a lexicographer and as the author of several remarkable articles on phraseology, collocations and co-occurrences and on the description of meaning. He was the chief editor, or one of the editors, of several dictionaries, some of which are highly original, particularly in their treatment of polysemy. Many people were hoping that he would eventually develop his views in a book, and this had been 'announced as "forthcoming" for many years'. 'Some people ... had given up hope that it would ever appear' (xv), but now, at last, after a period of preparation of sixteen years (215), what began as a 'disjointed collection of short essays and other fragments' has become 'a coherent text' (xv) of almost 500 pages.

The book is divided into thirteen chapters, five of which (chapters 3, 4, 5, 6 and 8) are based on previously published papers. They were all rewritten, but there is inevitably a certain amount of repetition that those readers who are already familiar with Hanks's work may find unnecessary but that will be useful to all those who are not. The book can be seen as a sort of hypertext, in which each point is first introduced in the general context of the theory and then more thoroughly developed in the following chapters. The present review reflects some of these repetitions.

Each chapter begins by a short (sometimes very short) abstract (except chapter 1 and chapter 13) and ends with a summary of the main points (except chapter 13). Oddly enough, there is no introduction. The Acknowledgments has the usual listing of the names of all those who helped Hanks develop his theory or write his text (Gilles-Maurice de Schryver played a decisive role in the encouragement of the author and also, apparently, in the shaping of the book), but it is mainly the story of how Hanks's ideas developed. He was influenced by John Sinclair and James Pustejovsky, and, to a lesser extent, by M.A.K. Halliday, Yorick Wilks, Charles Fillmore, Anna Wierzbicka and no doubt other linguists, but he is above all a lexicographer. He happened to be active when corpora became available for dictionary-making, when John Sinclair published his most important work and when prototype theory appeared in linguistics, and he was one of the first who believed that all three were important for lexicography and for semantics. His book is 'rooted in practical experience of monolingual lexicography' (xiv). His first job, from 1965 to 1970, was the edition of the Hamlyn *Encyclopedic World Dictionary* (1971), a dictionary that has been forgotten but was one of the sources of the *Collins English Dictionary* (1979), that Hanks also edited. He then worked with John Sinclair on the preparation of the *COBUILD Dictionary of the English Language* (1987), and that is when he began working with a corpus, 7 million words at first, then 18, and a lot more since then. In 1990 he moved to Oxford University Press, where he was chief editor of Current English Dictionaries. At the end of the decade he

produced the *New Oxford Dictionary of English* (1998; with a second edition called *Oxford Dictionary of English*), to this day the general dictionary of English in which the presentation of word meanings was most influenced by his ideas. In 2000 he worked with James Pustejovsky for a software company in Cambridge, Mass., and since the early 2000s he has been teaching at various universities: Brandeis, Berlin, Prague, Brno, Wolverhampton and Bristol. This, presumably, gave him more time for research and theory than he had ever enjoyed as a lexicographer, and the experience that he accumulated over the years eventually became the theory presented in this book (see also De Schryver 2010).

'Words and Meanings: the Need for a New Approach' (23 pages) is the first and certainly the most important chapter: it sums up the essence of Hanks's theory of norms and exploitations (TNE), 'a lexically-based, corpus-driven, bottom-up' (17) theory of language that matured after 'a lifetime of editing and writing dictionary definitions — attempting to account for the meaning of words, wrestling with the problem of word meaning' (7). Basically, the theory is simple: (i) words have normal, i.e. usual, typical, common, conventional, usually frequent, usages and meanings, to be distinguished from marginal, unusual, uncommon, atypical, non-conventional usages and meanings; (ii) any normal usage may vary in lexis, semantics and syntax while remaining basically the same; and (iii) the meaning of a word is determined by its environment, i.e. its valencies, its phraseologies, its collocations and more generally its context. The latter is not very original: many other linguists have said the same, in various ways, from Saussure to Firth to Sinclair, and it is also what Wittgenstein is remembered for in linguistics. Except that the linguist working with a large corpus has access to more contexts in less time, and can therefore draw conclusions that are more robust than those that were accessible before. And the first two points above are even more interesting: they are Hanks's conclusions after having worked with corpora that became larger and larger over the last three decades, and those conclusions were simply out of reach for linguists who did not have access to a corpus. Most preceding linguists did not have one, and it is not certain that they would have used it if they had, Hanks says (and he will say it again in chapter 12). Those who had one have tended to use it as a 'fish pond', in which to angle fish that fit their theory, while fish that did not were thrown back into the pond. Hanks, on the contrary, wants 'to find out what sort of fish are in the pond' (7): TNE is corpus-driven and bottom-up.

