MoCCA: A Model of Comparative Concepts for Aligning Constructicons

Arthur Lorenzi[1], Peter Ljunglöf[2], Benjamin Lyngfelt[2], Tiago Timponi Torrent[1,3], William Croft[4], Alexander Ziem[5], Nina Böbel[5], Linnéa Bäckström[6], Peter Uhrig[7], Ely Matos[1]

[1]Federal University of Juiz de Fora, [2]University of Gothenburg, [3]Brazilian National Council for Scientific and Technological Development – CNPq, [4]University of New Mexico, [5]HHU Düsseldorf, [6]Halmstad University, [7]University of Erlangen-Nuremberg

arthur.lorenzi@estudante.ufjf.br

Abstract

This paper presents MoCCA, a Model of Comparative Concepts for Aligning Constructicons under development by a consortium of research groups building Constructicons of different languages including Brazilian Portuguese, English, German and Swedish. The Constructicons will be aligned by using comparative concepts (CCs) providing language-neutral definitions of linguistic properties. The CCs are drawn from typological research on grammatical categories and constructions, and from FrameNet frames, organized in a conceptual network. Language-specific constructions are linked to the CCs in accordance with general principles. MoCCA is organized into files of two types: a largely static CC Database file and multiple Linking files containing relations between constructions in a Constructicon and the CCs. Tools are planned to facilitate visualization of the CC network and linking of constructions to the CCs. All files and guidelines will be versioned, and a mechanism is set up to report cases where a language-specific construction cannot be easily linked to existing CCs.

Keywords: construction, constructicon, comparative concepts

1. Introduction

Constructicons are digital collections of construction descriptions, in the sense and spirit of Construction Grammar. There are now such Constructicons available or under development for at least half a dozen languages (Lyngfelt et al., 2018), and more on the way. Some of them are developed for NLP purposes and others for language pedagogy – or both (Borin & Lyngfelt, forthc.; Ziem et al., forthc.). While these Constructicons are for the most part designed as monolingual resources, there is also ongoing work towards cross-linguistic application, or *multilingual constructicography*.

Previous efforts in Constructicon alignment have been designed as bilingual or trilingual comparisons (Laviola, 2015; Bäckström, Lyngfelt & Sköldberg, 2014; Lyngfelt et al., 2018). By using a constructional approach with tools from lexicography, the authors could discern close equivalents for most of the constructions investigated. However, the results showed a bias towards the source language, which was English, regarding both formal and functional properties. Furthermore, the method turned out to be far too time consuming to be feasible on a large scale and for more languages. Hence, the conclusion of these experiments is that, rather than construction by construction comparison, multilingual Constructicon alignment requires a language-neutral base of comparison.

In parallel, a series of workshops in Düsseldorf, Germany and Gothenburg, Sweden have been organized aimed at discussing possible methodologies for aligning constructions from the Brazilian Portuguese, English, German and Swedish Constructicons. The choice of Berkeley FrameNet frames as a possible comparative variable was obvious, not least because some Constructicons already link constructions and frames (Boas, Lyngfelt & Torrent, 2019). But it is clear that frames only link the meaning/function of constructions and the form could therefore not represent a comparative variable. So the choice also fell on the use of Croft's comparative concepts (Croft, 2022), even if this entails a new implementation for almost all Constructicons. In addition to the selected methodology, these workshops also focused on practical implementation; in particular, rules and processes were developed for aligning different constructions using comparative concepts and frames.

In this paper, we present one of the outcomes of these workshops – the analytical and technical guidelines for aligning constructions and Constructicons via MoCCA (Model of Comparative concepts for Constructicon Alignment). These guidelines are jointly developed and agreed upon by Constructicon-building teams (CBTs) henceforth referred to as the CBT consortium.

The overall idea of this enterprise is to connect constructions across and within languages using comparative concepts (CCs) as a shared base of comparison (Lyngfelt et al., 2022). The CCs provide language-neutral definitions of linguistic properties, and language-particular constructions may be linked to any and all CCs conforming to properties shared by the construction in question. Thereby the construction will also be connected to other constructions linked to the same CC.

The guidelines described here are primarily directed towards Constructicon-building Teams, but may of course be employed by any linguist wishing to connect or compare a particular set of constructions to other constructions within or across languages via comparative concepts.

