

A linguistic ontology for constructicography: the Research Constructicon and its ontology modules

Elodie Winckel, Peter Uhrig, Stephanie Evert

Friedrich-Alexander-Universität Erlangen-Nürnberg
Research Training Group *Dimensions of Constructional Space*
elodie.winckel@fau.de, peter.uhrig@fau.de, stephanie.evert@fau.de

Abstract

This paper introduces the Research Constructicon (RCxn), a project developed within the Research Training Group *Dimensions of Constructional Space*. The training group finances PhD projects in the framework of Construction Grammar (CxG), which views language as a network of form-meaning pairings. The RCxn is designed as a dynamic, community-driven resource that documents linguistic constructions while also capturing the research processes and findings associated with them. The project addresses three core dimensions: (1) the development of a modular ontology to represent constructions, their relationships, and the research surrounding them; (2) the implementation of a database populated by researchers' contributions; and (3) the creation of a web application to visualize and interact with the data. This paper focuses on our work to implement a rich ontology for the RCxn, which has to accommodate diverse research needs, from cross-linguistic comparisons to multimodal analyses, while ensuring flexibility and interoperability. We detail the modular design of the ontology, its alignment with semantic web standards (RDF/OWL), and the integration of existing ontologies (e.g., OLiA, FOAF). The RCxn's development is iterative, driven by feedback from our diverse group of PhD researchers.

Keywords: Constructicography, Semantic Web Technologies, Ontology

1. Introduction

This paper presents the Research Constructicon (hereafter: RCxn), a project funded by the German Research Foundation (Deutsche Forschungsgemeinschaft) as part of the Research Training Group *Dimensions of Constructional Space*. The latter is dedicated to research in Construction Grammar (CxG), a theoretical framework that views language as a network of form-meaning pairings referred to as “constructions”. According to Construction Grammar, constructions range from words and morphemes to complex syntactic patterns and discourse structures. The RCxn project aims to create a dynamic, community-driven resource that not only documents linguistic constructions but also captures the research processes and findings associated with them.

The project was divided into three interconnected dimensions. First, we developed an ontology for all the terms that will allow us build a database that documents individual constructions, their interconnections, and the research conducted on them. Accordingly, the ontology needs to cover a wide range of aspects, from the internal features of constructions to their relationships with other constructions and the scientific discourse surrounding them. Second, we used the ontology to create an RDF database that currently includes contributions from 13 researchers, with plans to expand to 30 or more by the end of 2027. Finally, we developed a web application to visualize and interact with the database content, which is now operational and accessible (<https://bdlweb.phil.uni-erlangen.de/RCxn/>).

While these three dimensions (ontology,

database, and web application) can be conceptualized as sequential steps, each one requiring the previous one to be implemented, their development has been iterative. The ontology continues to evolve as new needs emerge, influencing both the database structure and the web application's frontend.

In this paper, we focus on the first of these dimensions: the ontology that underpins the RCxn. We begin by outlining the theoretical motivations for adopting a network-based approach to constructionist research, emphasizing the advantages of Semantic Web Technologies. We then discuss the technical implementation of this ontology. Finally, we present the modular ontology we developed, where each module is tailored to address specific aspects of constructionist research, from documenting linguistic resources and their relationships to capturing the research process itself.

2. Motivation

A **reference constructicon** (hereafter: Cxn) is a collection of linguistic construction descriptions rooted in the framework of Construction Grammar, usually in the form of a databases accessible via a web application. A Cxn captures the full spectrum of linguistic units, including multi-word expressions, idiomatic phrases, and grammatical patterns. By contrast, a dictionary might catalog focus particles like *only* or *even* but not cleft constructions (e.g., *It was John who left*), which serve the same focalizing function. Cxns embrace the notion of a continuum of linguistic units (from words to phrases, sen-

tences, and even discourse patterns). This holistic approach aligns with the core tenets of Construction Grammar (Goldberg, 1995; Croft, 2001), which posits that language is organized as a network of constructions at all levels of abstraction.

In Germany, a German Cxn (FKD) has been developed at Heinrich-Heine-Universität Düsseldorf (FKD), and an English Cxn at Friedrich-Alexander-Universität Erlangen-Nürnberg (Herbst et al. 2023–; Herbst and Hoffmann 2024). The scope of FKD, for example, ranges from (semi-)lexicalized constructions like “NP pur” (‘NP in its purest form’; Babal 2026) to fully schematic constructions like nominal compounds¹.

