

Word classes in Radical Construction Grammar

William Croft, *University of New Mexico*

ABSTRACT

Word classes are taken to be essentialist categories; they serve as the building blocks for syntactic constructions, and are language universals (that is, cross-linguistically valid). But a word class is language-specific. This is because a word class is defined distributionally, by the occurrence of the words in a particular construction in a particular language. Hence it is not a building block for syntactic constructions. Moreover, a word class is not an essentialist category. A word class is a population: a spatiotemporally bounded set of historical entities—in this case, the occurrences of the words in question in language use in the language, interconnected by lineages of replication. Hence a word class cannot serve as a comparative concept for language universals. Instead, meaning and certain non-distributional properties of form can be used to compare word classes/populations defined by different constructions within and across languages, and reveal universals of grammar.

Keywords: Radical Construction Grammar, distributional analysis, essentialist category, population, comparative concept, language universals

1. Introduction

Radical Construction Grammar (Croft 2001, 2005a, 2013) is a construction-based theory of grammar that allows for the representation of both cross-linguistic and language-internal grammatical diversity. Radical Construction Grammar analyses conform to the following three basic principles:

- (1) Word classes and other syntactic structures are language-specific and construction-specific. What is universal (see §5) is patterns of variation in the verbalization of experience, represented for example in conceptual space.
- (2) The internal structure of the morphosyntactic form of constructions consists solely of the part-whole relation between construction roles (“slots”) and the whole construction. The complexity of constructions rests in the symbolic relations between the roles and their meanings/functions, and the rich semantic/functional structure expressed by the construction and its parts.
- (3) The morphosyntactic forms of constructions are language-specific, that is, there is potentially gradient variation of constructional form across and within languages.

Radical Construction Grammar (Croft 2001) and *Morphosyntax: Constructions of the World’s Languages* (Croft 2022) propose many specific analyses of constructions and

universals of constructional variation. Apart from those analyses, any analysis that conforms to the principles in (1)-(3) can be considered a Radical Construction Grammar analysis. Many if not most analyses of grammatical phenomena in the functional-typological approach conform to the principles of Radical Construction Grammar (although they are not always presented that way).

These three principles result from the observation of the great diversity of morphosyntactic form used to verbalize experiences within and across languages. They also result from a critique of methods of syntactic argumentation used in contemporary grammatical analyses that have their roots in the structuralist and generative traditions (for general overviews, see Croft 2009, 2010a; for critiques of specific analyses, see those references and Croft 2005b, 2007, 2010b; Croft and Van Lier 2012).

These assumptions involve deeply held—often unconsciously held—theoretical principles that clash with the empirical fact of diversity of grammatical form. Word classes are a classic illustration of this problem, and constitute the focus of this chapter as well as many of the publications cited above. Word classes play a quite different role in Radical Construction Grammar than they do in other grammatical theories. For this reason, after an initial discussion of word classes in Radical Construction Grammar in §§2-3, I will discuss the role of word classes in other grammatical theories in §§4-7, and the theoretical assumptions they serve: methodological opportunism, the skeleton model of language universals, the building block model of grammars and the essentialist theory of categories. In §§7-9, I describe the role of word classes in a theory like Radical Construction Grammar that does not make these theoretical assumptions.

2. Word classes and distributional analysis

Word classes are identified by distributional analysis. For example, English Adjectives¹ such as *tall* are defined by:

- (i) their occurrence as modifiers of nouns: *a tall tree*
- (ii) their occurrence as the complement of a copula *be* in predication: *That tree is tall*
- (iii) the fact that they inflect in a certain way (a morphological construction): *tall-er*, *tall-est*
- (iv) the fact that they can in turn be modified by certain degree expressions: *very tall*, *a little tall*).

In contrast, English Verbs such as *jump* are defined by:

- (i) their inflection for Tense and (in the Present) person in predication: *The deer jumped*
- (ii) their occurrence in a relative clause as modifiers of nouns: *the deer that jumped*
- (iii) the fact that they can occur in the Progressive or the Pluperfect: *the deer is jumping*; *the deer had jumped*

¹ In this chapter, names of word classes in specific languages are capitalized; see §8.

(iv) the fact that they can in turn be modified by certain other degree expressions: *the deer was jumping a lot*

At least as important is distributional absence. English Adjectives do not inflect for Tense and Person (**That tree tall(s)*), nor can they directly occur in the Progressive or Pluperfect (**That tree had tall(ed)*). English Verbs do not inflect for degree with *-er/-est* (**jump-er, *jump-est*), nor can they be modified by *very* (**The deer very jumped*). English Verbs cannot occur in their inflected or root form as modifiers of nouns (**The jumped deer was a doe*). These distributional facts, occurrence and non-occurrence, distinguish English Adjectives from English Verbs.

This type of argumentation, DISTRIBUTIONAL ANALYSIS, is universally used in syntactic argumentation for word classes, and for other morphosyntactic units, in structuralist, generative, functional and ‘theory-neutral’ analysis of a language. In the American structuralist tradition, distributional analysis was based purely on linguistic form. In other traditions, including generative grammar and typology, meaning or function may also play a role in defining constructions used in distributional analysis. The distribution of English Adjectives above makes reference to degree expressions, a category partly defined by meaning, and the *be* of predication, as opposed to other functions of *be*.

Distributional analysis is essentially OCCURRENCE IN A PARTICULAR ROLE IN A PARTICULAR CONSTRUCTION. For example, the distribution of English Adjectives described above can be defined as the occurrence of a word in the English Adjective word class in the underlined role in the following constructions:

- (i) Noun Modification (Attributive): [Art __ Noun]
- (ii) Copula Predication: [Sbj *be* __]
- (iii) Comparative Degree, Superlative Degree: [__ *-er*], [__ *-est*]
- (iv) Degree Modification: [*very/a little* __]

‘Construction’ is defined as in contemporary construction grammar (Fillmore, Kay and O’Connor 1988; Goldberg 1995; Croft and Cruse 2004): any morphosyntactic structure, complex or atomic (like a single word), syntactic or morphological (occurrence with an affix or other morphological operation), with elements of the construction being either substantive morphemes (like *be*, *-er*, or *a little*) or schematic elements (like Art, Noun, Sbj).

The definition given above acknowledges the role of constructions in distributional analysis—something which is not generally done. The constructions used to define word classes are called many different things instead of constructions: ‘criteria’ (Givón 2001a:49; Dixon 2010:38), ‘tests’ (McCawley 1998; Carnie 2013:47, 98-100), ‘evidence’, ‘phenomena’, ‘operation’, and ‘process’ (Mulder 1994:114). The words in a word class are said to have a particular grammatical or syntactic ‘distribution’ (Harris 1951:5; Carnie 2013:47), ‘behavior’ (McCawley 1998:186), ‘properties’ (McCawley 1998:18; Evans and Osada 2009:452; Schachter and Shopen 2007:2), ‘features’ (Amha 2001:89), ‘use’ (Jagersma 2010:268), or ‘function’ (Palmer 2009:94), instead of simply saying that they

occur in certain roles in certain constructions and not in others. As a result, an important point is often overlooked: DISTRIBUTIONAL ANALYSIS PRESUPPOSES THE PRIOR IDENTIFICATION OF THE CONSTRUCTIONS USED IN DISTRIBUTIONAL ANALYSIS.