Most linguists have relied on their own intuitions, like Fillmore or Apresjan, 'a recipe for self-fulfilling prophecies' (20), Hanks says, or on the acceptability judgements of native speakers, like Mel'čuk or Wierzbicka (or Cruse, but Cruse is never mentioned, not even in the References), but that did not allow them to distinguish the normal from the more or less abnormal. Acceptability judgements distinguish between what is possible and what is impossible, but that is not the main distinction, Hanks says: what matters most is the distinc-

tion between normal and 'less normal' utterances. Worse, judgements on one's own linguistic usage or on the usage of others are unreliable: they tend to focus on what is cognitively salient, i.e. striking, unusual, and to ignore what is socially salient, i.e. frequent, common (21); all lexicographers know that. Linguistics has neglected the more common usages and given too much attention to rare, unusual, marginal usages, for several reasons: because there are many, as the study of large corpora has shown (18), because they are cognitively salient, and because they make it possible to explore the limits of language, what the code accepts and what it doesn't, an implicit objective of much of modern linguistics. 'The linguist's tolerance of abnormality is unusually great' (20), Sinclair once wrote, and the consequences have been dramatic, Hanks says: it 'left linguistics drowning in a welter of imagined possibilities' (4).

What does Hanks find in his corpus? He finds evidence for the three points above: that every content word has a frequent usage, or a few frequent usages, that can therefore be considered normal, and less frequent, less normal usages; that every normal usage shows more or less important variations in lexis, semantics and syntax, and can be played with to produce special effects — so that the language users have a double competence, a competence to use words according to the norm and a competence to play with the norm, to exploit it; and that every meaning corresponds to a specific context, that can be more or less frozen. Often, though not always, one usage is clearly dominant. An example (given in chapter 13): in the British National Corpus, *spoil* is used in the sense, and in the pattern, of 'spoil an event that should be enjoyable' 60% of the time, the other usages being much less frequent ('spoil a view' 18%, 'spoil a child' 11%, 'food spoils' 3%, 'be spoiled for choice' 3%, 'spoil a paper ballot' 1%) (427).

TNE is a theory, in the sense that it aims at describing how language works, with lexis at its centre, but it begins as a practice, a method for the study of corpora in order to extract meanings based on textual evidence, not on intuition or acceptability judgements, to determine which are dominant, and to describe them. How does Hanks proceed? He starts from a word (a notion discussed in chapter 2; see below) and notes all the contexts in which the word is used, and in what form (though the influence of the variations of form on the use of a lexical item is one of the points that could have been developed further in the book). Those contexts can be reduced to patterns, with all their important elements, other content words and function words, arranged in a syntactic form — TNE is not only lexically based, corpus driven and bottom up; it is also pattern based. The analyst then proceeds to isolate those patterns that illustrate the normal uses of the word, for which there are a substantial number of similar, if not exactly identical, examples in the corpus. At this early stage, all the other uses are left on the back burner: they are either performance errors, in which case they will be ignored, or exploitations of the norm, i.e. rare uses based on normal patterns but differing from them in a way that is not governed by any discernible rule (see chapter 8); those will be examined later. They are

potentially interesting, because they are the inventive, the bold, the playful, deliberate or inadvertent uses of journalists, of fiction writers, of poets and other actors of linguistic creativity, most of which will be forgotten but some of which will survive and become new norms, but they are not the first priority. When the normal patterns have been listed, the lexical analyst tries to determine the meaning of the word in each pattern, surely the most difficult stage of the process, in which intuition and previous knowledge of the language and of the world play a more or less important role. The results are in the form of a list of all the normal patterns of the chosen word, with a meaning corresponding to each pattern, or vice versa, arranged according to their 'degree of normality'. This is radically different from what can be found in lexical bases such as, for example, WordNet, and in traditional dictionaries, where meanings are attributed to isolated words (19), where the meanings of polysemous words are listed chronologically, or according to some 'logic', and where all meanings are given equal weight, regardless of their normality.

Of course, the operation is not as simple as it looks, and the following chapters examine some the difficulties encountered by the lexical analyst. One of the first is the definition of 'normal' discourse: the usages that are retained for analysis must be authentic, but authenticity is not enough; they must also be natural (not stilted like many of the examples invented by linguists), usual, typical, common, conventional, frequent enough, and socially salient, i.e. mastered passively and actively by 'all' the users of the language community. Deciding what is entirely normal and what is less so may be difficult, because the boundaries are fuzzy. The main criterion is frequency: if a use is very frequent it must be considered normal. But a lower frequency does not designate a marginal use. For example, the Oxford English Corpus of 1.5 billion words has only six examples of *tell* in the sense 'The strain was beginning to tell', but Hanks decides that it must be considered a norm, because the meaning is distinct enough (16) — a case in which meaning, not frequency, is the basis of the decision.