While existing Cxns primarily serve as descriptive or reference resources, our project introduces a novel perspective: a **research constructicon** (RCxn) designed as a platform for documenting ongoing research in Construction Grammar. Currently, linguistic findings in this framework are predominantly published in research papers, lacking a centralized repository. Although initiatives like the DELPH-IN Consortium attempt to combine linguistic descriptions of constructions with varying levels of schematicity in a single implementation, their scope remains limited to specific communities, particularly those focused on formal syntax and semantics.

The RCxn aims to address this fragmentation by providing a community-driven tool for gathering, documenting, and sharing construction-based research. This platform will not only preserve the diversity of constructionist approaches but also enable researchers to build on existing work in a structured and transparent manner.

Building a RCxn presents unique challenges due to the inherently interconnected and dynamic nature of linguistic constructions, as well as the diversity of research in linguistics. One key complexity lies in capturing the relationships between constructions, such as inheritance hierarchies or sister constructions. These relationships are not merely descriptive but Construction Grammar posits that they have a cognitive reality. But comparative linguistics adds another layer of complexity, as the RCxn must facilitate cross-linguistic analyses of similar constructions. Finally, the diachronic dimension demands attention to how construction networks evolve over time, as historical shifts reshape inheritance links, functional roles, and formal properties. The network perspective is at the core of this kind of research: constructions are not isolated entities but part of a larger, dynamic system where their relationships define their behavior, variation, and evolution. The RCxn thus aspires to reflect this interconnectedness.

¹<https://framenet-constructicon.hhu.de/constructicon/constructionfamily?id=14>

The RCxn is not only focused on the network that constructions form with each other, though. One of its main goals is to provide a rich description of individual constructions and the research findings associated with them. For example, we aim to document phraseological research, which examines the interplay between compositionality and idiomaticity, showing how constructions range from fully transparent to highly conventionalized units. The challenges extend further into multimodality, where constructions span multiple formal levels (e.g., syntax, prosody and gesture).

3. Methodology

Development of the RCxn is guided by a considerable number of PhD projects associated with the Research Training Group (close to 30 completed and ongoing projects in total so far). Their work spans a wide range of phenomena—from cross-linguistic and diachronic comparisons to multimodal analyses of speech and gesture, each requiring different levels of formalism and granularity. This diversity mirrors the broader Construction Grammar community, making our group an ideal microcosm for developing a resource that can scale to wider use. To ensure the RCxn meets real-world research needs, we started by collecting construction descriptions from these researchers, analyzing their requirements for both the database structure and the frontend interface. This bottom-up approach is central to the project because it ensures that the project is grounded in the practical needs of its contributors.

One illustrative example is a PhD project in the domain of **phraseology**, which investigates constructions with a high degree of idiomaticity (idioms, or even proverbs) across languages (Rastegar, 2026). This work explores the factors impacting the acquisition of such idiomatic expressions in a second (or foreign) language. Especially, it aims to identify the impact of similarities of idioms (both formal and functional) in the first and second language. The idiom *to take the bread out of someone’s mouth* in English has formal as well as functional similarities with *nān-e kasi rā ājor kardan* (lit. ‘make someone’s bread a brick’, meaning ‘to deprive someone of their means of survival’) in Persian. On the other hand, the latter is formally similar but functionally different from *das ist ein hartes Brot* in German (lit. ‘it is a hard bread’, meaning ‘this situation is difficult’). For this research, the ability to link constructions across languages based on varying degrees of formal and functional similarity was critical.

Another **multimodality** project investigated whether gesture is an inherent component of certain constructions or whether phonology and ges-

ture operate as separate, simultaneously activated streams (labdounane, 2025). This work identified inherently multimodal constructions such as “This close to V-ing”. Its accompanying gesture (holding two fingers close together) appears to be entrenched alongside the verbal form. This finding suggests that the construction is truly multimodal, integrating both linguistic and gestural elements. For the RCxn, this posed a specific challenge: while it was necessary to include features for describing gestures, gesture could not be treated as a typical construction element. Unlike lexical or syntactic components, the duration and timing of gestures are more flexible.

A third project, situated in the domain of **computational linguistics**, examined how quantitative data from large corpora can be used to identify constructions in a (semi-)automatic manner (Patel, in prep). Central to this research is the use of coloprofiles, which are quantitative representations of a construction’s collocational behavior. A coloprofile captures the statistical distribution of lexical items or grammatical patterns that co-occur with a construction, providing insights into its formal and functional properties. For example, the construction “come + [noise/manner of movement] V-ing” is particularly frequent with verbs such as running, tumbling, hurting, and bursting, which, in turn, confer a connotation of sudden, uncontrolled action upon the construction. It was essential for the RCxn to integrate coloprofiles into its database and ensure they are accessible and visualizable on the web interface.

