

Norbert Corver, Jeroen Van Craenenbroeck\*, William Harwood,  
Marko Hladnik, Sterre Leufkens and Tanja Temmerman

# Introduction: The compositionality and syntactic flexibility of verbal idioms

<https://doi.org/10.1515/ling-2019-0021>

**Keywords:** verbal idioms, compositionality, syntactic flexibility, idiom formation

According to the Fregean principle of compositionality, the meaning of a complex expression is determined by the meanings of its parts and the rules used to combine them (cf. Partee 1984: 281). This principle seems flouted in the case of idiomatic expressions, which denote a meaning that is not straightforwardly derivable from (the combination of) the meanings of the individual lexical items that comprise that expression. A canonical example is *kick the bucket*, the meaning of which has nothing to do with either kicking or buckets; it simply means ‘to die’. Idiomatic expressions raise questions that are fundamental to the nature of syntax and semantics. Such questions form the focus of this special issue. They are addressed by the contributions to this volume through a range of theoretical, methodological, and empirical approaches, thus providing a multi-faceted answer.<sup>1</sup> Before introducing the individual papers, we first provide an overview of some of the major questions raised by idioms and tackled in this special issue.

---

<sup>1</sup> While the papers in this special issue approach the phenomenon of idiomaticity from a variety of different angles, theoretically they can all be broadly situated within the generative tradition. For other, e.g. constructivist, approaches to idioms and idiomaticity, see Jackendoff (1997), Croft and Cruse (2004), Evans and Green (2006), among many others.

---

**\*Corresponding author: Jeroen Van Craenenbroeck**, Faculteit Letteren, Blijde-Inkomststraat 21, KU Leuven, 3000 Leuven, Belgium, E-mail: [jeroen.vancraenenbroeck@kuleuven.be](mailto:jeroen.vancraenenbroeck@kuleuven.be)

**Norbert Corver**, Utrecht Institute of Linguistics-OTS, Trans 10, Utrecht University, 3512 JK Utrecht, Netherlands, E-mail: [N.F.M.Corver@uu.nl](mailto:N.F.M.Corver@uu.nl)

**William Harwood**, Kleemburg 18, 9050 Gentbrugge, Belgium, E-mail: [william.w.harwood@gmail.com](mailto:william.w.harwood@gmail.com)

**Marko Hladnik**, Stuytstraat 21, 3544 MP Utrecht, The Netherlands, E-mail: [mhladnik@gmail.com](mailto:mhladnik@gmail.com)

**Sterre Leufkens**, Utrecht Institute of Linguistics-OTS, Trans 10, Utrecht University, 3512 JK Utrecht, Netherlands, E-mail: [s.c.leufkens@uu.nl](mailto:s.c.leufkens@uu.nl)

**Tanja Temmerman**, Département des langues et lettres modernes (germaniques), Université Saint-Louis – Bruxelles, Kruidtuinlaan 43, 1000 Brussels, Belgium, E-mail: [tanja.temmerman@gmail.com](mailto:tanja.temmerman@gmail.com)

Non-compositionality has traditionally been taken to be the defining characteristic of idioms (Katz and Postal 1963: 275; Fraser 1970: 103; Chomsky 1980: 149; Cacciari and Tabossi 1988: 668; Fillmore et al. 1988: 504; Nunberg et al. 1994: 498; Marantz 1997: 207).<sup>2</sup> In addition, it has been argued that within the class of idioms, some are more compositional than others. Notably, Nunberg et al. (1994) postulate a class of compositional idioms, which they call *Idiomatically Combining Expressions* (ICEs).<sup>3</sup> In such idioms, both the verb and its arguments have an identifiable meaning that is independent from the meanings of other elements. For example, in the canonical example *spill the beans*, which means ‘divulge the secret’, *spill* refers to ‘divulge’, while *the beans* represents ‘the secret’. These elements are then combined into an idiom, by means of regular compositional rules. The second class of idioms distinguished by Nunberg et al. (1994) is that of *Idiomatic Phrases* (IdPs), involving idioms in which the elements are not separable in the idiomatic interpretation. For example, in the canonical IdP *kick the bucket* ‘to die’, no interpretational counterpart is available for either *kick* or *the bucket*, so that the idiomatic interpretation forms an undividable, non-compositional whole. Note that both ICEs and IdPs are non-compositional in the sense that they carry an idiosyncratic meaning that is unpredictable from the literal meaning of their parts. However, in ICEs, the idiomatically interpreted verb does combine compositionally with its equally idiomatically interpreted argument(s). In IdPs, on the other hand, there are no separate building blocks that make up the idiomatic meaning, and hence there is no composition at all.<sup>4</sup>