2. Comparative Concepts

Comparative concepts are the linguistic concepts used as the basis of cross-linguistic comparison in typology, although they were given this name only

recently (Haspelmath, 2010). Comparative concepts have been defined in terms of function (semantics, pragmatics) since Greenberg (1963) and Keenan & Comrie (1977). More recently, Haspelmath (2010) argues that some comparative concepts can be defined at least partly in terms of cross- linguistically valid properties of morphosyntactic form. Croft (2016, 2022) argues that comparative concepts are either completely functional, or are hybrids combining properties of function and form.

The CCs used in MoCCA are of five types: constructions, strategies (*see* 3.2.1), semantic content, information packaging and frames. The first four types, described in Croft (2022), are based on language typology. MoCCA uses an extension of the set of CCs presented by Croft (2022). Constructions and strategies are hybrid CCs, that is, pairings of form and function (see below). Semantic content and information packaging are purely functional CCs. The fifth type consists of the set of semantic frames defined in the Berkeley FrameNet 1.7 data release, as described in Ruppenhofer et al. (2016). Names of particular CCs, of any type, are written in boldface; when needed the CC type will be indicated within parentheses after the name.

Note that the language-neutrally defined constructions employed as CCs, henceforth CC-constructions, are not to be confused with language-particular constructions, which will here be called L-constructions. An L-construction, for example the English Polarity Question Construction (*Are you coming with us?*), is generally defined as the pairing of a language-specific form with a particular function.

In cross-linguistic comparison, particular constructions such as the **polarity question construction (cxn)** are compared first based on function: the set of form-function pairings across languages that express a function such as polarity questions (Weissweiler et al. 2024). The polarity question function is defined as **propositional content (sem)** packaged as an **interrogative (inf)** to which the interlocutor is expected to confirm, amend or disconfirm (Bolinger, 1978).

Cross-linguistic comparison starts from function because languages vary considerably in their morphosyntactic form in expressing functions such as the polarity question function. Although forms are language-specific, certain general properties of form can be defined cross-linguistically, for example, word order, prosody, a question particle, a special verb form, and so on. Thus, the English Polarity Question L-construction is an instance of the CC-construction **polarity question construction (cxn)** and uses the strategies of **word order (str)**—specifically, Subject-Auxiliary inversion—-and **prosody (str)**—specifically, final rise intonation.

L-constructions may be linked to one or more CCs, and related L-constructions may share some CCs but differ with respect to others. Thus, the CC links represent partial correspondences and should not be confused with equivalence. Also note that the CCs cannot cover all properties of all constructions in all languages. There will always be language-particular idiosyncrasies not covered by this alignment model.

In the remainder of this paper we present the analytical and technical guidelines for using MoCCA, as well as the methodology for reporting issues with the system.

3. Analytical Guidelines

The MoCCA analytical guidelines focus on the procedures and principles for associating comparative concepts, including frames, with constructions and, if applicable, construction elements (CEs). In this section, we present its first version (1.0).

3.1. The CC Network

The CCs, including frames, are presented as a database of related concepts. The Croftian CCs can be related using nine different relations. *Subtype*, *part*, *attribute*, *value* and *role/filler* are used for CCs of the same type. In addition, for construction CCs, there is a special part relation, *head,* for the head of the construction.

Strategies are related to construction in three different ways. All strategies are related to the construction whose form they describe with the *expression-of* relation. Two classes of strategies also make reference to another construction.

One class of strategies, recruitment strategies, recruit the form of a related construction. For example, the English Physical Sensation construction (*I have a headache/a cold*/etc.) recruits the form of the Presentational Possession construction (*I have a car*). The *recruited-from* relation links this strategy to its source construction.

In another class of strategies, the system of strategies, elements of the form of one construction are based on corresponding elements in another construction. For example, in the accusative alignment strategy, illustrated by transitive *She saw her* vs. intransitive *She was sleeping*, the form of the A (transitive subject) argument phrase is the same as the form of the S (intransitive subject) phrase, i.e. *she*, while the form of the P (transitive object) phrase is different, i.e. *her.* In this case, the intransitive construction serves as the model for the strategy for encoding the transitive construction's arguments (A=S and P≠S). The accusative alignment strategy therefore has a *modeled-on* relation to the intransitive construction.

Finally, the *function* relation is used to link semantic and information packaging CCs to CC-constructions. Frame CCs can be related via the *inheritance*, *subframe*, *perspective-on*, *precedes*, *using*, *causative-of*, *inchoative-of*, *metaphor* and *see-also* relations (for a more detailed explanation on FrameNet relations, see Ruppenhofer et al., 2016).

