

Chapter 21. Constructions in typological and cross-linguistic context

William Croft, University of New Mexico

1. Introduction

Construction grammar and typology appear to be quite different approaches to syntactic analysis, but in fact they share important fundamental assumptions. At the time of writing, however, the only major effort at combining the two approaches is Radical Construction Grammar (Croft 2001, 2022). Most construction grammar research has been done in a single language, usually one of the well-known European or East Asian languages, rather than a cross-linguistic analysis. Conversely, most typological analysis is not done in an explicitly constructional approach, although informally most typological analysis is fairly straightforwardly interpretable as constructional. Integrating the two approaches is fruitful for both construction grammar and typology.

2. Parallels between construction grammar and typological theory

Perhaps the single most fundamental tenet of construction grammar is that the basic unit of grammatical structure is a pairing of form and function, namely a construction (a symbolic unit in Cognitive Grammar).¹ ‘Form’ includes phonological, morphological and syntactic form. ‘Function’ includes meaning but also categories described as ‘discourse-functional’, ‘pragmatic’, or ‘information structure’; we will use the term ‘information packaging’ here (see §6). This approach contrasts with most other approaches to grammar (apart from Cognitive Grammar).² Those other approaches tend to be modular, that is, they are made up of distinct ‘levels’, each of which is self-contained and contains only one type of grammatical information: phonological, syntactic, semantic etc. Fillmore, Kay and O’Connor’s seminal paper in construction grammar (Fillmore, Kay and O’Connor 1988) argues that the syntactic, semantic, pragmatic and even phonological and prosodic idiosyncrasies of particular constructions means that highly general rules linking autonomous syntactic and semantic modules cannot capture the vast majority of a speaker’s knowledge about their language (see §9).

Constructions are basically the same type of unit, a form-function pairing, at all levels of complexity—from atomic units (words and morphemes) to phrases to complex sentences, and at all levels of schematicity—from completely substantive units such as *It takes one to know one* to completely schematic units such as the ditransitive construction [Sbj Verb Obj1 Obj2] (Goldberg 1995, 2006; Croft 2001). In other words, lexicon and syntax are not treated as separate components of the grammar; and idioms and other multiword expressions are syntactic structures as well as schematic structures without

¹ Some construction grammarians have argued that some schematic constructions may have only form, e.g. Fillmore (1999; for slightly different alternative analyses, see Goldberg 2006 and Croft 2009).

² Nontransformational theories such as Head-driven Phrase Structure Grammar that use attribute-value matrices (also called feature structures), where a single matrix can contain different types of grammatical information, are more similar to construction grammar in this respect. In fact, Sign-Based Construction Grammar has been described as a version of Head-driven Phrase Structure Grammar, albeit integrating elements of Berkeley Construction Grammar (Sag 2012:70).

any specific lexical content. Of course, there are differences between syntax and lexicon, but the differences can be attributed to the parameters of complexity (atomic vs. complex) and schematicity (substantive vs. schematic, not to mention a mix of the two as in most idioms). In contemporary construction grammar, unlike traditional grammar, ‘construction’ covers all form-function pairings.

Typology starts from the same two assumptions, but for different methodological reasons. Since Greenberg’s seminal word order paper (Greenberg 1966), the primary basis for cross-linguistic comparison in typology has been the equivalence in function of the morphosyntactic forms compared across different languages. That is, typology begins methodologically with a pairing of form and function. Typology starts from equivalence of function because the great diversity of linguistic forms precludes using form as the starting point for cross-linguistic comparison (see §5). Both construction grammar and typology use the form-function pairing to get a handle on the high degree of grammatical variation: variation within languages for most applications of construction grammar, and variation across languages for typology.

Typology treats the form-function pairing as basic, and the variation in formal encoding of the function as secondary. For example, for a desiderative/optative construction, one language may express this function syntactically, while another language expresses it lexically, with a single word: compare the single-word Mongolian expression in (1) with its periphrastic English translation (Janhunen 2012:155):

- (1) ir-eesai
 come-DES
 ‘**If only (s/he) would come!**’

Typologists are basically doing a constructional analysis, in the broad sense of ‘construction’ of contemporary construction grammar, when they do a ‘typology of relative clauses’ or a ‘typology of secondary predication’. Like construction grammarians, typologists usually end up identifying a family of closely related functions, and a family of morphosyntactic forms that express those functions rather than a single function and a single form.

3. The morphosyntactic structure of constructions in typological perspective

All human languages are general-purpose communication systems. They can be used to communicate any experience that a speaker wishes to convey, in order to achieve her goals in interpersonal interaction. Yet each language is unique, using its own set of constructions (including words) that arose in the speech community’s culture over its history (§11).

Morphosyntactic typology seeks to understand the relationship between general-purpose communication and culture-specific linguistic forms, by looking for patterns of forms encoding functions across languages. Due to the great diversity of morphosyntactic form across languages, only formal structure that is common to all constructions across languages is the part-whole relation between a construction (the whole) and the roles defined for its elements (the parts; see §7). By coincidence, this is the basic structure of a construction specified by the Berkeley FrameNet Constructicon project (Fillmore, Lee-

Goldman and Rhomieux 2012): the construction as a whole, and its construction elements (CEs).

In addition, typological diversity leads one to the conclusion that constructions and the categories that a construction's CEs define are language-specific—the result of historical processes in a speech community (§11). In fact, the categories that a construction's CEs define are also specific to that construction (§7). Both of these conclusions are controversial, even among typologists (although much typological practice can be interpreted as compatible with these conclusions). The controversy arises in part because the language-specific nature of constructional form and CE categories means that another basis is required to compare languages and develop universals of grammar (§5). These conclusions, drawn from typological patterns, are the foundational hypotheses of Radical Construction Grammar (Croft 2001).

4. Syntactic categories, distributional analysis and constructions

Many nonconstructional theories of syntax are built around systems of syntactic notation: a particular inventory of category labels for lexical and phrasal categories, as well as a particular inventory of types of relations between words and/or phrases in a sentence. Syntactic structures—that is, constructions—are defined in terms of a particular configuration of categories or elements with their relations to each other. The theoretical assumption between this notation can be called the 'building block model' (Croft 2013, to appear a): complex constructions are assembled from the 'building blocks' of the categories and relations in the inventory.

These inventories are intended to apply across all languages. They make up what is often called Universal Grammar. Some theoreticians allow languages to do without some of the categories and relations in the inventories (the "smorgasbord" or "cafeteria" approach to Universal Grammar). The universal applicability of the inventories of categories and relations can be called the 'skeleton model' (Croft to appear a): all languages share the same skeleton of morphosyntactic categories and relations, even if some languages leave some of them out.

There exist several of these systems of syntactic notation, also called 'frameworks', because linguists do not agree on what the proper inventory of categories and relations needs to be in order to analyze the structure of all constructions across all languages. Also, individual frameworks change over time, which leads to difficulties in understanding the syntactic notation of older versions of the frameworks, or frameworks that have been abandoned, as well as unfamiliar contemporary frameworks. Cross-linguistic diversity suggests that the assumption of Universal Grammar as a finite inventory of categories and relations should be discarded, and that a framework-free grammatical theory (Haspelmath 2010) is more suitable for a model of grammatical structure that applies to all languages.

The problem that arises is that it is not always obvious that a particular category is the correct description of a role in a construction, or the correct description of a construction as a whole, in another language. A common problem is whether or not a language has a category that is labeled 'adjective'. Many languages are said not to have adjectives. In some cases, linguists disagree as to whether a particular language has adjectives (Haspelmath 2012).

How are such problems supposed to be resolved? The answer, in contemporary terms, is supposed to be provided by syntactic argumentation. For example, (2)-(5) are intended to be arguments supporting the hypothesis that English has both adjectives and verbs:

- (2) Heather **sang**.
- (3) Heather is **tall**.
- (4) the **tall** woman
- (5) the woman that **sang**.

The first argument is that an adjective like *tall* takes a copula in (3), while a verb like *sang* does not, and in addition inflects for tense, as in (2). The second argument is that an adjective like *tall* is simply preposed to *woman* in (4), while a verb like *sang* is postposed and accompanied by the relativizer *that* in (5). These differences, and the unacceptability of switching the words around in (2)-(5) (**Heather talled; *the sang woman*, etc.), make the case that English has both adjectives and verbs.