Some challenges in designing the RCxn were about the representation of constructions in a general manner. For example, we needed to address how to capture the compositional aspects of a construction’s meaning, how to model the nesting of construction elements (e.g., specific properties of a preposition within a prepositional phrase), and how to represent the relationship between a construction A occurring within a construction B and construction B itself. Other challenges, however, were specific to individual projects.

4. Technical Background

To the best of our knowledge, all existing reference constructions rely on relational databases for their implementation. In contrast, our RCxn adopts Semantic Web Technologies, whose graph-based format is appealing for both theoretical and practical reasons. Theoretically, a **knowledge graph representation** aligns seamlessly with the principles of Construction Grammar, particularly as articulated by Diessel (2019, 2023).

Diessel’s model conceptualizes linguistic knowledge as a complex network comprising two interre-

lated dimensions. First, each individual construction (or linguistic sign) forms a network in its own right, as functional features are linked to semantic features, and construction elements are interconnected through sequential dependencies or valency relationships. Diessel refers to this internal organization as the “sign as a network”. Second, constructions themselves are embedded within a broader network, where schematic constructions serve as nodes in an inheritance hierarchy, and others are linked through derivational processes (e.g., a construction with a literal meaning with a formally similar construction expressing the metaphorical meaning) or shared similarities (e.g., phonetic similarities). This external organization is described as the “network of signs”. Knowledge graphs are uniquely suited to model both the internal structure of a construction (the “sign as a network”) and the relations between constructions (the “network of signs”), offering a theoretically grounded representation.

Our approach is further inspired by formalisms such as Head-Driven Phrase Structure Grammar (HPSG) and Sign-Based Construction Grammar (SBCG), where the internal structure of signs (e.g., a word, phrase, or construction) interconnects syntactic, semantic, and phonological properties within a single unit. For example, a verb’s description might link its argument structure (e.g., subject and object requirements) to the respective semantic roles of said arguments (e.g., agent, patient), capturing how form and meaning are integrated within the sign itself. This corresponds very closely to Diessel’s “sign as a network”. At the same time, HPSG and SBCG also model “networks of signs” through hierarchical relationships, such as inheritance hierarchies or type hierarchies. For example, the head-modifier construction inherit properties from a more general headed construction. Formally, these frameworks use attribute-value matrices, which can very straightforwardly be conceived as graphs (with features as edges and values as nodes).

Practically, the use of **RDF (Resource Description Framework)** offers several advantages. Its flexible schema allows for heterogeneous construction descriptions, where certain features can be omitted without implying absence or default values. This can be seen as a drawback to those interested in highly formalized representations. However, in the reality of capturing linguistic research, this flexibility is essential. Researchers bring diverse interests and analytical depths to their work: some may prioritize fine-grained syntactic or semantic details, while others focus on other aspects (like interconnection with other constructions). A rigid schema would risk excluding valuable contributions.

Additionally **Linked Open Data** principles allows

for interoperability and community collaboration. By developing and sharing this ontology, we enable other researchers to build upon our work, while also ensuring compatibility with related projects, such as the ACoLi Dictionary Graph (Chiarcos et al., 2020). Our approach builds on pioneering efforts in linguistic ontology development, including GOLD (General Ontology for Linguistic Description; Farrar and Langendoen 2010), OLiA (Ontologies of Linguistic Annotation; Chiarcos and Sukhareva 2015; Chiarcos et al.), and lemon (W3C OntoLex Community Group), which have demonstrated the value of Semantic Web Technologies for linguistic resource representation.

The ontology serves as terminological baseline for the database containing construction descriptions, which we refer to as the A-box.² Figures 1, 2 and 3 illustrate the representation of a construction description in our database, showing a simplified version of the English “come + [noise/manner of movement] V-ing” construction. This construction is used to express that someone or something moves towards a reference point in a specific manner.³ An example of this construction is *Sally came running excitedly into the room*, where someone (identified by the proper name *Sally*) moves toward the reference point (the inside of the room) in a very fast manner (denoted by the verb *run*). Figure 1 offers a visualization of the construction level. The construction consists of six elements: the (optional) subject, the verb *come*, a V-ing form of a verb expressing noise or manner of movement, an (optional) source, an (optional) trajectory and an (optional) goal. The IRI (Internationalized Resource Identifier) for the construction, `cx:comePLUSnoisemannerofmovementVing` (shortened in Figures 1 and 2 as `cx:CNM`) is related to its title, metadata, meaning and a sequence of its construction elements (see section 5.3.1). The IRI for the metadata (shortened to `cx:CNM_MD`) links to information about the creation date of the entry and the doctoral researcher who contributed this construction (`membr:Patel`). The researcher is in turn linked to their PhD project (`membr:Project_Patel`).