The distinction between ICEs and IdPs has so far been presented as a purely semantic one: the compositionality of a verbal idiom hinges on the decomposability of its meaning into a verbal part and one or more non-verbal entities. However, the extent to which an idiom can be decomposed depends on how exactly one paraphrases the figurative meaning of the expression, which is very

---

<sup>2</sup> Taken in a broad sense, all non-compositional word combinations may be seen as idioms, ranging from compound-like structures (*black and white film*) to entire sentences (*be that as it may*). This volume, however, will be mostly dealing with verbal idioms, i.e. idioms that consist of a verb and its arguments such as *kick the bucket* or *spill the beans*. As defining criteria of a verbal idiom, we propose that it must contain a lexical verb, have a non-literal interpretation, be able to interact with productive VP-external syntax, be comprised of lexical items that are found outside the context of the idiom, and be formed in a manner that obeys the regular syntactic rules of the language. See Harwood et al. (2016) for details and examples.

<sup>3</sup> Fillmore et al. (1988: 504–505) call these (partly) compositional idioms “encoding idioms”.

<sup>4</sup> So-called collocations (e.g. *to brush one’s teeth*, *to make a fuss*) constitute another class of ‘fixed expressions’. While collocations are lexically idiosyncratic (e.g. *to brush/#wash one’s teeth*), their subparts have a literal, non-figurative meaning, and the meaning of the collocation as a whole is derivable from the meanings of its parts. Thus, collocations are compositional. See Fellbaum (2015) for discussion.

much open to interpretation, and hence to subjectivity. For example, the idiom *to hit the road* can be paraphrased as ‘to leave’ but also as ‘to start on a journey’.<sup>5</sup> Fortunately, decomposability into smaller meaningful items is not the only available diagnostic. As noted by Wasow et al. (1983), Fillmore et al. (1988), Nunberg et al. (1994), and Everaert et al. (1995), among others, the semantic notion of compositionality strongly correlates with the syntactic notion of flexibility, i.e. the ability of idioms to undergo syntactic manipulations.

We can observe that ICEs are indeed able to undergo a number of syntactic manipulations without losing their idiomatic interpretation. For example, *spill the beans* allows for passivization, topicalization and modification, as demonstrated in (1).

- |     |                                               |                  |
|-----|-----------------------------------------------|------------------|
| (1) | a. The beans were spilled (by Bob).           | [Passivization]  |
|     | b. The beans, Bob has most certainly spilled. | [Topicalization] |
|     | c. Bob spilled the juiciest beans first.      | [Modification]   |

Furthermore, ICEs are relatively open to lexical variation, as demonstrated by (2), in which the object noun *beans* has been replaced by nouns that are similar in content.

- |     |                                |
|-----|--------------------------------|
| (2) | a. Come on, spill the details. |
|     | b. Come on, spill the news.    |
|     | c. Come on, spill the gossip.  |

If an IdP, on the other hand, undergoes syntactic transformations, the idiomatic reading is altogether lost and only the literal interpretation remains (cf. (3)).<sup>6</sup> In addition, non-compositional idioms are much more restricted when it comes to lexical variability, as illustrated in (4).