One consequence of the use of distributional analysis to identify word classes (and other grammatical categories and structures) is that DISTRIBUTIONALLY-DEFINED WORD CLASSES ARE LANGUAGE-SPECIFIC. The simple reason for this is that such word classes are defined by their occurrence in a particular role in a particular construction (or set of constructions), and those constructions are language-specific. English Adjectives do not occur in German constructions, and German Adjectives do not occur in English constructions. Although this property of word classes appears to be self-evident, it is in fact a highly contentious issue, because it is incompatible with certain assumptions about language universals and individual language grammars. This issue will be discussed in §§3-5.

Distributional analysis usually involves occurrence of forms in roles in multiple different constructions. In fact, it is usually considered a stronger argument for a particular theoretical analysis if one can argue for the same distribution in multiple constructions. However, a fundamental empirical fact of languages is that DIFFERENT CONSTRUCTIONS DEFINE DIFFERENT DISTRIBUTIONS. That is, the class of words defined by occurrence in a certain role in one construction is almost always different from the class of words defined by occurrence in a certain role in another construction (Croft 2001:34-36). For example, the four constructions used to define English Adjectives above define slightly different classes, as can be seen in the problematic cases below:

- (4) *Modification of a referent:*
 - a. This insect is alive.
 - b. *an alive insect

- (5) *Predication with a copula:*
 - a. An entire chapter is devoted to this problem.
 - b. *This chapter is entire.

- (6) *Degree inflections:*
 - a. tall-er, tall-est
 - b. *intelligent-er, *intelligent-est

- (7) *Degree modifiers:*
 - a. a very tall tree
 - b. *a very even number

In other words, the different constructions in (4)-(7) do not define a single word class of English Adjectives. Instead, they define a set of distinct but overlapping word classes.

These empirical observations were made by the American structuralists, who provide the most explicit and careful methodological discussion of distributional analysis. For

example, Bloomfield writes, ‘Form-classes are not mutually exclusive, but cross each other and overlap and are included one within the other, and so on’ (Bloomfield 1933:269). And Harris writes, ‘If we seek to form classes of morphemes such that all the morphemes in a particular class will have identical distributions, we will frequently achieve little success’ (Harris 1951:244). In a very large grammar of French containing 600 rules covering 12,000 lexical items, no two lexical items had exactly the same distribution, and no two rules (that is, constructions) had exactly the same domain of application (Gross 1979:859-60).

Hence, a consequence of the careful application of distributional analysis to word classes within a single language is that WORD CLASSES ARE ALSO CONSTRUCTION-SPECIFIC. This conclusion is at least as contentious as the conclusion that word classes are language-specific, because it is incompatible with other deeply or unconsciously held assumptions about the nature of individual language grammars. This issue will be discussed in greater detail in §§6-7.

3. Word classes and comparative concepts

If word classes are language-specific, then they cannot play the role in Radical Construction Grammar that word classes play in other grammatical theories. In other grammatical theories, word classes form the basis for a significant set of language universals. For those theories, word classes must be comparable across languages: that is, one should be able to treat English Adjectives and Lango Adjectives as instances of a cross-linguistic category of ‘adjective’, which has certain universal properties associated with it.

In such an approach, it makes sense to ask the question: “does a particular language have adjectives or not?” But this approach—word classes as cross-linguistic categories—is incompatible with distributionally defined word classes. In a strictly distributional definition of word classes, this question makes no sense. Word classes are defined by their occurrence in a particular role in a particular construction in a particular language. Hence a word class must be language-specific, as noted in §2. English Numerals are defined by their occurrence in the relevant slot in the English Numeral Modification Construction [___ N(-PL)]. Lahu Numerals do not occur in the English Numeral Modification Construction. Conversely, English Numerals do not occur in the Lahu Numeral Modification Construction.

For this reason, Radical Construction Grammar, in concert with a number of typologists (most recently, Dryer 1997a and Haspelmath 2010, but also many earlier typologists including Greenberg, as will be seen below), posits a different type of theoretical entity, which Haspelmath (2010) has christened a COMPARATIVE CONCEPT. A comparative concept, unlike a distributionally defined word class, is defined on a cross-linguistically valid basis.

The simplest type of comparative concept is a semantic or pragmatic concept; here we will call it a FUNCTIONAL COMPARATIVE CONCEPT. For example, a property, in the sense

of a unary valency, stable, inherent, gradable category, such as age, dimension and value (Dixon 1977) is a functional comparative concept. Words denoting property concepts can be identified in any language, and compared across languages. Functional comparative concepts were introduced into typology by Greenberg in his seminal paper on word order typology (Greenberg 1966:74).

Comparative concepts that are more interesting for typologists are those that combine both function and grammatical form (Haspelmath 2010). These can be called HYBRID COMPARATIVE CONCEPTS (Croft 2016a). Hybrid comparative concepts are possible because there are certain formal traits that can be cross-linguistically defined, i.e. not with respect to distributional occurrence in a role in a language-specific construction. For example, the order of elements, such as Adjective-Noun order in English vs. Noun-Adjective order in Zuni, is a cross-linguistically defined formal property. Another cross-linguistically defined formal property is zero vs. overt coding of a semantic category, such as zero coding of singular number in English vs. overt coding of singular number in Lithuanian. It is worth noting that cross-linguistically valid formal traits in morphosyntax always involve a functional category in their definition (Croft 2009:161-62). For example, the order of elements requires a functional definition of the elements that occur in that order, in our example the Adjective and Noun categories; and zero vs. overt coding is always coding of some semantic category.

Two types of hybrid comparative concepts are widely used in typology. The first is a CONSTRUCTION (Croft 2014, 2016a). A construction is a pairing of form and meaning/function, as in Construction Grammar (see §2). A construction as a comparative concept in typology is generally a construction that expresses a particular function. For example, a study of the typology of the “passive construction” (Siewierska 1985) or the typology of “intransitive predication constructions” (Stassen 1997) covers any morphosyntactic form that encodes the relevant function. A construction contrasts with a functional comparative concept, which is just the function, although the function serves as the basis for identifying form-function pairings that are instances of the construction.

The second widely used type of comparative concept in typology is a STRATEGY (Croft 2014, 2016a; the term is used early in modern typology; cf. Keenan and Comrie 1977 and Givón 1979). A strategy is a construction that expresses a particular function with a particular cross-linguistically valid formal trait. For example, property modification can be expressed with a prenominal strategy (Adjective-Noun) or a postnominal strategy (Noun-Adjective).

