Hans C. Boas / Benjamin Lyngfelt / Tiago Timponi Torrent

Framing constructicography

Abstract: Constructicography can be defined as a blend between Construction Grammar and Practical Lexicography, which aims at developing constructicons: repositories of form and function pairings in a language. In this paper, we present a comprehensive overview of this emerging field by (i) tracking the origins of both Frame Semantics and Construction Grammar and the repercussions of their intertwined developments to Computational Lexicography and Constructicography; (ii) comparing the impacts of the different degrees of interconnection between constructicons and framenets and (iii) discussing the possible applications of these resources. Also, we argue that Constructicography, while obviously building on the accumulated knowledge compiled by numerous Construction Grammar approaches to language, also contributes to its mother theory, since the effort to build coherent formalized computational resources forces constructionist analysis to go beyond describing families of constructions into the enterprise of describing a coherent construction grammar of a language.

Keywords: Frame Semantics, Construction Grammar, FrameNet, Constructicography, Constructicon

Hans C. Boas, University of Texas at Austin, 1 University Station, C3300, BUR 336, Department of Germanic Studies, Austin, Texas 78712, U.S.A., e-Mail: hcb@mail.utexas.edu
Benjamin Lyngfelt, University of Gothenburg, Box 200, Dept. of Swedish, 40530 Gothenburg, Sweden, e-Mail: benjamin.lyngfelt@svenska.gu.se
Tiago Timponi Torrent, Federal University of Juiz de Fora, Rua José Lourenço Kelmer, s/nº, Department of Portuguese Language and Linguistics, Juiz de Fora, Minas Gerais, 36036-900, Brazil, e-Mail: tiago.torrent@ufjf.edu.br

https://doi.org/10.1515/lexi-2019-0003

1 Introduction

This paper discusses the architecture of various FrameNet-based Constructicons to show how ideas and concepts from Constructicography and Construction Grammar mutually inform each other to enhance our understanding of the nature of language. We also aim to contribute to our understanding, among other things, of how the insights from Constructicography and Construction Grammar can be applied in the creation of foreign language teaching resources and computational techniques and tools. In doing so, we intend to shine more light on some of the following ideas outlined by Fillmore et al. (2012) in their seminal paper on the Berkeley FrameNet Constructicon:

> While building a Constructicon has different goals from those of designing a construction-based grammar of the language, the intention is that each construction will be represented in a way compatible with the development of a full grammar of the language (...). In some cases, we offer precise proposals for the treatment of a construction as it would appear in the grammar; in other cases the descriptions we present should be seen at least as organized observations about individual constructions, observations that need to be accounted for in a future complete grammar. In all cases we expect that the constructicon will contain useful information for advanced language pedagogy and that it will suggest new levels of expectation for parsing and other NLP activities. (Fillmore et al. 2012: 310)

The paper is structured as follows. In Section 2 we first give a short overview of how Construction Grammar and Frame Semantics evolved out of Fillmore's research in the late 1960 s and how the two theories developed since then. Then, we show how the idea for a Constructicon grew out of the work on the Berkeley FrameNet for English in the early 2000 s and we discuss how constructicography and Construction Grammar inform each other. In Section 3, we present constructicons that are more or less closely related to FrameNet, ranging from the Brazilian Portuguese constructicon (henceforth: ccn), which is FN-based, to the Russian ccn, which is only indirectly influenced by FrameNet via the Swedish ccn. We discuss how the various kinds of relations to FrameNet influence the way construction entries are compiled and how this differs from related work in Construction Grammar. Section 4 addresses the structure of constructicons in light of different theoretical and applied concerns, such as the question of how one can develop the resources from simple lists of construction entries into networks of constructions. In this connection, we also address the problems of coverage, relations between constructions, and interactions between constructions. In section 5, we discuss the relation between constructicography and construction grammar, return to the issue of constructional networks in light of this discussion, and provide an overview of how a constructicon can be useful for both human users and machines, in particular when it comes to building an empirical research program focused on investigating grammar. Section 6 presents our summary and concluding discussion.

2 From cases to frames and constructions

Before going into the details of how constructicons are compiled and used by construction grammarians, we first provide a brief overview of the origins of constructicons. We first discuss Fillmore's research on Case Grammar and Frame Semantics. Then, we show how FrameNet, a lexical database, has implemented the principles of Frame Semantics. Finally, we show how the idea of a constructicon grew out of research in Construction Grammar as well as the realization that certain limitations of a lexical FrameNet could only be overcome by extending the description and analysis of LUs at a level beyond the lexicon.

2.1 From Case Grammar to Frame Semantics to FrameNet

The origins of Construction Grammar and Frame Semantics can be traced back to Charles Fillmore's (1968) seminal paper *The Case for Case*, in which he proposed a set of so-called case frames. These case frames were representative of a verb's semantic valency and were supposed to allow linguists to figure out how they are mapped to syntax (similar to what became known later as so-called "linking" between semantics and syntax). Fillmore's original ideas were eventually abandoned, because the set of universal case frames (later known as semantic roles) ordered in a strict hierarchy turned out to be problematic (see Levin & Rappaport Hovav 2005 and Boas & Dux 2017).

During the 1970s and early 1980s, Fillmore further developed his ideas about semantic roles and their usefulness for structuring the lexicon (e.g., Fillmore 1975, 1977, 1978), eventually leading to the theory of Frame Semantics (Fillmore 1982, 1985a), which emphasized, among other things, the importance of cultural and world knowledge when describing the meaning of words.[1] Fillmore & Atkins (1992: 76–77) describe the main ideas underlying research in Frame Semantics as follows:

> A word's meaning can be understood only with reference to a structured background of experience, beliefs, or practices, constituting a kind of conceptual prerequisite for understanding the meaning. Speakers can be said to know the meaning of the word only by first understanding the background frames that motivate the concept that the word encodes. (Fillmore and Atkins 1992: 76–77)

In 1997, Fillmore founded the FrameNet project at the International Computer Science Institute in Berkeley, California, with the goal of applying semantic frames for the creation of an online lexical database documenting a variety of frame-semantic and syntactic information for the English lexicon (Fillmore & Baker 2010; Ruppenhofer et al.

1 For more details about how Frame Semantics and FrameNet grew out of earlier research by Fillmore during the 1960s and 1970s, see Boas & Dux (2017) and Boas (2018).

2017). In contrast to other lexical databases such as WordNet (Fellbaum 1998), which relies primarily on lexical relations, FrameNet (https://framenet.icsi.berkeley.edu) employs semantic frames to structure the lexicon using frame-semantic criteria (see Boas 2005b). Semantic frames are characterized in terms of Frame Elements (FEs), which are the participants (situation-specific semantic roles) that define the frame. For example, the FEs constituting the `Emptying` frame in FrameNet include Agent (the entity that does the emptying), Source (the region or container which is emptied of something), Instrument (the object with which the Agent empties the Source), etc.[2] Lexical units (LUs) are linguistic expressions (including all parts of speech and multi-word expressions) that evoke a given frame. LUs of the `Emptying` frame, for example, include the verbs *clean*, *degrease*, *drain*, *empty*, *flush*, *put*, and *strip* and nouns such as *decontamination* and *disarmament*.

In FrameNet, LUs can be thought of as specific senses of words (or multi-word expressions) evoking specific frames, which means that in FN polysemy is represented in terms of different senses of a word evoking different frames.[3] Consider, for example, the verb *clear*, which in FN consists of multiple LUs, each evoking a different frame, including `Emptying`, `Removing`, and `Verdict`. Semantic frames in FN are organized in a frame hierarchy that is organized in terms of frame-to-frame relations, which serve to capture relations across different frames. For example, Inheritance describes a relation in which a daughter frame inherits and further specifies information (including FEs) of a mother frame. In FN, the `Emptying` frame inherits from a higher level frame called `Container_focused_removing`. Other frame-to-frame relations include Perspective On, Precedes, Subframe Of, and Uses (see Petruck et al. 2004 and Ruppenhofer et al. 2016 for discussion).

The data contained in FrameNet are the result of an elaborate corpus-based workflow consisting of three main steps: (1) a team of lexicographers formulate a frame description (including definitions of FEs) and identify the LUs evoking the frame; (2) human annotators identify in extracted corpus data relevant sentences illustrating how LUs evoke a semantic frame and then annotate the FEs in those sentences; (3) Based on the annotations in the extracted corpus sentences and the frame definitions, lexical entries are compiled and stored in FN (for more details, see Fillmore & Baker (2010) and Boas (2013a)).

2 FN distinguishes between so-called core FEs, which are crucial for the understanding of the frame, and non-core FEs, which do not define the frame per se but provide additional information such as time, place, or manner.

3 In February 2019, FN contained 1224 frames, 10542 FEs, and 13,640 LUs.

flush.v

Frame: Emptying

Definition:

FN: clear of dirt, debris, etc., by running through with a fluid.

Frame Elements and Their Syntactic Realizations

The Frame Elements for this word sense are (with realizations):

Frame Element	Number Annotated	Realization(s)
Agent	(4)	CNI.-- (2) NP.Ext (2)
Cause	(1)	NP.Ext (1)
Source	(5)	PP[from].Dep (2) PP[of].Dep (2) AVP.Dep (2)
Theme	(5)	NP.Obj (5)

Figure 1: First part of lexical entry of *to flush* in the `Emptying` frame

Human users of the FN database typically access FN data by searching for specific semantic frames or LUs. When they search for a semantic frame, such as `Emptying`, they are first presented with the general frame description of the frame, followed (1) by a detailed list of FEs and their definitions, (2) the frame's frame-to-frame relations, and (3) a list of LUs that evoke the frame. Clicking on the lexical entry report of an LU such as *to flush* leads to a new page, which provides a definition of the LU, followed by a summary table listing how the FEs of the frame are realized syntactically (Figure 1) and a table summarizing the valence patterns for the LU (Figure 2) found in the corpus data.

Valence Patterns:

These frame elements occur in the following syntactic patterns:

Number Annotated	Patterns			
1 TOTAL	Agent	Source	Source	Theme
(1)	NP Ext	AVP Dep	PP[from] Dep	NP Obj
3 TOTAL	Agent	Source	Theme	
(1)	CNI --	PP[from] Dep	NP Obj	
(1)	CNI --	PP[of] Dep	NP Obj	
(1)	NP Ext	AVP Dep	NP Obj	
1 TOTAL	Cause	Source	Theme	
(1)	NP Ext	PP[of] Dep	NP Obj	

Figure 2: Valence table of *to flush* in the `Emptying` frame.

The valence table is very informative because it shows how specific combinations of Frame Elements are realized syntactically in terms of phrase type and grammatical function. Consider, for example, the valence table of *to flush* in Figure 2, which contains a total of three frame element configurations (FECs).[4] The first FEC and the third FEC only exhibit one syntactic realization in terms of phrase type and grammatical function, while the second FEC (Agent, Source, Theme) exhibits three distinct syntactic realizations.[5]

2.2 From Construction Grammar and FrameNet to the Berkeley constructicon

During the 1980 s, Fillmore and his associates at Berkeley began developing an alternative model of grammar that sought to overcome some of the shortcomings of the then prevalent Chomskyan generative-transformational paradigm. One of the main ideas was that the grammar of a language should not only focus on a few core phenomena, but that it should rather include all facets of grammatical phenomena, from the highly regular to (semi-)idiomatic phenomena. One of the first papers laying the groundwork for the approach that later became known as Construction Grammar (CxG) was Fillmore et al. (1988), which provided an analysis of the semi-idiomatic *let alone* construction. The late 1980 s and 1990 s saw the gradual emergence of various strands of CxG which all subscribe to a range of commonly shared principles, such as: (1) the basic organizing units of language are constructions (pairings of form with

4 The numbers in the left-hand column in Figure 2 refer to the total number of annotated corpus sentences for each FEC. Clicking on the number will display the annotated corpus sentence that forms the basis for the valence pattern.

5 The past 15+ years have seen the emergence of a variety of FrameNets for other languages, showing that semantic frames derived on the basis of English can also be re-used for the analysis of other languages (see Boas 2001 and 2002 for early exploratory work on re-using English frames for German). Spanish FrameNet started as the first large-scale FrameNet for a language other than English in 2002 (Subirats & Petruck 2003; Subirats 2009). Since then, FrameNets for other languages, including Japanese (Ohara et al. 2004, Ohara 2009), German (Burchardt et al. 2009), Swedish (Borin et al. 2010), Brazilian Portuguese (Salomão et al. 2013), and French (Candito et al. 2014) have been applying semantic frames derived on the basis of English to the description and analysis of the lexicons of their languages. See also the contributions in Boas (2009) for an overview. Besides general-domain multilingual FrameNets, there are also domain-specific FrameNet-type projects and databases dealing with specific aspects of the lexicon, such as the Kicktionary for soccer terminology in English, French, and German (Schmidt 2009), BioFrameNet covering biomedical terminology (Dolbey 2009), Bertoldi & Chishman (2011) for legal terminology in Brazilian Portuguese, the German Frame-based Online Dictionary, a learner's dictionary for English speakers learning German (Boas & Dux 2013, Boas et al. 2016), and m. knob (Multilingual Knowledge Base), a travel assistant providing tourist attraction recommendations, sentence translations and a Tourism lexicon for Brazilian Portuguese, Spanish and English (Diniz da Costa et al., 2018). For cross-linguistic applicability of semantic frames, see Boas (2019).

meaning/function); (2) there is no principled distinction between a so-called core and periphery; (3) there is no strict separation between syntax and the lexicon; (4) constructions are organized in terms of networks; and (5) grammar is non-derivational and non-modular (see Boas (2013b), Goldberg (2013) and Ziem & Lasch (2013) for discussion).[6]

Constructional research during the 1990 s and 2000 s focused for the most part on analyzing specific types of grammatical constructions, many of them semi-idiomatic, in order to better understand the nature of different types of constructions from the highly schematic constructions such as subject-auxiliary-inversion (Fillmore 1999, Goldberg 2006) to abstract meaningful constructions such as ditransitives (Goldberg 1995) and resultatives (Goldberg 1995, Boas 2003) all the way down to semi-idiomatic constructions such as *What's X doing Y* (Kay & Fillmore 1999) and idiomatic constructions such as *pain in the neck* (Nunberg et al. 1994). In a way, then, the constructional research of this period can be characterized as mainly focusing on individual case studies of particular (families) of constructions (see Goldberg & Jackendoff 2004).