- |     |                                           |                  |
|-----|-------------------------------------------|------------------|
| (3) | a. # The bucket was kicked (by Bob).      | [Passivization]  |
|     | b. # The bucket, Bob has gone and kicked. | [Topicalization] |
|     | c. # Bob kicked the painful bucket.       | [Modification]   |
- 
- |     |                             |
|-----|-----------------------------|
| (4) | a. # He kicked the tub.     |
|     | b. # He kicked the bin.     |
|     | c. # He kicked the disease. |

---

<sup>5</sup> See also Everaert (2010) for a critical evaluation of the use of compositionality as a way of distinguishing between ICEs and IdPs.

<sup>6</sup> Throughout this paper, a hash sign (#) is used to indicate loss of idiomatic interpretation.

As suitable as syntactic flexibility may look as a criterion to distinguish between ICEs and IdPs, however, there are a number of problems with it. First, the syntactic flexibility of a particular idiom is often established by means of a single-author judgement on the basis of constructed data, which leaves room for subjectivity. Secondly, there are idioms that allow some syntactic transformations, but not all. For example, Espinal and Mateu (2010: 1400–1403) point out that the idiom *V one's head off* behaves like an ICE in that it allows for pronominalization and ellipsis, but it is more like an IdP in resisting topicalization and modification. On this basis, some authors have argued against a two-way distinction between flexible and inflexible idioms, and in favor of a continuum of flexibility (Fraser 1970; Abeillé 1995; Dobrovolskij 1999; Espinal and Mateu 2010).

Of course, if there is a continuum of flexibility, there may equally be a continuum of compositionality, if these concepts represent the syntactic and semantic side of the same coin. Then this raises the question of where exactly idiomaticity should be located in the grammar and how the architecture of the grammar accommodates the phenomenon of idiomaticity. Perhaps surprisingly, most authors assume that it is not the syntax of idioms that is deviant. For example, Fellbaum (this issue) claims that idiomatic expressions follow the exact same syntactic rules as non-idioms. As such, the deviant interpretation of idioms arises post-syntactically at the syntax-semantics interface, or lexically in a separate lexical component like D(istributed) M(orphology)'s Encyclopaedia (Halle and Marantz 1993; Marantz 1997). The idea that idioms display 'normal' syntactic behavior is also hinted at in Everaert (2010: 92), who states, "the 'original' properties of words in idioms are never lost, always available, and it is this lexical transparency in idioms that needs to be accounted for in a structuring of the lexicon". For Everaert, these lexical properties (e.g. morphosyntactic features and co-occurrence restrictions) of idioms are defined at the level of the individual lexical items (atoms), which are part of what he calls the "narrow lexicon" (Everaert 2010: 87). The meaning of idioms is stored in a separate grammatical submodule (DM's Encyclopaedia), where actual words and phrases are listed, together with their conventional meaning properties.

This special issue contains a selection of papers presented at the 8th Brussels Conference on Generative Linguistics (BCGL8), entitled "The Grammar of Idioms", which was held at the Brussels Campus of KU Leuven on June 4–5, 2015, and was embedded within the research project "The Syntax of Idioms", jointly carried out by researchers from KU Leuven and Utrecht University. The overall goal of the issue is to increase our understanding of the nature of idioms by addressing the following questions:

1. Are all idioms (non-)compositional to the same extent? Can different classes of idioms be established on this basis? By what means can we establish the (degree of) compositionality of a particular idiom?
2. How similar or deviant in their syntax are idioms compared to non-idiomatic expressions? In other words, are idioms identifiable/definable in terms of their syntactic behavior (e.g. their resistance to syntactic manipulations)? Is there such a thing as ‘idiom syntax’ or ‘idiom grammar’, or are idioms just like non-idiomatic expressions when it comes to their syntactic behavior? Moreover, are all idioms alike in this respect or do they display variation?
3. Which parts of the grammar are involved in the formation and interpretation of idioms? Are they accessed by a post-syntactic lexicon, as in Distributed Morphology? Are they confined to specific domains of syntax? How do they interact with the PF and LF interfaces?