The facts in (2)-(5) are called arguments, evidence, tests, criteria in favor (or against) a syntactic analysis for a language (Croft 2022:9). This method of syntactic argumentation has a long history, and in mid-20th century American structuralism was called distributional analysis. These different terms obscure the basic fact: **syntactic argumentation is essentially the occurrence (or non-occurrence) of a word or phrase in a construction**. The constructions in (2)-(5) are described in (2')-(5'), with a gap to indicate the role or CE being used to define the categories 'adjective' and 'verb' in (2)-(5):

- (2') Intransitive (Declarative): Heather ___
- (3') Copular (Declarative): Heather is ___
- (4') Adjective Modification: the ___ woman
- (5') Subject Relative Clause: the woman that ___

Distributional analysis supports the construction grammar principle that constructions are grammatical units, because it makes essential use of constructions. In fact, distributional analysis presupposes the existence and identity of constructions. We will return to this point in §7.

The issue with distributional analysis for cross-linguistic comparison is that categories such as adjective and verb are determined for English using English constructions. When we turn to another language, say, Lao, its categories are determined using Lao constructions, not English constructions. English constructions do not exist in Lao. As a result, the syntactic argumentation does not carry over from English to Lao. One cannot simply assume that the categories in Lao are the same categories that are found in English, in this case, adjective and verb.

This would not be a problem if English, Lao and all other languages have exactly the same set of constructions, and the roles of those constructions define exactly the same categories. Then one could safely say that the inventory is universal. But that is not the case; compare (6)-(9) from Lao to (2)-(5) (all Lao examples from Enfield 2007, with page references given after the translation; numbers indicate tone; 0 = toneless):

- (6) saam3 khon2 taaj3
 three person **die**
 ‘Three people died.’ (273)
- (7) khuaj2 phen1 dam3
 buffalo his **black**
 ‘His buffalo is black’ (252)
- (8) lot1 dam3
 vehicle **black**
 ‘black car’ (263)
- (9) khon2 thii1 ñaang1
 person REL **walk**
 ‘the person who walks’ (246)

Lao uses seemingly the same construction for the translation equivalents of ‘die’ and ‘black’. Lao does have a copula construction, but it is used for translation equivalents of English nouns:

- (10) khaw3 pên3 nak0-hian2
 they COP CT.AGT-**student**
 ‘They were students.’ (285)

Also, while **the sang woman*, without a relativizer and preposed like an adjective, is unacceptable in English, Lao translation equivalents of English verbs may occur without a relativizer (and without a participial derivation, unlike the English translation):

- (11) maa3 haw1
 dog **bark**
 ‘barking dog’ (253)

Lao may have similar constructions to English (this is in fact not entirely clear; for example, there are two different copulas in Lao with different functions; Enfield 2007:284-88). But the distribution of words in those constructions is quite different.

Nevertheless, common practice is to assume that categories such as adjective, verb and copula in English can be carried over to other languages like Lao; and constructions are found that will fit what the analyst expects. Categories in another language such as Lao are equated with similar categories in English. To a lesser extent, constructions in another language are equated with similar constructions in English.

In practice, categories and constructions across languages are being equated by virtue of semantic similarity, despite the apparent use of distributional facts. This has been clear from the beginning of modern morphosyntactic typology, as shown by the following quote from Greenberg’s seminal word order paper:

It is here assumed, among other things, that all languages have subject-predicate constructions, differentiated word classes, and genitive constructions, to mention but a few. I fully realize that in identifying such phenomena in languages of differing structure, one is basically employing semantic criteria... The adequacy of a cross-linguistic definition of “noun” would, in any case, be tested by reference to its results from the viewpoint of the semantic phenomena it was designed to explicate. If, for example, a formal definition of “noun” resulted in equating a class containing such glosses as “boy”, “nose” and “house” in one language with a class containing such items as “eat”, “drink”, and “give” in a second language, such a definition would forthwith be rejected and that on semantic grounds. (Greenberg 1966:74)

But semantic similarity is not distributional analysis. That is the real problem. Distributional analysis is language-specific—determined by roles in constructions in a specific language. Semantics can serve as the basis for cross-linguistic comparison. But it is not directly related to distributions within a language.

5. Constructions as comparative concepts

This problem has engendered a lively debate in typology, on the nature of ‘comparative concepts’ (Haspelmath 2010), that is, concepts that are valid for cross-linguistic comparison (see also Croft 2003:6-19). From Greenberg (1966) onward, typologists have assumed that the basis for cross-linguistic comparison must be semantic. There are also properties of morphosyntactic form that are cross-linguistically valid, such as word order or the presence vs. absence of morphemes coding a function (Haspelmath 2010). However, these properties of morphosyntactic form are not distributional.

Construction grammar provides a means to resolve this problem. Constructions, including words, are pairings of form and meaning. Hence, when a typologist is comparing, say, intransitive predication constructions across the world’s languages (Stassen 1997), they are comparing constructions that express the same function, and looking at how they are similar or different in morphosyntactic form. Thus, ‘construction’ can serve as a comparative concept: any morphosyntactic form in any language that expresses a particular function (Croft 2022:17).³

Typologists regularly use another type of comparative concept to describe different types of morphosyntactic form found across languages to express a function: ‘strategy’ (see for example, Keenan and Comrie 1977; Givón 1979; Stassen 1997). A strategy is a subclass of a construction: those constructions that express a particular function but employ a particular type of morphosyntactic form (Croft 2022:19).

Strategies can be classified into three broad types (Croft 2022:19-22). Encoding strategies are simply different ways of encoding function into morphosyntactic form, including word order or presence vs. absence of a morpheme (or morphemes) encoding a particular function. Examples include the word order strategy for the polarity question construction in English (*Are you married?*) vs. an overt question particle strategy in the

³ The particular function is described by a compound term, e.g. ‘intransitive predication construction’. The unmodified term ‘construction’ refers to any form expressing any function in any language, as in construction grammar.

polarity question construction in K'ichee' (formerly known as Quiché) Mayan (*Lā at culanic?*; Mondloch 1978:21). These strategies, like all strategies, are defined in terms of cross-linguistically valid properties of form, not distributional facts.

A more complex type of strategy is a system of strategies for two or more constructions. In a system of strategies, the strategies are defined in terms of similarities and differences in morphosyntactic form between the constructions in a language. For example, alignment strategies are defined in terms of the relationship between the encoding of the arguments in a transitive construction to the one argument in an intransitive construction. In the accusative alignment of English, the transitive subject (labeled A by typologists) is encoded in the same way as the intransitive subject (S), and the transitive object (P) is encoded in a different way (the accusative). In the ergative alignment of Yuwaalaraay, it is the P argument that is encoded in the same way as the S argument (absolutive), and the A argument is encoded in a different way (the ergative; Williams 1980:36):

(12) The snake [nominative] bit the man [accusative].

(13) The woman [nominative] didn't run.

(14) *ɖuyu-gu nama ɖayn-Ø yi:-y*
 snake-ERG that man-ABS bite-NFUT
 'The snake [ergative] bit the man [absolutive].'

(15) *wa:l nama yinar-Ø banaga-ɲi*
 NEG that woman-ABS run-NFUT
 'The woman [absolutive] didn't run.'

The third and most important type of strategy (§10) is the recruitment strategy. In the recruitment strategy, the function of the construction in question is expressed by recruiting the morphosyntactic form of a related construction. For example, the presentation of an internal physical sensation such as hunger recruits the copular strategy of property predication in English (*I'm cold*), but in French it recruits the 'have' strategy of the French presentational possession construction (*J'ai froid*, lit. 'I have cold').⁴ More generally, English recruits the form of the property predication construction, while French recruits the form of the presentational possession construction.

6. Constructional function and organization in cross-linguistic perspective

Constructions and strategies as comparative concepts form the basis of typological analysis. Croft (2022) is a large-scale survey of constructions of the world's languages and the most common strategies used to express those constructions. The constructions are organized by function. The analysis of function is based on the hypothesis that

⁴ The names and definitions of constructions and strategies are taken from Croft (2022). The glossary of comparative concepts from Croft (2022) is open access at: <https://www.cambridge.org/highereducation/books/morphosyntax/1AAB4F5F9C553F675170DCA3F03F82E2/resources/glossary/E54930706FCF7E5865A4FD76F7983DE1> (accessed January 11, 2023).

function can be broken down into two dimensions: semantic content and information packaging.