The patterns that correspond to meanings reveal lexical sets, i.e. groups of words that can be used — and are used — in a particular position of a particular pattern with a word in a particular meaning. For example, the direct objects of *fire* in one of its meanings (*gun*, *rifle*, *pistol*, *revolver*, *machine gun*, etc.) are a lexical set. In another meaning, the objects of *fire* are human beings, more precisely employees, another lexical set. The words of a lexical set are united by a semantic type, and semantic types can be placed in an ontology: for example, the direct objects of *fire*[1] are firearms, firearms are artefacts, etc. (13). This will be developed in chapter 5.

Chapter 2, 'What is a Word?' (40 pages), explores the concept of 'word' together with the neighbouring concepts of 'type', 'token', 'lemma', 'lexeme', 'phraseme' or 'multiword expression', and 'lexical entry'. The question is what constitutes a unit for a theory in which the lexicon plays the central role. The discussion does not end with a clear conclusion recommending terms or con-

demning others (29), but in the rest of the book Hanks uses mostly *word* and *lexical item*, though the latter is 'used in different ways by different writers' (29); *lexical unit* is used on page 389 to describe Fillmore's work but is not defined and is absent from the index. The chapter ends with a discussion of neology and of proper names that could have been shorter. Neology concerns mostly nouns, less often adjectives, rarely verbs and virtually never function words (42). The lexical types that are created are mostly terms (i.e. lexical items with a precise meaning that has been stipulated by the specialists of the domain) and multiword expressions, two categories that are not well covered in traditional dictionaries. The pages on proper names (33 *sq*.), though not absolutely necessary in the general architecture of the book, are interesting because proper names are rarely mentioned in books on semantics. Hanks is a specialist: readers may remember that he edited a dictionary of first names and a dictionary of surnames, and that the *Hamlyn* and the *Collins* that he edited both had proper names, contrary to the mainstream tradition of general dictionaries in the twentieth century in England. Proper names are numerous in many sorts of texts, and will be more and more numerous in monitor corpora: 'In some large lexical databases, aiming at full coverage of a language, over 70% of the lexical entries already are proper names, and this percentage continues to increase' (35). They cannot be neglected, because they carry information, real-world knowledge that is part of the competence in a language, and future lexicographers, Hanks argues, will have to include more of them and describe them more fully. This is not pursued in the rest of the book.

Chapter 3, 'Do word meanings exist' (19 pages), is a new version of a paper published in *Computers and the Humanities* in 2000 that has been widely regarded as an important contribution to the literature. Hanks argues that words do not have meanings in isolation, that they only have meaning potentials, and that these potentials are activated when the words are used. Meanings, as a consequence, are best seen as events rather than as entities: they take place in contexts of space and time. The meaning potential of a word is made up of semantic components — the 'semes' of other linguistic schools, but Hanks does not use the word. For many words, one semantic component is salient, i.e. it is activated most of the time, if not all the time, but for others 'different combinations ... are activated in different contexts' (83). Readers may think of the semantic portrait of *game* by Wittgenstein, but there are many other examples. The identification of these components and the exploration of the ways in which they combine in different contexts (82) are among the major tasks of lexicographers, and they have not done very well so far, again because they have tended to define words in isolation.

Hanks then returns to the relation between meaning and context. One question is how far the lexical analyst should cast his net when trying to determine the meaning of a word on the basis of its contexts of use. 'In the overwhelming majority of cases, a correct meaning can be assigned to a keyword on the basis of clues in its immediate environment' (81), but in other cases

it is necessary to consider a wider context, particularly for nouns (see also chapter 5). This is out of the reach of the corpus analyst working with a KWIC list but it can be retrieved more or less easily in all modern corpora. Normally, consideration of the context and of phraseology points to the right meaning, but the operation is not 'a magic bullet': it solves many problems of ambiguity, but not all (82). For example, the two meanings of *check* ('inspect' and 'cause to slow down or stop') are used in the same syntactic pattern, and in some instances of use they even co-exist (76), i.e. the same context allows the two meanings, not one or the other but the two together. Polysemy, Hanks concludes, is much more complex than its presentation as a flat list of distinct meanings in most dictionaries suggests.

