Gilquin, G. 2026. Collostructions. In H. Nesi & P. Milin (eds) *International Encyclopedia of Language and Linguistics, Third Edition. Volume 12* (pp. 40-44). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-323-95504-1.00963-7

Collostructions

Gaëtanelle Gilquin
UCLouvain, Belgium

Abstract
Collostructions, which are the product of collostructional analysis, show how certain words are associated with a given slot in a construction. This article explains the link between collostructions and collocations, describes the different types of collostructional analysis and how collostructions are retrieved from corpora. It also points to the existence of collostructions at different levels of analysis, in different languages, time periods, etc. and discusses their possible exploitation for pedagogical purposes.

Keywords
Collexeme; Collexeme analysis; Collo-profile; Collostructicon; Collostructional analysis; Construction grammar; Corpus linguistics; Covarying collexeme analysis; Distinctive collexeme analysis; Phraseology

Key points
- Corpora reveal that certain words are preferentially associated with given slots in constructions, forming what are known as 'collostructions'.
- Collostructions help establish the meanings of constructions and their contexts of use and can be compared across languages, language varieties, time periods, etc.
- The collostructicon, an inventory of collostructions, may be a useful pedagogical resource for learners, whose collostructions often differ from those produced by native speakers.

Introduction

Corpus linguistics, through the systematic analysis of large collections of authentic language, has contributed to the development of phraseology and has revealed the existence – and recurrence – of various types of phraseological phenomena. Among these are the so-called 'collostructions'. In the publication that introduced this concept, Stefanowitsch & Gries (2003) defined 'collostruction' as "the combination of a collexeme and a collostruct", where 'collexemes' are "[l]exemes that are attracted to a particular construction" and a 'collostruct' is "a construction associated with a particular lexeme" (Stefanowitsch & Gries, 2003, p. 215). In accordance with the theoretical framework of construction grammar that inspired Gries and Stefanowitsch, constructions should be understood here as "conventionalized pairings of form and function" (Goldberg, 2006, p. 3). For example, the ditransitive construction in English has a certain form, which can be represented as [$SUBJ_{agent}$ V $OBJ_{recipient}$ OBJ_{theme}], and a function, which can be roughly described as expressing the transfer of some theme to a recipient by an agent. Corpus analyses have shown that the verb that is most strongly associated with the ditransitive construction is *GIVE* (Stefanowitsch & Gries, 2003, p. 229), which grants a special status to the following collostruction: [$SUBJ_{agent}$ *GIVE* $OBJ_{recipient}$ OBJ_{theme}].

From collocations to collostructions

The term 'collostruction' is a blend of 'collocation' and 'construction'. The notion of collocation, which refers to the association between words that often occur together, has been around for quite some time. Using corpus query tools, the 'collocates' of a target item can be extracted automatically from a certain span of words to the left and/or to the right of the target item. While this produces a rich output from which interesting tendencies can be derived, it should be borne in mind that collocates may display all sorts of syntactic relationships with the target item and that some may be less relevant to the grammatical patterning of the item.

Altenberg & Granger (2001) highlighted this as one of the limitations of collocational analysis. They noted that the collocates of the verb *MAKE* automatically extracted from corpora were "words which occur frequently in the span of *MAKE*, not all of which are 'constructional' collocates of *MAKE*" (Altenberg & Granger, 2001, p. 185). They illustrated this by means of examples (1) and (2), which both include the collocate *argument*.

(1) Yet, to **make** an even stronger *argument*, these proponents for adoptive parents should state why this is a bad court decision and find other examples. (Corpus example taken from Altenberg & Granger, 2001, p. 187)

(2) The goal in this type of *argument* is to **make** the public aware of the truths of nuclear power. (Corpus example taken from Altenberg & Granger, 2001, p. 187)

In (1), the word *argument* functions as the direct object of *MAKE* and its presence points to the association of *MAKE* with 'speech' or 'verbal communication' collocates (Altenberg & Granger, 2001, p. 178), together with other collocates such as *claim*, *statement* or *comment*. The occurrence of *argument* in (2), on the other hand, does not contribute to this tendency, because in this case, *argument* is not the direct object of *MAKE*. Collocational analysis, however, does not take syntactic structure into account and, therefore, cannot distinguish between (1) and (2).