The two-dimensional analysis of function originated from a typological analysis of parts of speech (noun, verb, adjective; Croft 1991, 2001). Semantics—nouns denote persons and things, verbs actions, and adjectives properties—has been treated as an unreliable guide for the analysis of parts of speech, since for example a noun may denote an action (*running*) or a property (*height*). However, distributional analysis is not suited to the cross-linguistic analysis of parts of speech either, for the reasons given in §4.

For parts of speech, the information packaging dimension are the propositional acts (Searle 1969; Croft 1991) of reference, predication and modification. These are of course categories that are already used in grammatical description, particularly for reference grammars of indigenous languages, but they are not usually considered to be information packaging functions. The nature of information packaging can be illustrated by the table of examples of propositional act functions and semantic classes in Table 1 (Croft 2022:13, Table 1.1):

	reference	modification	predication
object	<i>the sharp thorns</i>	<i>the thorn's color</i>	<i>It is a thorn.</i>
property	<i>sharpness</i>	<i>the sharp thorns</i>	<i>Those thorns are sharp.</i>
action	<i>I said [that the thorns scratched me]. the [scratching of the thorns]</i>	<i>the thorns [that scratched me] the thorns [scratching me]</i>	<i>The sharp thorns scratched me.</i>

Table 1. Packaging of semantic classes in different propositional act functions.

Table 1 illustrates three basic principles of the information packaging. First, any semantic class can be packaged as reference, predication, or modification. Second, certain combinations are “privileged” or prototypical: specifically, reference to objects, predication of actions, and modification by properties. Their special status is reflected in the minimal overt coding of the prototypical combinations in contrast to the nonprototypical combinations. Typologically, the universal pattern is that the nonprototypical combinations are encoded by at least as many morphemes as the prototypical combinations (Croft 2003:183-88). For instance, Lao conforms to the universal pattern even though property predication is zero coded just like action predication (see examples (5) and (6)). Third, the precise coding in a particular language is a convention of the speech community. It is a convention of Lao that it has recruited the action predication construction’s form (namely, zero coding) for property predication. Conversely, it is a convention of English that it has not done so, instead most likely recruiting the copula strategy of object predication (this convention is ancient in the history of English).

These properties of information packaging are precisely the properties of construal or conceptualization, a fundamental process in cognitive semantics (Talmy 1977; Langacker 1987:116-37; Croft and Cruse 2004, chapter 3; Croft 2012:18). Construal of an

experience serves the interlocutors’ goals in discourse. This is why any concept can be construed in almost any way. Construal is constrained by the nature of reality. This is why some construals are more common than others, and encoded in a “privileged” way. Finally, construal is constrained by cultural convention. This is why, for example, English and Lao use different strategies (zero vs. copula) for property predication, even though both languages have both strategies available: it is a fact of their linguistic history.

Croft (2007, 2022) argues that all verbalization of experience involves some sort of information packaging of semantic content. In fact, syntax often conforms to information packaging more closely than the semantic content of the experience (hence the primary organization of the 2022 textbook in terms of information packaging).

For example, Table 2 illustrates the information packaging options of degrees of salience represented by subject, object and oblique argument encoding with respect to different semantic roles (Croft 2022:175, Table 6.2). Again, any semantic role can be encoded as just about any grammatical role, but there is a “privileged” encoding of agent as subject, patient or theme as object, and other roles as obliques (this is the basic, “active” or “direct” voice). English does not typically encode agents like objects, but the voice systems of Algonkian languages and Philippine Austronesian languages do.

	core		oblique
	subject	object	
agent	<i>The protesters sprayed green paint on the sidewalk.</i> <i>The director presented the watch to Bill.</i>	—	<i>Green paint was sprayed on the sidewalk by the protesters.</i> <i>Bill was presented with the watch by the director.</i>
patient, theme	<i>Green paint was sprayed on the sidewalk by the protesters.</i> <i>The watch was presented to Bill by the director.</i>	<i>The protesters sprayed green paint on the sidewalk.</i> <i>The director presented the watch to Bill.</i>	<i>The protesters sprayed the sidewalk with green paint.</i> <i>The director presented Bill with the watch.</i>
goal, recipient	<i>The sidewalk was sprayed with green paint.</i> <i>Bill was presented with the watch by the director.</i>	<i>The protesters sprayed the sidewalk with green paint.</i> <i>The director presented Bill with the watch.</i>	<i>The protesters sprayed green paint on the sidewalk.</i> <i>The director presented the watch to Bill.</i>

Table 2. Packaging of semantic roles in events as either subject, object or oblique.

The typological patterns of the argument encoding of different semantic types of events, different alignment systems, basic voice and derived voice (passive-inverse, applicative and causative) can be largely accounted for by distinguishing semantic participant roles and degrees of topicality/salience of referents in those roles (Croft 2022, chapters 6-9).

Another example is the construal of various semantic relations between events as a figure-ground or asymmetric relation (adverbial subordination) vs. a complex-figure or symmetric relation (coordination), illustrated in Table 3 (Croft 2022:466, Table 15.1):

Semantic relation	Adverbial construction	Coordinate construction
Anterior	<i>He washed the car before driving to the party.</i>	<i>He washed the car and drove to the party.</i>
Posterior	<i>He drove to the party after washing the car.</i>	<i>He washed the car and drove to the party.</i>
Overlap	<i>He washed the car while the sun was still shining.</i>	<i>The sun was shining and he was washing the car.</i>
Cause	<i>She went to bed because she was exhausted.</i>	<i>She was exhausted and (so) went to bed.</i>
Purpose	<i>I will grab a stick to defend myself.</i>	<i>I will grab a stick and defend myself.</i>
Apprehensional	<i>I grabbed a stick lest he attack me.</i>	<i>Grab a stick or he will attack you.</i>
Means/Positive Circumstantial	<i>He got into the army by lying about his age.</i>	<i>He lied about his age and got into the army.</i>
Negative Circumstantial	<i>She carried the punch into the living room without spilling a drop.</i>	<i>She carried the punch into the living room, and/but she didn't spill a drop.</i>
Additive	<i>In addition to having your hand stamped, you must show your ticket stub.</i>	<i>You have to have your hand stamped and show your ticket stub.</i>
Substitutive	<i>We barbecued chicken at home instead of going out to eat.</i>	<i>We didn't go out to eat, and/but barbecued chicken at home.</i>
Subtractive	<i>He did all the problems correctly except he missed the proof on the last one.</i>	<i>He did all the problems correctly but he missed the proof on the last one.</i>
Conditional	<i>If you do that, the terrorists have won.</i>	<i>Murphy, you do that and the terrorists have won, ...</i>
Concessive	<i>Although John had no money, he went into this expensive restaurant.</i>	<i>John had no money, but he went into this expensive restaurant (anyway).</i>

Table 15.1. Packaging of semantic relations between events as either coordination or (adverbial) subordination.

Yet again, either construal of the event semantic relation is possible in English for all of the event relations, and contribute to the explanation of their cross-linguistic variation (Croft 2022, chapters 15, 17).

There are several major information packaging parameters. Most important is the reference-predication-modification distinction described above. Within reference, referents are packaged in terms of information status, particularly accessibility and identifiability. Within predication, dependent arguments are packaged in terms of topicality, as noted above. Within modification, modifiers serve to particularize the referent, that is, starting from the type description provided by the head noun, modifiers subcategorize the type, select an instantiation of the referent, and situate it in physical space and mental space (Croft 2007; 2022:103-13). Finally, relations between events may be construed asymmetrically (a figure-ground construal) or symmetrically (a complex figure construal; Croft 2001, chapter 9). This list likely does not exhaust all information packaging construals. Nor are the different construals entirely discrete. There is typological evidence for a reference-modification continuum and a modification-predication continuum (Croft 2022:130-54, 443-48).

The two-dimensional analysis of function has proven useful to account for typological patterns and indicates that the function side of a construction consists of these two dimensions. Equally importantly, the analysis of function has important consequences for the organization of the constructicon, the inventory of constructions of a language. Most construction grammars primarily consider the organization of the constructicon in terms of a taxonomic hierarchy or lattice of the morphosyntactic structures of constructions. However, constructions are also organized in the constructicon in terms of function. In the approach described here, both semantic content and information packaging organize the constructicon.