Of course, constructions and strategies are broad categories subsuming many hybrid comparative concepts, just as there are many different functional comparative concepts. More specific comparative concepts, such as ‘numeral modification construction’ or ‘prenominal modifier strategy’, have been defined by typologists. Haspelmath suggests that comparative concepts are created by typologists for their ‘usefulness’ (Haspelmath 2010). But their ‘usefulness’ is ultimately founded on an empirical basis, namely the existence of language universals that require the relevant comparative concepts to be formulated, such as Greenberg’s Universal 18: ‘When the descriptive adjective precedes

the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise' (Greenberg 1966:86; see also Croft 2019).

4. Cross-linguistic word classes and methodological opportunism

Comparative concepts are not word classes—that is, they are not distributionally defined in terms of roles in language-specific constructions. However, word classes play an important role in other grammatical theories. Specifically, word classes such as 'noun', 'verb' and 'adjective' are considered to be cross-linguistic categories as well as language-specific categories in other grammatical theories. How do word classes serve as cross-linguistic categories as well as language-specific categories in those theories (see also Baker and Croft 2017)?

One approach is to define cross-linguistic categories in essentially semantic terms. The most consistent approach in this respect is that of Cognitive Grammar (Langacker 1987a,b, 2008). Langacker offers conceptual definitions of 'noun', 'verb' and all other categories that play a role in his general theory of language. In the two closely related functional theories Role and Reference Grammar (Van Valin and LaPolla 1997) and Functional Grammar (Dik 1997), the basic theoretical categories are predicates and arguments/terms, which are functional categories; they are subdivided into semantically defined subclasses. Functionally defined categories are analogous to functional comparative concepts. But their relationship to distributionally defined language-specific word classes is unclear.

A second approach is to posit abstract formal categories that represent word classes, and to argue that the abstract formal categories are—or, sometimes, are not—manifested directly in word classes in specific languages. These abstract formal categories are defined in terms of their role in the abstract formal structures posited in the grammatical theory. The most consistent approach in this respect is generative grammar, particularly Baker (2003). One could also include here the less theoretically-oriented practice of many typologists and field linguists who assume that 'noun', 'verb' and 'adjective' are cross-linguistic categories, and then ask whether a particular language has 'adjectives' or not. Abstract formal categories are not comparative concepts in the sense defined in §3. But their relationship to distributionally defined language-specific word classes is unclear as well.

Radical Construction Grammar argues that the only way that traditional word classes could be identified as cross-linguistic categories is through an inconsistent and selective use of distributional analysis which is called there *METHODOLOGICAL OPPORTUNISM* (Croft 2001, chapter 1). A simple example of two closely related languages, English and German, illustrates the point (Croft 2007:411-12, 416-17). German Adjectives are defined distributionally by the fact that they index (agree with) the Number and Case values of the German Noun that they modify. English Adjectives do not index the English Nouns that they modify; the main distributional definition is their occurrence in the relevant role in the English Noun Phrase.

Now, ‘English Adjective’ and ‘German Adjective’ are language-specific categories. They are defined in terms of certain English and German constructions respectively. Hence, at one level, it is completely arbitrary that they are both called ‘Adjective’. But they are considered to be instantiations of the same word class ‘Adjective’ for the purposes of cross-linguistic comparison in many theoretical approaches, including generative grammar. From the point of view of distributional analysis, the equating of German Adjectives and English Adjectives is opportunistic. Not only do English Adjectives not occur in German and German Adjectives not occur in English, but the constructions used in distributional analysis are totally different.

One could use the “same” constructions for distributional analysis across the two languages. The only way to do that is by comparative concepts, in particular of indexation of particular semantic categories such as cardinality (number) and participant role (case) in modification. But the only modifiers in English that index any feature of the head are the English Demonstratives *this/these* and *that/those*. So this approach would equate German Adjectives with English Demonstratives (only approximately, since English Demonstratives index Number but not Case). It is obvious that the reason English Adjectives and German Adjectives are called ‘Adjective’ is functional: the two word classes both include words that correspond to property concepts, and the words are also used as modifiers in referring expressions. The use of distributional analysis is inconsistent and selective, designed to identify two word classes as the ‘same’ across languages for other reasons—based on function in this case, but based on abstract theoretical reasons in other theories. Or they are based on a preconception of whether the language has Adjectives at all.

5. Cross-linguistic word classes and the skeleton model of language universals

Dryer (1997a) poses the problem of cross-linguistic categories with respect to grammatical relations, but notes parallels with phonemes and parts of speech, that is, the word classes discussed here. He suggests four things that grammatical theories might propose that exist (Dryer 1997a:116-17, as adapted in Croft 2001:32):

- a. categories and relations in particular languages
- b. similarities among these language-particular categories and relations across languages
- c. functional, cognitive and semantic explanations for these similarities
- d. categories and relations in a cross-linguistic sense

The assumption of (d), the existence of cross-linguistic categories and relations, reflects a hidden assumption about the nature of language universals, which can be called the SKELETON MODEL of language universals. The skeleton model presupposes that language universals are of the form “All languages have X”, where “X” is a particular grammatical category or structure; “X” constitutes the “bones” of the skeletal model. Another way to describe the skeleton model is the widely held assumption that a language universal is only something “which is found in all languages”, as described by an anonymous

reviewer. Some theories describe the set of language universals in this sense as “Universal Grammar”.

In fact, most adherents to the skeleton model allow that not everything that is a language universal in this sense may be found in a language. For example, for many linguists, including typologists who adhere to the skeleton model, ‘adjectives’ do not need always to be present. (This has been called the “cafeteria” model: a language can pick and choose its categories from Universal Grammar.) But enough of the skeleton has to be instantiated in a particular language grammar, and language universals in the skeleton sense must be instantiated in enough languages, to be considered “universal”.

A consequence of the skeleton model is the assumption that the theoretical entities that are language universals in this sense are of the same type as the entities in particular language grammars—e.g., word classes. In the skeleton model, language universals are a subset of what is found in particular language grammars. Particular language grammars also have many arbitrary patterns that “flesh out” the skeleton of structure provided by language universals. But particular language grammars always include instantiations of the language universals (the “bones”). For example, for many linguists, the skeleton includes ‘noun’ and ‘verb’: all languages are assumed to have nouns and verbs in the skeleton sense.

In contrast to the skeleton model, Dryer argues that in a theory of cross-linguistic universals, one does not need (d); one only needs (c) (Dryer 1997a:139). In fact, trying to establish (d) impedes the explanation of cross-linguistic patterns (ibid., 140).

Radical Construction Grammar, and much of modern typology, follows Dryer and does not adopt the skeleton model. In Greenbergian typology, the vast majority of language universals, and certainly most of the interesting language universals, are universals of patterns of cross-linguistic variation. In Greenbergian typology, a language universal is a proposition that is true of all languages. The proposition need not be of the form “All languages have X”. In particular, Greenberg proposed implicational universals of the form “If a language has Y, then it also has X”. Universal 18, presented in §3, is an implicational universal. Universal 18 is not the description of a particular grammatical structure that may or may not be found in a language, but it is true of all languages. Hence there is no need in Radical Construction Grammar, or in typology, for cross-linguistic categories such as ‘noun’ or ‘adjective’. Language universals are empirical generalizations that reflect ‘functional, cognitive and semantic explanations’ for word class variation—Dryer’s (c).² But language universals in the typological sense are based on patterns of word class variation that are systematically, not opportunistically, defined. The role of word classes in uncovering these types of universals and their explanation will be discussed in §§7-9.