The emergence of CxG as an alternative model of language is interesting for our discussion for at least two reasons. First, the idea that the lexicon and syntax should not be strictly separated has been proposed by research in lexicography (see Hanks 2013) and phraseology (see Cowie 1998). Similarly, research in Frame Semantics since the 1970 s (Fillmore 1975, 1978) proposed that syntax and the lexicon should not be confined to separate modules. In the mid-1980 s, Fillmore envisioned an expanded version of lexical entries capable of explicitly licensing syntactic patterns as the following quote illustrates:

> If new-style lexical entries for content words were to be seen instead as constructions capable of occupying particular higher-phrase positions in sentences and included both the needed semantic role and the needed specifications of structural requirements (...), we could see such structures as providing expansions of their existing categories. (Fillmore 1985b: 84)

Second, Fillmore (1988: 37) also suggested that describing and analyzing grammar on the basis of constructions could lead to an inventory of constructions governed by a set of principles regulating how such constructions interact: "The grammar of a language can be seen as a repertory of constructions, plus a set of principles which govern the nesting and superimposition of constructions into or upon one another."[7] The idea of a "repertory of constructions" is picked up in later work by Jurafsky (1991: 18), who coins the term "constructicon" in analogy to the term "lexicon."

Given these connections between Construction Grammar and Frame Semantics, Fillmore (2008a) proposes the extension of lexicographic work by FrameNet to also

6 See the various contributions on the different strands of CxG in Hoffmann & Trousdale (2013).

7 Interestingly, Pattern Grammar (Hunston & Francis 2000) makes many similar proposals regarding the relationship between what has traditionally been called the lexicon and syntax.

cover grammatical constructions. In a year-long pilot project, FrameNet researchers investigated the feasibility of extending FrameNet's analytical and technical apparatus to allow for a systematic analysis of grammatical constructions in corpus examples. Based on a modified version of the FN database and the FN annotation software, FrameNet researchers identified, analyzed, and annotated constructions in a very similar way as LUs (for details, see Fillmore 2008a, Fillmore et al. 2012, Boas 2017, and Lee-Goldman & Petruck 2018), resulting in construction entries for about 75 constructions. The English constructicon has since been followed by related constructicon projects for Brazilian Portuguese, German, Japanese, Russian and Swedish (each presented in a chapter in Lyngfelt, Borin et al. 2018). See Section 3 below for more details.

2.3 The relation between Construction Grammar and constructicography

The discussion in the previous section has shown that the concept of construction is at the center of both the theoretical framework of CxG as well as the more applied endeavor of constructicography (parallel to lexicography or grammaticography). CxG and constructicography both focus on exploring and analyzing constructions, but perhaps in different ways as the following quote from Lyngfelt (2018a: 2) illustrates:

> Thus, 'constructicon' now exhibits the same kind of polysemy as the related notions 'grammar' and 'lexicon': a theoretical notion of a linguistic system, on the one hand, and a corresponding descriptive resource, on the other. (...) Practical constructicon development may be characterized as a blend between construction grammar and lexicography, which we call constructicography.

But what are the theoretical and practical implications of having two distinct yet related approaches towards investigating grammatical constructions and the lexical items occurring in them? To answer this question, consider the ways in which research in CxG evolved in the 1980 s and 1990 s, namely as a series of in-depth case studies demonstrating the need for a holistic approach that goes beyond narrowly focused investigations of linguistic phenomena that the generative-transformational paradigms of that period saw as the central and interesting objects of observation, because they belonged to what was then thought of as the "core" (grammar), as opposed to the "periphery." Starting with Fillmore et al.'s (1988) seminal study on the *let alone* construction numerous other constructional analyses (for an overview please see the contributions in Hoffman and Trousdale 2013) have shown time and again that much of what was thought to be analyzable in terms of abstract syntactic phenomena actually turned out to be irregular or idiomatic to different degrees.

In a similar but more comprehensive way, Fillmore & Kay's (1993) unpublished Construction Grammar Coursebook (which articulated holistically for the first time the approach later known as Berkeley Construction Grammar) (see Fillmore 2013) and Goldberg's (1995) seminal book on argument structure constructions (which laid the

groundwork for what today is known as Cognitive Construction Grammar) (see Boas 2013b) both present collections of important case studies on a range of different constructions in a coherent fashion. While the formalisms and motivations between the two approaches differed in certain ways, both approaches demonstrated that CxG as a research paradigm was to be regarded as a serious competitor to other frameworks such as the Minimalist Program, Lexical-Functional Grammar, and Head-driven Phrase Structure Grammar. In addition, constructional research on languages other than English demonstrated that the concepts and ideas of CxG developed on the basis of English were also applicable to the analysis of other languages, though with certain specific limitations (see, for example, Croft (2001), Iwata (2008), and the contributions in Boas (2010)). However, during the 1990s and 2000s, most constructional research continued to focus primarily on case studies of individual constructions or constructional families, providing very detailed analysis of specific constructions. But the bigger picture was missing: How were constructions related to other constructions? How did constructions interact with each other when it comes to licensing utterances? What is the relation between lexical and constructional interaction, and, more specifically, how can the productivity of constructions be constrained in order to avoid the licensing of unacceptable utterances?

As shown in Section 2.1, Jurafsky (1991) and Fillmore (2008a) proposed the idea of a constructicon, including a lexicon, as a collection of construction entries documenting constructions of various types. In a way, then, one can say that the idea of a constructicon can be traced back to two distinct yet related backgrounds. Jurafsky's proposal can be seen as a preliminary proposal for capturing and storing the many different construction analyses that emerged as case studies during the 1980s and early 1990s. Fillmore's (2008a) idea for a constructicon, related to Jurafsky's, comes out of the systematic analysis of the English lexicon that led to the many entries in the Berkeley FrameNet database since 1997.

The corpus-based workflow underlying FrameNet produced a wealth of data that could not be systematically accounted for only at the lexical level. Based on this data and the commitment to the idea that there is a continuity between grammar and the lexicon, Fillmore proposed an extension to (lexical) FrameNet in order to cover structures that went beyond lexical units. According to Fillmore (2008a: 1), this extension would show more clearly how "each lexical item carries with it instructions on how it fits into a larger semantic-syntactic structure, or, alternatively, how semantic-syntactic structures are to be built around it." The Berkeley constructicon for English was a one-year long effort to create a prototype constructicon database parallel in structure to lexical FrameNet, consisting of construction entries that were derived on the basis of annotated corpus data. The construction entries compiled by the Berkeley FrameNet group are structured in a similar way as the lexical entries of FrameNet.[8]

8 Contrary to lexical FrameNet, whose frames are organized in a frame hierarchy with frame relations, the prototype Berkeley Constructicon for English was originally conceived as a pure list of con-

To illustrate, consider the `let_alone` construction (see Fillmore et al. 1988), which sets up two propositions as points on a single pragmatically-determined scale. The construction entry for let_alone consists of a prose description of the construction, together with a definition of the Construction Evoking Element (CEE, parallel to the frame-evoking target LU in lexical FrameNet) and the Construction Elements (CEs). The entry specifies that the construction is evoked by the multi-word conjunction *let alone*, which serves as the CEE, as the annotated example sentence *[$_{Context\_of\_focus}$ None of these arguments is] [$_{First\_conjunct}$ notably strong] [$_{CEE}$ let alone] [$_{Second\_conjunct}$ conclusive]* illustrates. The proposition that includes the First_conjunct is pragmatically stronger than the proposition that includes the Second_conjunct, and so the truth of the proposition that includes the First_conjunct entails the truth of the proposition that includes the Second_conjunct. The construction entry also includes (1) definitions (with examples) for the CEs Context_of_Focus, First_conjunct, Second_conjunct, and Trigger, (2) a set of annotated corpus examples illustrating the distribution of *let_alone* in context, and (3) a realization table (similar to the valence table in lexical FrameNet) showing how the individual CEs may be expressed syntactically. In Section 3 we return to a more in-depth discussion of the design of construction entries in order to show how construction entries may be organized in networks of different types.

As pointed out above, the construction entries in the Berkeley Constructicon are the result of a corpus-based workflow and they are organized only in terms of an alphabetical list (no network structure comparable to that of the FrameNet hierarchy). Unlike the different formalisms proposed by various proponents of CxG, e. g. Berkeley Construction Grammar (Fillmore & Kay 1993, Fillmore 2013), Cognitive Construction Grammar (Goldberg 1995), Sign-based Construction Grammar (Sag 2012) or Embodied Construction Grammar (Bergen & Chang 2013), the structure of these entries do not exhibit any specific commitment to a particular representation format or formalism.

This means that the Berkeley constructicon entries are, in a way, agnostic towards the different motivations and goals of the substrands of the various types of CxG (e. g. computational implementation, typology, psychological reality, etc.) and instead focus on providing an empirically based inventory of constructions that can be documented when conducting a full-text analysis of a corpus. The resulting construction entries can be thought of as similar in status to the periodic table of elements, which lists the specific properties of each element like hydrogen, oxygen, and calcium. The construction entries can thus be thought of as simple entries without any further specifics of how the constructions combine with other constructions to license utterances or how they are organized in terms of a constructional network (but see Sections 4 and 5 below for different ways of studying the interactions between construc-

struction entries without any specifics about relations between constructions. See Section 4 for more details regarding relations between constructions and how they can be captured in different ways.

tions). These tasks may then be further explored by research in CxG, which not only determines how individual constructions combine with other constructions to license utterances or how they are organized in terms of a constructional network, but whose insights will then also provide additional information that may be useful for refining existing construction entries.

3 FrameNet-based Constructicon(s)

Ten years after Frame Semantics had already made its way into the domain of computational language resources – framenets – the Berkeley FrameNet team started a pilot project to bring Construction Grammar in as well, as described in the previous section. The so-called "Beyond the Core project (BTC)" started in 2008 as "an attempt to augment the FrameNet lexicon with constructional information" (Lee-Goldman & Petruck 2018).[9] Since then, other constructicons have been developed but not necessarily as an augmentation of a framenet. In this section, we present a variety of such efforts, which are categorized in a continuum ranging from framenet-derived to framenet-influenced resources. To build such a framenet-relatedness continuum, we take into consideration both the analytical grounding of the constructicon on a framenet and the software infrastructure it relies on. We also discuss how the different levels of interconnection between constructicons and framenets plays a role in shaping the analyses the first can deliver, while also creating the opportunity to expand the coverage of the framenet to which they are linked.

3.1 FrameNet-derived Constructicons

The group of framenet-derived constructicons include, of course, the original English Ccn, but also the Japanese Ccn and the Brazilian Portuguese Ccn, both of which are to some degree distant from the polar position occupied by the English Ccn in the frame-relatedness continuum.

The English Ccn was conceived as an exploratory study aimed at augmenting Berkeley FrameNet, and has mostly remained as such for the past decade. It was created as an extension of the existing Berkeley FrameNet software apparatus to also cover the domain of multiword expressions and extra-lexical material. The motivation behind this decision was, as pointed out by Fillmore (2008a), the striking par-

9 By 2008, Berkeley FrameNet was already expanded into German (Boas 2002/2005a), Japanese (Ohara et al. 2003) and Spanish (Subirats & Petruck 2003) (see the contributions in Boas (2009) for an overview), and Framenet initiatives were being started for Brazilian Portuguese (Salomão 2009) and Swedish (Borin et al. 2010).

allels between (frame-based) lexical and constructional description and annotation (see Table 1). Given this context, the English Ccn was developed in a way that the construction entries in the database are composed of: (a) a name; (b) a prose definition; (c) a set of one or more construction elements (CEs); and (d) prose definitions of the CEs. The parallel with how frames are defined in Berkeley FrameNet is, indeed, striking. Once a construction was defined in the database, it was possible to annotate sentences featuring constructs licensed by it, by applying CE – and also GF and PT – labels to it.

Table 1: Lexical and constructional description and annotation compared (Fillmore 2008a)

Lexical FrameNet	Constructicon
Frame descriptions describe the frames and their components, set up FE names for annotation, and specify frame-to-frame relations; lexical entries are linked to frames, valence descriptions show combinatory possibilities, entries link valence patterns to sets of annotated sentences.	Constructicon entries describe the constructions and their components, set up construction elements (CEs, the syntactic elements that make up a construct), explain the semantic contribution of the construction, specify construction-to-construction relations, and link construction descriptions with annotated sentences that exhibit their type.
The FEs are given names according to their role in the frame, and provide labels for the phrases in the annotations that give information about the FE.	The CEs are named according to their function in the constructs, they provide the labels on words and phrases in annotated sentences.
The syntactic properties - grammatical functions and phrase types - are identified for all constituents that realize frame elements.	Phrase types are identified for constituents that serve as CEs in a construct; for constructions that are headed by lexical units, grammatical function labels will also be relevant.
Example sentences are selected that illustrate the use of the lexical units described.	Example sentences are selected and annotated for the ways they illustrate the use of the construction.
Annotations identify the LU, the FEs, and the GFs and PTs of the segments marked off,	Annotations contain labels for the CEs and identify, for lexically marked constructions, the relevant lexical material.
Valence patterns are identified, and linked to the annotations.	Varieties of construct patterns are identified and linked to the annotations.
Frame-to-frame relationships are documented and displayed in a separate resource.	Construction-to-construction relationships are identified and (will eventually be) displayed

Some construction descriptions in the English Ccn feature information about the frame evoked by the construction, which is sometimes present also in the fact that the CEs of some constructions are coincidental with the FEs of the frame they evoke, such as in the `way_neutral` cxn – note, in Figure 1, the information that this construction evokes the Motion frame inserted right below its name. Other entries may also supply information about inheritance relations with other constructions, such as the `comparison_inequality` cxn – note, in Figure 2, the information that this

construction inherits the comparison cxn inserted right below its name. Nevertheless, those pieces of information are not explicitly modeled in the database, that is, there is no database level connection between constructions and frames in the English Ccn, nor between one construction and another it inherits.[10]

Although not under active development, the English Ccn remains a model for other ccn endeavors, to varying degrees, and also serves as a hub for multilingual development (cf. Lyngfelt, Torrent et al. 2018, Ziem et al. (this volume)).