The CC network is a directed acyclic graph (DAG), with arcs between CCs of the same type always having the direction of *child* to *parent*, *part* to *whole*, *attribute* to CC, *value* to *attribute* and *role/filler* to CC. Relations between different CC types are always directed to CC-construction. When building these relations, clusters of the same CC type are analyzed as taxonomic trees (*e.g.* the modification construction tree). When all trees of all five different

CC types are combined via the *expression-of*, *recruited-from*, *modeled-on* and *function* relations, the network becomes a DAG.

It is also possible that even within a CC cluster, there may be multiple, alternative taxonomies. This stems from the fact that language phenomena can be analyzed in different ways and to represent it, CCs can have multiple parents. When possible, however, multiple parents are always avoided as a way to make the network clearer.

To avoid inconsistencies, a set of constraints were devised and implemented to validate the state of the network. These validators, among other things, check whether the CC type constraints for each relation are respected and that no CC node is isolated. A more specific constraint, considered for CC-sem and CC-inf, would be: if X is a *value* of Y and Y is an *attribute* of Z, then X must be a *subtype* of Z.

In its current state, the network consists of 2286 CCs (of which 1222 are frames) and 3547 relations between them. The 1672 relations between non-frame CCs are all new and were manually created for MoCCA.

Previous experience with the Global FrameNet Shared Annotation task (Torrent et al., 2018; Giouli et al., 2020) reveals that users will often find the need to expand the model by adding new CCs or revising the existing ones. This is not a choice teams will be allowed to make on the fly, since it would compromise the alignment. Nonetheless, mechanisms are proposed for dealing with cases where teams cannot find a perfectly matching CC for the construction or construction element under analysis (see Section 5).

3.2. Associating CCs with Constructions

Associating CCs with constructions requires consideration of both function and form of constructions. This is not a simple task. For example, English uses a special pronoun for reflexive meaning for all persons, e.g. *I cut myself*, while Brazilian Portuguese and Swedish use the ordinary transitive object pronouns for 1st and 2nd person: Swedish *Jag skar mig* 'I cut myself' (lit. *I cut me*).

In contrast, for reciprocal meaning English uses a special multi-word expression: *We saw each other*; but some verbs express reciprocal meaning without any special form: *We met*. Brazilian Portuguese uses the same pronouns that are used for reflexive meaning: *Nos vemos amanhã na cidade* 'We'll see each other downtown tomorrow'; or a special form *um ao outro*, with or without the reflexive pronoun. Swedish uses a special form different from its reflexive pronoun, not unlike English: *Tvillingarna avskyr varandra* 'The twins detest each other'; but in some cases use a special verb form in *-s*: *De träffas och talas vid* 'They meet (each other) and talk (to each other)'.

This many-to-many mapping between form and function across languages requires us to compare form and function partly independently.

The association between CCs and L-constructions is guided by the following four principles, which will be discussed in detail in the subsections below.

1. **Application:** Link L-constructions to both a CC-construction and one of the CC-construction's CC-strategies.

2. **Generality:** Link an L-construction to the CC-construction at the lowest relevant level of generality in the CC-construction taxonomy.

3. **Constructional inheritance:** If there are multiple L-constructions in a taxonomic hierarchy, link them to the CC-constructions in the CC-construction taxonomic hierarchy that most closely matches the L-construction hierarchy.

4. **Analytical targets:** Link L-construction CEs to corresponding CC-constructions where the latter exist, and to CC-strategies where the L-construction CE is introduced by a strategy.

3.2.1. Application

In linguistic typology, function is the primary basis for cross-linguistic comparison, because function can be compared directly across languages. A typology of reciprocal constructions analyzes the variation in form for the expression of reciprocal meaning. This is an onomasiological approach to the analysis of constructions. Hence, “reciprocal construction” in a cross-linguistic sense is any morphosyntactic form expressing a particular function in any language.

Morphosyntactic form is language-specific, but some morphosyntactic properties can be defined in cross-linguistic terms. Variation in form across languages can be classified into morphosyntactic *strategies*, such as special pronoun form or a special verb form (Keenan and Comrie, 1977; Croft, 2022).

Thus, an L-construction in a Constructicon for a specific language includes a specification of a particular function (in cross-linguistic terms, a CC-construction) and its language-specific morphosyntactic form (a CC-strategy). An L-construction should therefore be linked to both a CC-construction and a CC-strategy of that CC-construction in the network.