² In fact, universals such as Universal 18 are typically not exceptionless. However, even when they are not, they may still manifest the underlying functional, cognitive and semantic explanation that does apply to all languages (see Dryer 1997b).

In Radical Construction Grammar, universal patterns of variation—that is, similarities and differences among language-particular categories, Dryer’s (b)—in the realm of parts of speech are explained in terms of certain combinations of semantic concept categories and information-packaging functions, such as modification by property concepts. These are comparative concepts, not word classes. These universal patterns of variation do not require positing a cross-linguistic category of ‘adjective’. Universal patterns of variation in the occurrence of indexation in modification constructions, for example, can be characterized in terms of implicational universals defined over semantic categories, e.g. ‘If a property concept modifier indexes the referent, then the deictic concept modifier does as well’.

6. Language-specific word classes and the building block model of grammar

The preceding sections show that a major difference between Radical Construction Grammar and other theories is the nature of language universals. In particular, word classes are not doing the work of providing a skeleton for language universals. But the role of word classes in the analysis of the grammars in specific languages in Radical Construction Grammar also differs from other theories. In the debates over the relationship between language universals and language-specific word classes, there are hidden assumptions about specific language grammars as well.

It is generally assumed that there is a small number of word classes in a language. The major word classes are Noun, Verb and (in some cases) Adjective; in addition a larger but not enormous set of minor word classes is posited. It is also generally assumed that the word classes are mutually exclusive and form an exhaustive partition of the words in a language, although most grammatical descriptions allow for some polycategoriality or polysemy, that is, membership in more than one word class. In particular, word classes are GLOBAL in the sense of being shared across grammatical constructions of the language (Croft 2001:45).

Many reference grammars include chapters on word classes near the beginning of the grammar. Some reference grammars are entirely organized around word classes. With this assumption, it seems more plausible to consider the small number of word classes to be universal in the sense of being shared across languages, and hence part of the skeleton model of language universals.

Radical Construction Grammar argues that this is not what is found in the grammars of specific languages. As noted in §2, the word class defined by occurrence in a role in one construction will not in general be the same as the word class defined by occurrence in a role in another construction in the same language. A consistent, strictly distributional analysis will lead to a very large number of word classes. Moreover, the word classes will overlap everywhere and will therefore not form a partition of the words in the language. In other words, word classes are not global; they are specific to the construction that defines them.

How is the result of consistent, strict distributional analysis reconciled with the assumption of a small number of (largely) mutually exclusive word classes that partition the words of a language? Again, Radical Construction Grammar argues that the only way that word classes can be rendered global is through an inconsistent and selective use of distributional analysis, namely methodological opportunism. One selects just one construction and ignores the other constructions, or a handful of constructions with strongly overlapping distribution and ignores the differences in distributions of the individual constructions (and any other constructions that are not used). For example, the differences in distribution of English Adjectives in (4)-(7) in §2 are ignored in positing a construction-independent, global word class of Adjective in English.

There is a hidden assumption here as well. The motivation for seeking a small set of word classes that is shared across grammatical constructions of a language is the BUILDING BLOCK MODEL of grammar. In the building block model, word categories are atomic, primitive units of grammatical analysis and structure. Constructions are built out of word categories. For example, a simplified characterization of the English Noun Phrase construction is [Dem Num Adj N], or [Dem [Num [Adj N]]]. There are of course other modifiers in the English Noun Phrase, but including them would simply show that there are more building blocks necessary to define the English Noun Phrase construction.

Radical Construction Grammar rejects the building block model along with its prerequisite, global word classes. There is a circularity of argumentation in the building block model of word classes. Word classes are defined by their occurrence in constructions (see §2). But once they are defined by constructions, word classes are assumed to be atomic primitive units, and constructions are defined as being built out of combinations of word classes (Croft 2001:34-37). In other words, identification of constructions is presupposed in order to define word classes, but then constructions are defined in terms of word classes.

The way out of this circular reasoning that is taken by other grammatical theories is an a priori assumption that word classes as building blocks are “already there”, and distributional analysis is merely a way of “discovering” their existence (Croft 2001:10-11). But, as with the assumption of cross-linguistic word classes, linguists do not agree on what the building blocks are. So they opportunistically select just the distributional facts that support their a priori assumptions. When two linguists disagree, there is no way one can resolve the disagreement because both linguists are being methodologically opportunistic: both are right for the distributions they present in their arguments, and both are wrong due to the distributions they ignore or dismiss. These a priori language-specific word classes are also the a priori cross-linguistic word classes we observed in the preceding sections.

Instead, Radical Construction Grammar follows the consequence of consistent, strict distributional analysis and recognizes that constructions, not word classes, are the primitive units of grammatical analysis. Word classes are defined in terms of roles in constructions. They are not the building blocks out of which constructions are made. There is a many-to-many mapping of the words of a language to roles in the

constructions of the language. Overlapping word classes are the starting point of grammatical analysis of particular languages. This approach will be illustrated in §§8-9, but first we turn to another hidden assumption about word classes and their role in grammatical theory.

7. Language-specific word classes as populations rather than essentialist categories

In Radical Construction Grammar, distributionally defined language-particular word classes are completely different from the comparative concepts with which language universals are described. Comparative concepts are essentialist categories, while language-particular word classes are populations in the sense of the neo-Darwinian synthesis of evolutionary biology.

Of these two, the nature of comparative concepts is the more familiar. Comparative concepts like humanness (the semantics of animacy), cardinality (the semantics of number and numerals) or linear order (a cross-linguistically definable temporal property of morphosyntactic form) are defined by essential properties. These essential properties allow us to identify a linguistic construction (form-meaning pairing) as instantiating—or not instantiating—the comparative concept, in any language. This is, of course, the “classical” definition of a category as a kind: a grouping of individuals by virtue of a set of properties that all and only the individuals possess. This type of grouping of individuals goes under many names; for example, Dahl (2016:428) calls it ‘Universal’. Here we use the term ‘essentialist’, which describes the defining feature of this type of category and is also the term used in evolutionary biology (Mayr 1982:256; Hull 1976, 1988:215-16).

Individuals, by contrast, are spatiotemporally bounded, historical entities: they exist only in a particular space and time. The individuality of a quartz crystal is its unique spatiotemporal existence: it was formed at a point in time, it exists in some location for some period of time, and it will pass out of existence when it is destroyed. In other words, individuals are historical entities. A spatiotemporally bounded set of historical entities (individuals) is also a historical entity (an individual): this is a POPULATION in the biological sense (Ghiselin 1974; Hull 1976; Mayr 1982:272-75).