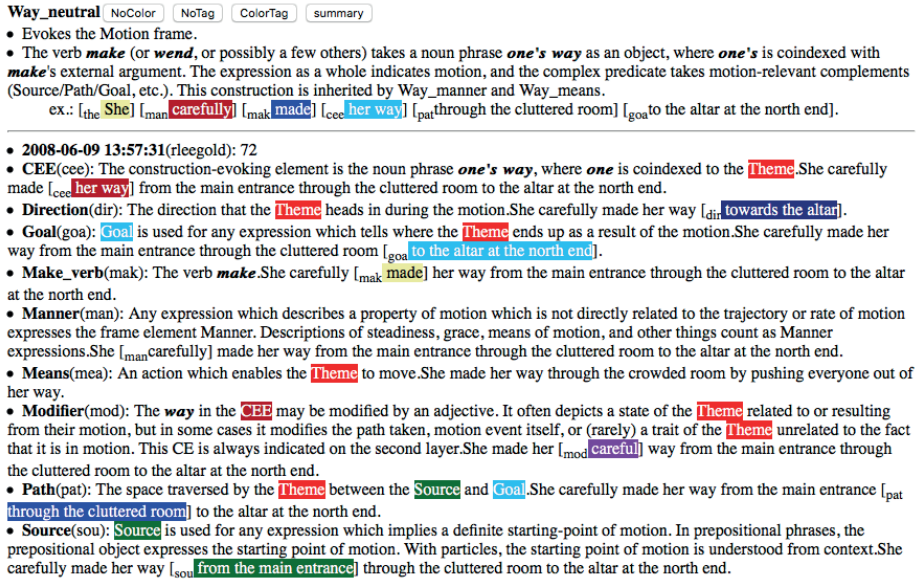

Figure 1: The `way_neutral` cxn in the English Ccn.

The Japanese FrameNet Ccn (Ohara 2018) is an extension of the Japanese FrameNet, currently developed mainly to be useful for linguists, but in the longer term intended as a resource for language education and NLP applications.Using the same software infrastructure originally devised for the lexicon, the Japanese Ccn is yet to be integrated into the Japanese FrameNet Lexicon at database level. However, the Japanese Ccn advances in the discussion of how frames and constructions relate to each other. Ohara distinguishes two types of frame-evoking constructions – those evoking semantic frames and those evoking interactional frames – and three types of non frame-evoking frames. Constructions evoking *semantic* frames are those whose

10 The reasons behind these absences and also some ideas on what features the English Ccn would have, if the project had moved forward, are discussed in Lee-Goldman and Petruck (2018).

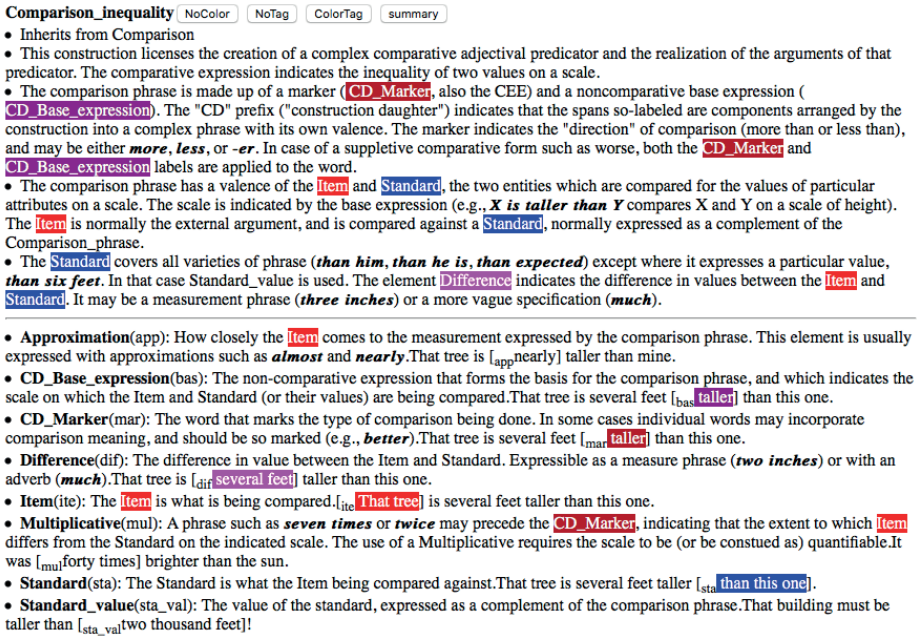

Figure 2: The `comparison_inequality` cxn in the English Ccn.

meaning can be accounted for in terms of framenet-like frames – such as the `comparative_inequality` cxn, which is illustrated in Figure 3.[11] As indicated in the Interpretation field of the entry, this cxn arguably evokes the `Comparison_inequality` frame.[12]

11 Throughout this paper, all non-English construction entries have have been translated into English for the sake of comprehension, either by the authors of the original work that is cited or by the authors of this paper. Likewise, we often use English names for cxns in other languages. Please note that this naming practice is not intended to indicate any presumptions of cross-linguistic identity, let alone universality, but is only used for convenience.

12 The `Comparative_inequality` frame does currently not exist in Berkeley FrameNet, although there is a superordinate `Evaluative_comparison` frame (for a discussion of comparative constructions and frames, see Hasegawa et al. 2010).

The `comparative_inequality` construction
{ [ENTITY1] [ENTITY2] [FEATURE] }

M	Clause
D1	NP. May include the phrase *no hoo* (Lit. 'the NP's side')
D2	NP, accompanied by a case marker *yori*
D3	a plain adjective
Interpretation	Evokes the `Comparative_inequality` frame that reports inequalities between two *Entities* as arguments of a plain adjective.

{ [ENTITY1 *kore* (*no hoo*) *ga*] [ENTITY2 = CEE *are yori*] [FEATURE *nagai*] }
this GEN side NOM that than long
'This is longer than that.'

Figure 3: The `comparative_inequality` cxn in Japanese. (Ohara, 2018: 157)

In addition to semantic, framenet-like frames Ohara also assumes *intectional* frames (cf. Fillmore 1982), to account for the meaning of constructions such as the `te_linkage` cxn. The interpretation of this construction reads as follows: "Two clausal conjuncts report two events and two events exhibit temporal sequentiality. The construction evokes the `*Relevancy` frame, in which the Speaker construes the two reported events to be somehow relevant." Given the types of constructions just listed, it is possible to state that, although the Japanese Ccn is inspired by the English Ccn, it moves away from it to the extent that it includes other possibilities for accounting for the meaning – or function, or *interpretation* – of constructions.

As for non frame-evoking frames, Ohara (2018) discusses three different types: those that are compositionally interpretable, those whose more elaborated constructions evoke frames, and those omitting repetitive position-specific constituents. For the constructions that are compositionally interpretable, e. g. the `modifier_head` cxn, this field will be filled by something like an NP, with a Modifier AP modifying a Head NP (Ohara, 2018: 152). There are also constructions, such as the auxiliary `V_te_iru` cxn, which does not evoke a frame in its more abstract version, but whose inherited constructions do evoke frames. As for the constructions omitting position-specific components, such as the `gapping` cxn, the interpretation proposed will be filled by information such as "each non-final conjunct is missing some material that is present in the former conjunct", being, therefore, more formal-oriented.[13]

13 Various types of frame-bearing and non frame-bearing cxns are also discussed in Lyngfelt, Bäckström et al. (2018, section 5), in relation to the Swedish constructicon. We will return to this discussion in section 4.2.

The last of the framenet-derived constructicons is both a descriptive resource for human users and a machine-readable database, used for constructional parsing (Matos et al. 2017) and machine translation (Tavares, 2018). It Brazilian Portuguese Ccn (Torrent et al. 2014; Torrent, Matos et al. 2018), which brings both theoretical and implementational changes to the original framework. On the theoretical side, the Brazilian Portuguese Ccn embraces the principle of continuity between grammar and the lexicon, meaning, among other things, that in FrameNet Brasil, frames are not part of the lexicon infrastructure, but are conceived as a semantic network which can be accessed by means of purely lexical or highly schematic constructions, in addition to all the different types in between these two – such as coining patterns, for instance. Given this premise, a series of implementation efforts have been made in the database structure so as to reflect those principles. As an example, consider the `comparison_inequality` cxn – Figure 4 – which is the Brazilian Portuguese equivalent to the English entry presented in Figure 2, and licenses sentences such as (1) and (2).[14]

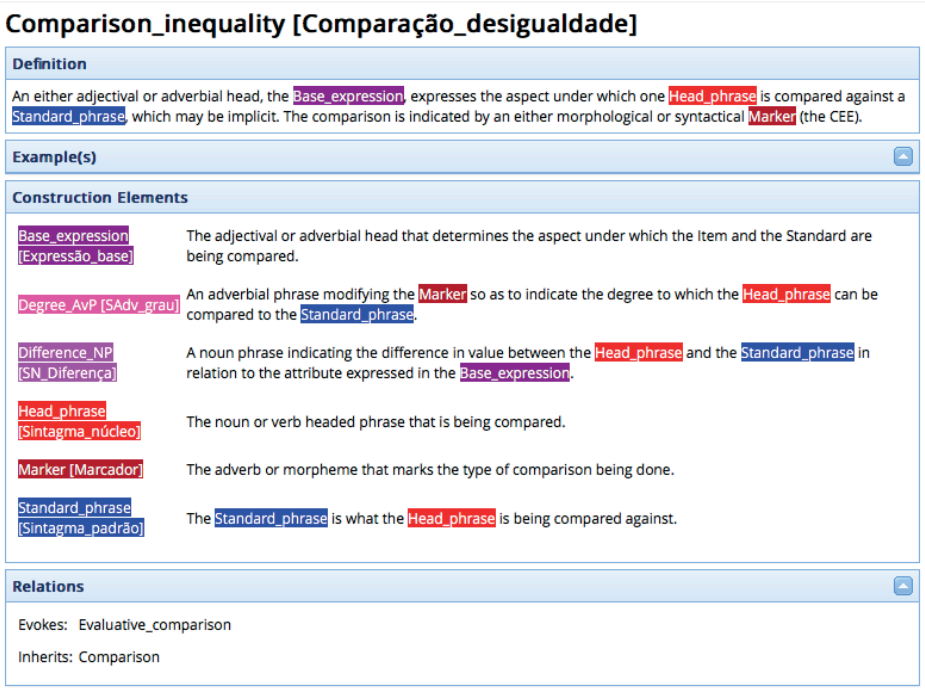

Figure 4: The `comparison_inequality` cxn entry in the Brazilian Portuguese Ccn.

14 For a discussion on the notion of constructional equivalence, see Lyngfelt, Torrent et al. (2018).

(1) [Aquele carro]$_{Head\_phrase}$ é [muito]$_{Degree}$ [mais]$_{Marker}$ [caro]$_{Base\_exp}$
That car is much more expensive
[que]$_{Marker}$ [o nosso]$_{Standard\_phrase}$
than the ours
That car is much more expensive than ours.

(2) Maria [corre]$_{Head\_phrase}$ [duas vezes]$_{Difference\_NP}$ [mais]$_{Marker}$
Maria runs two times more
[rápido]$_{Base\_expression}$ [que]$_{Marker}$ [eu]$_{Standard\_phrase}$
fast than I
Maria runs two times faster than I do.

At first look, the entries in Figures 2 and 4 look quite similar. However, both the analytical guidelines and the infrastructure behind them is quite different. First, note that the CEs in the Brazilian Portuguese entry are primarily defined by their formal properties and syntactic relations. This alone leads to a reduction in the number of CEs, compared to the English entry. Note that neither the Standard_value nor the Multiplicative CEs are present in the Brazilian Portuguese equivalent entry. Nevertheless, when there is a connection between a CE and some semantic import associated with it, this information is still stored in the resource. This is made possible because, differently from what happens in the English Ccn, in the Brazilian Portuguese resource, there are structured database level connections between the constructions and the frames they may evoke. For the `comparison_inequality` cxn, the frame evoked is `Evaluative_comparison`, and the mapping between CEs and FEs is shown in Figure 5.

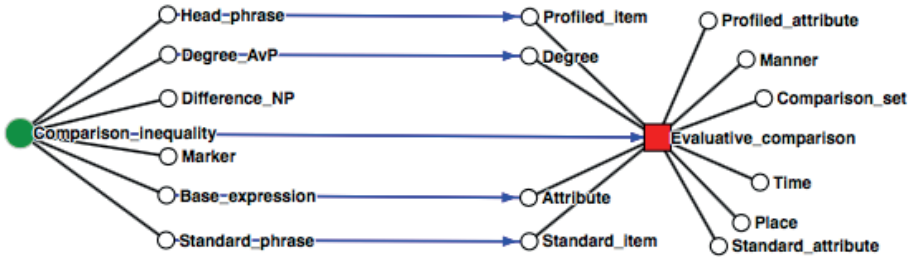

Figure 5: The evokes relation held between the `comparison_inequality` cxn and the `Evaluative_comparison` frame in FrameNet Brasil.

Second, there are also database-level connections between the `comparison_inequality` cxn and the `comparison` cxn it inherits. Because the internal structure of the mother construction – that is, the CEs in it – can be paired with the internal structure of the daughter, the connections go down to this level where applicable, see Figure 6. In the Brazilian Portuguese Ccn, inheritance is conceived as full inheritance (see Torrent, Matos et al. 2018 for discussion), meaning that relations of this type are posited when the daughter construction has all the information from the mother construction and more (Kay & Fillmore 1999).

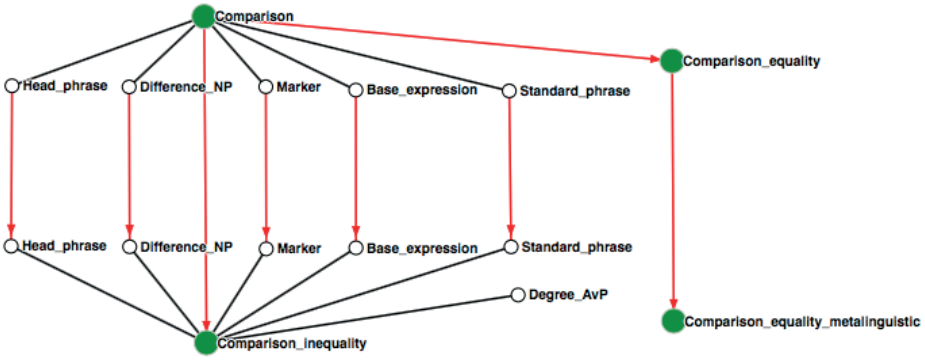

Figure 6: Inheritance relations between comparative constructions in FrameNet Brasil.