Construction grammarians often take a semasiological approach, examining different functions expressed by a language-specific morphosyntactic form, such as the special reflexive pronoun form being used for reciprocal meaning in Brazilian Portuguese. For cross-lingual constructicon alignment, these correspond to two distinct CC-constructions, where one of the constructions has recruited the form of the other (a recruitment strategy).

3.2.2. Generality

When choosing the CC to be associated with a construction or CE, the most specific one that is applicable should be used.

The CCs in the linking model are organized in a network. Thus, linking an L-construction to a CC also connects it to related CCs of different generality and, indirectly, to associated L-constructions. This feature

may somewhat compensate for differences in granularity between L-construction entries in different Constructicons, through a graph structure identifying the closest corresponding target construction. Linking at too high a level of generality, however, may overgenerate and create less accurate connections. Therefore, one should try to find the most specific CC possible for any given construction or CE.

Sometimes this means that the best solution is to link to two or more co-hyponym CCs rather than a single hyperonym. If a supertype CC captures more subtypes than the ones relevant for characterizing the construction under analysis, then, the relevant subtypes should be associated with the construction, instead of the supertype.

3.2.3. Constructional Inheritance

Language-particular Constructicons may also be organized in inheritance networks. When looking at the network of CCs, it is important to consider the CC-construction's degree of generality/specificity in relation to that of the L-construction under analysis in the Constructicon.

For Constructicons that model construction inheritance, CCs and frames should be associated only once in an inheritance chain, at the adequate level of generality. For example, if a Constructicon has a general construction for relative clauses and four other constructions for subtypes of relative clauses, the more general **relative clause construction (cxn)** CC should be associated to the more general L-construction, while the subtypes – such as **anaphoric head (cxn)** and **free relative clause constructions (cxn)** – should be associated to the daughter L-constructions.

3.2.4. Analytical Targets

In the proposed linking model, CCs can be associated with constructions, CEs or both. If the Constructicon in question models constructional constituency in a way that allows for direct association of CEs to CC-constructions in a *part* relation to the CC-construction linked to the whole L-construction, CEs can be manually associated; they cannot be automatically derived from the *part* relation. If not, all applicable CCs should be associated at the level of the construction.

If the CC-construction already has CCs of its parts, they can be linked to L-construction CEs. For example, in linking the English Adjective Modification Construction, illustrated by *very large turkey*, the CCs **adjective modification construction (cxn)**, **adjective attributive phrase (cxn)**, and **referent expression (cxn)** are applicable to the whole referring phrase, its modifier (*very large*) and its head (*turkey*) respectively.

It is important to distinguish when the CE is associated with a CC-construction in a *part* relation, or a part of the CC-strategy used by the L-construction. For example, the English Finite Complement Clause L-construction, as in *Sally said that she ate the leftovers*, has four language-specific CEs: the matrix complement- taking predicate or CTP (*said*), the matrix Subject argument phrase (*Sally*), the complementizer (*that*), and the complement (*she ate the leftovers*). The matrix CTP and its dependent arguments and complement are CEs of the CC-construction **complement clause construction (cxn)**. The complementizer, on the other hand, is introduced by the CC-strategy **complementizer (str)**.

4. Technical Guidelines

This section of the guidelines aims to provide Constructicon-building teams (CBTs) with information on the requirements for implementing the alignment between Constructicons. They cover issues concerning the database format, tools and versioning.

4.1. Database format

Considering the need to preserve the autonomy of different CBTs on how to organize and manipulate their data, MoCCA is split into a main database file for CCs and language specific files. These files follow the “keep it simple” principle, i.e. they are easy to read, both by humans and computer algorithms, and contain only information relevant for the Constructicon alignment.

4.1.1 The CC Database File

The first file, referred to as CC Database File, comprises the set of comparative concepts agreed upon and provided by the consortium and their relations. Since the CCs are the main features used to align constructions from different projects, this file should be treated as somewhat static. Changes are expected, but should not be drastic or as fast as other data, as they can potentially change the alignment of all Constructicons. The main content of the file is its version and the CCs themselves. Each CC entry must contain an unique, persistent CC ID and the CC's type, name, definition and relations. To represent relations, each CC entry must include attributes for each relation type described in Section 3.1. These attributes must contain a list of CC IDs with which the CC relates via that type. If that list is empty, the attribute may be omitted. A YAML-like schema for this file looks like this (a **?** indicates an optional field):

- CC Database File version
- List of CCs:
    - CC ID
    - CC Type
    - CC Name
    - CC Definition
    - Relation r_i: list of CC IDs **?**