Population thinking emerged with the neo-Darwinian synthesis in evolutionary biology (see Mayr 1982, chapter 6 for a brief history). Population thinking resolved serious problems with the essentialist theory of a biological species by abandoning the essentialist theory. Biological species cannot be defined in terms of a set of essential properties. There are species that share seemingly essential structural traits, yet do not interbreed (Hull 1988:104). Above all, the “essential” traits of species evolve over time and may disappear. The “essential” traits of species therefore cannot define the species.

Instead, species are defined as a reproductively isolated population of organisms. As such, a species is a historical, spatiotemporally bounded individual: ‘Just as the name “Gargantua” denotes a particular organism from conception to death, “*Gorilla gorilla*” denotes a particular segment of the phylogenetic tree’ (Hull 1988:215). Populations may

be a grouping of individual entities, but they are very different type of grouping from an essentialist category. Since a species is itself an individual, a species name is a proper name, not the name of a type (Ghiselin 1974), and the relation between the constituent organisms and the population is not a type-instance (token) relation, but more like a part-whole relation (Hull 1976).³

What sort of “category” is a distributionally defined language-particular word class? Dahl suggests that single values of a language-particular category such as the Future Tense ‘tend to be Individual-like’, but questions this conclusion for language-particular, distributionally defined word classes such as Adjective (Dahl 2016:429).

In Radical Construction Grammar, distributionally defined language-particular word classes are populations. Radical Construction Grammar is a component of the evolutionary framework for understanding language presented in Croft (2000). In the evolutionary framework, speech communities are populations, and so are the linguistic entities dependent on them: languages, utterances, and the structural parts of utterances. A speech community is a population of (relatively) communicatively isolated speakers (Croft 2000:17-19); it is a spatiotemporally bounded, historical entity. A language is the population of the actual utterances produced by the speech community (Croft 2000:26)—hence Radical Construction Grammar is a completely usage-based theory of language. All of the familiar “problems” of defining languages are found in defining species, and can be addressed in the population theory of a language (Croft 2000:13-20; the “problems” are mostly due to the gradual process of speciation/language birth, and the incompleteness of reproductive/communicative isolation).

The population of utterances that constitutes a language forms a population not only by virtue of being produced by members of a speech community. Single-word utterances are replications of the words that make them up. That is, words are replicated from prior uses of that word. Multiword utterances are the result of the recombination of linguistic units replicated from prior uses of those units. The linguistic units that are replicated and recombined are the constructions and the words that fill the roles in the constructions. The replication process is of course mediated by speakers, who replicate and recombine words and constructions and other linguistic units in language use, based on their knowledge about their language—the utterances they have been exposed to, and used themselves. A speaker’s grammar is also a spatiotemporally bounded individual—the speaker’s lifetime knowledge about her language—and so is the collection of the grammatical knowledge about their language of all the members of the speech community.

In contrast, structuralist and generative theories of language and language structure are essentialist. Words and constructions exist as abstract entities, not bound to time or space; the same applies to the rules that govern their combination (not REcombination, which

³ The population theory of species is still contested in evolutionary biology (Mayr 1982:276,279; Hull 1988:213, fn. 2). Dahl proposes using a modified essentialist theory for language-specific categories, adopted from biologists who do not accept the population theory of species (Dahl 2016:435-36; Dahl does not discuss the population theory).

implies replication and thus historical existence). Grammars in either the sense of an idealized speaker-hearer's knowledge or in the sense of an abstract description of grammatical structures and rules are also essentialist. It is likely that the widespread essentialist representation of a grammar, the sentences generated or sanctioned by the grammar, and the parts of those sentences including word classes, underlies arguments that language-specific word classes are the same kind of thing as comparative concepts, which are defined in essentialist terms (see above).

The Radical Construction Grammar definition of a word class is an example of the population definition of linguistic units, including words, phrases and constructions. In order to explain the population definition of word classes, we will start with the definition of a single word, such as English *heart*. In an essentialist theory, English *heart* is a single individual abstract unit with a particular form and meaning, possibly multiple meanings, as in a dictionary entry. In the essentialist view, the word has as an essential property its word class, as asserted in introductory linguistics and syntax textbooks in various approaches, generative and otherwise (e.g., Carnie 2013:44; Fabb 2005:11; O'Grady et al. 1997:164; Finegan 2007:35).

In the population view, English *heart* is a population of uses of the word, replicated through the lifetimes of speakers of English and, thanks to the overlap of generations of speakers, through the history of the language. Thus, a word as a population has a historical (temporal) as well as a spatial dimension. Of course, for a single speaker, the most relevant uses are those in that speaker's direct experience. But the uses that speaker is exposed to are replications of prior uses outside the speaker's immediate spatiotemporal experience, and those prior uses influenced the uses that the speaker experiences. The uses of *heart* in the English speech community form a LINEAGE, or rather a set of intertwining lineages of replications of prior uses. These lineage structures may not seem very relevant to the analysis of a common word such as *heart*. However, the lineage of uses is more obviously central to the understanding of a linguistic term such as *adjective* or *universal*, and is reflected in citations in the linguistic literature referring to prior uses of the term, tracing the lineage of its use and the evolution of its meaning.

The same is true of constructions. A construction such as the English Progressive construction, described in essentialist terms as [SBJ *be* VERB-*ing* ...] is a population of its uses. The population is defined by the replication of the construction by speakers of the English speech community. Unlike a simple word, a construction also involves recombination: recombination of the construction, forms of *be*, the morpheme *-ing* and the fillers of the roles: the English Passive Verbs that have occurred in the construction and the English Passive Subject phrases that have occurred in the construction—and the latter in turn are recombinations of the elements of the Subject phrase. Since the English Progressive construction involves recombination, an utterance replicating the Progressive construction constitutes multiple intertwining lineages.

Now we may provide a Radical Construction Grammar definition of word classes as populations. Word classes are distributionally defined. They are therefore language-

specific and construction-specific. As such, a word class is a population: the population of elements that have filled, and will fill, a particular role in the replications of a particular construction (i.e., in specific utterances), which also forms a population. The names for word classes, like other language-specific linguistic entities, are proper names. Hence the convention adopted in typology and in Radical Construction Grammar to capitalize the names of language-specific word classes (and constructions) accurately captures their nature as populations.

As a population, a word class is completely different from a comparative concept, which is a kind, defined by the essential traits of the kind. Word classes are “made up of” particular linguistic entities—individuals in a language. Comparative concepts are also “made up of” particular linguistic entities—individuals within and across languages. But the relation between a population and its constituent entities, and the relation between a kind and its instantiations, are totally different. The first is a relation between a spatiotemporally bounded individual and its component parts (the replications); the second is a relation between a universal kind and instances that share the essential traits of the kind (Hull 1976).⁴

8. The role of word classes in Radical Construction Grammar: language-internal and cross-linguistic variation

Sections 3-7 focused largely on what word classes in Radical Construction Grammar are not, because what they are not is linked to deeply or unconsciously held assumptions about the nature of words, grammar, and language universals by advocates of other linguistic theories. Word classes are not essentialist categories; they are populations defined by occurrence—and recurrence—in a particular role in a particular construction. Word classes are not building blocks for grammatical structure; they are the product of recombination in the replication of the constructions that define them. Finally, word classes as spatiotemporally bounded populations cannot be used to provide a skeleton of grammatical categories for Universal Grammar; language universals must be built on comparative concepts, which are kinds, that is, essentialist categories.