Finally, a set of (soft) constraints model the internal properties of the construction, both in terms of constituency and also in terms of other relevant restrictions applied to the CEs, as in Figure 7. For the `comparison_inquality` cxn, these constraints model: (a) the constructions licensing each CE, indicated by "cxn_"; (b) the relative order of the daughter signs, indicated by "bef_", when a given CE must antecede another, or "mee_", when a given CE must antecede and be adjacent to another; and (c) the lexical material filling in a CE, indicated by "lex_".[15]

15 For a discussion of all the possible constraints in the Brazilian Portuguese Ccn, see section 4.1.

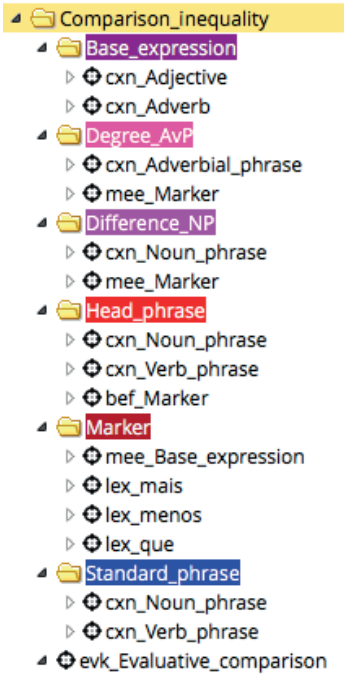

Figure 7: Soft constraints applied to the `comparison_inequality` cxn.

Although the Brazilian Portuguese Ccn implements a set of innovative relations and changes the role of the network of frames in a framenet from the background meaning structure evoked by lexical items to that recruited by any sort of linguistic material, it is still built as part of FrameNet Brasil, and is, therefore, framenet-based. In the next section we present resources that do refer to a framenet, but are not based on one.

3.2 The Middle Ground

The German and the Swedish Ccns stand in the middle of the framenet-relatedness continuum, but for different reasons. The German Ccn (Boas & Ziem 2018, Ziem et al. (this volume)), in its current stage, follows the path of the original Beyond the Core project (Fillmore 2008), by analyzing families of constructions, some of which would be the German equivalents to the constructions in the English Ccn.[16] Also, sim-

16 The German Ccn aims to implement a full grammar of German (see Boas & Ziem 2018), but, in its initial stage, it has been focusing on contrastive analyses that take the entries in the English Ccn as points of comparison.

ilarly to the entries in the English Ccn in which a frame is evoked, the CEs in the entries of the German Ccn may be defined primarily in terms of their semantics, with no database connection between the constructions and the frames in FrameNet.[17] This aspect makes the German Ccn very similar to the Berkeley project. However, the German Ccn includes a set of different kinds of information and analytical steps in its infrastructure that makes it less dependent on a framenet. First, aside from the construction definition, the CEs and Construction Evoking Element[18], already present in the English Ccn, the German Ccn also lists Correlated Elements, or CorEs, which are defined as "a word, or a string of words, that co-occurs with a construction in such a way that it enhances, or supplements, a (semantic, pragmatic, discourse-functional, syntactic) property of a construction" (Boas & Ziem 2018: 216). It also makes a distinction between internal and external CEs, which is not – at least consistently – made in the previous three resources described in section 3.1. Figure 8 depicts a summary of the `comparison_inequality (vergleich_ungleichheit)` cxn in the German Ccn and illustrates the basic structure of each entry.

Beyond the differences in structure, the methodology for building the German Ccn involves a pipeline (Boas & Ziem 2018: 216, 217) that incorporates software not present in the constructicons discussed in section 3.1 above. First, German corpora are surveyed for typical instances of the construction to be analyzed. From this initial set of instances, the properties of the construction (both formal and functional) are preliminarily determined, including the tentative CEs. Second, the examples are automatically parsed for parts of speech, phrase types and grammatical functions. Third, semantic annotation of the examples is conducted manually with WebAnno. Fourth, using a locally developed tool, the Construction Analyzer, the incremental output of the previous phases is converted into the sort of annotation made in the English Ccn. This tool also identifies syntactic realization patterns of constructions, CEs and CEEs. Finally, construction entries are compiled by analysts, based on the evaluation and interpretation of the results produced in the previous four phases.

The German Ccn is currently only available for research purposes, although different kinds of applications are envisioned further down the road.

17 Where applicable, construction entries in the German Ccn have, in the *Frames* field, links to the Berkeley FrameNet frame reports.

18 At least for some constructions, the German Ccn uses the CEE category differently from what was defined for the English Ccn, since instead of applying the CEE tag to an existing CE, in the German Ccn an element can be created solely as the CEE.

Vergleich_Ungleichheit

Definition

The construction Comparison_Inequality belongs to the comparison constructions and displays the inequality of two values on a scale.
The comparison expression consists of a MARKER that simultaneously functions as a CEE and a DIMENSION. Only these two CEs are licensed by the construction as internal construction elements. The MARKER, realized as the suffix -er attached to the DIMENSION, indicates the type of comparison (here: inequality and not equality).
The comparison expression also includes the external construction elements COMPARED and REFERENCE. These are the units that are compared on a scale with respect to certain values of specific attributes. The scale is indicated by the DIMENSION. The DIMENSION therefore indicates more precisely how the element in focus (COMPARED) exceeds an existing value (REFERENCE). The DIMENSION can either be realized as an adjective (in connection with a copula verb) describing a property or as an adverb (in connection with all other verbs) describing the exercise of an activity more precisely.

Type

Comparison

Construction Elements

Internal Construction Elements

MARKER	The MARKER indicates the type of comparison involved. At the same time, it functions as a CEE and expands the dimension into a comparative by adding the -er suffix.
DIMENSION	This mandatory CE names the common property whose value is referred to in the comparison. It can be either an adjective (associated with a copula verb) of a property, or an adverb (associated with all other verbs) of an activity.

External Construction Elements

APPROXIMATION	This optional CE is realized directly before the CEE or directly before SPECIFICATION, if available. It indicates that the value of the COMPARED on the scale does not yet completely exceed that of the REFERENCE_VALUE. The emphasis is not on equality, but on the approximate exceedance of it.
DIMENSION	This mandatory CE names the common property whose value is referred to in the comparison. It can be either an adjective (associated with a copula verb) of a property, or an adverb (associated with all other verbs) of an activity.
COMPARED	This mandatory CE is compared with the REFERENCE_VALUE in relation to a DIMENSION
REFERENCE_VALUE	This mandatory CE is a fixed value on a scale that is used for comparison.
SPECIFICATION	This optional CE is always realized before the CEE and indicates how far apart the values of the COMPARED and REFERENCE_VALUE CEs comparisons are. The specification can be realized either by an exact unit of measurement or by a more general indication of how clear, much etc..

Figure 8: The `comparison_inequality` cxn in the German Ccn.

As for the Swedish Ccn, differences to the resources described in 3.1 are more prominent. Although it does have pointers to the frames in Swedish FrameNet and also features entries whose CE names resemble FEs, the two resources are not fully integrated to the point that descriptive choices made for the Swedish Ccn impact the development of the Swedish FrameNet or vice versa, as is the case for the Brazilian Portuguese Ccn. Furthermore, the cxn entries in the Swedish Ccn contain several other types of information (cf. Lyngfelt, Bäckström et al. 2018: 82ff.). The entries may be displayed in either simple or extended mode, where the simple mode only presents the core features *Name*, *Definition*, *Structure sketch* and annotated *Examples*, as illustrated by the `comparison_inequality(jämförelse.olikhet)` cxn in Figure 9:

comparison_inequality – *Hon är smartare än de flesta* 'She is smarter than most'	
Definition	A [phenomenon]Theme differs from an [object of comparison]Corresponding with respect to a certain [property]Property. The object of comparison may be implicit.
Structure	[Adj|Adv$_{komp}$ (än^{2} XP)] | [*mer|mindre* Adj|Adv (än^{2} XP)] 'Adj|Adv$_{comp}$ (than XP)] | [*more|less* Adj|Adv (than XP)'
Examples	[Det där trädet]Theme är [[högre]Property [än]Subjn [det här]Corresponding]jämförelse.olikhet 'That tree is higher than this one'
	Jag uppfattade [henne]Theme som [[mer intelligent]Property [än]Subjn [jag själv]Corresponding]jämförelse.olikhet 'I considered her more intelligent than myself'
	[Det andra projektet]Theme fungerade [[mindre bra]Property]jämförelse.olikhet 'The other project worked less well'

Figure 9: The Swedish `comparison_inequality` cxn entry in simple display mode, transposed into English

The extended mode, on the other hand, includes up to 16 different fields of information (where applicable), including feature analyses of the CEs; information about lexically fixed elements (*Keywords*, cf. CEEs) and common slot fillers (*Common words*), both of which are linked to corresponding entries in several lexical resources; connections to frames and to corresponding English cxn entries;[19] and meta information about grammatical *Category* and *Type* of cxn, as well as inheritance relations. Types in the Swedish Ccn – as well as in the German Ccn – are sets of cxns united by some salient property that is shared (cf. Section 4 below). For instance, the example cxn in Figure 9 belongs to the type *jämförelse* 'comparison', along with other comparative cxns. The Swedish Ccn is intended as a multi-purpose resource – for linguistics, language education and language technology – and different features of information are relevant for different types of users and applications. A particular aim is to account for cxns of relevance for L2 acquisition of Swedish.

It is also worth pointing out the fact that, despite the differences highlighted above, the four constructicons somehow related to the original English Ccn elaborate further on the structure of the resource, adding and connecting different types of information. Note that, for instance, all of them invest in broader and deeper definitions of CEs. In the Japanese Ccn (Figure 3) daughter signs are sometimes named after their semantic import, but there are fields in the entry (D1, D2...) specifying their formal properties. In the Brazilian Portuguese Ccn (Figure 4), CEs are defined by their formal properties which are modeled in the database via soft constraints and relations between constructions and frames. In the German Ccn (Figure 8), the Constructional Analyzer summarizes syntactic properties of CEs, extracted from annotation. And in the Swedish Ccn, each CE may be defined for the following additional subfields:

19 The `jämförelse.olikhet` cxn is connected to both the `Evaluative_comparison` frame (by an active link) and to the English `comparison_inequality` cxn (only by reference).

- *Category*: the morphosyntactic description of the CE
- *LU*: the specification of a given LU in the lexical infrastructure of *Språkbanken*[20]
- *Role*: general semantic roles such as Agent, Undergoer, Result
- *Grammatical Function*: the grammatical function of the CE
- *Other*: any relevant property of the CE not captured by the other four features

As we will point out in Section 4.1, relating pieces of information in a constructicon is key for developing a coherent database. We now turn to the other extreme of the framenet-relatedness continuum: that inhabited by constructicons which were developed without any connection to a framenet, but are still somehow influenced by one.

3.3 Framenet-influenced Constructicons

The other end of the framenet-relatedness continuum is occupied by three different resources. The first two of them appear on the continuum because they refer to some analytical category related to Frame Semantics. The last one is on the continuum only because it shares software infrastructure with the Swedish Ccn, and, therefore, may have been influenced by modeling decisions made for this resource.

The first effort in this group of initiatives is the Embodied Construction Grammar (ECG) Analyzer (Bryant 2008). Developed to be a constructional parser of sentences, the ECG Analyzer makes use of a constructicon of English defined according to the theoretical background of ECG (Bergen & Chang 2013). This is to say that construction meanings are defined in terms of schemas, not frames. However, the proximity between frames and schemas is not coincidental. Moreover, recent work in ECG (Dodge et al. 2017) has made use of Berkeley FrameNet valence patterns as means to extract argument structure constructions from annotated corpora. Dodge et al. (2017) report on the extraction of the `cause_motion` argument structure cxn shown in Figure 10 from the valence patterns of the LUs in the `Cause_motion` frame in FrameNet.

In the representation, the meaning attribute of the construction has Cause_motion as its value. This stands for the Cause_motion schema in ECG, which, for this particular work, was automatically extracted from the `Cause_motion` frame in FrameNet. Dodge & Petruck (2014) demonstrate how FrameNet frames can be related to – and automatically transformed into – ECG schemas, highlighting the analogies between frames and schemas.

20 The Swedish Constructicon is integrated with other resources in the infrastructure of *Språkbanken* ('The Language Bank of Swedish'), including the Swedish FrameNet and other lexical resources to which the Keywords and Common words are linked (Lyngfelt, Bäckström et al. 2018: 42ff.).

```
construction Cause_motion_pattern_10
    subcase of ArgumentStructure

contructional
    constituents
        v: Verb
        pp-into: PP-Into
        np: NP

meaning: Cause_motion
    constraints
        self.m ⬚ v.m
        profiledParticipant ⬚ self.m.agent
        self.m.goal ⬚ pp-into.m
        self.m.theme ⬚ np.m
```

Figure 10: The `cause_motion_argument_structure` cxn in ECG

Also making use of Berkeley FrameNet valence patterns, Perek & Patten (2019) propose the development of a comprehensive constructicon for English from the COBUILD Grammar Patterns (Francis et al. 1996, 1998). The authors developed a semi-automated methodology that surveys Berkeley FrameNet data for verbal valence patterns and then matches them to the grammar patterns in the COBUILD verb entries, associating entries and frames. Afterwards, annotators manually correlate data from the entries which were not correctly identified with the missing frames. Another valency-based constructicon approach is being developed from the Erlangen Valency Patternbank (Herbst, this volume).

Finally, the Russian Ccn (Janda et al., 2018) aims to complement existing resources for Russian by documenting constructions which are not covered by other resources such as, for example, the Russian FrameBank (Lyashevskaya & Kashkin 2015). In the FrameBank resource, lexical items are described in terms of their valence and syntactic frames, together with dictionary-like definitions. However, according to Janda et al. (2018), important aspects of Russian grammar are not covered, both in terms of language documentation and pedagogy. The Russian Ccn is primarily designed to be useful for (foreign) language education.

The Russian Ccn uses the infrastructure developed for the Swedish Ccn, but does not include any reference to any framenet. Therefore, it's only marginally influenced by a framenet, since only the software apparatus used bears some relation to one of the other constructicons on the framenet-relatedness continuum.

