Verb patterns, noun collocations, and grammatical metaphors

Patrick Hanks & Sara Može
University of Wolverhampton

This paper presents certain aspects of the *Pattern Dictionary of English Verbs* (PDEV), a corpus-driven lexical resource (in progress), which is freely available at www.pdev.org.uk. These aspects concern phraseology, meaning, noun-verb relationships, and grammatical metaphor.

 A dictionary such as PDEV was first proposed by Hanks and Pustejovsky in 2005, using the Corpus Pattern Analysis (CPA) procedure described in Hanks (2004). On this basis, in 2012-15 AHRC funded a project at the University of Wolverhampton for researching development of such a dictionary.

PDEV consists of two linked databases. The first consists of sample concordances for verbs in the British National Corpus (BNC), each use in the samples being tagged with a pattern number. The second database is a set of detailed and explicit verb patterns, which also states the 'implicature' or meaning of each pattern. Additionally, a hierarchical ontology of semantic types was developed (see Jezek and Hanks, 2010). Applied to collocating nouns, these semantic types are used to identify and disambiguate patterns of verb use.

So far, analyses of 1400 verbs have been completed. Another 400 analyses are awaiting finalization. In BNC, some 5400 lemmas are tagged as main verbs (a figure commensurate with the number of verb entries in the *Concise Oxford Dictionary*). Thus, the links between phraseology and meaning for approximately one third of all English verbs have been systematically analysed.

PDEV shows, among other things, that, whereas words in isolation tend to be highly ambiguous, patterns of verb use are almost always unambiguous.

Associated research (Hanks, 2012) showed that, whereas statistically significant collocates of verbs are almost always found in lexicogrammatical patterns, statistically significant noun collocations do not necessarily occur in patterns. For example, co-occurrence of *bathroom* with *shower* selects either the 'apparatus' or the 'human activity' meaning of *shower,* distinguishing it from the 'weather event' meaning. However, collocating nouns can co-occur anywhere in proximity. The Oxford English Corpus has the following examples among others:

1. The master bedroom has an en suite *bathroom* and *shower*.
2. Slowly, I pushed the covers off and went to the *bathroom* for a quick *shower*.
3. Grace took a *shower* in the guest *bathroom* while Kelsie lazily stretched in bed.

Turning back to verbs, we discuss PDEV's analysis of the verb *club.* The three patterns are:

#	%	Pattern & Primary implicature
1.	52.94%	[[Human 1]] club [[Human 2 | Animal]] (to {death})
		[[Human 1]] hits [[Human 2 | Animal]] with a large, heavy stick or bat, resulting in serious injury or {death} of [[Human 2 | Animal]]
2.	26.47%	[[Human_Group]] club [NO OBJ] {together} ({to/INF [V]})
		[[Human_Group]] jointly combine their efforts and resources in order to achieve {to/INF [V]}
3.	20.59%	[[Human]] go clubbing [NO OBJ]
		[[Human]] frequents night clubs, in order to relax and have fun
		[[Human]] engages in the nightclub's activities, such as dancing, socializing, and consuming drinks

Pattern 1 here is a grammatical metaphor. Quite a few verbs are grammatical metaphors, typically based on nouns denoting animals (e.g. *dog, hound, weasel, ferret, hare, wolf*; see Deignan 2005: 48, 153); others on nouns denoting artefacts (*club, knife, gun*), vehicles (*ship, cycle*), and body parts (*eye, elbow, knee*). Most of these verbs are used mainly or exclusively with transitive or intransitive prepositional particles ('prepositions' and 'adverbs' in traditional terminology), i.e. they are phrasal verbs.

This simple difference between the lexicogrammar of nouns (**nominal groups**) and the lexicogrammar of verbs (**predicators**) yields a theoretical basis for semantic analysis and computational processing of texts. In turn, this has considerable potential for pedagogical applications. Many computational-linguistic researchers in distributional semantics are working towards using collocations for meaning identification and processing. A corpus-driven theory of meaning and use (Hanks, 2013), associated with PDEV's empirical analyses of data, provides a solid foundation for such research.

References

Deignan, Alice (2005): *Metaphor and Corpus Linguistics*. Amsterdam and Philadelphia: John Benjamins.

Hanks, Patrick (2004): 'Corpus Pattern Analysis'. In *Proceedings of Euralex 2004.* Lorient, France.

Hanks, Patrick (2012): 'How people use words to make meanings: semantic types meet valencies'. In A. Boulton and J. Thomas (eds.) *Input, Process and Product: Developments in Teaching and Language Corpora*. Brno, CZ: Masaryk University Press.

Hanks, Patrick (2013): Lexical Analysis: Norms and Exploitations. Cambridge, MA: MIT Press.

Hanks, Patrick and James Pustejovsky (2005): 'A Pattern Dictionary for Natural Language Processing'. In *Revue française de linguistique appliquée* 10(2).

Jezek, Elisabetta, and Patrick Hanks (2010): 'What lexical sets tell us about conceptual categories.' In *Lexis: E-journal in English lexicology. 4: Corpus Linguistics and the Lexicon*.