Radical Construction Grammar, and typology in general, proceeds by the inductive analysis of empirical linguistic data. Empirical linguistic research forms generalizations over samples of the populations of words, constructions and utterances produced by speakers of the language, and seeks to explain those generalizations. Freed from requiring word classes to be essentialist kinds, building blocks of grammars, and the skeleton for language universals, Radical Construction Grammar uses patterns of distribution of words to uncover language universals.

⁴ Dahl (2016) and Gil (2016) suggest that different dialects, varieties and even languages may have the “same” category, e.g. the Perfect in different English dialects or the Relative Case in Eskimo-Aleut languages (Dahl 2016:430). Gil extends this idea to code-mixing and borrowing. What these categories have in common is a shared lineage (language contact can lead to category lineages “jumping” languages). Lineages are historical entities, so they are not comparative concepts of the type discussed in §3. Language-specific categories are single branches of shared lineages, though there are complex issues in deciding whether there is sufficient communicative isolation to define a separate language lineage, as noted above.

Radical Construction Grammar is based on distributional analysis. In section 2, it was observed that, empirically, word classes defined by distributional analysis are construction-specific and language-specific. In the grammatical description of a single language, distributional analysis produces a many-to-many mapping between words and constructions, or more precisely, between words and a particular role in a construction. This many-to-many mapping does not “clump” words into a small number of global (construction-independent) word classes.

Taking this view means taking a less restrictive perspective on the representation of word classes. Word classes are not mutually exclusive. However, in most grammatical theories, overlap is generally restricted to a taxonomic hierarchy (a tree), or a set of shared features, which leads to a lattice. A lattice quickly becomes unreadable when every construction determines a different distribution. More significantly, it does not easily capture generalizations linking overlapping distributions. In order to capture these generalizations, we must also find a basis for cross-linguistic comparison of word classes.

Constructions are defined in terms of both form and function. Typically this is done informally, not least because distribution is commonly not described as occurrence in constructions (§2). In construction grammar, distributional analysis is defined explicitly in terms of words occurring in roles in complex constructions. Words and complex constructions are both pairings of form and function. The function of words and (complex) constructions provides a basis for cross-linguistic comparison, because functions are comparative concepts. Thus, words can be compared in terms of translation equivalents for the relevant word sense, and constructions can be compared in terms of equivalent functions, such as predication and modification. Methodological opportunism is avoided by comparing distributions across languages of semantically equivalent words in functionally equivalent constructions. Finally, hybrid comparative concepts allows for the comparison of formal morphosyntactic properties of constructions.

In other words, language-particular distributional facts can be viewed not just as parts of language populations, but also as instantiations of essentialist comparative concepts, defined by function and also form. Typology uncovers language universals manifested by language-particular distributional facts. Functional, cognitive and semantic explanations for language universals (see §5) account for significant facts of languages as historical entities (see §7): constraints on distributional patterns, and constraints on paths of evolution of words and constructions.

For example, English uses different constructions with different strategies for predication of action concepts, property concepts and object concepts:

- (8) a. Donna sings.
- b. Donna **is** tall.
- c. Donna **is an** Alabaman.

These constructions share the same function, predication, and so they can be compared. One can also compare predication constructions across languages—‘predication construction’ is a comparative concept. We can then examine distribution patterns for words expressing action concepts, property concepts and object concepts across languages as well as within languages.

In this case, there is an implicational universal across languages with respect to the strategies used for predication, specifically the use of a separate morpheme (overt coding) vs. absence of such a morpheme (zero coding) for predication of action, property and/or object concepts: a predication construction for any semantic type on the hierarchy Action < Property < Object uses at least as many morphemes as predication of any semantic type to the left on the hierarchy (Croft 1991:130; Stassen 1997:127; Pustet 2003). In English, overt coding is found in property and object predication; see (8b-c). In Mandarin, overt coding is found in object predication only (Li and Thompson 1981; but see below). In Makah, no overt coding is used for action, property or object predication (Jacobsen 1979).

The language-particular word classes defined by the zero coded and overtly coded predication constructions are different. Based on the predication constructions in each language, one might say that English “distinguishes Predicated Adjective and Predicated Verb word classes”, Mandarin “has a single Predicated Non-Nominal word class”, and Makah “has a single Predicate word class”. But Mandarin, Makah and English all instantiate the language universal given at the beginning of this paragraph. Specifically, this universal and other universals of predication, modification and reference constructions indicate that action concepts are prototypically predications, and that property concepts are prototypically modifiers and object concepts are prototypically referring expressions (Croft 1991, 2001).

Yet this hierarchy is already manifested in English alone. English has both overt coding and zero coding strategies for its predication construction. Most object predication constructions involve two morphemes, *be* and *a*, whereas property predication involves just one morpheme, *be*. In fact, most typological universals involve comparison of patterns of distribution—variation in use—of multiple constructions in a language, and equivalents of those same constructions across languages. In other words, GRAMMATICAL VARIATION WITHIN A LANGUAGE AND GRAMMATICAL VARIATION ACROSS LANGUAGES ARE GOVERNED BY THE SAME UNIVERSAL PATTERNS AND PRINCIPLES (Croft 2001:107).

If we drill deeper into language use in a single language, this principle continues to hold. In some languages, for example Lango, constructions exhibiting different strategies for modification—zero, the Attributive Particle, or the combination of Attributive Particle and Relative Pronoun—are used in both property and action modification. Hence there is no difference in the distribution of the different modification constructions in categorical terms. However, the more overtly coded Attributive+Relative combination, using two morphemes instead of one, is less normal for property concepts but preferred for action concepts (Croft 2001:78-80, from Noonan 1992 and personal communication), as expected from the universal that property concepts are prototypically modifiers. In Mandarin (and Cantonese; Li and Thompson 1981:143; Matthews and Yip 2004:158;

Croft 2020, Lecture 3 and references cited therein), the degree modifier for property concepts is frequently used in predication without denoting intensification; that is, it is grammaticalizing into overt coding for property predication.

9. Word classes, the semantic map model, and exemplar semantics

In Radical Construction Grammar, construction-specific word classes can be compared within and across languages using the meanings of words and the functions of constructions. This comparison reveals universal patterns of variation in distribution. The preceding section gives simple cases where implicational hierarchies familiar from typology capture the universals.