The contributions to this special issue all address one or more of these questions. The paper “How flexible are idioms? A corpus-based study” by Christiane Fellbaum takes a clear position with respect to the first and second issue, by stating that in appropriate contexts, speakers produce idiom variations consistent with the rules of freely composed language. On the basis of a corpus study of English and German, Fellbaum shows that syntactic manipulations of all (verbal) idioms occur in everyday speech, contrary to claims that only ICEs allow such syntactic transformations. The data discussed in this paper suggest that the syntactic behavior of all idioms is equivalent to that of non-idiomatic language.

Berit Gehrke and Louise McNally take a semantic approach in their paper “Idioms and the syntax/semantics interface of descriptive content vs. reference”. The article starts from syntactic accounts of verb-object idioms involving a determiner that either is merged higher than the combination of the verb and its nominal object, or is late-merged, and they show that this causes problems for semantic compositionality. Gehrke and McNally point out that this idiom puzzle is part of a more general problem: verbs as a rule select for the descriptive content of nouns, also outside of idioms. The authors develop an analysis of the syntax-semantics interface that uses two levels of representation, dissociating the descriptive content from referential information. The paper is especially relevant in the light of the third set of questions raised above, as it achieves in the semantics something that syntactic accounts set out to do, thus offering a new perspective on the architectural division of labor in idiom construction.

Another paper that is interesting in light of the second and third set of questions is “Against the parallelism between the NP and the clause: Evidence from idioms”, by Carlo Cecchetto and Caterina Donati, which also locates the ‘deviancy’ of idioms in the semantics rather than the syntax. The authors start

out from the constituency theory of idiom formation that states that a category can only receive an idiomatic interpretation if it forms a constituent at some point in the derivation. They defend this theory for the verbal domain, but show how it wrongly predicts that constructions in which the determiner and the noun can form an idiom to the exclusion of a prepositional phrase selected by the noun should not be possible.<sup>7</sup> Cecchetto and Donati provide ample evidence that such idioms do exist in Italian, and, moreover, that they are not rare. The analysis they propose relies on the idea that nouns do not take complements the way verbs do and that PPs that follow nouns are not real arguments. Instead, these PPs can be thought of as adjuncts and can thus be late merged (i. e. become part of the syntactic representation at a later point in the derivation). Finally, they discuss idiomatic DPs where the noun and the PP receive an idiomatic reading to the exclusion of the determiner, and present several pieces of evidence that these idioms are the output of the morphological component. As such, a uniform theory of structure formation in the syntax of the nominal domain can be maintained.

Julia Horvath and Tal Siloni focus on a different part of the architecture of idioms in their paper “Idioms: The type-sensitive storage model”, in particular the question of how they are stored in the mental lexicon. They show how two types of idioms can be distinguished on the basis of their ability to undergo diathesis alternations, i. e. the possibility to appear as a verbal passive and/or a corresponding active transitive. The first type, clausal idioms, can be unique to a single diathesis (e.g. idioms that only occur as a verbal passive) and are argued to be stored in one piece. On the other hand, phrasal idioms can undergo diathesis alternations, and they are stored as subentries of another lexical entry. Thus, the paper goes into questions (2) and (3) by postulating a different storage strategy for two types of idioms, distinguishable on the basis of their syntactic flexibility.