In the following section, we discuss some of the implications of possibilities related to the relative position of each constructicon on the framenet-relatedness continuum.

3.4 Discussion

Given the different configuration possibilities presented for each of the resources discussed so far, the framenet-relatedness continuum can be drawn as shown in Figure 11. Note that there are two lines on the continuum: one that goes from the pole where the frame database is integrated with the constructicon to the one where only indirect references to some sort of framenet structure is made; and another that connects the Swedish and the Russian resources due to the fact that the latter shares the software infrastructure developed for the first. The Brazilian Portuguese and the Japanese constructicons have advanced in integrating the frame database into the constructicon beyond the point achieved by the original English resource. Therefore, the first are closer to the left extremity of the continuum than the latter.

Figure 11: The framenet-relatedness continuum

The degree to which a constructicon is related to a framenet may play an important role in shaping the way the analyses are carried out. Of course, the pre-existence of the extensive descriptive work needed to build a framenet goes a long way in helping constructicographers account for the semantic import of – non-lexical – constructions, specially because, as pointed out repeatedly in the literature, framenet annotation is constructionally inspired (Ruppenhoffer et al. 2016; Torrent & Ellsworth 2013). It also allows for the implementation of the foundational notion in CxG that – at least some – constructional meaning is defined in terms of frames.

Nevertheless, it may also bring some bias to the analyses. Because framenets were developed as lexical resources, parts of the model are intrinsically grounded in the way concepts are lexicalized in the target language. Not only distinctions in frame and frame element definitions can be lexicographically grounded; even frame to frame relations are used in framenets to account for lexical alternations, such as the Causative_of and Ichoative_of relations, for example. Therefore, if, on the one hand, constructicography can benefit from framenets, constructicon building efforts are a great opportunity for broadening the scope of frames and of the relations held between them in a framenet, since they provide an additional source of linguistic evidence, which includes but also goes beyond lexical items.

The idea is that, if constructions are brought into play as linguistic evidence supporting frame definitions and frame to frame relations, the resulting network of frames can become even more comprehensive, in the sense that language users can use different strategies to recruit the frames need for interpreting a given sentence. In a multilingual analysis setting, framenets that rely not only on lexical evidence can

become better sources of alignment between languages. In a preliminary report on the Multilingual FrameNet Shared Annotation Task, Torrent, Ellsworth et al. (2018) present and discuss the frame similarity metric calculated between translated sentence pairs that have been annotated for English and Brazilian Portuguese using the Berkeley FrameNet 1.7 Data Release. They demonstrate that differences in the part of speech each language uses to lexicalize a concept trigger chain reactions in the way frames are evoked across translations, although both sentences in the pair still create the background for the invocation of the same gist frames, that is, both sentences in the pair convey the same basic meaning.

Aside from the potential to contribute to the development of framenets covering a broader range of language phenomena, and the descriptive value and applied uses of the various ccn resources themselves (see above), it is worth stressing that constructicography, regardless of the relation of the resulting resource with a framenet, contributes to the development of the Construction Grammar community. This is so because, in an attempt to systematize the constructions in a language, practitioners in the field face the need to go beyond the study of compartmentalized families of constructions and integrate those families together into a larger resource. Constructicographers must make decisions on the basic principles to be followed in their resources and be consistent about their application. Hence, besides the commitment with constructions all the way down, meaning that fine-grained analyses of small sets of constructions still find room in constructicography, the field is also committed with constructions all the way up (cf. Lyngfelt 2018b), meaning that the more general language-property-like features of a grammar must also be addressed in the resource. These issues are further addressed in sections 4 and 5.

4 The structure of the constructicon

The standard theoretical conception of a constructicon is a network, typically depicted as an inheritance network (e. g. Fillmore & Kay 1993, Sag 2012, Goldberg 2013, Hilpert 2014, Lyngfelt 2018a). Although the internal structure of this network is still highly understudied, the basic notion as such is assumed generally across construction grammars. In constructicography, however, the structure of a descriptive constructicon resource is a somewhat different issue; and the CxG idea of a network may not be the best approach for a constructicon database, due to limited coverage and depending on the purpose of the resource; different approaches may be more or less well suited for different applications.

Looking at the existing constructicon developments, there are three basic types of structuring principles currently employed: (1) list structures, (2) relational networks, and (3) categories/types. These are by no means mutually exclusive. Not only are they often combined in various ways, but there is also some conceptual overlap between networks and categories.

List structures in this context refer exclusively to alphabetically ordered lists (as opposed to lists ordered by, e. g., usage frequency or search frequency). This is the ordering principle most similar to traditional lexicography and also how the current ccn projects started out. It remains the basic structure of the English and the Russian ccns and is the default format of presentation in the ccns for Brazilian Portuguese and Swedish. Although there is clearly some structural hierarchy behind cxn labels such as `adjective_as_nominal.abstract`, `adjective_as_nominal.anaphoric` and `adjective_as_nominal.people`, these cxns are not otherwise grouped in any way, but listed alphabetically in the same straight list as all other cxns in the English ccn (http://www1.icsi.berkeley.edu/~hsato/cxn00/21colorTag/index.html).

While simple and intuitive, alphabetical lists are not as well suited for constructicography as for lexicography, mainly for two reasons: Firstly, it is not always obvious what to call a cxn and hence not where to place or find it in the list. Even very well-known cxns may be known under different names, such as the `comparative_correlative` or *`the_X-er_the_Y-er`* cxn. Secondly, the level of generalization is not *a priori* given. How would a user know whether to look for, say, a `conditional_clause` cxn or a `subordinate_clause` cxn? For such reasons, as well as an aim to follow CxG theory and also test its assumptions empirically (see below), many ccn projects are under development from mere lists of cxn descriptions to more elaborate structures.

Relational networks correspond to the theoretical notion of a constructicon as a network. They are employed in framenets and to some extent in current constructicon initiatives, especially as regards inheritance relations. Thus, specific cxns are modeled to inherit properties from more general ones, forming a network of inheritance relations. This is a quite dynamic structuring principle, in the sense that one can take any point in the network as the point of departure, and for that particular cxn see both which cxns it inherits from and which cxns inherit from it (see Figure 6).

Uniformly, these networks assume multiple inheritance, so that `infinitival_relative` cxns may inherit properties from both `relative_clause` cxns and `infinitival_clause` cxns (or `infinitival_verb_phrase` cxns, depending on the analysis). Some models, e. g., the Brazilian Portuguese ccn, assume full inheritance, where "When one construction inherits another, the first contains all the information of the second and – in the nonvacuous case – more" (Kay & Fillmore 1999: 7). Others, like the Swedish ccn, allow default inheritance (or common inheritance), which means that all properties are inherited *unless overridden by specifications in the inheriting cxn* (Lyngfelt, Bäckström et al., 2018: 98 f.). For instance, the Swedish counterparts to `adjective_as_nominal` cxns are treated as `noun_phrase` cxns, thus inheriting general NP properties – with the notable exception of not containing a nominal head.

In addition to inheritance, other relations between cxns may also be posited, notably horizontal links, (also called subpart links, Goldberg 1995: 78 f., Hilpert 2014: 62 f.), which hold between cxns containing the same kind of element, e. g. a `reflexive` or a `gerund` CE. In principle, such relations may also be construed in terms of inheritance, since they concern cxns inheriting a given property from the same cxn. We will return to such relations in section 4.1 below.

From a theoretical perspective, a network structure is attractive not only in that it conforms to CxG theory but also because it serves as a test of essentially untested assumptions. When you need to actually propose a network of cxns defined according to a set of systematic features, and then incorporate new (sets of) cxns into this network, you need to (a) define constructional properties consistently across very different types of cxns, and (b) establish and define points/cxns in the network through which other cxns are connected, i. e., account for relationally central cxns in a way that holds across the whole network. This means a practical test of the language-as-a-network-of-constructions idea as such, and also pushes towards coverage of previously neglected areas. Notably, the traditional constructionist focus on the so-called periphery of language has, somewhat ironically, led to a tendency to leave other areas of language to the periphery of CxG theory. When constructing a full network, however, you can no longer pick and choose what cxns to work on, but end up dealing with precisely those issues that tend to be neglected in CxG. One may of course argue that ccn resources should instead stick to the focus on periphery tradition, thus filling a descriptive gap by covering patterns that are neglected in other kinds of linguistic resources, and to some extent this is what current ccns do. Nevertheless, a theoretical approach that claims relevance for language as a whole, making assumptions regarding the whole alleged grammar-lexicon continuum, eventually has to live up to those claims, and developing a ccn resource into a network of cxns is one way of working these things out.

While theoretically attractive, relational networks can only partially cover the structure of the current ccns, mainly because none of the databases has reached sufficient empirical coverage to be construed as a (global) network. Groups of cxns may form limited, local networks, but, as of yet, they are only islands in the motley collection of disparate and still mostly unconnected cxns.

Categories, or **types**, are thematic sets of cxns grouped according to some shared property (or set of properties). Such properties may concern form as well as function, and particular cxn elements as well as the cxn as a whole. For instance, `resultative` cxns all have a resultative meaning component, `reflexive` cxns include a reflexive morpheme, `compound` cxns share a certain kind of word structure, etc. Assuming multiple categorization, a `reflexive_resultative` cxn may be categorized as both resultative and reflexive, as well as an ASC (argument structure cxn) and possibly a `verb_phrase` cxn (if you consider ASCs to be phrasal, cf. Boas 2014b for discussion). All ccns assume basic grammatical categories such as phrase types, and particularly the Swedish ccn includes a type system consisting of a wide variety

of categories.[21] In a ccn linked to a framenet, frames may also function as a kind of category by gathering all cxns that evoke the same frame or family of frames, at the same time relating them to lexical units evoking the same frame.

Categories/types are fixed foci in the structure, as opposed to the more fluent structure of relational networks, in which any given cxn may be taken as the point of departure. From a user's perspective, categories provide ready-made groupings, whereas networks seem more versatile; although the inheritance relations themselves are fixed, they may be viewed from any point in the network. In a system of multiple categorization, however, where the same cxn may belong to several different types, users can form their own categories by combining types. Furthermore, categories and inheritance relations often capture the same set of cxns: Whenever a category also corresponds to a cxn, which is often the case, all cxns belonging to that category may also be regarded as inheriting properties from the corresponding superordinate cxn. On the other hand, there are also categories that consist of sets of cxns without a common superordinate cxn. For example, the Swedish ccn includes a category (or *type*) for time expressions, i. e. various cxns expressing temporal relations. It is a relevant category from the perspective of L2 education, since the diverse set of idiomatic time expressions in Swedish is difficult to master for an L2 learner, but there is no superordinate time cxn of which these cxns are subtypes. Another example is the set of polarity-sensitive cxns. Consequently, while inheritance relations and categories/types sometimes yield the same groupings, each of the systems also capture structural aspects not easily covered by the other.

4.1 Relations within and between constructions, and beyond

When building a constructicon as a resource covering the full grammar of a language, constructicographers will be confronted with the need to posit relations within and between constructions from the very beginning. Because constructicons are implemented as databases, even the definition of CEs, that is, the basic constituency of constructions, will call for some sort of modeling. That is to say, for instance, that, if a CE is a NP, at some point the constructicon will include a NP construction that will be linked (or at least referred to) by that CE. The Japanese and Swedish ccns refer to constructions licensing the daughter signs of other constructions in specific fields (such as the D1...DX fields in the Japanese Ccn and the category field

21 A constructional *type* in this system may be defined by "any salient property shared by a group of constructions" (Lyngfelt, Bäckström et al. 2018: 58). Types may thus be characterized by functional properties (causative, comparison), specify overall formal structure (compound, coordination), concern particular construction elements (expletive, reflexive), etc.

in the Swedish Ccn). The Brazilian Portuguese Ccn and the ECG Analyzer, in turn, implement this analytical aspect in the form of constraints applied to the CEs – see Figures 7 and 10.

Besides accounting for constructional constituency, other constraints and relations can be implemented in a constructicon so as to both enrich the database structure and also enhance overall analytical coherence in the resource. In the formal pole, constraints can be applied to CEs so as to account for:

- the relative order of the constituents: the Before and Meets constraints used in both the Brazilian Portuguese Ccn (see Figure 7) and in the ECG Analyzer determine the order of the CEs and also whether there might or not be some sort of intervening material between them;
- the specification of the lexical material filling a slot: the Brazilian Portuguese Ccn uses this constraint to implement the notion of CEE at the database level, specifying, for example, the lexical manifestation of the Marker CE in the `comparion_inequality` cxn (Figure 7);
- restrictions on the inflectional properties of the CEs: also in the Brazilian Portuguese Ccn, this constraint can be applied to CEs (or to their heads) to determine that a given item must be either singular or plural. Tavares (2018), in modeling the `indefinite_noun_quantification` cxns in Brazilian Portuguese, demonstrates that, for constructs denoting small quantity, such as the one in (3), the noun being quantified must be in the singular form.

(3) [pitada]$_{\text{Noun\_1}}$ [de inveja]$_{\text{de\_Prep\_Noun\_2}}$
Pinch of envy
a pinch of envy

On the meaning pole, relations and/or constraints can be used to model frame and schema evocation, and semantically driven slot filling restrictions and correlations. Both the Brazilian Portuguese Ccn and the ECG Analyzer have relations connecting constructions and their constituents to frames/schemas and their elements, respectively. In the former, this is implemented via the evokes relation shown in Figure 5, while in the latter it is modeled via the meaning keyword in Figure 10. Nonetheless, frame/schema evocation is not the only meaning aspect of constructions that may be accounted for in a constructicon. Many constructionist analyses define semantic slot filling restrictions for CEs. In Brazilian Portuguese, for example, each of the `indefinite_noun_quantification` constructions specifies constraints for the nouns filling in the Noun_1 slot. Those restrictions can be modeled either in terms of lists of lexical items, or, if the restriction is more general, in terms of frames or families of frames whose LUs can figure in the Noun_1 slot. Moreover, as it is usually the case, when a given slot filling constraint is posited for one of the CEs, a correlated constraint should also be posited for another CE. For example, when the Noun_1 slot is filled with the lexical item *poço.n 'well'*, the nominal head of the de_Prep_Noun_2 CE

must be one of the nominal LUs in the Mental_property frame, licensing constructs such as (4) (Tavares 2018).