Collostructional analysis, by contrast, takes constructions as a starting point and, hence, structure into account. The input for a collostructional analysis is a list of words that occur in certain slots of specific constructions. In the case of the verb *MAKE* mentioned above and for the type of structure that Altenberg & Granger (2001) were interested in, for instance, only nouns found in the object slot would be included in the analysis ([*MAKE* **OBJ**]). This means that the results of a collostructional analysis help describe, not a word as such (*MAKE*), but a construction ([*MAKE* OBJ]). For one and the same word, different collostructional analyses could be carried out. For *MAKE*, one could also carry out a collostructional analysis of phrasal verbs by focusing on the particles combining with *MAKE* ([*MAKE* Prt]) or a collostructional analysis of verbal causative constructions by focusing on the infinitive verbs used with *MAKE* ([SUBJ *MAKE* OBJ V_{inf}]; see Gilquin, 2006), among other analyses.

Searching for collostructions

There are three main types of collostructional analysis (see Stefanowitsch & Gries, 2003; Gries & Stefanowitsch, 2004a, b):

i. collexeme analysis, which examines one slot in one construction (e.g. the particle slot in [*MAKE* Prt]);
ii. (multiple) distinctive collexeme analysis, which compares one slot in two (or more) similar constructions (e.g. the particle slot in [*MAKE* Prt] and [*TAKE* Prt]);
iii. covarying collexeme analysis, which investigates two slots in one construction (e.g. the combination of verb and particle in [V Prt]).

For all three types, the first steps of the analysis involve retrieving the instances of the construction(s) from the corpus data and, then, identifying the lexemes that occur in the slot(s) under study. These steps require corpora with detailed syntactic annotation (such as parsed corpora) or, if such corpora are not available, quite some manual work. From there, scripts are available (e.g. Gries, 2004/2022; Flach, 2021) that compute the 'collostruction strength' of the lexemes, that is, a statistical measure of association that indicates the degree of attraction (or repulsion) between a lexeme and a construction.

Table 1, for example, lists the ten verbs with the highest collostruction strength for the infinitive slot of the *MAKE* causative construction ([SUBJ *MAKE* OBJ V_{inf}]) in academic writing data from the British National Corpus (see Gilquin, 2015, p. 263 for a longer list). Each of them corresponds to a collostruction that is particularly strong: [SUBJ *MAKE* OBJ *feel*], [SUBJ *MAKE* OBJ *seem*], [SUBJ *MAKE* OBJ *appear*], etc.

Verbs (raw freq.)	**Coll. strength**
feel (21)	45.594185
seem (25)	45.023448
appear (23)	44.476489
work (14)	24.832698
look (10)	17.327026
think (10)	16.533777
vanish (5)	13.104694
laugh (3)	9.052522
refer (5)	7.730256
sound (3)	7.500258

Table 1. Top ten collexemes in the V_{inf} slot of the *MAKE* causative construction [SUBJ *MAKE* OBJ V_{inf}], with raw frequency and collostruction strength (Gilquin, 2015, p. 263)

Lists of collexemes such as those in Table 1 are usually examined with a view to determining certain semantic fields that are associated with the construction. For [SUBJ *MAKE* OBJ V_{inf}], it turns out that non-volitional verbs are strongly attracted to the construction, especially descriptive verbs such as *seem*, *appear*, *look* and *sound*. This gives information about the meaning of the construction as a whole. The presence of non-volitional verbs is thus not compatible with the coercive meaning that is often thought to be typical of causative *MAKE* (cf. *He made me do it*). Instead, the main meaning of the construction can be described as "to cause a process that is not directly dependent on the CAUSEE" (Gilquin, 2006, p. 41), as exemplified by (3).

(3) It **made** Callaghan *seem* untypically fragile and lacking in confidence. (Corpus example taken from Gilquin, 2015, p. 262)

Different types of collostructions

Constructions in the framework of construction grammar exist at different levels of abstraction. Accordingly, collostructions can be more or less abstract. The example given above for the ditransitive construction, [$SUBJ_{agent}$ *GIVE* $OBJ_{recipient}$ OBJ_{theme}], is fully abstract, save for the collexeme *GIVE* itself, being made up of phrasal patterns. The example [SUBJ *MAKE* OBJ *feel*] is partially lexically filled, as it includes the verb *MAKE* in addition to the collexeme *feel*. The collostruction [*accident waiting to happen*], as an instance of the construction [N *waiting to happen*] (see Stefanowitsch & Gries, 2003, p. 219), is fully lexically filled, with only lexical items in addition to the collexeme *accident*.