The papers in this special issue add to the existing literature on idioms in two important ways. First of all, they bring more breadth and depth to the data that form the basis for generalizations about idioms. In the literature on idioms, it is too often the case that conclusions are drawn on the basis of English data only, and typically even on a limited set of prototypical examples.<sup>8</sup> This special issue overcomes this imbalance by including data from German (Fellbaum; Gehrke and McNally), Italian, French (Cecchetto and Donati), Hungarian (Gehrke and McNally), and Hebrew (Horvath and Siloni). Secondly, while a lot of the literature

<sup>7</sup> For discussion of the constituency of idioms, see also Marantz (1984) and Lebeaux (2000).

<sup>8</sup> Some notable exceptions include Abeillé (1995) for French, Everaert (1995) for Dutch, Bresnan (2001: 10–14) for Dyrbal, Grégoire (2009) for Dutch, and Kim (2014) for Korean and Japanese.

on idioms makes use of constructed data and single-author judgments, this special issue contains a paper (Fellbaum) that is based solely on corpus research, thus broadening the methodological variety as well as the data diversity.<sup>9</sup> This special issue is similar in nature to Everaert et al. (1995) in taking an interdisciplinary and multi-methodological approach. It could be seen as a successor to that volume, one that includes the most recent findings in idiom research, as well as adding new crosslinguistic data from both elicitation and corpus studies.

Although it would be too optimistic to expect that the papers in this special issue all converge on the same conception and analysis of idioms, certain general tendencies nevertheless emerge. For example, there seems to be a growing consensus that idioms display more regular grammatical behavior than is often assumed. If this is true, it makes the question of what makes idioms (*to kick the bucket*, *to spill the beans*) distinct from non-idiomatic expressions (*to eat a banana*) a non-syntactic one. In this vein, several linguists have proposed that idioms can only be defined and understood in relation to a language community (Hockett 1956: 222–223; Fillmore et al. 1988: 504; Nunberg et al. 1994: 492; Everaert 2010: 79–81). Members of such a community assign and agree on a particular meaning for a fixed combination of words. As such, idioms are conventionalized linguistic expressions, and knowledge of idioms involves two things: (a) knowledge of the lexical properties of the individual lexical atoms which together form the idiomatic expression (in a syntactically regular way) and which are reflected in syntactic structure, and (b) knowledge of the meaning of the entire expression as agreed upon (by convention) by the members of the linguistic community in which the idiom is used. Finally, the syntactic regularity and flexibility of idioms might in turn provide a way of differentiating them from other types of fixed phrases, such as formulaic expressions (*How are you?*), sayings (*better late than never*), collocations (*to brush one's teeth*), and proverbs (*Birds of a feather flock together*) (see also Harwood et al. 2016).

**Acknowledgements:** We would like to thank Eefje Boef, Alexis Dimitriadis, Kristel Doreleijers, Loes Koring, the participants of BCGL8 in Brussels (June 4–5, 2015), the members of CRISP and UiL-OTS, and the anonymous reviewers of all contributions to this special issue for their constructive remarks, insightful questions, and useful comments.

The research project “The Syntax of Idioms” (KU Leuven / Universiteit Utrecht) was funded by FWO (G.A049.12N) and NWO (326-78-001).

---

<sup>9</sup> For earlier corpus-based studies of idioms, see Moon (1998), Fellbaum (2007), Geyken (2007), Grégoire (2009), and the PARSEME project <https://typo.uni-konstanz.de/parseme/>.