(4) [poço]$_{Noun\_1}$ [de nervos]$_{de\_Prep\_Noun\_2}$
well of nerves
nervous wreck

Clearly, modeling all those aspects of constructions has a considerable impact on the amount of work involved, as well as on the time and effort needed to build a constructicon. To alleviate this issue, inheritance relations (see Figure 6 in section 3.1) can be posited so that more general properties of the constructions can be assigned to the root nodes of families of constructions, while the bottom nodes are described for the more specific aspects of each subconstruction. Modeling inheritance relations in a constructicon also reinforces analytical coherence, and it allows for an implementation of the commitment to constructions all the way down, as well as all the way up. For the `indefinite_noun_quantification` constructions, for instance, the information that all these constructions evoke the Quantified_mass frame, as well as the constructional type of the CEs and their order, is modeled only once in the mother construction, from which these properties are inherited by the 14 daughter constructions (Tavares 2018).

4.2 Crosslinguistic considerations

Multilingual constructicography involves all the intricacies of multilingual lexicography (cf. Adamska-Sałaciak 2010: 387), and then some. In addition to the well-known problems with establishing functional equivalents in the source and target languages – where full equivalence in all aspects is rare if not impossible (e. g. Farø 2004) – multilingual constructicography also has to consider formal differences between the corresponding cxns (Bäckström, Lyngfelt & Sköldberg 2014; Lyngfelt, Torrent et al. 2018).[22] Some languages employ a question particle to express what others mark by word order and/or prosody, case morphology in one language corresponds to prepositions or word order restrictions in another, etc. Hence, the mapping does not always concern entities of the same basic category; and even when it does, the relevant properties may be distributed differently. For example, although both

22 Hence, the problem of mapping cxns between languages concerns not only linking but also representation, i. e. how to represent the structure of the cxns in a format applicable to both the source and the target languages. Since linguistic categories are not consistent across languages, developing such a metalanguage is a challenge indeed. We will not, however, go into the representation problem here, instead focusing on the linking issues. For a discussion, see Lyngfelt, Torrent et al. (2018: 280–287).

Swedish and Spanish motion cxns are typically expressed by verb phrases, most Swedish cxns encode manner of motion on the verb and path by adverbials, whereas their Spanish counterparts rather specify path by the verb and manner by adverbials (e. g. Talmy 2000). In many prospective applications of multilingual constructicography, such as translation or language education, such differences need to be handled.

Constructicography is also more sensitive to issues of granularity and categorization, partly due to variation between languages – regarding form, function or both – partly because there are so many different ways to conceptualize abstract patterns. Even where the languages are fairly similar, the cxns may be grouped differently due to different editorial decisions in different ccn projects, especially since the inevitable lumper-splitter issues concern both formal and functional distinctions and the same set of distinctions may yield different taxonomies depending on the relative ordering of formal and functional properties. In an electronic database, such differences may be handled by cross-categorization, at least to some extent, but the fundamental complexity of constructional categorization remains a source of discrepancies.

Nevertheless, pilot studies comparing English construction entries to Swedish and Brazilian Portuguese, respectively, show quite promising results (Bäckström et al. 2014, Laviola 2015, Lyngfelt, Torrent et al. 2018). These studies were able to establish close counterparts to all but a few of the English constructions in both Swedish and Brazilian Portuguese, although with varying degrees of functional equivalence and formal similarity (cf. also Boas & Ziem 2018). The comparisons were unidirectional, meaning, on the one hand, that they were based on the categorization for English and, on the other hand, that they only investigated mappings from English to the two target languages and not to what extent the same relations would hold in the opposite direction. For example, the closest Swedish equivalent to the English `let_alone` cxn is `för_att_inte_tala_om` (lit. 'for to not speak of'), but the closest English equivalent to the latter cxn is probably *not to mention*. To establish accurate links one would have to make comparisons in both directions between all the languages involved, and since even unidirectional comparisons are very time-consuming, a large-scale multilingual resource based on such direct comparisons between cxns does not seem to be realistically feasible. Therefore, a different linking approach is required, preferably a system that

- can link constructions by groupings as well as directly
- is (sufficiently) language-neutral
- offers more than one linking strategy.

An obvious possibility for framenet-related constructicons is to connect cxns across languages via frames, an approach that has been fairly successful in multilingual lexicography (e. g. Boas 2009). It has long been common practice in CxG to represent (at least some aspects of) constructional meaning in terms of Frame Semantics, and, as mentioned in section 3.1 above, many ccn entries are already linked to existing

FrameNet frames. Thus, to the extent that semantically related cxns in different languages can be linked to the same semantic frames, they can thereby be linked to each other. Such a connection requires (a) a meaning similarity that may be characterized in terms of a frame, (b) a frame entry capturing this meaning, and (c) that the same frame definition is employed for the languages in question.

However, as also noted in Section 3.1, while some cxns may be considered *frame-bearing*, analogously to frame-evoking words (Fillmore et al. 2012: 325), others have a less straightforward relations to frames, and still others can hardly be ascribed any frame-like meaning at all (cf. Lyngfelt, Bäckström et al. 2018, Ohara 2018). For example, while a *ditransitive* cxn arguably evokes the Giving frame (or some more specific frame inheriting from it), it is hard to envision any frame being evoked by a *gapping* cxn (as in *Some would take them to the top half of the vale, and others the bottom half*, Fillmore et al. 2012: 327). In between, there are various cxns being more or less related to frames in different ways, such as *imperative* cxns, which, while obviously related to the Request frame, hardly evoke it the same way words like *order* and *command* do; rather than referring to requests, like these words do, *imperative* cxns are used to perform them.

There are also cases where the meaning may well be characterized in terms of a frame, but no such frame description yet exists. This is a matter of coverage, on the one hand; and a question of what a frame is (or should be), on the other. Hence, in addition to the 'cognitive' or 'semantic' frames that current framenets predominantly consist of, Fillmore (1982) and Ohara (2018) propose 'interactional' frames, "having to do with how we conceptualize what is going on between the speaker and the hearer, or between the author and the reader" (Fillmore 1982: 117), as illustrated by the Japanese *te_linkage* cxn in Section 3.1.

In summary, some cxns may be linked via frames and others may not, whether due to properties of the cxns or due to a lack of appropriate frame entries. Furthermore, the frame relation is limited not only to the meaning/function of the cxns but also to certain aspects of the constructional meaning/function. It is therefore desirable to also develop other means to connect cxns in different languages. One possibility is to use *comparative concepts* of the kind employed in language typology (e. g. Haspelmath 2010, Croft 2016). These are theoretical constructs designed specifically for the purpose of comparing linguistic structures across languages, without any presumption of cross-linguistic descriptive categories. Hence, CCs should be well suited for connecting cxns across languages. Since typological comparisons are usually concerned with quite general linguistic properties, existing CCs typically correspond to more general patterns than the highly specific cxns that constitute the majority of the entries in the current constructicon resources. Thus, they represent features shared by groups of cxns, presumably inherited from more general cxns, which also means that, by multiple inheritance, the same cxn may relate to more than one CC.

As a first exploration of this possibility, a pilot study (Lelie 2019) was conducted to compare the set of construction types and grammatical categories in the Swedish

constructicon to an extensive set of CCs defined in Croft (2018).[23] Since the types and categories are defined by general properties shared by groups of cxns, they are conceptually somewhat similar to CCs, except for being language specific, and they conveniently represent a large number of cxn entries. Lelie found that most of the Swedish types and categories match fairly well with Croft's CCs, the categories less so than the types, which is not surprising given that grammatical categories vary across languages. Non-matches were mostly due to lack of coverage in the source material, and a few partial mismatches are explained by specific properties of Swedish. For example, the Swedish type 'passive' includes `impersonal_passive` cxns, which are not covered by Croft's language-neutral passive CC. Otherwise, matching problems had less to do with language specific features and more to do with the Swedish type system still being under development. In general, the results from Lelie's (2019) pilot study are promising.

In combination, frames and comparative concepts (and possibly other features) seem to provide a versatile linking system for a multilingual constructicon infrastructure. For some purposes, meaning correspondences (frames) are more relevant, for others, morphosyntactic properties (CCs). These linking strategies may also be combined, so that cxns that match in terms of both frame and CC and/or more than one CC are more closely connected than pairings based on only one of these features. In fact, some of the CCs in Croft (2018) are semantic concepts, basically equivalent to semantic frames. How well such a cross-lingual linking system would work in actual practice, however, remains to be tested.[24]

5 Other considerations

As discussed in Section 2 above, constructional research during the 1980 s and 1990 s was mainly concerned with providing in-depth analyses of specific types of constructions, showing that a theory of language should not only focus on a particular area of language, but instead on the entirety of language. At the same time, these constructional analyses were framed within the larger framework of a theory of CxG, i. e. they also paid attention to broader theoretical questions surrounding concepts, including division between the lexicon and syntax, motivation, the status of constructional networks, the connection between form and meaning, the role of metaphor and metonymy, and the interplay of different types of linguistic information. Goldberg's (1995) seminal work on argument structure constructions is a prime example of such con-

23 We are grateful to Bill Croft for letting us (specifically Ben Lyngfelt) use his unpublished manuscript.

24 A computational infrastructure for such a multilingual linking system is under development by Matos (in prep.).

structional research developing a theory of language based on constructionist concepts and methodologies.

But it was not until the Berkeley FrameNet's "Beyond the Core" project (Fillmore 2008) of building a prototype constructicon that constructionists really thought about a coherent approach towards building a constructicon for a broad variety of different types of constructions. In a way, the constructicon building effort can be seen as an application of CxG to the in-depth description (and analysis) of the constructions of a language, in this case English. In other words: the theory and principles of CxG provided the foundation for in-depth descriptions of constructions. This early effort systematically studied how different types of constructions could be described with a uniform corpus-based methodology and formalism that resulted in the formulation of specific construction entries for a wide range of different types of constructions and not just argument structure constructions or other types of partially filled (semi-) idiomatic constructions. Thus, the early constructicon-building efforts undertaken by the FrameNet project can be seen as laying the foundation for what is now known as constructicography. At the same time, however, the constructicon-building workflow only focuses on identifying and describing different types of constructions, eventually resulting in construction entries listed in the constructicon in alphabetical order. It is important to point out that beyond in-depth descriptions of different types of constructions, the Berkeley prototype constructicon does not make any claims regarding broader principles underlying the organization of a constructicon or how different types of constructions interact with each other in order to license particular utterances.[25]

What now is the relationship between CxG and constructicography? A comparison of the efforts of constructional researchers from the 1980 s to the 2000 s to establish an alternative theory of language with the efforts of constructional researchers to implement constructional principles to develop a constructicon post 2008 can perhaps be characterized as follows: Originally, constructional research inspired constructicography and now constructicography can also inform constructional research. For example, as research focuses more and more on a particular group of interrelated constructions, such as the resultative construction (Goldberg 1995, Boas 2003, Goldberg & Jackendoff 2004, Boas 2005b) or the subject-auxiliary-inversion construction (Fillmore 1999, Goldberg 2006), constructicographers arrive at a greater number of construction entries for constructions that are somehow related in terms of their form and meaning. In the original conception of the Berkeley constructicon, these constructions would only appear in an ordered alphabetical list. But keeping a broader constructional approach in mind that also seeks to inform the overall theory of CxG, constructicographers are now in a position to provide detailed information regarding the different types of relations between constructions that share similarities. The

25 On the different approaches see the chapters in Lyngelt, Borin et al. (2018) and Boas (2019).

insights resulting from these case studies can then provide, for example, insights into the nature of different types of constructional networks, thereby contributing to the overall theory of CxG. In other words: extensive constructicographic research is in essence applied research in CxG, eventually feeding directly into our overall understanding of how to account for a language with CxG. This mutual beneficial relation is similar to the preferred interplay between lexicography and linguistic theory.

The exchange of ideas between constructional researchers interested in a broader theory of CxG and those concerned with building a constructicon can follow multiple formats. For example, when constructicographers first formulate a construction entry, they are often constrained by time and resource limitations, as well as concerns for user-friendliness and applicability, which means that a construction entry may represent a compromise between coverage and in-depth description of a construction (similar to a dictionary entry that is often not at the level of detail that we perhaps would like it to be). This means that such construction entries should only be regarded as approximations that may serve as a starting point for further research by constructional researchers interested in examining such constructions in more detail. These further investigations then allow constructional researchers to refine construction entries and to determine how individual constructions are related to other constructions, how they interact with other constructions, and how they may serve to license utterances in different configurations. The insights from this research can then be used to refine construction entries. This iterative process is potentially open-ended since constructional researchers may always come across new data that may be used to further augment and refine existing construction entries.[26]

Once constructicons get further developed towards cxn networks rather than lists of cxn entries, the distinction between theoretical CxG and applied constructicography gets less clear-cut. Developing a coherent network structure not only requires compatible cxn descriptions, and thus a consistent treatment of categories and descriptive tools, it also requires an articulate model of how cxns interact. The network structure as such does not necessarily have to amount to one single network – it may well consist of a set of smaller networks – but these would still need to fit together, indirectly by their instantiating constructs if not by direct network connections. Sooner or later, one has to go beyond cxn descriptions and take on the question of how constructs licensed by cxns may be combined into larger structures. In addition to basic unification, one then needs a way to handle coercion phenomena,

26 There are at least two other ways in which research in CxG and constructicography may mutually inform each other. First, constructicographers start at zero and conduct full-text annotation of a corpus to determine which construction entries are needed to license all sentences in a corpus. This exhaustive methodology provides a valuable empirical basis for research in CxG since one needs to determine how constructions interact in order to license the multitude of sentences. Second, constructicographers base their formulation of construction entries on existing research in CxG together with corpus evidence. See Boas (in press) for more details.

by mechanisms sufficiently restricted not to over-generate, reliable mechanisms to treat metaphor and metonymy, etc. In short, this means building a coherent construction *grammar*.

Such an undertaking goes well beyond applying CxG research to descriptive practice, since the questions at hand remain vastly understudied within CxG. Even back in 1988, Fillmore presented a two-fold CxG conception of language: "a repertory of constructions, *plus a set of principles which govern the nesting and superimposition of constructions into or upon one another*" (Fillmore 1988: 37, emphasis added). Yet, most work in CxG has not ventured beyond the repertory part, even though a comprehensive repertory cannot be attained without a working model of the combinatory mechanisms as well. Consequently, developing a constructicon resource into a coherent network structure means breaking new ground in CxG research. The application will then not only serve as a practical test of the theory, but also entail substantial elaboration of the theory itself. Can such a goal really be combined with the ambition to develop constructicons as user-friendly resources made up of simple cxn entries? We believe it can – and should. A coherent structure is a requirement of any comprehensive database, and all of the internal structure does not have to be displayed in the user interface, as is also the case for lexicographic resources (see L'Homme and Cormier 2014 for discussion).