Bybee (1985:118) proposes different strengths of connections between word forms in her typological, usage-based, network theory of the organization of the lexicon. The primary networks of lexical relations that she examines are inflectional paradigms and inflection classes. Her principles carry over to syntactically complex constructions in the constructicon, since the lexicon is part of the constructicon (Croft and Cruse 2004:303-4). Relations between words may be phonological, semantic or morphological. Phonological relations are the weakest but they are not nonexistent, as lexical competition in phonological space indicates. Semantic relations are much stronger, even in the absence of phonological relations. Bybee's morphological relations involve parallel phonological and semantic relations, that is, morphological relations are symbolic relations.

For complex syntactic constructions, then, the strongest relations should be parallel morphosyntactic and semantic relations. In other words, constructions with similarly structured construction elements (CEs) whose semantic structure is parallel will have the strongest relations. This extension of Bybee's model applies specifically to paradigmatic relations between constructions, organized by similarity of meaning, form or both.

One example of a semantic relation leading to an analogical reformation of a syntactic paradigm is given in (16) (Croft and Cruse 2004:320):

(16)		<i>Declarative:</i>	<i>Imperative</i>	<i>Prohibitive</i>
	<i>Verbal</i>	a. He jumped.	c. Jump!	e. Don't jump!
	<i>Nonverbal</i>	b. He is brave	d. Be brave!	f. Don't be cruel!
				< g. Be not cruel!

The nonverbal prohibitive construction changed in form in the history of English to conform with the verbal prohibitive construction due to its close semantic similarity. This is despite the fact that the nonverbal prohibitive now has an otherwise anomalous structure: two auxiliaries in a row, the first of which (*Don't*) is semantically inconsistent with a stative predicate. This also indicates that *Don't* has been reanalyzed as a simple prohibitive morpheme. Another relevant factor is that the token frequency of nonverbal prohibitives is probably much lower than that of verbal prohibitives. The unique constructional form was weakly entrenched and thus was susceptible to replacement.

Bybee does not represent more schematic morphological constructions (e.g. a morphological template of inflections of a word such as English past [Verb-*ed*]) as separate nodes in a taxonomic network—the typical representation in construction grammar. Instead, she represents a more schematic construction as a pattern of similarity relations between individual word type such as the English regular past *walked* ~ *talked* ~ *stalked* or the partially productive irregular past *snuck* ~ *struck* ~ *strung* ~ *spun* ~ *hung* (Bybee 1985:130). Bybee's representation is characteristic of an exemplar-based model (see §§8, 11).

7. Linking constructional analysis in a single language and cross-linguistic constructional analysis

Cross-linguistic constructional analysis proceeds on the basis of comparative concepts (§5). Comparative concepts can be divided into two broad types. The first are constructions, which can be of any morphosyntactic form, and are defined purely in terms of their function. The second are strategies, which are subtypes of constructions that are also specified to belong to a particular type of morphosyntactic form. The use of constructions and strategies as comparative concepts provides the basis for cross-linguistic comparison and the formulation of typological universals. Yet constructions and strategies are not defined by distributional analysis, the foundation of syntactic analysis in a single language. This poses a problem in relating typological universals of language to the grammars of individual languages.

Before we address this problem, we must take a closer look at distributional analysis of single languages. The same problems in using distributional analysis across languages also arise in distributional analysis in a single language. Syntactic categories that are defined distributionally, that is, via their occurrence in a role in a construction, are not identical from one construction to the next.

For example, an English Adjective can be defined by four different constructions (Croft 2022:8):

- (17) a. *Modification*: a **tall** tree
 b. *Predication*: That tree is **tall**.
 c. *Comparative/Superlative Inflection*: **tall-er**, **tall-est**
 d. *Degree Modification*: very **tall**, a little **tall**

But not all words usually described as Adjectives occur in all four constructions (Croft 2022:10):

- (18) a. *Modification*: *an **alive** insect
 b. *Predication*: This chapter is **entire**.
 c. *Comparative/Superlative Inflection*: ***intelligent-er**, ***intelligent-est**
 d. *Degree Modification*: *a very **even** number

A somewhat more complex example is provided by the distribution of direct object phrases (Croft 2001:35-36). Most direct objects can occur in postverbal position without a preposition, and as the subject of a corresponding passive:

- (19) a. *Prepositionless postverbal phrase*: Jack kissed **Janet**.
 b. *Passive subject phrase*: **Janet** was kissed by Jack.

But some prepositionless postverbal phrases cannot occur as a passive subject, and some passive subjects cannot occur as prepositionless postverbal phrases:

- (20) a. Jack weighs **160 pounds**.
 b. ***160 pounds** is weighed by Jack.
- (21) a. *Claude Debussy lived **this house**.
 b. **This house** was lived in by Claude Debussy.

As with the cross-linguistic mismatches in distribution, the common approach is to ignore the distributions that do not match what the analyst expects. But this is empirically inadequate: the anomalous facts are simply ignored, or some exception feature is introduced. This was called ‘methodological opportunism’ in Croft (2001:30, 41). Nevertheless, the anomalous English facts fit into cross-linguistically more widespread phenomena, such as the association of passive subject with some degree of affectedness (thus disfavoring (20b) but favoring (21b)).

The assumption behind methodological opportunism is the building block model of syntactic structure described in §3.1. Constructions are built up out of combinations of syntactic categories and relations that supposedly exist independently from constructions. But the distributional method that identifies syntactic categories and relations assumes we have already identified the constructions that are supposedly defined by those categories and relations. This is a circular argument. Radical Construction Grammar argues that instead of preserving the inventory of categories and relations, constructions should be treated as basic, and categories and relations as derived from the roles in those constructions (i.e. the CEs; see Croft 2001, chapter 1 for a fuller discussion).

Distributional analysis is valid, but it presupposes the existence and identifiability of the constructions used in the analysis. The question arises: how does one identify constructions? Constructions consist of elements (the CEs), the meaning of the CEs, and the meaning of the construction as a whole. A construction is defined by its form plus its function. Constructions are like any conceptual category as described in cognitive psychology and cognitive linguistics: some constructional categories are more discrete than others; constructional categories have more prototypical and more peripheral members; constructional categories vary in their cue-validity (Croft 2001:51-53).

The distributional analysis of the syntax of a language is a gigantic many-to-many mapping between constructions on the one hand, and the words or larger units which serve as fillers for each role/CE in each construction on the other hand. There is no shortcut via a small inventory of categories and relations and a set of constructions built out of those categories and relations. The many-to-many mapping serves as the starting point for reconstructing (construction) grammar on a sound within-language and cross-linguistic basis.

There is a crucial link between language-specific constructions and the distributions they define, and constructions and strategies and comparative concepts. The distributional mapping involves form-function pairings for both constructions and CEs. That is, function is part of the many-to-many distributional mapping. Function is also a comparative concept, namely the basis for the comparative concept of constructions.

One can therefore compare constructions expressing the same function, such as property predication, across languages—not to mention different constructions expressing the same function in a single language. And one can compare the construction elements, that is, the constructions (including words) that fill the equivalent roles in the construction, in terms of their function—for example which property concept words fill the relevant role in the property predication construction in each language—not to mention different property predication constructions in a single language.

In other words, distributional analysis can be done across languages as well as within languages, by using the functions of the constructions and the CEs to align constructions, both across and within languages. Strategies can also be used to align constructions that share morphosyntactic structures that can be validly defined across languages.

A classic example of this methodology is Keenan and Comrie's (1977) typological analysis of noun phrase accessibility, that is, which argument roles of the predicate in a relative clause construction can be relativized, that is, serve as the head of the relative clause. Keenan and Comrie use a semantic definition of 'relative clause' to compare constructions across languages. Relative clause constructions are used to relativize different sets of argument roles across languages. Many languages have multiple relative clause constructions, each of which is used to relativize a set of argument roles. In many cases, the sets of argument roles for different relative clause constructions overlap. Keenan and Comrie discovered that the argument roles that can be relativized can be arranged in an Accessibility Hierarchy, namely subject < object < oblique < possessor (this is a modified version of the Accessibility Hierarchy; see Comrie 1989:164). Keenan and Comrie propose universals of the distribution of relativizable argument roles for relative clause constructions. In addition, they propose universals of distribution for particular relative clause strategies. Specifically, they argue that the more explicitly the

relative clause strategy encodes the relativized argument role, the lower on the hierarchy are the argument roles it can relativize (this formulation is also from Comrie 1989).