Chapter 4, 'Prototypes and norms' (28 pages), examines phraseology, i.e. the more or less frozen patterns that help define a meaning. Patterns have been neglected by dictionaries in general, particularly by English dictionaries — German, Czech and Modern Greek have done better. Lexicographers have 'tended to focus on conventions of meaning and to neglect conventions of phraseology' (104), Hanks writes, perhaps because phraseology is more difficult to identify and describe. Yet patterns are important, as chapter 1 has made clear, because they are keys for the identification of meaning. They are what the language user has to analyze and interpret to give each word its right meaning and understand the message. The lexicon is 'a store of shared beliefs or meaning potentials, each of which is associated with one or more phraseological norms' (87). Patterns are what the lexical analyst sees in a corpus, and frequent patterns suggest common usages. 'Any sizable corpus will usually show a very large number of very similar uses of each word — similar not only in terms of syntactic construction but also in terms of preferred collocations', Hanks writes again (91). A norm is identified by grouping similar corpus lines together around a phraseological prototype; note the word *similar*, meaning that the phraseologies that are grouped together are not necessarily identical; they are formally different manifestations of the same basic pattern, and 'judgment is required to decide what counts as "similar"' (92). Once this has been done, the frequency of each pattern can be calculated, and that will help in the identification of normal usages: '... unusual uses are put on one side for later analysis ..., and then the number of corpus lines in each group is counted to discover the comparative frequency of each pattern in the sample' (92). As we have seen in chapter 1, some rare usages must also be 'recognized as patterns, despite their rarity, because they have distinctive meanings' (92; an example is given in chapter 7; see below). When a pattern has been identified as a norm, with its variations (alternations and exploitations), 'the next step is to associate it with a meaning' (95). Patterns normally serve to determine meanings, and, in some rare cases, meanings serve to identify patterns; the apparent contradiction is probably unavoidable.

The number of patterns associated with each word is highly variable. For 'most words, just one or two patterns are salient (socially salient, i.e. frequent),

while the other patterns are less common' (95), as we have seen for *spoil*. The verb *climb* has two basic meanings (and is presented accordingly in the *Oxford Dictionary of English*), 'clamber' and 'ascend', and a number of other usages that are much less common. Each basic meaning is associated with groups of normal subjects, normal objects and normal adjuncts, lexical sets that correspond to semantic types but play a syntagmatic role: 'Different lexical sets in different syntactic roles can alter the meaning of the target word' (105). These sets are fuzzy: for example, some words may belong to one set with one verb but not with another verb with a very close meaning. Lexical sets are 'prototypical in character' (105), with words that are more central than others, because they are found in the lexical sets of all, or almost all, related words.

Chapter 5, 'Contextual dependency and lexical sets' (31 pages), continues the description of the operation in which the analyst collects patterns, defines lexical sets for each syntactic slot and then determines the meaning of the central word for each pattern: the 'identification of normal complementation patterns by the corpus analyst, not only in terms of valencies but also in terms of lexical sets, is an essential step in determining a word's meaning' (113). The 'semantics of each word in a language is determined by the totality of its complementation patterns', Hanks writes (113), echoing Firth's pronouncement: 'You shall know a word by the company it keeps' (Firth 1968: 179). For a verb, the essential elements to consider are subjects, objects and adverbials, and for a noun they are the words that are typically found in its wider context, even those that have no syntagmatic relationship to it. Corpus evidence shows 'what patterns of usage are normal, central, and typical and ... which patterns are the most frequent'. They are those that should be retained for description: the corpus analyst is 'concerned with the regular and the normal, not with the boundaries of linguistic possibility' (115), Hanks writes again. But, as we have seen, the identification of regular patterns is only the first step in the identification of meaning: 'corpora provide direct evidence for patterns of usage, but only indirect evidence for meanings' (116). For many words, one pattern is highly dominant; *spoil*, again, or the verb *urge*: 61% of its uses are in the pattern of 'a person urging another person to do something', while 'a person urging a steed or another person onward or upward (or in some other direction)' accounts for only 3.5% (117). Dictionaries never say what meanings are dominant, and in what proportions; all meanings are given equal weight.

Hanks then returns to the construction of lexical sets. They must be assembled with great care, he says: if they are too narrow or if they are too broad they are not much use in the identification of meaning. For example, one can urge a horse on, or a stallion, or a camel (or, I suppose, an ox, an elephant, etc.). Or a car, the corpus says, in a few examples that are clearly an exploitation of the basic pattern. But even if one such example describes a Ford Sierra being urged on over difficult terrain by its driver it would be unwise to include *Sierra* in the lexical set of the nouns that can be objects of the verb *urge*, because then any name could be a member as well (118). Similarly, it would be point-

less to include *John* or *Sylvia* in the lexical set of the objects of the verb *fire* in the sense of 'fire an employee'.