The usefulness and applications of constructicons go beyond the – yet very relevant – purpose of systematizing constructional analysis so as to allow for continuity and reproducibility in the field, as well as for building the basis for an empirical research program focused on investigating grammar, as pointed out before. On the one hand, if one focuses on human users, constructicons have the potential to support language pedagogy, as they can provide coherently organized information about language structures, their functions and, if the resource is somehow aligned with another constructicon, how they relate to structures in another language.

On the other hand, if machine users are the focus, constructicons can serve as language models, as lexicons have been doing for the past two decades. Because the kinds of structures in a constructicon tend to be more complex than those in a lexical resource, the kind of information modeled in a constructicon is even harder to derive from purely raw-data based machine learning techniques, such as Embeddings, Hidden Markov and Vector Space Models. As an example, consider the pair of Brazilian Portuguese sentences in (5–6):

(5) $[\text{Maria}]_{\text{Subj}}$ $[[\text{quebrou}]_{\text{V}}$ $[\text{as nozes}]_{\text{DObj}}]_{\text{Pred}}$
Maria break.PST.3SG the walnuts
Maria cracked the walnuts open.

(6) $[\text{O celular}]_{\text{Subj}}$ $[[\text{quebrou}]_{\text{V}}$ $[\text{a tela}]_{\text{DObj}}]_{\text{Pred}}$
The cell phone break.PST.3SG the screen
The screen on my cell phone broke.

Note that, as pointed out by Almeida (2016), although both sentences present the same [Subj [V [DObj]]] structure, their meaning is quite different. The construct in (5) is licensed by the `active_direct_transitive` cxn, in which the Subject CE is paired with the Agent FE in the `Transitive_action` frame, and the Direct_object CE is paired with the Patient FE. The one in (6), in turn, has both the Subject and the Direct_object CEs paired with the Entity FE in the `Undergoing` frame, since both the phone – the whole – and the screen – the profiled part – were affected by the event of breaking. Sentences like (6) are licensed by the `split_argument` cxn in Brazilian Portuguese, which is characterized by the evocation of the `Undergoing` and the `Part_whole` frames, despite the fact that its formal pole is identical to that of the `active_direct_transitive` cxn.

A purely data-driven approach would certainly fail to recognize instances of the `split_argument` cxn at corpora, since the very high frequency of constructs licensed by the `active_direct_transitive` cxn and the coincidental form of both constructions in the phrase structure level – which is usually the analytical level used by automatic parsers – would preclude the system from "seeing"the difference between the two patterns. However, if there is an explicit machine readable model that can be fed into the machine learning process, and if this model coherently states the whole-part relation held between the Subject and the Direct Object, then, model-based learning techniques (Winn et al. forthcoming) could be used to identify constructs licensed by the `split_argument` cxn in corpora. Taking this reasoning one step further and into the multilingual setting, (partially) aligned constructicons can be used for machine translation either as a source of parametrized learning corpora or as a repository of rules to be used in a post-editing system, in a way similar, although more complex, to how computational lexicons are used for terminology injection (Arcan et al. 2017), for example.

6 Conclusions

In this paper, we discussed how Constructicography and Construction Grammar can mutually inform each other, and how they (may) relate to FrameNet in various way. To do that, we have:

(a) reviewed the common history shared by Frame Semantics and Construction Grammar, both based on Fillmore's earlier research on Case Grammar;
(b) presented current constructicon development efforts, classifying them along a framenet-relatedness continuum;
(c) discussed alternatives for structuring constructicons and their implications to the application of the resulting resources in human- and machine-oriented tasks.

In light of these discussions we may address the pros and cons of integrating a constructicon with a framenet. While there is a long tradition of combining CxG with Frame Semantics, it is by no means necessary to do so. Any ccn requires some way to represent meaning, and sooner or later the ccn will have to be integrated with, or include, some kind of lexicon; but neither of them really needs to be based on FrameNet. One major reason for the existing connections is simply that the first constructicon was developed at the same place and by more or less the same research group as the original FrameNet. Also, not only in the English case but also for the majority of the ccns, there was a framenet in place before a ccn project was initiated and, hence, making use of existing structures instead of starting from zero had many practical benefits. A more principled benefit of the common "heritage" is that both resources derive from the same school of thought and that the respective theories they are based on were developed in tandem and designed to be compatible.

At the same time, however useful, the framenet influence also restricts the format of the cxn descriptions, or at least adds a lexically oriented bias. These restrictions, or biases, are perhaps not that severe, since construction grammarians have chosen to represent constructional meaning in terms of Frame Semantics long before there was a framenet. Still, a framenet structure is not designed to handle relations beyond those between a head (or the cxn as a whole) and its direct arguments. Therefore, additional features are needed to capture a more complex structural hierarchy, such as the relation between a filler and a gap in a long-distance dependency relation (see Sag 2010 on filler-gap cxns).

Furthermore, the value of connections to a framenet depends on the basic purpose of the constructicon. In the case of the English, Japanese and Brazilian Portuguese ccns, there is no question of whether, or why, to connect to a framenet; the frame-based lexicon and ccn are treated as two integral parts of the same effort. In other cases, the framenet is more of an external resource and the ccn developers may choose to what extent, if any, integration with a framenet is desirable. The Swedish ccn, for example, is connected with several lexical resources in the LT infrastructure of Språkbanken, the Swedish framenet being only one among many, and the pivot resource connecting all of them is another lexicon, called SALDO (Lyngfelt, Borin et al. 2018). In the Swedish case, the ccn was deliberately made structurally independent from the framenet, in order to avoid the cxn analyses being restricted by framenet considerations. Nevertheless, many cxn entries are linked to frames, in fact more than half of them, but only when this is considered relevant from the point of view of the ccn.

One major benefit of connecting a ccn to a framenet is that one thereby will also establish ties to ccns for other languages, partly because these usually have a framenet-oriented design, partly for the possibility of using frames to relate cxns across languages. This introduces something of a dilemma. On the one hand, the lexicalist bias of a frame-related ccn structure may be reduced if the influence between framenets and ccns is bidirectional and not just a one-way dependency. That way, the

framenet can be adapted to better suit the needs of the ccn. On the other hand, the potential of using frames as a tool for connecting cxns across languages depends on a fixed set of frames, which means that adapting framenets to language-specific constructional properties would reduce the cross-linguistic applicability of the frames. It should, however, be possible to get the best of both worlds. Adapting language-specific framenets to corresponding ccns does not necessarily preclude using (most likely the English) FrameNet as a pivot resource in a cross-linguistic infrastructure, as long as there is a robust system for relating the language-specific framenets to this pivot. Fortunately, there is an ongoing collaboration between different research groups to develop a multilingual framenet infrastructure to this very effect (Gilardi & Baker 2018, Torrent, Ellsworth et al. 2018).

References

Adamska-Sałaciak, A. (2010). Examining equivalence. *International Journal of Lexicography* 23(4), 387–409.

Almeida, V. G. (2016). *Identificação Automática de Construções de Estrutura Argumental*. [Automatic Identification of Argument Structure Constructions]. M.A. thesis, Department of Linguistics, Federal University of Juiz de Fora, Juiz de Fora, Brazil.

Arcan, M., Turchi, M., Tonelli, S., & Buitelaar, P. (2017). Leveraging bilingual terminology to improve machine translation in a CAT environment. *Natural Language Engineering* 23(5), 763–788.

Bäckström, L., Lyngfelt, B., & Sköldberg, E. (2014). Towards Interlingual Constructicography. On correspondence between constructicon resources for English and Swedish. *Constructions and Frames* 6, 9–32.

Bergen, B., & Chang, N. (2013). Embodied construction grammar. In T. Hoffman & G. Trousdale (eds.). *The Oxford handbook of construction grammar* (pp. 168–190). Oxford: Oxford University Press.

Bertoldi, A. and Rove Luize de Oliveira Chishman (2011) Developing a frame-based lexicon for the Brazilian legal language: The case of the Criminal_Process frame. *AICOL* 2011, 256–270.

Boas, H.C. (2002). Bilingual FrameNet Dictionaries for Machine Translation. In González Rodríguez, M., and C. Paz Suárez Araujo (eds.), *Proceedings of the Third International Conference on Language Resources and Evaluation*. Las Palmas, Spain. Vol. IV, 1364 – 1371.

Boas, H.C. (2003). *A constructional approach to resultatives*. Stanford: CSLI Publications.

Boas, H.C. (2005a). Semantic Frames as Interlingual Representations for Multilingual Lexical Databases. *International Journal of Lexicography* 18(4), 445–478.

Boas, H. C. (2005b). From Theory to Practice: Frame Semantics and the Design of FrameNet. In: Langer, S. & D. Schnorbusch (eds.), *Semantik im Lexikon* (pp. 129–160). Tübingen: Narr.

Boas, H. C. (Ed.) (2009). *Multilingual FrameNets: Methods and Applications*. Berlin/New York: Mouton de Gruyter.

Boas, H.C. (Ed.) (2010). *Contrastive Studies in Construction Grammar*. Amsterdam/Philadelphia: John Benjamins.

Boas, H.C. (2013a). Wie viel Wissen steckt in Wörterbüchern? Eine frame-semantische Perspektive. *Zeitschrift für Angewandte Linguistik* 57, 75–97.

Boas, H.C. (2013b). Cognitive Construction Grammar. In T. Hoffmann and G. Trousdale (eds.), *The Oxford Handbook of Construction Grammar* (pp. 233–254). Oxford: Oxford University Press.

Boas, H.C. (2014a). Zur Architektur einer konstruktionsbasierten Grammatik des Deutschen. In Lasch, Alexander & Ziem, Alexander (Eds.), *Grammatik als Netzwerk von Konstruktionen. Sprachwissen im Fokus der Konstruktionsgrammatik* (= Sprache und Wissen 15) (pp. 37–63). Berlin: de Gruyter.

Boas, H.C. (2014b). Lexical and phrasal approaches to argument structure: Two sides of the same coin. *Theoretical Linguistics* 40.1–2, 89–112.

Boas, H.C. (2017). Computational Resources: FrameNet and Constructicon. In B. Dancygier (Ed.), *The Cambridge Handbook of Cognitive Linguistics* (pp. 549–573). Cambridge: Cambridge University Press.

Boas, H.C. (2018). Zur Klassifizierung von Konstruktionen zwischen 'Lexikon' und 'Grammatik.' In S. Engelberg, H. Lobin, K. Steyer, and S. Wolfer (Eds.), *Wortschätze. Dynamik, Muster, Komplexität* (pp. 33–50). Berlin/Boston: De Gruyter.

Boas, H.C. (2019). Zur methodologischen Grundlage der empirischen Konstruktikographie. In D. Czicza, G. Diewald & V. Dekalo (eds.), *Konstruktionsgrammatik VI* (pp. 237–262). Tübingen: Stauffenburg.

Boas, H.C. (to appear). A roadmap towards determining the universal status of semantic frames. In R. Enghels and M. Jansegers (eds.), *New approaches to contrastive linguistics*. Berlin/Boston: de Gruyter.

Boas, H.C. & Dux, R. (2013). Semantic frames for foreign language education: Towards a German frame-based dictionary. *Veridas On-line. Special Issue on Frame Semantics and its Technological Applications*. 82–100. http://www.ufjf.br/revistaveredas/edicoes- 2013

Boas, H.C. & Dux, R. (2017). From the past into the present: From case frames to semantic frames. In *Linguistics Vanguard* 2017. 1–14. DOI: 10.1515/lingvan-2016-000

Boas, H. C., Dux, R. & Ziem, A. (2016). Frames and constructions in an online learner's dictionary of German. In S. De Knop and G. Gilquin (eds.), *Applied Construction Grammar* (pp. 303–326). Berlin: de Gruyter.

Boas, H. C. & Ziem, A. (2018). Constructing a constructicon for German: Empirical, theoretical, and methodological issues. In Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 183–228). Amsterdam: John Benjamins.

Borin, L., Dannélls, D., Forsberg, M., Gronostaj, M. T., & Kokkinakis, D. (2010). Swedish FrameNet++. *Swedish Language Technology Conference* 2010.

Burchardt, Aljoscha, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Padó, & Manfred Pinkal. (2009). Using FrameNet for the semantic analysis of German: annotation, representation, and automation. In: Hans C. Boas (ed.), *Multilingual FrameNets: Methods and Applications* (pp. 209–244). Berlin/New York: Mouton de Gruyter.

Bryant, J. E. (2008). *Best-fit Constructional Analysis*. Ph.D. Dissertation in Computer Sciences. University of California Berkeley. Berkeley.

Candito et al. (2014). Developing a french framenet: Methodology and first results. In *Proceedings of the 9th international conference on Language Resources and Evaluation (LREC 2014)*, Reykjavik, Iceland.

Cowie, A. P. (Ed.) (1998). *Phraseology. Theory, Analysis, and Application*. Oxford: Clarendon Press.

Croft, W. (2001). *Radical Construction Grammar.* Oxford: Oxford University Press.

Croft, W. (2016). Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology* 20, 377–393.

Croft, W. (2018). *Morphosyntax*. Draft manuscript, University of New Mexico. November 2018.

Diessel, H. 2015. Usage-based construction grammar. In E. Dabrowska & D. Dvjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 296–322). Berlin/New York: De Gruyter.

Diniz da Costa et al. (2018). FrameNet-Based Modeling of the Domains of Tourism and Sports for the Development of a Personal Travel Assistant Application. In: *Proceedings of the Eleventh Interna-*

tional Conference on Language Resources and Evaluation (LREC 2018), 6–12, Miyazaki, Japan: ELRA.

Dodge, E. & Petruck, M. (2014). Representing Caused Motion in Embodied Construction Grammar. In *Proceedings of the 2014 ACL Workshop on Semantic Parsing*, 39–44. Baltimore, Maryland.

Dodge, E., Trott, S., Gilardi, L. & Stickles, E. (2017). Grammar Scaling: Leveraging FrameNet Data to Increase Embodied Construction Grammar Coverage. IN *The AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding Technical Report SS-17–02*. Palo Alto, CA: AAAI Publications, v.17, p.154–162.