Since, according to construction grammar, words and morphemes are constructions, collostructions can be examined at the word- or morpheme-level too. Hartmann (2014), for example, performs a collostructional analysis of German *ung*-nominalization, studying the association between verbs and the suffix *–ung*. His analysis reveals the predominance of collostructions such as [*Ordnung*] ('order') or [*Meinung*] ('opinion') and the rarity of collostructions such as [*Machung*] ('making') or [*Nehmung*] ('taking').

As Hartmann's (2014) study illustrates, collostructions can be investigated in any languages, provided corpora are available to investigate them. In addition to English and German, Chinese (e.g. Shan & Lin, 2023), Dutch (e.g. Levshina et al., 2011) and French (e.g. Lauwers & Tobback, 2020) are among the languages that have regularly been studied using collostructional analysis. Contrastive collostructional analysis is even possible, as long as equivalent constructions can be identified in the different languages. Hartmann & Mikkelsen (2024) compare collostructions expressing the future with *going to* and *will/shall* in English and with *skal/vil* and *kommer til å* in Norwegian. While collostructional analyses normally have to be carried out independently for each of the languages involved, Gilquin (2015) explores the possibility of relying on translation to carry out a distinctive collexeme analysis directly comparing English and French constructions.

Collostructions have also been compared at different periods. This diachronic perspective can help gain a better understanding of how constructions evolve over time. Hilpert (2008) shows that between the 16th and 20th centuries, the collostructions of [*will* V] changed from including mainly verbs denoting intentional actions (e.g. [*will declare*], [*will destroy*]) to including more verbs expressing involuntary actions (e.g. [*will understand*], [*will remember*]). This points to a phenomenon of grammaticalization, whereby *will* was originally a marker of intentionality before becoming a marker of future time reference. Focusing on a shorter time span, Rautionaho & Fuchs (2021) compare the collostructions of the English stative progressive in 1994 and 2014. They show that, overall, the stative progressive has not become more frequent during this 20-year span, but that specific collostructions such as [*BE hoping*], [*BE wondering*] and [*BE expecting*] have become more prominent, which they suggest "could be the cause of the popular impression of the continuing spread of stative progressives" (Rautionaho & Fuchs, 2021, p. 35).

While, in the subtitle of their seminal paper, Stefanowitsch & Gries (2003) refer to "the interaction of words and constructions", collostructional analysis has extended beyond words. In Gilquin (2013), the focus shifts to "the interplay between senses and constructions". Instead of lists of words, the input consists of lists of word senses, e.g. *go_operate* and *go_become* for the verb *GO*. This sense-based collostructional analysis provides more specific semantic information about the causative constructions under study. It reveals, for example, a contrast

between [SUBJ *HAVE* OBJ *done_clean*] and [SUBJ *GET* OBJ *done_deal*], illustrated by (4) and (5). This difference would have gone unnoticed in a more traditional kind of collostructional analysis, since both collostructions include the past participle *done*. In Rautionaho & Fuchs (2021), it is semantic classes of verbs (e.g. affective, relational) that serve as input for the collostructional analysis.

(4) She took her to **have** her nails *done*. (Corpus example taken from Gilquin, 2013, p. 130)

(5) Who's picking up Laura's point which is that you **get** all the contractual arrangements *done*? (Corpus example taken from Gilquin, 2013, p. 130)

Another interesting extension of collostructional analysis is the inclusion of gestures. Uhrig (2022) proposes the concept of 'crossmodal collostruction', which has to do with the degree of association between words (or, more generally, constructions) and gestures. His study of verbs of throwing shows, for instance, that *PITCH* is more likely to be uttered with a gesture than *LOB*. Kuryu (2025) conducts a similar kind of analysis for [ADV *and* ADV] constructions of the reduplicative type (e.g. *over and over*, *again and again*) and the oppositive type (e.g. *back and forth*, *up and down*). The findings reveal that each construction type is associated with distinct gestures, namely gestures with cyclic movements (and more precisely 'Cyclic-Sagittal') for reduplicative adverbial constructions and gestures with bidirectional movements for oppositive adverbial constructions.

Pedagogical exploitation of collostructions

As should be clear from what precedes, collostructions offer useful insights into constructions – their phraseological preferences, their meanings, their diachronic evolution, their language-specific features, etc. One area to which collostructional analysis also has an important contribution to make is pedagogy. Through formal instruction, language learners are often taught how to produce well-formed constructions, but they may not be aware of the lexemes that are attracted to the variable slots of these constructions. For example, advanced learners of English are usually quite good at forming the passive construction, but they often fail to realize that some verbs are much more likely to be used in the passive construction than others, and textbooks rarely include information about such aspects (see Gilquin & Granger, 2021).