8. Methods to analyze cross-linguistic distribution patterns and the organization of functions

There is so much variation in distribution within and across languages that typologists use a variety of techniques to visualize the patterns in the variation, and hence infer patterns from that variation. The main technique used is the semantic map model. Another useful technique that is used is multidimensional scaling.

The semantic map model represents functions as points or nodes in a network (graph structure; Haspelmath 2003, Croft 2001). For example, each lexical concept can be represented as a node in the network. A construction's distribution is the set of nodes, e.g. lexical concepts, which occur in the relevant role in the construction. The distribution of constructions across nodes is not arbitrary: nodes representing concepts that occur in the same construction are semantically related. (This is not always true, for historical or accidental reasons; one must account for noise in the empirical data of cross-linguistic syntactic distribution.) If the distribution patterns of enough constructions are included in the analysis, a network can be constructed that represents semantic relations between concepts that are manifested in syntactic, morphological or lexical distributions. Many typologists use the term 'semantic map' for both the conceptual network, and the mapping of a construction's distribution onto the conceptual network. Croft (2001:92-98) distinguishes the two, using 'conceptual space' for the former and 'semantic map' only for the latter.

A very simple example of the semantic map model was developed to characterize ergativity (Comrie 1978; see §5). Typologists accommodate both ergative and accusative alignments in a single representation. A, S and P roles are represented as distinct nodes. Accusative languages group A and S against P; ergative languages group A against S and P. The relations between A, S and P are captured in the graph structure in Figure 1. The circles indicate the categories found in ergative languages (dashed) and accusative languages (solid).

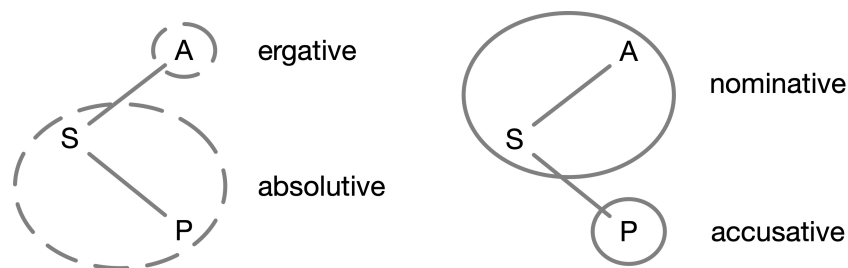


Figure 1. Semantic maps of ergative and accusative systems in the same conceptual space of A, S and P semantic role clusters.

Typologists use semantic maps to account for variation in syntactic, lexical and morphological categories. Semantic maps replace universal essentialist building-block categories such as “noun”, “verb” and “adjective”, “subject” and “object” and so on. The

principle linking universal semantic relations among concepts to constructions is the Semantic Map Connectivity Hypothesis: a construction's distribution must cover a connected subgraph of the conceptual space (Croft 2001:96).

The conceptual space—the graph structure linking nodes representing functions—is hypothesized to be universal, and so can function as a comparative concept. A semantic map—the connected subgraph representing the set of functions of a single construction in a single language, i.e. its distribution—is language-specific. Much variation in categories is allowed, but the categories can be compared across constructions and across languages by the ways in which they overlap in conceptual space. The structure of the conceptual space and the semantic relations between the concepts that are represented in the conceptual space are a major part of the functional organization of the construction for any language.

However, it is difficult to apply the semantic map model to a large number of concepts and/or a large number of languages. There is so much variation in construction distributions across languages, especially when there are a large number of concepts, that it is impossible to produce a coherent graph structure representation of the conceptual space manually. (There is now an algorithm to do it computationally [Regier, Khetarpal and Majid 2013], but a large conceptual graph structure is still difficult to visualize easily.) Some linguists, including Levinson et al. (2003) and Croft and Poole (2008) use multidimensional scaling to automatically generate conceptual spaces from large and complex datasets of cross-linguistic variation in the mapping from constructional form to meaning.

Multidimensional scaling (MDS) represents a conceptual space as a Euclidean geometrical space, that is, a continuous space, rather than a graph structure of discrete nodes. A function is a point on the space. Conceptual relations between points in the same space are represented directly as Euclidean distance. MDS (unlike principal component analysis or correspondence analysis) reduces all the variability of the data to a fixed number of dimensions. The most useful number of dimensions is determined by fitness statistics. Thus patterns in complex variation in distributions are easy to visualize.

The use of MDS allows one to plot points representing a much larger number of much more fine-grained situation types than a small number of crude semantic categories like 'property' and 'action' or 'transitive agent' and 'transitive patient'. Examples include the Bowerman-Pederson set of 73 spatial relations pictures (Levinson and Wilkins 2006:570-75), or Dahl's (1985) tense-aspect questionnaire with 250 sentence contexts.

When data on the cross-linguistic expression of these large sets of fine-grained situation types is plotted by MDS, it challenges the assumption behind the graph structure representation that semantic categories are discrete. If semantic categories were discretely conceptualized in language use, one would expect points representing instances of those classes to clump together in tight clusters. But that is not what is found (Croft 2010a). Instead, points are distributed across broad areas of conceptual space. This indicates that speakers are sensitive to fine-grained situation types, and make grammatical distinctions throughout the conceptual space. Moreover, the conceptual space really is continuous: the dimensions of the space are straightforwardly interpretable. For example, the 'IN' part of the conceptual space of spatial relations, based on data from nine languages in Levinson et al. (2003), is a continuum based on degree of enclosure of the figure by the ground (Croft 2010a:12-13).

9. Syntactic relations, semantic relations and symbolic relations in constructions

Typology also has implications for the internal morphosyntactic structure of constructions. In §3, it was stated that cross-linguistic evidence indicates that the internal morphosyntactic structure of a construction is simply a set of part-whole relations between the construction as a whole and the CEs that make up the construction. In addition, there are also the symbolic relations between the construction's form and its function, and each CE's form and its function. This hypothesis has not been a major topic of discussion in typology. Representations of syntactic structure are not typically encountered in either typological studies or thereference grammars that serve as the empirical base for typological research.

Most formal syntactic theories, and some construction grammars, posit syntactic relations between CEs rather than part-whole relations between each CE and the construction as a whole. These syntactic relations are represented in different ways: as constituents, as dependencies between words, and as sometimes quite complex combinations of relations. Construction grammars include symbolic relations, but modular theories do not, since syntax and semantics reside in separate modules.

Thus, there are three possible types of relations involving a construction's morphosyntactic form: (i) part-whole relations from a CE to the construction as a whole (called 'roles' to distinguish them from other relations); (ii) syntactic relations between CEs; and (iii) symbolic relations from a construction to its meaning, and from the construction's CEs and their meanings.

What typological evidence is there which would argue for the elimination of syntactic relations? First, arguments for constituency or dependency relations, and for categories such as head and dependent or argument and adjunct, are based on distributional analysis (Croft 2001:185-90 and chapter 7). Distributional analysis is language-specific and so cannot serve as the basis for comparative concepts of syntactic relations. Other evidence against syntactic relations comes from considering models in which syntactic relations are preserved, but instead either symbolic relations or roles (part-whole) relations are eliminated. Here we briefly summarize the basic types of evidence; see Croft (2001, chapter 5) for a fuller discussion.

The elimination of symbolic relations takes us back to the modular model of grammar: there is syntactic structure, including relations, in one module, and semantic structure in another module. Even in a modular model, there must be some linking or interface rules between modules to get the syntax-semantics mapping right. Typically, these rules appeal to general iconic principles that match a syntactic structure to the corresponding semantic structure.

However, there are many examples of noniconic mappings in grammatical constructions around the world that undermine general linking rules, or lead to complicated workarounds in comparison to simply positing direct symbolic relations between CEs and semantic structure (Croft 2001:206-20). For example, in many languages, the quantifier "floats" away from the referring phrase it modifies, and modifies the verb, as in Akimel O'odham (formerly known as Pima; Munro 1984:273):

- (22) hegai 'uuvi 'o vees ha-ñeid hegam ceceoj
 that woman 3.AUX **all** them-see **those men**
 'The woman saw all the men.'