Dolbey, A. (2009). *Bioframenet: a framenet extension to the domain of molecular biology.* Unpublished dissertation, UC Berkeley.

Farø, K. (2004). Hvornår går man over åen efter vand? Idiomatiske ækvivalensproblemer i leksikografi og leksikologi. Med dansk og tysk som eksempel ['When do you cross the river for water? Issues of idiomaticity and equivalence in lexicography and lexicology. The example of Danish and German']. *LexicoNordica 11*, 85–108.

Fellbaum, C. (1998). *WordNet*. Cambridge, Mass.: MIT Press.

Fillmore, C.J. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in Linguistic Theory* (pp. 1–90). New York: Rinehart & Winston.

Fillmore, C.J. (1975). An alternative to checklist theories of meaning. *Proceedings of the First Annual Meeting of the Berkeley Linguistic Society (BLS)*, 123–131.

Fillmore, C.J. (1977). Scenes-and-frames semantics. In A. Zampolli (ed.), *Fundamental Studies in Computer Science* (pp. 55–88). Dordrecht: North Holland Publishing.

Fillmore, C.J. (1978). On the organization of semantic information in the lexicon. *Papers from the Parasession on the Lexicon, Chicago Linguistic Society*, 148–173.

Fillmore, C. J. (1982). Frame Semantics. In The Linguistic Society of Korea (Ed.), *Linguistics in the Morning Calm* (pp. 111–137). Seoul: Hanshin Publishing Company. Reprinted in D. Geeraerts (Ed.), *Cognitive Linguistics: Basic Readings* (pp. 373–400). Berlin & New York: Mouton de Gruyter.

Fillmore, C. J. (1985a). Frames and the Semantics of Understanding. *Quaderni di Semantica* 6, 222–254.

Fillmore, C. J. (1985b). Syntactic intrusions and the notion of grammatical construction. In *Annual Meeting of the Berkeley Linguistics Society*, 73–86.

Fillmore, C. J. (1988). The mechanisms of "Construction Grammar". *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistic Society* (BLS), 35–55.

Fillmore, C. J. (1999). Inversion and constructional inheritance. In G. Webelhuth, J.-P. Koenig and A. Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation* (pp. 113–128). Stanford: CSLI Publications.

Fillmore, C. J. (2008a). Border conflicts: FrameNet meets construction grammar. In *Proceedings of EURALEX13* (pp. 49–68). Barcelona: Universitat Pompeu Fabra.

Fillmore, C. J. (2008b). Wherewithal: eine verborgene Konstruktion. In A. Stefanowitsch & K. Fischer (Eds.), *Konstruktionsgrammatik II: Von der Konstruktion zur Grammatik* (pp. 279–283). Tübingen: Stauffenburg.

Fillmore, C. J. (2013). Berkeley Construction Grammar. In T. Hoffmann and G. Trousdale (eds.), *The Oxford Handbook of Construction Grammar* (pp. 111–132). Oxford: Oxford University Press.

Fillmore, C. J. & Beryl T.S. Atkins. (1992). Toward a Frame-based Lexicon: The Semantics of RISK and its Neighbors. In: A. Lehrer & E. Kittay (eds.), *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organization* (pp. 75–102). Hillsdale: Erlbaum.

Fillmore, C.J. & Baker, C. (2010). A frames approach to semantic analysis. In B. Heine, & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 313–340). Oxford: Oxford University Press.

Fillmore, C. J., Lee-Goldman, R., & Rhomieux, R. (2012). The FrameNet Constructicon. In H.C. Boas, & I. Sag (Eds.), *Sign-based Construction Grammar* (pp. 309–372). Stanford: CSLI Publications.

Fillmore, C. J. & Kay, P. (1993). *Construction Grammar Course Book*. UC Berkeley: Department of Linguistics.
Fillmore, Charles J., Paul Kay, and Mary O'Connor. (1988). Regularity and idiomaticity in grammatical constructions: The case of 'let alone.' *Language* 64, 501–538.
Francis, G., Hunston, S. & Manning, E. (1996). *Collins COBUILD Grammar Patterns 1: Verbs*. London: HarperCollins.
Francis, G., Hunston, S. & Manning, E. (1998). *Collins COBUILD Grammar Patterns 2: Nouns and Adjectives*. London: HarperCollins.
Gilardi, L. & Baker, C. F. (2018). Learning to Align across Languages: Toward Multilingual FrameNet. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018), Miyazaki, Japan: ELRA, 13–22.
Goldberg, A. E. (1995). *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
Goldberg, A. E. (2006). *Constructions at Work.* Oxford: Oxford University Press.
Goldberg, A. E. (2013). Constructionist approaches. In Th. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 15–31). Oxford: Oxford University Press.
Goldberg, A.E. & Jackendoff, R. (2004). The English resultative as a family of constructions. *Language* 80.3, 532–568.
Hanks, P. (2013). *Lexical Analysis*. Cambridge, Mass.: MIT Press.
Hasegawa, Y., Lee-Goldman, R., Ohara, K. H., Fujii, S. & Fillmore, C. J. (2010). On expressing measurement and comparison in English and Japanese. In H. C. Boas (Ed.), *Contrastive Studies in Construction Grammar* (pp. 169–200). Amsterdam: John Benjamins.
Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86, 663–687.
Herbst, Th. (this volume).
Hilpert, M. (2014). *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.
Hunston, S., and G. Francis (2000). *Pattern Grammar: A corpus-driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: John Benjamins.
Iwata, S. (2008). *Locative alternation*. Amsterdam/Philadelphia: John Benjamins.
Janda, L. A., Lyashevskaya, O., Nesset, T. Rakhilina, E. & Tyers, F. M. (2018). A constructicon for Russian: filing in the gaps. In Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 165–181). Amsterdam: John Benjamins.
Jurafsky, D. (1991). *An On-line Computational Model of Human Sentence Interpretation: A Theory of the Representation and Use of Linguistic Knowledge*. Doctoral dissertation. University of California, Berkeley.
Kay, P. & Fillmore, C. J. (1999). Construction grammar and linguistic generalizations: The *what's X doing Y?* construction. *Language*, 75, 1–33.
Laviola, A. B. (2015). Frames e Construções em Contraste: uma análise comparativa português-inglês no tangente à implementação de constructicons ['Frames and Constructions in Contrast: a Portuguese-English comparative analysis in regards to the implementation of constructicons']. MA Thesis, Federal University of Juiz de Fora.
Lee-Goldman, R., & Petruck, M. R. L. (2018). The FrameNet constructicon in action. In: Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 19–39). Amsterdam: John Benjamins.
Lelie, M. (2019). Typer och kategorier som länkningssystem inom flerspråkig konstruktikografi ['Types and categories as a linking system in multilingual constructicography']. BA thesis, Dept. of Swedish, University of Gothenburg.
Levin, Beth & Rappaport-Hovav, M. (2005). *Argument realization*. Cambridge: Cambridge University Press.

L'Homme, M. C.; Cormier, M. C. (2014). Dictionaries and the digital revolution: A focus on users and lexical databases. *International Journal of Lexicography*, v. 27, n. 4, p. 331–340.

Lyashevskaya, O., & Kashkin, E. (2015). FrameBank: a database of Russian lexical constructions. In M. Yu. Khachay, N. Konstantinova, A. Panchenko, D. I. Ignatov, & G. V. Labunets (Eds.). *Analysis of images, social networks and texts*. Fourth international conference, AIST 2015, Yekaterinburg, Russia, April 9–11, 2015, Revised selected papers (= Communications in computer and information science, Vol. 542) (pp. 337–348). Springer.

Lyngfelt, B. (2018a). Introduction: Constructicons and constructicography. In Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 1–18). Amsterdam: John Benjamins.

Lyngfelt, B. (2018b). *Constructions all the way up? The case of clausal patterns in Swedish.* Presentation at the 10th International Conference on Construction Grammar (ICCG10) in Paris, July 2018.

Lyngfelt, B., Bäckström, L., Borin, L., Ehrlemark, A., Rydstedt, R. (2018). Constructicography at work. Theory meets practice in the Swedish constructicon. In Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 255–302). Amsterdam: John Benjamins.

Lyngfelt, B., Borin, L., Ohara K. & Torrent, T. T. (Eds.) (2018). *Constructicography: Constructicon development across languages*. Amsterdam: John Benjamins.

Lyngfelt, B., Torrent, T. T., Laviola, A., Bäckström, L., Hannesdóttir, A. H. & da Silva Matos, E. E. (2018). Aligning constructicons across languages: A trilingual comparison between English, Swedish, and Brazilian Portuguese. In Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 255–302). Amsterdam: John Benjamins.

Matos, E. E. (in prep.). Comparing Constructions Crosslinguistically using Comparative Concepts. Federal University of Juiz de Fora.

Matos, E. E., Torrent, T. T., Almeida, V. G., Silva, A. B. L., Lage, L. M., Marcao, N D., Tavares, T. S. (2017). Constructional Analysis Using Constrained Spreading Activation in a FrameNet-Based Structured Connectionist Model In The AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding Technical Report SS-17-02. Palo Alto, CA: AAAI Publications, v.17, p. 222–229.

Nunberg, G., Sag, I. A. & Wasow, T. (1994). Idioms. *Language* 70, 491–538.

Ohara, K. H. (2009). Frame-based contrastive lexical semantics in Japanese FrameNet: The case of risk and kakeru. In: H.C. Boas (Ed.), *Multilingual FrameNets: Methods and Applications,* 163–182. Berlin/New York: Mouton de Gruyter.

Ohara, K. H. (2018). Relations between frames and constructions. In Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 141–163). Amsterdam: John Benjamins.

Ohara, K. H., Fujii, S., Sato, H., Ishizaki, S., Ohori, T., Suzuki, R. (2003). The Japanese FrameNet Project: A Preliminary Report. In *Proceedings of Pacific Association for Computational Linguistics* (PACLING'03), 249–254. Halifax, Canada. August, 2003.

Ohara, K. H., Fujii, Seiko, Ohori, Toshio, Suzuki, Ryoko, Saito, Hiroaki, and Shun Ishizaki. (2004). The Japanese FrameNet Project: An introduction. In: *Fourth International Conference on Language Resources and Evaluation (LREC 2004). Proceedings of the Satellite Workshop "Building Lexical Resources from Semantically Annotated Corpora"*, 9–11.

Perek, F. & Patten, A. (2019). *Towards a constructicon using patterns and frames*. *International Journal of Corpus Linguistics* 24.3, 354–384. University of Birmingham.

Petruck, M. R. L., Fillmore, C. J., Baker, C., Ellsworth, M. & Ruppenhofer, J. (2004). Reframing FrameNet data. In *Proceedings of the 11th EURALEX International Congress*, Lorient, France, 405–416.

Ruppenhoffer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R. & Scheffczyk, J. (2016). *FrameNet II: Extended Theory and Practice*. Berkeley: ICSI, available at: http://framenet.icsi.berkeley.edu.

Ruppenhofer, J., Boas, H.C., & Baker, C. F. (2017). FrameNet. In P. A. Fuertes-Olivera (ed.), *The Routledge Handbook of Lexicography* (383–398). New York: Routledge.

Sag, I. A. (2010). English Filler-Gap Constructions. *Language* 86, 486–545.

Sag, I. A. (2012). Sign-Based Construction Grammar: An Informal Synopsis. H. C. Boas & I. A. Sag (Eds.), *Sign-Based Construction Grammar* (pp. 69–202). Stanford: CSLI Publications.

Salomão, M. M. M. (2009). FrameNet Brasil: um trabalho em progresso. *Calidoscópio* 7, 171–182.

Schmidt, T. (2009). The Kicktionary – A multilingual lexical resource of football language. In H.C. Boas (ed.), *Multilingual FrameNets: Methods and Applications* (pp. 101–134). Berlin/New York: Mouton de Gruyter.

Subirats, C. (2009). Spanish FrameNet: A frame-semantic analysis of the Spanish lexicon. In H.C. Boas (ed.), *Multilingual FrameNets: Methods and Applications* (pp. 135–162). Berlin/New York: Mouton de Gruyter.

Subirats, C. & Petruck, M. R. L. (2003). Surprise: Spanish FrameNet. *International Congress of Linguists. Workshop on Frame Semantics*, Prague (Czech Republic), July 2003.

Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge, MA. MIT Press.

Tavares, T. (2018). *Requisitos para a modelagem de padrões de cunhagem e construções semi-produtivas no constructicon da FrameNet Brasil com foco no fomento ao desenvolvimento de tradutores automáticos* [Requirements for modeling coining patterns and semi-productive constructions in the FrameNet Brasil Constructicon focusing on the development of automatic translators]. Ph.D. Dissertation in Linguistics. Universidade Federal de Juiz de Fora, Brazil.

Torrent, T. T. & Ellsworth, M. J. (2013). Behind the Labels: Criteria for Defining Analytical Categories in FrameNet Brasil. *Veredas* 17(1), 44–65.

Torrent, T. T., Ellsworth, M., Baker, C. F. & Matos, E. E. (2018). The Multilingual FrameNet Shared Annotation Task: a preliminary report. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018), Miyazaki, Japan: ELRA, 62–68.

Torrent, T. T., Lage, L. M., Sampaio, T. F., da Silva Tavares, T., & da Silva Matos, E. E. (2014). Revisiting border conflicts between FrameNet and Construction Grammar: Annotation policies for the Brazilian Portuguese Constructicon. *Constructions and Frames* 6, 34–51.

Torrent, T. T., Matos, E. E., Lage, L., Laviola, A., da Silva Tavares, T., de Almeida, V. G., & Sigiliano, N. (2018). Towards continuity between the lexicon and the constructicon in FrameNet Brasil. In Lyngfelt, Borin, et al. (Eds.), *Constructicography: Constructicon development across languages* (pp. 107–140). Amsterdam: John Benjamins.

Winn, J., Bishop, C. M., Diethe, T., Guiver, J. & Zaykov, Y. (forthcoming). *Model Based Machine Learning*. Available at: http://mbmlbook.com/index.html

Ziem, A. & Lasch, A. (2013). *Konstruktionsgrammatik*. Berlin/Boston: de Gruyter.

Ziem, A., Flick, J., Sandkühler, Ph. (this volume). The German Constructicon Project: Framework, Methodology, Resources.