Useful in this respect are collostructional analyses applied to the comparison of native and learner collostructions (e.g. Gilquin, 2012; Deshors, 2016), as they make it possible to highlight differences between native speakers' and learners' collostructions. Thus, while collostructions such as [SUBJ *MAKE* OBJ *seem*] or [SUBJ *MAKE* OBJ *appear*] are very typical of the *MAKE* causative construction, as noted earlier, learners of English tend to prefer collostructions like [SUBJ *MAKE* OBJ *be*] or [SUBJ *MAKE* OBJ *become*] (Gilquin, 2012). Such results can be used as a starting point for pedagogical applications, to teach learners the associations that characterize constructions in the target language and that they are not aware of.

Herbst (2018, 2020) introduces the concepts of 'collostructicon' (by analogy with 'constructicon', which refers to the inventory of constructions making up a language) and 'colloconstruction grammar', thus stressing the importance of collostructions. He argues that "a collostructicon could be a central part of (learners') constructicon" (Herbst, 2018, p. 18). In the CASA (constructionist approach to syntactic analysis) model (Herbst & Hoffmann, 2018), constructions are presented with 'collo-profiles', which list the items that are found in specific

slots of the constructions, often with an indication of their frequency or strength of association. These collo-profiles are included in the CASA-based *Construction Grammar of the English Language* (Herbst & Hoffmann, 2024) and the online CASA-Constructicon of the English Language (www.constructicon.de), both of which constitute valuable pedagogical resources. Figure 1, reproduced from Herbst & Hoffmann (2024, p. 120), represents the *WILLING*$_{etc}$-*TO*-INFINITIVE CONSTRUCTION, made up of an adjective (Adj) followed by a postmodifier (postmod) taking the form of a *to*-infinitive clause. The figure shows some of the lexemes occurring in the adjective slot. The size of the adjectives indicates how common they are in the construction. Such figures can be used to teach collostructions or produce pedagogical materials that help students practise collostructions (e.g. gap-fill exercises, as suggested by Ziem & Feldmüller, 2023). This should help learners develop and enrich their collostructicon.

WILLING$_{ETC}$-*TO*-INFINITIVE CONSTRUCTION		
Describing a POTENTIAL ÆFFECTOR'S ATTITUDE towards an ACTION		
POTENTIAL ÆFFECTOR	ATTITUDE	ACTION
contextual anchor	Adj	postmod
	prepared willing keen eager, etc.	*to*-infinitive clause whose unrealized subject argument is provided by the contextual anchor

Figure 1. CASA representation of the *WILLING*$_{etc}$-*TO*-INFINITIVE CONSTRUCTION (Herbst & Hoffmann, 2024, p. 120)

Conclusion

Like collocations, collostructions provide useful information about phraseology, semantics, contexts of use, etc. Unlike collocations, however, this information targets a specific slot having a certain role in a construction (e.g. object, agent). From 2003 when the concept was first introduced by Gries and Stefanowitsch until today, we have witnessed a gain in popularity and diversification of the languages and domains to which collostructional analysis is applied. While the types of corpora (ideally, parsed corpora) and the amount of manual work (to identify the lexemes in the slot of a construction) that are typically required to carry out a collostructional analysis may have prevented it from becoming as widespread as collocational analysis, the various developments that it has known over the last few years point to its vitality. Moreover, the potential of large language models to help with the retrieval of collostructions has started to be explored (see Ziem & Feldmüller, 2023), which means that we may expect to hear even more about collostructions in the future.

References

Altenberg, B. & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, *22*(2), 173–194.

Deshors, S. C. (2016). Inside phrasal verb constructions: A co-varying collexeme analysis of verb-particle combinations in EFL and their semantic associations. *International Journal of Learner Corpus Research*, *2*(1), 1–30.

Flach, S. (2021). Collostructions: An R implementation for the family of collostructional methods. Package version v.0.2.0. https://sfla.ch/collostructions/

Gilquin, G. (2006). The verb slot in causative constructions. Finding the best fit. *Constructions*, SV1-3/2006. https://constructions.journals.hhu.de/article/view/445

Gilquin, G. (2012). Lexical infelicity in English causative constructions: Comparing native and learner collostructions. In J. Leino & R. von Waldenfels (Eds.), *Analytical causatives: From "give" and "come" to "let" and "make"* (pp. 41–63). München: Lincom Europa.