One could argue that 'all' describes not just a collection of individuals (the men), but alternatively describes a collective or summative act of seeing. Thus, there is an iconic mapping between *vees* 'all' as a verbal modifier and its meaning.

However, this argument does not apply to the following example (Munro 1984:275):

- (23) vees ñei 'ant heg heñ-navpuj ha-maakaika
all see 1SG.AUX ART **my-friends** their-doctor
 'I saw the doctor of all my friends.'

In (23), I see just one person, the doctor; it seems implausible that seeing a single doctor can be construed as a collective or summative event.

Other noniconic cases of form-function mapping include possessor ascension, where like quantifier float the semantic possessor is a dependent of the verb but the possessor has nothing to do with the action; anomalous agreement, where a word agrees with a referent that doesn't have any direct relationship to the concept denoted by the agreeing word; and "raising", "tough-movement" "passive" and "clause collapsing" constructions where a referent that is a participant in a dependent event but has nothing to do with the main clause event is expressed as a dependent argument of the main clause event. These cases are numerous enough that symbolic relations offers a better analysis than syntactic relations. This is of course the basic argument for construction grammar: there are many specialized and idiosyncratic constructions that cannot be captured by general rules of syntactic form and general rules linking form to function.

The other alternative is the elimination of syntactic roles, leaving both syntactic relations and symbolic relations. This would be a construction grammar model that lacks syntactic roles but retains syntactic relations. It is not clear if any current variant of construction grammar is an instance of this model, but it is a logical possibility.

Here, the problem is that syntactic relations impose more structure than actually exists in some cases (Croft 2001:220-33). A syntactic relation implies that there are two CEs between which the relation holds. Sometimes there are more than two elements, which are strictly ordered, such as the three arguments of the ditransitive construction in Kilivila (Senft 1986:110):

- (24) eseki luleta yena guyau
 he.give his.sister fish chief
 'The chief gives his sister the fish.'

There are four possible ways to link together the three CEs in the absence of the part-whole role. If one had roles linking the arguments to the construction as a whole, then there is a single analysis of Verb, Indirect Object, Direct Object, Subject as roles in the ditransitive construction. Other ordering phenomena such as second position auxiliaries are also simpler to represent in terms of the role of the auxiliary in the clause as a whole.

The greater problem is the absence of one of the units in a syntactic relation. This is common with flagging (also called case marking) and indexation (also called agreement). A flag—adposition or case affix—encodes the syntactic relation. But there are many constructions in many languages in which the flag is optional or absent. For example in differential object marking, some object phrases take an adposition while others do not, or take it variably. The simpler analysis is that the object phrase is CE related to the construction as a whole; it does not matter whether or not the flag is present.

Indexation poses the most serious problem. Indexation is traditionally called “agreement”, because the traditional analysis is that the so-called agreement marker agrees with another element in the construction. But in many cases, the element that the CE supposedly agrees with is simply not there, or it is there only optionally. Again, the simpler analysis is that the “agreeing” CE is related to the construction as a whole, and also has a symbolic link to its referent. For this reason, many typologists have replaced the term ‘agreement’ with ‘indexation’, which implicitly indicates the symbolic relation.

These cross-linguistic phenomena all present evidence that syntactic relations are not necessary, and in fact require the positing of syntactic structure which is not there in many constructions. Nor can syntactic relations plus general linking or interface rules capture what symbolic relations in a construction easily do. The result is a much simpler syntactic structure: the construction as a whole, plus the CEs and their relations to the construction as a whole. Most important however are the symbolic relations of the construction as a whole and its individual CEs to a rich representation of the construction’s function. These symbolic relations are of course what is most distinctive about construction grammar.

10. Constructions, typology and diachrony

Both construction grammar and typology include language change, that is, diachrony, in their theoretical scope. Indeed for many typologists, ultimate explanations for typological universals are largely diachronic (Greenberg 1979). Language change in construction grammar is too large a topic to cover here. Instead, I will focus on developments in typology that converge on a constructional approach to processes of morphosyntactic change, and what they imply about the structure and organization of constructions.

Recruitment, one of the general types of strategies described in §5, is defined in fundamentally dynamic terms: speakers recruit the morphosyntactic form of another construction to express a particular function. Recruitment is in fact the fundamental morphosyntactic strategy; encoding strategies and systems of strategies are further stages after recruitment.

A speaker recruits a new form to express a related function. As time goes on, the recruited form adapts to its new function. For example, in French, the presentation of a physical sensation allows for degree modification, which is not possible for the source construction of presentational possession (Croft 2001:115):

- (25) J’ai très froid
I’ve **very** cold
‘I’m very hungry.’

- (26) *J'ai très une voiture
I've have **very** a car

If the form of the source construction is fused or eroded, as in many grammaticalized constructions, then the source of the recruited form is difficult or impossible to identify, and we think of it as simply a way to encode the function of the construction. For example, if it weren't for orthography, we might not be able to identify the source of the spoken form of the *going to* future construction of English:

- (27) a. I am traveling to deliver the letter.
b. I am **going to** deliver the letter.
c. I'm **gonna** deliver the letter.
d. I'm [ənə] deliver the letter.

If the morphosyntactic form of the source construction is replaced, then the source of the recruited strategy cannot be identified. For example, the word order inversion of the polarity question strategy (*Will he leave?*) is a relic of the verb inversion (*Leave he?*) of earlier stages of English, but the strategy was replaced for verbs in general by the *do* strategy (*Did he leave?*; Bybee and Thompson 1997).

The situation is more complex for a system of strategies, where a variety of processes lead to similarities and differences between the forms of the two constructions. For example, Chung (1977:15-16) suggests the following historical scenario for the emergence of ergative alignment in Pukapukan :

- (28) na patu mātou i te tamaiti
PST hit we ACC the child
'We hit the child.'
- (29) kai-na loa na tamaliki e te wui aitu pau
eat-PASS EMP the.PL children AGT the PL spirit **done**
'The children were all eaten by the spirits.'
- (30) lomilomi ai e tana wawine ma na tana lua tulivae ia
massage PRN **ERG** his woman and the.PL boy two knee that
'The wife and the children massaged his two knees.'

Example (28) illustrates the older active transitive construction: A follows the verb, and in turn is followed by the P argument using the accusative preposition. Example (29) is the passive construction, which forms a system with the active transitive construction in that the encoding of the P argument of the passive (the "passive subject") is the same as the A argument of the intransitive, but the verb form is different (it has the passive suffix *-na*), and the A argument is also expressed differently than in (28), using the oblique preposition *e*. Example (30) shows both loss of the passive suffix from the verb and a change in word order. The change in word order is possibly by analogy to (28), or simply an adaptation to its new function, indicating that the oblique *e*-marked argument is now construed as a core argument of the transitive verb (the ergative).

Attested recruitment strategies are another important piece of evidence for the structure of the conceptual space. Recruitment is diachronic evidence for relations between points in the conceptual space (van der Auwera and Plungian 1998). Croft (2022) describes many different recruitment strategies, from which a space of constructional changes can be developed (Croft, in prep.).

The best studied set of diachronic paths in conceptual space are the semantic changes in grammaticalization (major overviews include Lehmann 2015; Heine, Claudi and Hünnemeyer 1991; Hopper and Traugott 2003). There is now a vast literature on grammaticalization. Kuteva et al. (2019) gives an extensive inventory of semantic changes in grammaticalization, although intermediate steps in the paths of change are not generally given.

Grammaticalization theorists originally focused on the emergence of a single morpheme to encode a “grammatical function”, typically an inflectional category. Example (27) illustrates this process for the development of *be going to* into a future marker in English. Most grammaticalization theorists now consider grammaticalization to be a phenomenon affecting constructions as a whole (e.g. Bybee 2003:146; Himmelmann 2004). The development of an inflectional category marker involves not only the narrowing of possible periphrastic expressions, say of futurity, to a single one (*be going to*), but also the schematization of the category with which the inflectional marker is associated (for example, *it's gonna rain* is not an agentive process, unlike the purpose clause originally associated with *be going to*). Hence, at the very least, grammaticalization involves both the recruitment of a word or phrase to become a morpheme encoding a grammatical inflectional category, and the generalization of the word or phrase to which that inflectional category applies. (Generalization, of course, is just successive recruitment of additional words/phrases to combine with the incipient inflectional morpheme.)