Some semantic sets are unified not only by their semantics, the meanings of their words, but also by their 'semantic prosody', a concept introduced by Sinclair (1991) and Louw (1993) to designate the positive or negative connotation of some words. An example used by Sinclair is the verb *set in*: it has a marked preference for subjects such as *rot*, *decay*, *malaise*, *despair*, *decadence*, *impoverishment*, *infection*, *prejudice*, etc., all with a negative semantic prosody, and therefore *set in* can also be said to have a negative semantic prosody (124) — note the word *preference*, meaning that other subjects are possible, but marginal. For nouns, the analyst must identify 'statistically significant collocates in the environment of the target word', and as we saw earlier such collocates 'do not necessarily have to be in a structured relationship' (134) with the word. Incidentally, a large corpus reveals things that could not be noted before. For example, Hanks notes that in his corpus the word *spider* co-occurs significantly with the word *bath*, but this seems to be restricted to English; other languages do not show the same co-occurrence (135). Why? Is it because there are fewer spiders in other communities, or fewer baths, or because people are less afraid of spiders, or because foreign spiders do not like baths, or because a spider in a bath is so common in other countries that nobody mentions it, or for another reason? Such correlations in the wider contexts are potentially interesting for language teachers.

The identification of the relevant elements in each relevant pattern is not an easy operation: 'they have to be teased out, often painstakingly and slowly' (141). Hanks concludes that much remains to be done: 'procedures have to be developed for distinguishing relevant features from mere noise. Appropriate levels of generalization have to be chosen at every step, for every pattern of every word' (141).

Chapter 6, 'Norms change over time' (27 pages), starts from a well-known fact: word meanings change with time, so that we sometimes find it difficult to interpret a word used in an old text. In TNE these changes are part of the evolution of norms. To study them, the lexical analyst needs a historical corpus with 'examples of "everyday" texts ... as well as great works of literature' (145). That is not easy, because 'the work of great writers of the past tend to have survived, whereas the mundane, everyday use of a language ... has left fewer traces behind' (158). Interesting examples are the words *enthusiasm* and *condescension* as used by Jane Austen. Hanks has discovered that *enthusiasm* was slowly changing from negative to positive semantic prosody in the eighteenth century when Austen was writing, while *condescension* was moving in the opposite direction, from positive to negative. With TNE the analyst can try to reconstruct the system at a given period of a given dialect, with its primary norms, its secondary norms, its alternations and its exploitations, and to understand how the whole system evolved (see chapter 10). In passing, Hanks discusses a view of literature that has ancient origins but became popular for a

while in universities in the late twentieth century, according to which as soon as it has been written the text acquires a life of its own and belongs exclusively to the reader, not to the author. There is some truth in this, Hanks says, but in its extreme version it is a 'first step down a road leading to the absurdity of Humpty Dumpty's position' (a word means whatever I want it to mean), 'a solipsistic universe ... in which other minds, other worlds, learning, and scholarship — the whole social consensus of meaning in language — all count for nothing. In such a universe, language itself becomes meaningless' (156). How refreshing, for those of us who have seen the devastation produced in the minds of young students by some quasi-fanatical advocates of the more extreme versions of that approach.

Chapter 7, 'Three types of alternation' (38 pages), details the variations that can take place in normal phraseologies. Alternations are of three types: a pattern can vary in lexis (*clutching / grasping at straws*), semantic-type (*treating injured people / injuries / injured legs*, etc.) and syntax (*He broke the window / The window broke*) and yet remain the same pattern. Chapter 8, 'Exploitations' (40 pages), is about the distortions of a norm that 'create new meanings ad hoc and ... say old things in new ways' (211), that Hanks chooses to call 'exploitations'. The difference may have been unclear when Hanks started discussing the two phenomena, but it clarifies when he writes: 'alternations are regular, second-choice elements within an overall pattern, exploitations are typically dynamic, creative, or graphic choices within the boundaries of possible language use' (212). Exploitations are typically infrequent, i.e. not socially salient, but created to produce a strong effect, i.e. cognitively salient (214), so that the 'distinction between alternation and exploitation is in many cases one of frequency, coupled with semantic or rhetorical effect' (216). Indeed, frequency is not always enough, as we have seen. In one example (given in chapter 10), Hanks decides that 'The industry is scratching its head' is an exploitation of a normal pattern of the verb *scratch*, not an alternation, because the two main words (*industry* and *head*) show 'incoherence' (291). Of course, here again, as often in linguistics, there is no sharp division between the two phenomena, alternation and exploitation, because normality is a cline: 'Some uses of words are more normal than others' (214).