Figure 2 describes further two of the construction elements, the mandatory elements 2 and 3, and should therefore be seen as an expansion of two nodes of the previous graph. Both are linked to their formal description, a node of type `rcxn:SlotForm`, that links to other constructions.

²The A-box (where A stands for assertional knowledge), is stored separately and therefore uses different prefixes than the ontology, also called T-Box (where T stands for terminological knowledge).

³The full entry can be viewed online at https://bdlweb.phil.uni-erlangen.de/RCxn/app_entries/construction/comePLUSnoisemannerofmovementVing.

Element 2 is indeed an instance of the “come” construction and Element 3 of the “Verb+ing” construction.

The simplified description of this construction uses classes and properties defined by the Constructicography module (prefix `rcxn:`; section 5.3.1) of our ontology, by the RDF ontology (prefix `rdf:`; <http://www.w3.org/1999/02/22-rdf-syntax-ns#>), by the FOAF ontology (prefix `foaf:`; <http://xmlns.com/foaf/0.1/>) and by the DCMI Metadata Terms ontology (prefix `dcterms:`; <http://purl.org/dc/terms/>). The prefixes `cx:` and `membr:` refer to the A-box of the RDF database, where the IRIs for constructions (and their features) and for researchers (and their research questions) are stored, respectively.

To keep the visualization of Figures 1 and 2 comprehensible, we omitted many details of the construction description. The complete A-box contains the description of the other construction elements and their word order, the examples that illustrate the construction, research data such as coloprofile for Element 3, and further description of the researcher and their research question. Figure 3 displays fragments of the code in the A-box that correspond to what is shown in the visualization.

5. The Ontology

Having outlined the theoretical motivations and technological foundations of our project, we now turn to the core contribution of this paper: the modular ontology designed to be used in the RCxn. These modules share a common purpose (the RCxn’s database) and are designed to document both linguistic resources and linguistic research.

5.1. Modular Ontology

Some of the modules in our ontology document **linguistic resources** while addressing foundational questions about the formal and functional properties of constructions. For instance:

- How many construction elements (e.g., slots, lexical items, or morphological markers) comprise a given construction?
- What is the meaning of the construction?
- Under which pragmatic conditions can this construction be used?

Complementing this, some of the modules are targeted at documenting **linguistic research**, while focussing on the scientific process and findings associated with each construction. Key questions here include:

- What is the coloprofile of the construction in a particular corpus?