Other universal patterns of distribution are more complex, and different methods are used to capture the patterns. The basic method is the SEMANTIC MAP MODEL, also with a long lineage in typology (for overviews, see Croft 2003:133-39; Haspelmath 2003). In the semantic map model, word meanings occurring in distributionally defined word classes are mapped in a conceptual space structured according to variation in distribution within and across languages. For example, one can organize word meanings in a conceptual space according to their distribution in different predication constructions, as is done by Stassen (1997). The classic semantic map model uses a graph (network) structure for the conceptual space: word meanings that occur in the same construction(s) are nodes joined by edges (links) in the graph (for an algorithm for constructing the graph, see Regier et al. 2013; for fitness statistics for the algorithm, see Croft to appear).

For even more complex patterns of distribution, multidimensional scaling (MDS) can be used (Croft and Poole 2008; Croft 2010c, to appear; for an implementation for use in linguistics, see Timm 2020). MDS uses a Euclidean spatial model: word meanings that occur in the same construction(s) are spatially near each other. Otherwise it is constructed in the same way as the traditional graph structure model: organize meanings in a conceptual space such that meanings of words that occur in the same overlapping construction-specific distributions are “closer” to each other, either in terms of paths through the graph or in terms of Euclidean distance. For example, Rogers (2016) uses the distribution of 49 object, property and action concept words in reference, modification and predication constructions in 11 languages to reveal universal patterns of variation in so-called “parts of speech”.

Semantic map model research using the graph-based or multidimensional scaling methods has used progressively more fine-grained meanings. For example, Rogers’ MDS analysis of property concepts confirms that different subclasses of property concepts form a scale from “more nouny” to “more verby” “adjectives” (terms are in scare quotes because these are not word classes as building blocks, but a way to convey the scalar quality of the language universal; see also Dixon 1977; Wetzler 1992:242; Stassen 1997:168-69; Croft 2001:96-97; Ye 2021).

The logical conclusion to this process is to use nonlinguistic stimuli, such as the Bowerman-Pedersen spatial relations pictures (found in Levinson and Wilkins 2006:570-

75) that cover the spatial dimensions roughly covered by English *at*, *in* and *on*. Empirical research in eliciting words using nonlinguistic stimuli, such as the adpositions used to express these spatial relations in nine languages by Levinson et al. (2003), reveals even finer-grained variation in distribution across languages than was previously realized. Distributional variation should not be thought of as being organized into a network of discrete concepts but rather as multiple continuous dimensions of meaning or function (Croft 2010c). The dimensions of meaning/function in conceptual space underlie typological language universals (see §5); language-specific word classes “cut” these dimensions into constructionally-defined categories constrained by the conceptual dimensions.

Even this research does not fully capture the nature of variation in distributional patterns. This research assumes that any particular meaning/function, even a very finely-defined specific function, is expressed by one word or construction in a language. The reality of course is that speakers vary in their choice of word or construction for a particular experience, even including the spatial situations represented in the Bowerman-Pedersen stimuli (the Levinson et al. study abstracts away from within-language speaker variation). This is a truism: different speakers, and even the same speaker at different times, verbalize an experience in different ways. Variation in verbalization of scenes in the Pear Film (Chafe 1980), for example, is ubiquitous (Croft 2010d).

In order to integrate variation in verbalization to the theory of word classes, one must shift from an essentialist view of word classes, in which a word is a possible filler of a constructional role, to the population view of word classes described in §7, in which actual frequencies of occurrence are what matters. And one must look at the distribution of word plus construction across the situations mapped in conceptual space.

This is a far more complex task than simply recording occurrences of words in constructions. But preliminary studies suggest it is a fruitful approach. A study of the 20 verbalizations of scenes in the Pear Film found in Chafe (1980) shows that variation in verbalization is the source of grammatical, lexical and constructional change found in language histories and across languages (Croft 2010d). A follow-on study of the same data showed that the frequency distribution of words and constructions varied according to well-known semantic dimensions (Croft 2020, Lecture 9). For example, animacy and alienability governs preference for definite article vs. 3rd singular possessive for recurrent reference, degree of control over an event governs preference for subject vs. oblique realization of the experiencer in unintended human events, and direct manipulation and individuated theme vs. indirect causation and less individuated theme governs preferences in the choice of verb in the application (‘putting’) argument structure construction.

The conclusion that can be drawn from these studies is that VARIATION IN LANGUAGE USE (VERBALIZATION) AS WELL AS GRAMMATICAL VARIATION WITHIN A LANGUAGE AND GRAMMATICAL VARIATION ACROSS LANGUAGES ARE ALL GOVERNED BY THE SAME UNIVERSAL PATTERNS AND PRINCIPLES (Croft 2010d, 2020). Distributional patterns of words and constructions are best thought of as probability distributions of use over

conceptual space (Croft 2020:271). These probability distributions are directly manifested in language use, that is, the verbalization of experience. These patterns of variation get conventionalized in grammatical patterns of the speech community. Conventions change over time, and as speech communities split, the patterns of variation in language use come to be reflected in typological universals of cross-linguistic variation (Croft 2016b).

References

- Amha, Azeb. 2001. *The Maale language*. Leiden: CNWS Publications.
- Baker, Mark C. 2003. *Lexical categories: verbs, nouns and adjectives*. Cambridge: Cambridge University Press.
- Baker, Mark and William Croft. 2017. Lexical categories: legacy, lacuna, and opportunity for functionalists and formalists. *Annual Review of Linguistics* 3.179-97.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.
- Carnie, Andrew. 2013. *Syntax: a generative introduction*, 3rd ed. Chichester: Wiley-Blackwell.
- Chafe, Wallace (ed.). 1980. *The Pear Stories*. New York: Ablex.
- 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. 2005a. Logical and typological arguments for Radical Construction Grammar. *Construction Grammar(s): Cognitive and cross-language dimensions* (Constructional Approaches to Language, 3), ed. Mirjam Fried and Jan-Ola Östman, 273-314. Amsterdam: John Benjamins.
- Croft, William. 2005b. Word classes, parts of speech and syntactic argumentation” [Commentary on Evans and Osada, “Mundari: the myth of a language without word classes”]. *Linguistic Typology* 9.431-41.
- Croft, William. 2007. Beyond Aristotle and gradience: a reply to Aarts. *Studies in Language* 31.409-30.
- Croft, William. 2009. Methods for finding language universals in syntax. *Universals of language today*, ed. Sergio Scalise, Elisabetta Magni and Antonietta Bisetto, 145-64. Berlin: Springer.
- Croft, William. 2010a. Ten unwarranted assumptions in syntactic argumentation. *Language usage and language structure*, ed. Kasper Bøye and Elisabeth Engberg-Pedersen, 313-50. Berlin: Mouton de Gruyter.
- Croft, William. 2010b. Pragmatic function, semantic classes and lexical categories [commentary on Smith, “Pragmatic functions and lexical categories”]. *Linguistics* 48.787-96.
- Croft, William. 2010c. Relativity, linguistic variation and language universals. *CogniTextes* 4.303 [<http://cognitextes.revues.org/303/>]
- Croft, William. 2010d. The origins of grammaticalization in the verbalization of experience. *Linguistics* 48.1-48.
- 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. 2014. Comparing categories and constructions crosslinguistically (again): the diversity of ditransitives [Review article on *Studies in ditransitive constructions*:

- a comparative handbook, ed. Andrej Malchukov, Martin Haspelmath and Bernard Comrie]. *Linguistic Typology* 18.533-51.
- Croft, William. 2016a. Comparative concepts and language-specific categories: theory and practice. *Linguistic Typology* 20.377-93.
- Croft, William. 2016b. Typology and the future of cognitive linguistics. *Cognitive Linguistics* 27.587-602.
- Croft, William. 2019. Comparative concepts and practicing typology: on Haspelmath's proposal for "flagging" and "(person) indexing". *Te Reo* 62.116-29.
- Croft, William. 2020. *Ten lectures on construction grammar and typology*. Leiden: Brill. (Revised version of *Ten lectures on construction grammar and typology*, Foreign Language Teaching and Research Press, Beijing, 2012.)
- Croft, William. 2022. *Morphosyntax: constructions of the world's languages*. Cambridge: Cambridge University Press.
- Croft, William. To appear. Using quantitative methods for semantic maps. To appear in *Zeitschrift für Sprachwissenschaft*, special issue on "The future of mapping: new avenues for semantic maps research", ed. Athanasios Georgakopoulos and Stéphane Polis.
- 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.
- Croft, William and Eva Van Lier. 2012. Language universals without universal categories [commentary on Sandra Chung, "Are lexical categories universal? The view from Chamorro"]. *Theoretical Linguistics* 38.57-72.
- Dahl, Östen. 2016. Thoughts on language-specific and crosslinguistic entities. *Linguistic Typology* 20.427-37.
- Dik, Simon C. 1997. *The theory of Functional Grammar, part 1: the structure of the clause* (2nd rev. ed.), ed. Kees Hengeveld. Berlin: Mouton de Gruyter.
- Dixon, R. M. W. 1977. Where have all the adjectives gone? *Studies in Language* 1.19-80.
- Dixon, R. M. W. 2010. *Basic linguistic theory, vol. 2: grammatical topics*. Oxford: Oxford University Press.
- Dryer, Matthew S. 1997a. Are grammatical relations universal? *Essays on language function and language type*, ed. Joan Bybee, John Haiman and Sandra A. Thompson, 115-143. Amsterdam: John Benjamins.
- Dryer, Matthew S. 1997b. Why statistical universals are better than absolute universals. *CLS 33: Papers from the Panels*, ed. Kora Singer, Randall Eggart & Gregory Anderson, 123-45. Chicago: Chicago Linguistic Society.
- Evans, Nicholas and Toshiki Osada. 2009. Mundari: the myth of a language without word classes. *Linguistic Typology* 9.351-90.
- Fabb, Nigel. 2005. *Sentence structure* (2nd ed.). London: Routledge.
- Fillmore, Charles J., Paul Kay and Mary Catherine O'Connor. 1988. Regularity and idiomatity in grammatical constructions: the case of *let alone*. *Language* 64.501-538.
- Finegan, Edward. 2007. *Language: its structure and use*, 5th edition. Belmont, Calif.: Thomson Wadsworth.

- Ghiselin, Michael T. 1974. A radical solution to the species problem. *Systematic Zoology* 23.536-44.
- Gil, David. 2016. Describing languoids: when incommensurability meets the language-dialect continuum. *Linguistic Typology* 20:439-62.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Givón, Talmy. 2001. *Syntax, volume I*. Amsterdam: John Benjamins.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago 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. (Reprinted in Joseph H. Greenberg, *On language*, Stanford University Press, Stanford, 1990, 40-70.)
- Gross, Maurice. 1979. On the failure of generative grammar. *Language* 55.859-885.
- Harris, Zellig S. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.
- 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.
- Hull, David L. 1976. Are species really individuals? *Systematic Zoology* 25.174-91.
- Hull, David L. 1988. *Science as a process: an evolutionary account of the social and conceptual development of science*. Chicago: University of Chicago Press.
- Jacobsen, William H., Jr. 1979. Noun and verb in Nootkan. *The Victoria Conference on Northwestern Languages*. (British Columbia Provincial Museum Heritage Record No. 4.) Victoria, B. C.: British Columbia Provincial Museum, 83-155.
- Jagersma, Abraham Hendrik. 2010. *A descriptive grammar of Sumerian*. Ph.D. dissertation, Universiteit Leiden.
- Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63-99.
- Langacker, Ronald W. 1987a. Nouns and verbs. *Language* 63.53-94.
- Langacker, Ronald W. 1987b. *Foundations of Cognitive Grammar, vol. I: theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2008. *Cognitive grammar: a basic introduction*. Oxford: Oxford University Press.
- 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.
- Li, Charles and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley and Los Angeles: University of California Press.
- Matthews, Stephen and Virginia Yip. 1994. *Cantonese: a comprehensive grammar*. London: Routledge.
- Mayr, Ernst. 1982. *The growth of biological thought: diversity, evolution, inheritance*. Cambridge, Mass: Belknap Press.
- McCawley, James D. 1998. *The Syntactic Phenomena of English* (2nd ed.). Chicago: University of Chicago Press.

- Mulder, Jean Gail. 1994. *Ergativity in Coast Tsimshian (Sm'algyax)*. (University of California Publications in Linguistics, 124.) Berkeley: University of California Press.
- Noonan, Michael. 1992. *A grammar of Lango*. Berlin: Mouton.
- O'Grady, William, Michael Dobrovolsky, and Mark Aronoff. 1997. *Contemporary linguistics: an introduction* (3rd ed.). New York: St. Martin's Press.
- Palmer, Bill. 2009. *Kokota grammar*. (Oceanic Linguistics Special Publications, 35.) Honolulu: University of Hawai'i Press.
- Pustet, Regina. 2003. *Copulas: universals in the categorization of the lexicon*. Oxford: Oxford University Press.
- Regier, Terry, Naveen Khetarpal and Asifa Majid. 2013. Inferring semantic maps. *Linguistic Typology* 17.89-105.
- Rogers, Phillip. 2016. Illustrating the prototype structures of parts of speech: a multidimensional scaling analysis. MA thesis, University of New Mexico.
- Schachter, Paul and Timothy Shopen. 2007. Parts of speech systems. *Language typology and syntactic description, vol. 1: clause structure*, 2nd ed., ed. Timothy Shopen, 1-60. Cambridge: Cambridge University Press.
- Siewierska, Anna. 1985. *The passive: a comparative linguistic analysis*. London: Croom Helm.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Timm, Jason. 2020. MDS for Linguists.
https://github.com/jaytimm/mds_for_linguists, accessed 1 September 2020.
- Van Valin, Robert D. Jr. and Randy J. LaPolla. 1997. *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- Wetzer, Harrie. 1995. Nouniness and verbiness: a typological study of adjectival predication. Ph.D. dissertation, University of Nijmegen.
- Ye, Jingtang. 2021. Property words and adjective subclasses in the world's languages. PhD dissertation, Leipzig University.