The chapter continues with a review of the main tropes that can be used to exploit a norm: metonymy, synecdoque, zeugma, oxymoron, understatement, euphemism, etc. Here Hanks gets a bit carried away by his enthusiasm (in the modern sense), as when he discusses cases of hyperbole, irony or sarcasm, whose role in the exploitation of a norm is not obvious (236). The use of tropes in the creative use of language brings Hanks back to the question of ambiguity that he has discussed in many of his publications, because it is one of the main points where he differs from other linguists: 'Ambiguities are plentiful in the literature of linguistics, because they are based on invented examples isolated from any real context of utterance, but corpus evidence shows that genuinely baffling ambiguity arises in remarkably few cases' (243). In most cases, the

context leaves no doubt as to which meaning is being used.

The study of exploitations continues in chapter 9, 'Intertextuality: Literature and the exploitation of norms' (32 pages), with an exploration of the role of noteworthy language users, famous authors among them, but also journalists, teachers, lawyers, etc., in the evolution of word meaning through their exploitations of norms, and how their creations often become norms in their turn. Most of the chapter is about Shakespeare and the Bible, of course, and there is a very good — though more important for lexicography than for lexical analysis — passage about the lexical innovations of James Joyce in *Finnegan's Wake* (275 *sq*.). Should they be included in dictionaries? What are the best criteria for inclusion *vs* exclusion? How can they be defined if there is only one occurrence, or even if they are used by only one author? The *Oxford English Dictionary* has a few of these *hapax legomena*, but it is difficult to see what criteria were used to select them and reject the others.

Chapter 10, 'Word and pattern meaning: A complex linguistic gestalt' (21 pages), contains 'portraits' of words of varying complexity for the lexical analyst, *scratch*, *throw* and a few others. A norm can give rise to an exploitation that can become a secondary norm and/or be used as a basis for another exploitation, and so on, creating a web of, in some cases, extreme complexity. This raises the question of how the language users store such complexities, and how they can pick the right interpretation when the word is used (298).

Chapter 11, 'Meaning, philosophy of language, and anthropology' (42 pages) and Chapter 12, 'The role of the lexicon in linguistic theory' (62 pages, by far the longest chapter), could have been placed earlier in the book, but undoubtedly they are the two chapters that students (and teachers) of linguistics will need to have read carefully and keep for future reference. In chapter 11, Hanks notes that 'many of the developments that have most fundamentally affected our understanding of meaning in language took place in philosophy of language and anthropology rather than in linguistics' (306), but then proceeds to describe the work of linguists such as Ogden and Richards, Bar-Hillel, Wierzbicka as well as that of philosophers and anthropologists, Aristotle, Wilkins, Leibniz, Wittgenstein, Grice, Austin, Rosch, Putnam, Kripke, and others. Now that large corpora are available, Hanks writes, containing plentiful evidence of language as it is normally used, 'an immense task of sifting lies ahead, to determine which linguistic hypotheses can be maintained satisfactorily and which must be modified or abandoned' (307). One problem with what philosophers have written about language and meaning is that they have failed to distinguish between 'meaning in naturally occurring language and the stipulated scientific meaning of a rigorously defined concept' (311; Hanks does not say how the definition of a scientific term can be rigorous if the words used in its definition have fuzzy meanings, but that is a minor point in the book). For most philosophers, and for many linguists, 'the vagueness and fuzziness of meaning of ordinary words ... was an imperfection' (313), and it was not until the 1960s that they 'began to realize that far from being a minor imperfection, [the] fuzzi-

ness of word meaning is a central design feature of natural language, contributing a flexibility that allows existing words to be applied to new situations and to be exploited in new and interesting ways' (336). Many linguists, and most philosophers of language have tried to describe an idealized language that did not suffer from the imperfections of real discourse. They 'spent enormous amounts of time and effort speculating about possible but implausible sentences' (307), failing to distinguish between possible usage and normal usage. Their work is 'strewn with examples of self-fulfilling theoretical prophecies, in which bizarre examples are first invented, then judged to be acceptable (according to the researcher's intuitions), and then presented as evidence for conclusions about some aspect of the nature of language or linguistic rules. However, bizarre examples are conducive to bizarre theories' (307).