## References

- Abeillé, Anne. 1995. The flexibility of French idioms: A representation with lexicalized tree adjoining grammar. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: Structural and psychological perspectives*, 15–42. Hillsdale, NJ: Lawrence Erlbaum.
- Bresnan, Joan. 2001. *Lexical-functional syntax*. Malden, MA & Oxford: Blackwell.
- Cacciari, Cristina & Patrizi Tabossi. 1988. The comprehension of idioms. *Journal of Memory and Language* 27. 668–683.
- Chomsky, Noam. 1980. *Rules and representations*. New York, NY: Columbia University Press.
- Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Dobrovol'skij, Dimitrij. 1999. Haben transformationelle Defekte der Idiomstruktur semantische Ursachen? In Nicole Fernandez Bravo, Irmtraud Behr & Claire Rozier (eds.), *Phraseme und typisierende Rede*, 25–37. Tübingen: Stauffenburg.
- Espinal, M. Teresa & Jaume Mateu. 2010. On classes of idioms and their interpretation. *Journal of Pragmatics* 42(5). 1397–1411.
- Evans, Vyvyan & Melanie Green. 2006. *Cognitive linguistics: An introduction*. Mahwah, NJ: Lawrence Erlbaum.
- Everaert, Martin. 1995. Auxiliary selection in idiomatic constructions. In Léa Nash & Georges Tsoulas (eds.), *Actes du premier colloque Langues et Grammaire*, 103–117. Paris: University Paris–8.
- Everaert, Martin. 2010. The lexical encoding of idioms. In Malka Rappaport Hovav, Ivy Sichel & Edith Doron (eds.), *Syntax, lexical semantics and event structure*, 76–98. Oxford: Oxford University Press.
- Everaert, Martin, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.). 1995. *Idioms: Structural and psychological perspectives*. Hillsdale, NJ: Erlbaum.
- Fellbaum, Christiane (ed.). 2007. *Idioms and collocations: From corpus to electronic lexical resource*. Birmingham: Continuum.
- Fellbaum, Christiane. 2015. Syntax and grammar of idioms and collocations. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax - Theory and analysis. An international handbook*, vol. 1, 777–802. Berlin & Boston, MA: De Gruyter Mouton.
- Fillmore, Charles J., Paul Kay & Mary C. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language* 64. 501–538.
- Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6. 22–42.
- Geyken, Alexander. 2007. The DWDS corpus: A reference corpus for the German language of the 20th century. In Christiane Fellbaum (ed.), *Idioms and collocations: From corpus to electronic lexical resource*, 23–41. Birmingham: Continuum.
- Grégoire, Nicole. 2009. *Untangling multiword expressions: A study on the representation and variation of Dutch multiword expressions*. Utrecht: Utrecht University dissertation.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & S. Jay Keyser (eds.), *The view from building 20*, 111–176. Cambridge, MA: MIT Press.
- Harwood, William, Marko Hladnik, Sterre Leufkens, Tanja Temmerman, Norbert Corver & van Craenenbroeck Jeroen. 2016. Idioms: Phasehood and projection. Ms Utrecht University & KU Leuven. <http://lingbuzz.auf.net/lingbuzz/002951>.

- Hockett, Charles. 1956. Idiom formation. In Morris Halle, Hugh McLean, Horace G. Lunt & Cornelis H. Van Schooneveld (eds.), *For Roman Jakobson: Essays on the occasion of his sixtieth birthday, 11 October 1956*, 222–229. The Hague: Mouton.
- Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.
- Katz, Jerrold & Paul Postal. 1963. Semantic interpretation of idioms and sentences containing them. *Quarterly Progress Report of the MIT Research Laboratory of Electronics* 70. 275–282.
- Kim, Kyumin. 2014. Phases and idioms. Ms. University of Ottawa.
- Lebeaux, David. 2000. *Language acquisition and the form of grammar*. Amsterdam & Philadelphia, PA: John Benjamins.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: MIT Press.
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2). 201–225.
- Moon, Rosamund. 1998. *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Clarendon Press.
- Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70. 491–538.
- Partee, Barbara. 1984. Compositionality. In Fred Landman & Frank Veltman (eds.), *Varieties of formal semantics: Proceedings of the 4th Amsterdam Colloquium, 1982*, 281–311. Dordrecht: Foris Publications.
- Wasow, Thomas, Ivan A. Sag & Geoffrey Nunberg. 1983. Idioms: An interim report. In Shiro Hattori & Kazuko Inoue (eds.), *Proceedings of the XIIIth international congress of linguistics*, 87–96. Tokyo: Nippon Toshi Center.