Grammaticalization is a cover term for diachronic processes that lead to complex constructions containing new grammatical inflections. It is taken to contrast with lexicalization, another cover term for processes that lead to new lexical items. Less attention has been paid to processes that lead to constructions containing new word-formation morphemes (this could be called ‘derivationalization’; see Brinton and Traugott 2005:51; Traugott and Trousdale 2013:160-77). Traugott and Trousdale (2013) generalize this entire set of diachronic processes as constructionalization, the emergence of new constructions.

Grammaticalization research seems to have shifted decisively in a constructional direction. A constructional form is recruited from a neighboring function in conceptual space, to be used to primarily convey a grammatical (inflectional) meaning. The form increases in frequency of use in its new function. Its distribution in other constructions, and the distributional patterns of the categories/roles that make it up, adapt to its new function. Its conventionalization leads to phonological reduction and erosion, and greater rigidity and restriction in syntactic structure.

11. Constructions as lineages defining populations

The description in §3.2 of recruitment as a strategy in the analysis of cross-linguistic constructional patterns treats recruitment as a shift between the morphosyntactic forms of

two discrete functions: the form conventionally used for one function is recruited for a semantically similar function. The two functions are assumed to be close in conceptual space. (There are also recruitment for figurative uses such as metaphor and metonymy. These are different yet grammatically relevant dimensions of conceptual space.) In §8, however, it was argued that multidimensional scaling analyses indicate that conceptual space is more continuous, and that speakers are sensitive to very fine-grained distinctions between situations.

This hypothesis is further confirmed by an analysis of verbalization, the process of morphosyntactically expressing experiences, in a set of twenty English Pear Stories narratives (Croft 2010b; data from Chafe 1980). These narratives are verbalizations of a short film that speakers observed and then were asked to describe to an experimenter. When one looks at multiple verbalizations in a controlled setting, it is obvious that there is immense variation in verbalization: no one describes the same scene in exactly the same way. This fact is of course not surprising, but its consequences for the representation of grammar have not been drawn out fully.

The analysis of variation in verbalization in the Pear Stories indicates that alternative morphosyntactic constructions used by speakers are potential incipient sources of grammaticalization and lexical semantic change (Croft 2010b). For example, the basic English verb for application events is *put*. Speakers used a number of other verbs than *put* to describe scenes in the Pear film with application events, including *throw*, *toss*, and *stuff*. These verbs are common etymological sources of ‘put’, for example Ancient Greek *bállō* ‘throw’ > Modern Greek *vazo* ‘put’, Latin *mittere* ‘let go, throw’ > French *mettre* ‘put’, and Old English *potian* ‘thrust, push’ > Modern English *put* (data from Buck 1949). Croft (2010b) argues that grammaticalization and lexical semantic change originate in variation in verbalization.

A close comparison of the Pear Stories scenes also shows that subtle differences in scenes are systematically represented by frequency differences in the use of variant forms. For example, in scenes where an event that is not intended by a human participant is verbalized, events that are more likely to be under the control of the person are more likely to verbalize the person as subject in the argument structure construction, while events less likely to be under the control of the person are more likely to verbalize them with another participant as the subject, in a gradient scale (Croft 2010b:30, 2021:264-65).

Typological comparison and verbalization in a single language both indicate that very fine-grained distinctions between situations being verbalized are grammatically relevant, determining both variation and change in the frequency of use of different constructions for subtly different scenes. That is, a speaker’s grammatical knowledge must include knowledge of a very fine-grained set of situation types, ranged along continuous dimensions of conceptual variation. The mapping between form and function—i.e. a construction—can be conceived of as a probability distribution of constructions across conceptual space (Croft 2021:271). These probability distributions are inferred from the frequency distributions of alternative forms—by the speaker learning and using language, not just by the linguist analyzing the patterns after the fact. Finally, language change involves a gradual shift of these probability distributions over time, even over the speaker’s lifetime.

This analysis of constructions can be situated in a broader theory of language change rooted in language use. Every experience that is verbalized is unique (Croft 2007:348-49).

Of course, speakers construe an experience for the purposes of communication (see §4). But this is not easy:

The hearer cannot read the speaker's mind. The hearer of an utterance, like the speaker, has his own alternative construals of the scene potentially available to him, and cannot be certain of the precise construal intended by the speaker. The speaker's choice of words and constructions are based on her prior exposure to and use of those words and constructions in other communicative acts, and are chosen for her intentions in the current situation. But the hearer's knowledge of the words/constructions is based on his own past exposure to and use of them, which is different from the speaker's. Moreover, no two experiences are identical and so any choice of words and constructions will not precisely characterize the construal of the experience being communicated anyway. Thus there is a fundamental indeterminacy in the construal of a scene and its interpretation in a communicative act. (Croft 2010b:11-12)

In other words, every time I speak, I am recruiting morphosyntactic constructions (words and complex structures) that have been used for prior experiences that are not identical to the one that I am currently trying to verbalize. The success of my communication is always uncertain; and that uncertainty leads to variation in verbalization, and can lead to gradual shifts in verbalization choices that may lead to language change (see Croft 2000:99-114).

Croft (2000) develops this basic observation into an evolutionary framework for language change. Another way of describing language use is that every time I speak, I am replicating words and complex constructions (and also sounds) that I have heard or previously used myself. In the evolutionary framework, language use does not simply involve a set of tokens that a speaker stores, in an exemplar-based representation of grammatical constructions which represents subtle differences in the meaning expressed in a particular occasion of use as well as subtle variations in formal structure. The grammatical construction has a temporal structure as well, defined by the chains of replications of the construction, which are called lineages (Croft 2000:32-34, 2021:284-85, to appear a, b).

The result is that a language-specific grammatical construction is actually a population (in the biological sense; Hull 1988) of tokens of the construction that are defined by the history of uses of that construction, that is, intertwining lineages of replications. The English word *cat* is a Noun not primarily because the concept is conceptualized in a particular way, but because speakers of English replicate *cat* and other English Nouns in a particular constructional role when they replicate a construction in an utterance such as *I don't wanna see the little cat* [Santa Barbara Corpus]. The English category Noun is the population defined by those intertwining replications (Croft to appear b).

This is another difference between language-specific categories and constructions on the one hand, and comparative concepts on the other. Language-specific categories and constructions are historical entities, embodied in utterances and bounded in space and time by their use in the speech community. Of course, each occasion of use is unique in terms of form and function, but they are bound together by the speaker's act of

replication in communicative interaction. In contrast, comparative concepts are ahistorical, essentialist categories of formal structure, meaning and construal that linguists use to analyze the historically situated instances of constructional use and their evolution, diversification, and extinction over time.

In an exemplar-based model, a speaker has knowledge about a construction in her language based on the experiences she has verbalized herself, as well as the verbalizations of her interlocutors, in her lifetime. They represent fragments of the lineages of the construction population in the speech community, but they form the basis of the speaker's future verbalizations—that is, replications—using the construction.

12. Conclusion

The tremendous amount of morphosyntactic diversity in the world's languages indicates that only minimal assumptions can be made about the morphosyntactic structure of constructions. The morphosyntactic structure of constructions consists of only the construction as a whole, its construction elements (CEs), and the part-whole relation, or role, of CEs in constructions. CEs are defined by their distribution in the construction they belong to, that is, categories defined by CEs are construction-specific.

Instead, construction function, and the symbolic relations between a construction and its CEs and their functions, are the locus of explanation for typological universals of constructions. Constructions can be compared cross-linguistically by functional equivalence, and also by the morphosyntactic strategies that they employ. Functions are organized in a largely continuous conceptual space, and the inventory of constructions is organized by relations in the conceptual space.

The morphosyntactic relations among constructions is the result of the strategy of recruiting a morphosyntactic form from a functionally related construction. Every experience that is verbalized by a speaker is unique, so every act of replicating a construction is recruiting a previous use of that construction and hence construing the current experience as an instance of the previous experience. Constructions are populations of utterances using that construction and the collection of experiences that they verbalized. In an exemplar-based constructional model, a speaker's knowledge of a construction is the population of uses that she has been exposed to, which forms the basis of her future uses of that construction.