In Chapter 12, Hanks continues the presentation of the work of his predecessors, only linguists this time, situating the theory of norms and exploitations in relation to other theories of language, past and present (347). Having noted that the importance of the lexicon has been underestimated by a vast majority of linguists until the late twentieth century, he reviews the work of Humboldt, Saussure, Trier (and other advocates of semantic field theory), Apresjan, Mel'čuk, Chomsky, Jackendoff, Meyer, Bresnan, Pustejovsky, Langacker (and other cognitivists), Fillmore, Firth, Halliday, Sinclair, Hoey, Stubbs, etc. He stresses, again, that many of them did not use, or refused to use, a corpus and therefore based their theories on invented examples that led them to debatable conclusions. One can feel his pleasure when he discusses, for example, Chomsky, acknowledging his importance but at the same time arguing that his theories are far removed from the reality of discourse: wrong objective (how to construct well-formed sentences), wrong method (introspection), constant hesitations and re-adjustments over the last fifty years, and eventually results that have not kept their promises in the clarification of how language works.

Hanks's review of the literature in those two chapters is good reading and will prove extremely useful. His culture is immense, as befits the editor of the monumental *Lexicology: Critical Concepts in Linguistics*, published in 2008, and his views are challenging because they are those of a linguist with strong convictions against which the work of his predecessors can be measured. Sometimes, however, one is left wondering how exactly the work being described relates to TNE, or even if it does at all. Could Hanks, again, have been carried away by his enthusiasm? How, for example, can Ogden and Richards' semiotic triangle be useful in the lexical analysis of a corpus? (329); How can Wittgenstein's well-known discussion of the meaning of *game* be useful to TNE? (325); How does Wierzbicka's treatment of *game* 'stand up in the light of corpus evidence'? (326); What use are Austin's notions of locution, illocution and perlocution for Hanks's work? (334), etc. No doubt Hanks could answer these questions, but I wish he had been more explicit in the book. Other (admittedly minor, in the context) questions are not answered: Why does Wilkins's work 'seem very odd to modern readers'? (314); Why has Leibniz's immense influ-

ence on thinking about language been 'often based on misunderstanding of his work'? (316); How exactly was Firth influenced by Malinowski? (335); etc.

Chapter 13, 'The broader picture' (21 pages), sums up the main points of TNE: Making 'predictions about probable usage is much more useful than speculating about the boundaries of possibility' (415); TNE is about 'rules for using words, rather than [about] rules for constructing sentences' (416); 'The normal meanings of a word can be extrapolated from statistical study of a large number of its uses' (410); 'The difficulty lies in achieving just the right level of generalization' (411); Most of the time, though not always, 'different patterns activate different meanings' (an exception is *hazard*, as in 'hazard one's life' and 'hazard a guess') (413). The final pages examine the fields in which TNE could be useful: the semantic web, natural language-processing, artificial intelligence, language learning, language teaching, and of course computational lexicography. Hanks mentions his own *Pattern Dictionary of English Verbs*, saying that 'at the time of writing (August 2010), approximately 12.5% of the *PDEV* is complete, after five years of work. At the current rate of progress if there is not a substantial injection of funds enabling the project to recruit a professional lexicographic staff, it will not be completed until 2045, when the author will be 105 years old' (427). Good luck, Patrick.

It is difficult not to be convinced when Hanks says that words have normal usages and meanings, that normal usages vary in lexis, semantics and syntax, that the meaning of a word is determined by its environment and that words in isolation only have meaning potentials. What could the unconvinced say? They could note that the method raises a chicken-or-egg question. In TNE, the meaning of a word is determined by its contexts, but, as we have seen in one or two instances, there are cases where Hanks has to admit that it is the presence of a word that makes it possible to interpret the context. A (marginal?) example: in 'Doctors treating Michael Gibson', the 'sense of the named entity ("Medical Patient") is activated in this context by the verb *treat*' (178). More precisely, the chicken-or-egg question is between semantics and syntax. Does meaning determine syntax, or does syntax determine meaning? Hanks argues that meaning is too vague and variable to be the basis of the operation but his analyses sometimes work the other way round.