4.1.2 The Linking Database Files

The second file type in MoCCA stores the linking between a Constructicon and the comparative concepts. A Linking Database File uniquely identifies a Constructicon among all others and for that reason must contain an alphanumeric identifier for that Constructicon, its name, version and the ISO 639-3 code of its language. This file must contain an explicit indication of which CC Database File version was used for the linking process. Its main content is a list of L-constructions. Each L-construction entry must have an ID, name and description and a list of associated CC IDs. The L-construction ID does not

need to be universally unique, i.e., it only needs to be unique for the Constructicon in question. The list of CC IDs is restricted to IDs from the CC Database File in its version specified by the linking file.

When applicable, the L-construction entry should also specify the ID of its parent L-construction and a list of its CEs in the Constructicon. The required data for CE entries mirrors that of L-construction entries: CE ID, name and description, parent CE ID and a list of CC IDs. Optional data for a Linking File includes name and descriptions in English and up to 3 example constructs. A representation of this complete schema in a YAML-like format looks like this:

- Constructicon ID
- Constructicon Name
- Constructicon Version
- CC Database File Version
- MoCCA Guidelines version
- Language ID (ISO 639-3)
- List of constructions:
    - L-construction ID
    - L-construction Name
    - L-construction Name (en) **?**
    - L-construction Description
    - L-construction Description (en) **?**
    - Parent L-construction ID **?**
    - List of CC IDs (linking to the CC file)
    - List of Examples [0-3] **?**
    - List of L-construction CEs:
        - CE ID
        - CE Name
        - CE Name (en) **?**
        - CE Description
        - CE Description (en) **?**
        - Parent CE ID **?**
        - List of CC IDs (linking to the CC file)

4.2. Tools

Relating constructions to comparative concepts is the only way in which data from different projects can be connected. To make this process easier, faster and inconsistency-free, it is possible to develop a linking tool. This can be a web interface or API with which users (*i.e.* Constructicon developers) can easily link their constructions and CEs to CCs from MoCCA.

Another possible useful tool could use the linked databases from all existing Constructicons to show how different constructions are related between different languages. This can be done via different kinds of visualizations of the underlying CC graphs.

4.3. Versioning

To increase compatibility and preserve the ability of projects to work at their own pace, all of the files and guidelines previously discussed need to be versioned. Every database or tool built based on the CC or the Linking Databases needs to explicitly include the version of those files that was used as part of the metadata. In the case of the Linking Database Files, the Constructicon projects are expected to update their version according to the changes made and also provide the version of the analytical guidelines that were followed. When changing the CC Database File version used by a Linking Database File, documentation will be provided to guide the automatic or manual update to a new version, depending on the changes made to the CCs by the CBT consortium.

5. Reporting Issues

This section presents guidelines for situations where it is not possible to adhere to the four principles in Section 3.2 or any other problem arises. In such cases, teams should report an issue. This reporting can be done using GitHub's Issues system at the appropriate repository under our organization[1].

A special type of issue are situations where the L-construction will not appear to fit into the function defined for any of the CC-constructions in the CC network, or the strategies defined for the relevant CC-construction. To fix that, a new CC must be proposed.

In the case of an apparent missing CC-construction, the report should include a proposed name and definition, the semantic content and information packaging CCs that define the proposed CC-construction, and taxonomic and/or partonomic relations between the proposed CC-construction and existing CC-constructions in the network.

In the case of an apparent missing CC-strategy, the report should include the name and definition and the ID of the CC-construction which the CC-strategy expresses, what type of strategy is involved, and in the case of a system strategy or a recruitment strategy, what system the strategy is part of, and what the source construction for the recruitment strategy is.

For more general problems or issues with other principles, the report must describe the situation and if possible, propose a change. This description should indicate whether the issue pertains to one of the four principles in Section 3.2 and where it relates to, *i.e.* guidelines or one of the database files. The solution, if present, must propose a change to that part of MoCCA.

In all cases, the CBT consortium will analyze the issue report and implement the solution that best handles the case while minimizing the impact to the full alignment.

6. Conclusion

In this contribution, we have introduced a model for aligning Constructicons based on comparative concepts (Croft 2016, 2022). More than 2,000 CCs, including all FrameNet 1.7 frames, have been collected in the MoCCA database and are made available to the research community together with a set of guidelines that define the process of aligning constructions from different Constructicon projects. In addition to these analytical and technical guidelines, a process for reporting issues and suggesting amendments has been outlined.