Gilquin, G. (2013). Making sense of collostructional analysis: On the interplay between verb senses and constructions. *Constructions and Frames*, *5*(2), 119–142.

Gilquin, G. (2015). Contrastive collostructional analysis: Causative constructions in English and French. *Zeitschrift für Anglistik und Amerikanistik*, *63*(3), 253–272.

Gilquin, G. & Granger, S. (2021). The passive and the lexis-grammar interface: An inter-varietal perspective. In S. Granger (Ed.), *Perspectives on the L2 phrasicon: The view from learner corpora* (pp. 72–98). Bristol: Multilingual Matters.

Goldberg, A. E. (2006). *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.

Gries, S. Th. (2004/2022). Coll.analysis 3/4. A program for R.

Gries, S. Th. & Stefanowitsch, A. (2004a). Extending collostructional analysis. A corpus-based perspective on "alternations". *International Journal of Corpus Linguistics*, *9*(1), 97–129.

Gries, S. Th. & Stefanowitsch, A. (2004b). Covarying collexemes in the *into*-causative. In M. Achard & S. Kemmer (Eds.), *Language, culture, and mind* (pp. 225–236). Stanford: CSLI.

Hartmann, S. (2014). Constructing a schema: Word-class changing morphology in a usage-based perspective. *Yearbook of the German Cognitive Linguistics Association*, *2*(1), 235–252.

Hartmann, S. & Mikkelsen, O. (2024). Future constructions in English and Norwegian: A contrastive corpus study. *Languages in Contrast*, *24*(2), 170–196.

Herbst, Th. (2018). Is language a collostructicon? A proposal for looking at collocations, valency, argument structure and other constructions. In P. Cantos-Gómez & M. Almela-Sánchez (Eds.), *Lexical collocation analysis: Advances and applications* (pp. 1–22). Cham: Springer.

Herbst, Th. (2020). Constructions, generalizations, and the unpredictability of language: Moving towards colloconstruction grammar. *Constructions and Frames*, *12*(1), 56–95.

Herbst, Th. & Hoffmann, Th. (2018). Construction grammar for students: A constructionist approach to syntactic analysis (CASA). *Yearbook of the German Cognitive Linguistics Association*, *6*, 197–218.

Herbst, Th. & Hoffmann, Th. (2024). *A construction grammar of the English language: CASA – A constructionist approach to syntactic analysis*. Amsterdam: John Benjamins.

Hilpert, M. (2008). *Germanic future constructions. A usage-based approach to language change*. Amsterdam: John Benjamins.

Kuryu, D. (2025). Crossmodal collostructional analysis of English [ADV and ADV] constructions: Multimodal constructions or crossmodal collostructions? *Language and Cognition*, *17*, e39.

Lauwers, P. & Tobback, E. (2020). Predicational and (quasi-)specificational constructions with fronted adjectival predicates in French. *Transactions of the Philological Society*, *118*(3), 468–499.

Levshina, N., Geeraerts, D. & Speelman, D. (2011). Changing the world vs. changing the mind: Distinctive collexeme analysis of the causative construction with *doen* in Belgian and Netherlandic Dutch. In F. Gregersen, J. Parrott & P. Quist (Eds.), *Language variation – European perspectives III: Selected papers from the 5th international conference on language variation in Europe (ICLaVE 5), Copenhagen, June 2009* (pp. 111–122). Amsterdam: John Benjamins.

Rautionaho, P. & Fuchs, R. (2021). Recent change in stative progressives: A collostructional investigation of British English in 1994 and 2014. *English Language and Linguistics*, *25*(1), 35–60.

Shan, W. & Lin, Z. (2023). Collostructional analysis on Chinese modal verb construction *neng bu neng* + VP. *Review of Cognitive Linguistics*, *21*(2), 351–376.

Stefanowitsch, A. & Gries, S. Th. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, *8*(2), 209–243.

Uhrig, P. (2022). Hand gestures with verbs of throwing: Collostructions, style and metaphor. *Yearbook of the German Cognitive Linguistics Association*, *10*(1), 99–120.

Ziem, A. & Feldmüller, T. (2023). Dimensions of constructional meanings in the German Constructicon: Why collo-profiles matter. *Yearbook of the German Cognitive Linguistics Association*, *11*(1), 203–226.