One could also say that much of what Hanks says is not really new. This is partly true, but partly only. Hanks shows chapter after chapter that the analysis of word meaning in a large corpus has already uncovered many aspects of language use that had never been discussed, and that more can be expected in the future. One could even say that Hanks's theory solves only those problems of identification and definition of word meaning that could already be solved by more traditional methods, if not by pure and simple intuition. But, again, that is not really fair: many of Hanks's word studies end with conclusions that differ from what is on offer in traditional dictionaries. It is true that even after a careful study of all contexts, the linguist may be left with a residue of cases that resist analysis. Hanks himself readily admits that TNE does not always work

smoothly. He knows that in some cases the analysis of the corpus must be aided by 'introspection and a minimal amount of editorial art' (148). But this may be because the problem has no solution: 'in many cases, there is no single "correct" answer to the lumper/splitter debate' (180); or it may be because the method still needs to be improved. Much remains to be done, Hanks writes, to find effective procedures for selecting relevant information in the context, to explain the 'rules' of semantic-type alternations and of exploitations, to discover the rules of ellipsis (198 *sq*.), etc.

One of the central characters in Hanks's book is the dictionary. Most of his mentions are to point at the weaknesses of existing general dictionaries: they fail to distinguish normal and less normal usages, they fail to adequately describe the contexts of use, they are weak on phraseologies, etc. But for him the objective of the general dictionary is clear: it is first of all to record common, frequent, normal usages. Most modern lexicographers will agree, not only in England: in the age of the electronic corpus, that is what they want to do, and that is what a dictionary is for, they say. Perhaps, but the users need most help on the marginal usages that are found in the work of past and present fiction writers, poets or journalists and in the discourse of specialists of a science or a technique. In fact, one wonders, apart from foreign students and the specialists of lexical databases and their various applications, who needs the more common entries in a dictionary. What does Hanks think? Can the same dictionary attain the two objectives, or do we need two types of general dictionary? Not to mention a third objective, telling the users not how most people speak and write but how people should speak and write, or how the best users, the élite, speak and write. Should this traditional type be abandoned? These questions are not central for Hanks in this book, but I still wish he had said more on them.

Lexical Analysis is remarkably well produced, as indeed one would expect from the MIT Press. I found only eight (small) errors in its almost 500 pages: 188, 226, 246, 274 (*Duchamps* should be *Duchamp*), 362 ('grincer les dents' should be 'grincer des dents', at least in my dialect), 371, 404, 423. Also, on page 40, the reader is referred to examples '(6)–(8)' that I could not find, probably a relic of the preceding version of chapter 2.

Lexical Analysis is a big book, it is not cheap but it is exceptionally good value. It is theoretical and practical, it is limpid and extremely subtle, it is highly personal and can also be used as a manual for students. It is not always easy to read, but the main ideas are repeated in different chapters and the message is clear. It is 'intended to serve as part of the foundation for future empirical research in many language-related disciplines: corpus linguistics, cognitive linguistics, computational linguistics, historical linguistics, philosophy of language, and language teaching, among others. It offers a new foundation for a variety of practical tasks such as computation of meaning in ordinary language, dictionary making, and textbook writing' (22). *Lexical Analysis* is indeed an important book, and Hanks is right to be ambitious: his book is a must for spe-

cialists of several disciplines. TNE is not only convincing; it seems obvious, and many people will regret that Hanks took such a long time before developing it and offering it in book form. Who could now envisage exploring and describing word meaning in the traditional way? *Lexical Analysis* makes us all want to try TNE, and no doubt many have already tried. Every semanticist, every lexicologist, every lexicographer will want to know how far the theory applies to his/her discipline and to his/her own language. It will be interesting to see how TNE influences semantics and how it changes the dictionaries and lexical databases of English and other languages in the near future.

References

De Schryver, G.-M. (Ed.). 2010. *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks.* Kampala, Uganda: Menha Publishers.

Firth, J.R. 1968. A Synopsis of Linguistic Theory, 1930–1955. Palmer, F.R. (Ed.). 1968. *Selected Papers of J.R. Firth, 1952–9*: 168-205. London: Longman.

Hanks, P. 2000. Do Word Meanings Exist? *Computers and the Humanities* 34: 205-215. Also in Fontenelle, T. (Ed.). 2008. *Practical Lexicography: A Reader.* Oxford Linguistics. Oxford/New York: Oxford University Press.

Hanks, P. (Ed.). 2008. *Lexicology: Critical Concepts in Linguistics.* 6 volumes. London/New York: Routledge.

Louw, B. 1993. Irony in the Text or Insincerity in the Writer? — The Diagnostic Potential of Semantic Prosodies. Baker, M., G. Francis and E. Tognini-Bonelli (Eds.). 1993. *Text and Technology: In Honour of John Sinclair.* Amsterdam: John Benjamins.

Sinclair, J.M. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Henri Béjoint
Centre de Recherche en Terminologie et Traduction
Université de Lyon
Lyon
France
Henri.Bejoint@univ-lyon2.fr