[1] https://github.com/comparative-concepts

While the model presented here is aimed at Constructicon-building teams (CBTs), we encourage linguists working with other languages and constructions to also consider linking their resources to MoCCA as a way to compare their work to existing Constructicons. By expanding the pool of languages linked to the comparative concepts, we can further improve the comparison model.

MoCCA can also form the basis of schemas and guidelines for other types of annotations, both within specific languages and across languages. It is also a useful resource for projects working with interlingua or other computational methods that could leverage the network of CCs. Finally, it could also serve in any type of contrastive work or as a resource for teaching.

Acknowledgements

MoCCA builds upon the foundation of previous work and from the discussions held at the Constructicon Alignment Workshop 2022, in Gothenburg (https://www.globalframenet.org/caw2022). We are grateful for the exchanges with Elodie Winckel, Kyoko Ohara, Thomas Herbst and Valentina Zhukova.

Authors acknowledge the support of the Graduate Program in Linguistics at the Federal University of Juiz de Fora. Research collaboration leading to MoCCA was funded by CAPES/STINT grant 99999.009910/2014-00 and CAPES/PROBRAL grant 88887.144043/2017-00. Lorenzi's research was funded by CAPES PROBRAL PhD exchange grant 88887.628831/2021-00 and CAPES PROEX PhD Grant 88887.816228/2023-00. Torrent is an awardee of the CNPq Research Productivity Grant number 315749/2021-0.

References

Bäckström, L., Lyngfelt, B., & Sköldberg, E. (2014). Towards interlingual constructicography: on correspondence between constructicon resources for English and Swedish. *Constructions and Frames*, *6*(1), 9-33.

Boas, H. C., Lyngfelt, B., & Torrent, T. T. (2019). Framing constructicography. *Lexicographica*, *35*(2019), 41-85.

Bolinger, D. (1978). Yes—no questions are not alternative questions. In *Questions* (pp. 87-105). Dordrecht: Springer Netherlands.

Borin, L., Lyngfelt, B. (Forthcoming). Framenets and constructicons. *The Cambridge Handbook of Construction Grammar*, ed. by Mirjam Friend and Kiki Nikoforidou. Cambridge University Press.

Comrie, B. (2014). Noun Phrase Accessibility and Universal Grammar 1. In *Universal Grammar (RLE Linguistics A: General Linguistics)* (pp. 3-45). Routledge.

Croft, W. (2016). Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology*, *20*(2), 377-393.

Croft, W. (2022). *Morphosyntax: constructions of the world's languages*. Cambridge University Press.

Giouli, V., Pilitsidou, V., & Christopoulos, H. (2020). Greek within the Global FrameNet Initiative: Challenges and Conclusions so far. In *Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet* (pp. 48–55). European Language Resources Association.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*, *2*, 73-113.

Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, *86*(3), 663-687.

Laviola, A. (2015). *Frames e Construções em Contraste: uma análise comparativa português-inglês no tangente à implementação do constructicons*. Master's thesis, Federal University of Juiz de Fora.

Lyngfelt, B., Borin, L., Ohara, K., & Torrent, T. T. (Eds.). (2018). *Constructicography: Constructicon development across languages* (Vol. 22). John Benjamins Publishing Company.

Lyngfelt, B., Torrent, T. T., Matos, E. E. S., & Bäckström, L. (2022). Comparative concepts as a resource for multilingual constructicography. *Valency and constructions. Perspectives on combining words* (Meijerbergs arkiv för svensk ordforskning, 46), 100-29. Göteborg: Meijerbergs institut för svensk etymologisk forskning, Göteborgs universitet.

Ruppenhofer, J., Ellsworth, M., Schwarzer-Petruck, M., Johnson, C. R., & Scheffczyk, J. (2016). *FrameNet II: Extended theory and practice*. International Computer Science Institute.

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* (pp. 62-68). European Language Resources Association.

Weissweiler, L., Böbel, N., Guiller, K., Herrera, S., Scivetti, W., Lorenzi, A., Melnik, N., Bhatia, A., Schütze, H., Levin, L., Zeldes, A., Nivre, J., Croft, W., & Schneider, N. UCxn: Typologically Informed Annotation of Constructions Atop Universal Dependencies. arXiv. https://doi.org/10.48550/arXiv.2403.17748

Ziem A., Willich A., & Michel, S. (Eds.). (Forthcoming). *Constructing Constructicons*. John Benjamins.