References

- Brinton, Laurel J. and Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.
- Buck, Carl Darling. 1949. *A dictionary of selected synonyms in the principal Indo-European languages*. Chicago: University of Chicago Press.
- Bybee, Joan L. 1985. *Morphology: A study into the relation between meaning and form*. Amsterdam: John Benjamins.
- Bybee, Joan L. 2003. Cognitive processes in grammaticalization. *The new psychology of language*, vol. 2, ed. Michael Tomasello, 145-67. Mahwah, N. J.: Lawrence Erlbaum Associates.
- Bybee, Joan L. and Sandra A. Thompson. 1997. Three frequency effects in syntax. *Proceedings of the 23rd Annual Meeting of the Berkeley Linguistics Society*, ed. Matthew L. Juge & Jeri O. Moxley, 378-88. Berkeley: Berkeley Linguistics Society.
- Chafe, Wallace (ed.). 1980. *The pear stories*. New York: Ablex.
- Chung, Sandra. 1977. On the gradual nature of syntactic change. *Mechanisms of syntactic change*, ed. Charles Li, 3-55. Austin: University of Texas Press.
- Comrie, Bernard. 1978. Ergativity. *Syntactic typology*, ed. Winfrid Lehmann, 329-94. Austin: University of Texas Press.
- Comrie, Bernard. 1989. *Language universals and linguistic typology* (2nd ed.). Chicago: University of Chicago Press.
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- Croft, William. 2000. *Explaining language change: an evolutionary approach*. Harlow, Essex: Longman.
- Croft, William. 2001. *Radical Construction Grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William. 2003. *Typology and universals*, 2nd edition. Cambridge: Cambridge University Press.
- Croft, William. 2007. The origins of grammar in the verbalization of experience. *Cognitive Linguistics* 18.339-82.
- Croft, William. 2009. Constructions and generalizations [review of Goldberg, *Constructions at work*]. *Cognitive Linguistics* 20.157-66.
- Croft, William. 2010a. Relativity, linguistic variation and language universals. *CogniTextes* 4.303 [<http://cognitextes.revues.org/303/>]
- Croft, William. 2010b. The origins of grammaticalization in the verbalization of experience. *Linguistics* 48.1-48.
- Croft, William. 2012. *Verbs: aspect and causal structure*. Oxford: Oxford University Press.
- Croft, William. 2013. Radical Construction Grammar. *The Oxford Handbook of Construction Grammar*, ed. Graeme Trousdale and Thomas Hoffmann, 211-32. Oxford: Oxford University Press.
- Croft, William. 2021. *Ten lectures on construction grammar and typology*. Leiden: Brill.
- Croft, William. 2022. *Morphosyntax: constructions of the world's languages*. Cambridge: Cambridge University Press.

- Croft, William. To appear a. Word classes in Radical Construction Grammar. *The Oxford handbook of word classes*, ed. Eva Van Lier. Oxford: Oxford University Press.
- Croft, William. To appear b. Philosophical reflections on the future of construction grammar, or, Confessions of a radical construction grammarian. *Constructions and Frames*.
- Croft, William. In preparation. Diachronic typological universals. *The Wiley-Blackwell companion to diachronic linguistics*, ed. Adam Ledgeway, Edith Aldridge, Anne Breitbarth, Katalin É.Kiss, Joseph Salmons and Alexandra Simonenko. Hoboken, New Jersey: Wiley.
- Croft, William and D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Croft, William and Keith T. Poole. 2008. Inferring universals from grammatical variation: multidimensional scaling for typological analysis. *Theoretical Linguistics* 34.1-37.
- Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.
- Enfield, N. J. 2007. *A grammar of Lao*. (Mouton Grammar Library, 38.) Berlin: Mouton de Gruyter.
- Fillmore, Charles J. 1999. Inversion and constructional inheritance. *Lexical and constructional aspects of linguistic explanation*, ed. Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol, 113-128. Stanford, Calif.: CSLI Publications.
- Fillmore, Charles J., Paul Kay and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: the case of *let alone*. *Language* 64.501-538.
- Fillmore, Charles J., Russell R. Lee-Goldman and Russell Rhomieux. 2012. The FrameNet constructicon. *Sign-Based Construction Grammar*, ed. Hans C. Boas and Ivan A. Sag, 309-72. Stanford, Calif.: CSLI Publications.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
- Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Grammar*, ed. Joseph H. Greenberg, 2nd edition, 73-113. Cambridge, Mass: MIT Press.
- Greenberg, Joseph H. 1979. Rethinking linguistics diachronically. *Language* 55.275-290.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. *The new psychology of language, vol. 2*, ed. Michael Tomasello, 211-42. Mahwah, N. J.: Lawrence Erlbaum Associates.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86.663-87.
- Haspelmath, Martin. 2012. How to compare major word-classes across the world's languages. *Theories of everything: in honor of Edward Keenan* (UCLA Working Papers in Linguistics, 17), ed. Thomas Graf, Denis Paperno, Anna Szabolcsi and Jos Tellings, 109-30. Los Angeles: UCLA. (Available at <http://www.linguistics.ucla.edu/faciliti/wpl/issues/wpl17/wpl17.html>.)

- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: University of Chicago Press.
- Himmelmann, Nikolaus. 2004. Lexicalization and grammaticalization: opposite or orthogonal? *What makes grammaticalization—a look from its fringes and its components*, ed. Walter Bisang, Nikolaus Himmelmann and Björn Wiemer, 19-40. Berlin: Mouton de Gruyter.
- Hopper, Paul and Elizabeth Traugott. 2003. *Grammaticalization* (2nd edition). Cambridge: Cambridge University Press.
- Hull, David L. 1988. *Science as a process: an evolutionary account of the social and conceptual development of science*. Chicago: University of Chicago Press.
- Janhunen, Juha A. 2012. *Mongolian*. Amsterdam: John Benjamins.
- Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63–99.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog and Seongha Rhee. 2019. *World lexicon of grammaticalization, second edition*. Cambridge: Cambridge University Press.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, vol. I: theoretical prerequisites*. Stanford: Stanford University Press.
- Lehmann, Christian. 2015. *Thoughts on grammaticalization, third edition*. Berlin: Language Science Press. (First edition, *Arbeiten des Kölner Universalien-Projekts*, 48, Köln: Institut für Sprachwissenschaft, 1982. Revised edition published by LINCOM Europa, München, 1995.)
- Levinson, Stephen C., Sérgio Meira, and the Language and Cognition Group. 2003. ‘Natural concepts’ in the spatial topological domain—adpositional meanings in crosslinguistic perspective: an exercise in semantic typology. *Language* 79.485-516.
- Levinson, Stephen C. and David Wilkins (eds.). 2006. *Grammars of space: explorations in cognitive diversity*. Cambridge: Cambridge University Press.
- Mondloch, James L. 1978. *Basic Quiché grammar*. (Institute for Mesoamerican Studies, Publication 2.) Albany: Institute for Mesoamerican Studies.
- Munro, Pamela. 1984. Floating quantifiers in Pima. *Syntax and semantics 16: The syntax of Native American languages*, ed. Eung-Do Cook and Donna B. Gerds, 269-87. New York: Academic Press.
- Regier, Terry, Naveen Khetarpal and Asifa Majid. 2013. Inferring semantic maps. *Linguistic Typology* 17.89-105.
- Sag, Ivan. 2012. Sign-Based Construction Grammar: an informal synopsis. *Sign-Based Construction Grammar*, ed. Hans C. Boas and Ivan A. Sag, 69-202. Stanford, Calif.: CSLI Publications.
- Searle, John R. 1969. *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Senft, Gunter. 1986. *Kilivila: The language of the Trobriand Islanders*. Berlin: Mouton de Gruyter.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Talmy, Leonard. 1977. Rubber sheet cognition in language. *Papers from the Thirteenth Regional Meeting, Chicago Linguistic Society*, ed. Woodford A. Beach et al., 612-628.

- Traugott, Elizabeth Closs and Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- van der Auwera, Johan and Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2.79-124.
- Williams, C. J. 1980. *A grammar of Yuwaalaraay*. (Pacific Linguistics, B74.) Canberra: Australian National University.