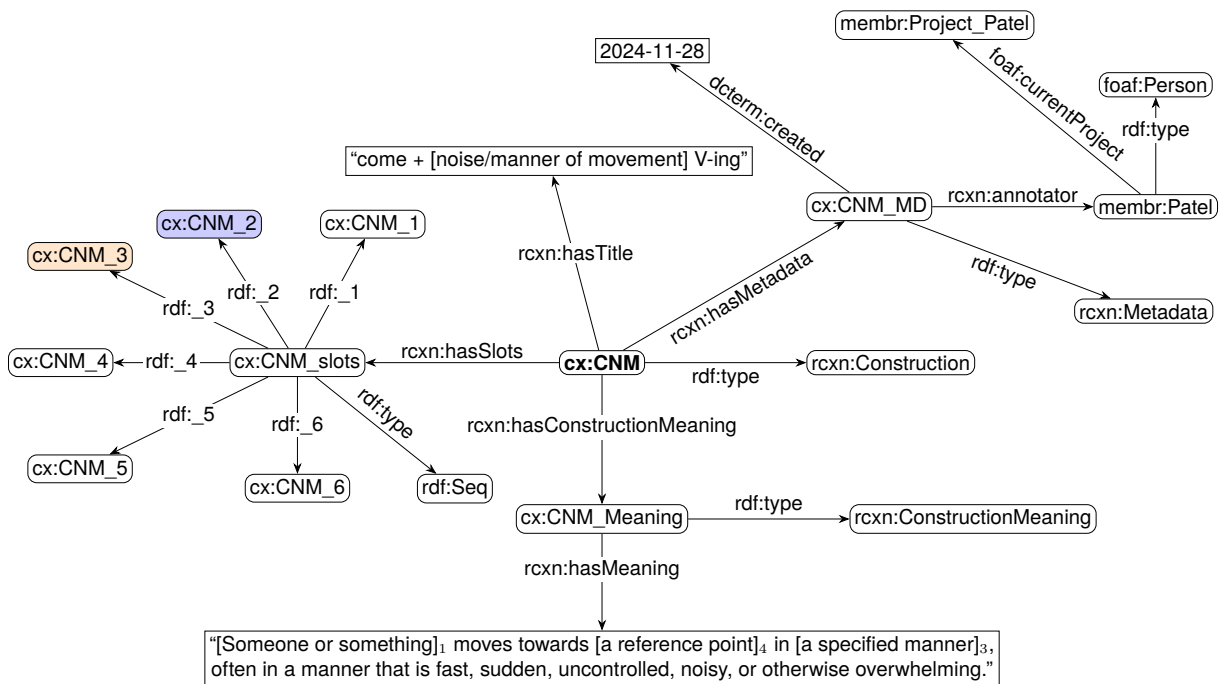


Figure 1: A simplified graph representation of the “come + [noise/manner of movement] V-ing” Construction in the Research Construction. Nodes with round edges denote IRI objects, while nodes with sharp edges denote literal objects.

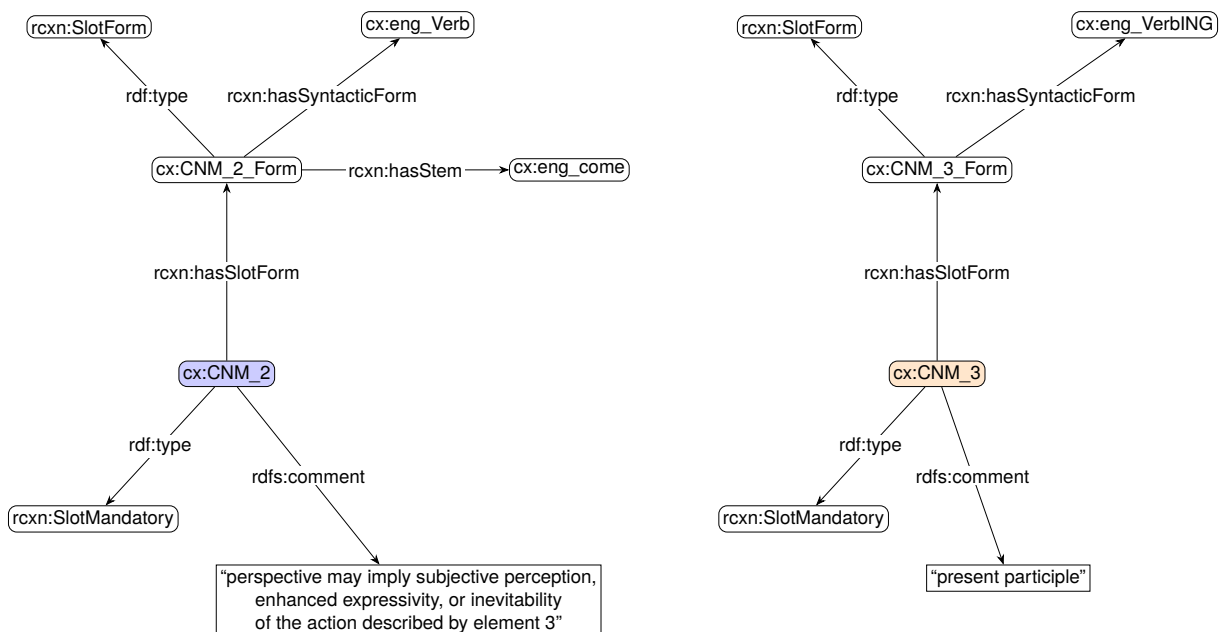


Figure 2: A simplified graph representation of the second and third construction elements of the “come + [noise/manner of movement] V-ing” Construction. Nodes with round edges denote IRI objects, while nodes with sharp edges denote literal objects.

```

membr:Patel a foaf:Person ;
  foaf:currentProject membr:Project_Patel ;
  [...] .

cx:comePLUSnoisemannerofmovementVing a rcxn:Construction ;
  [...]
  rcxn:hasConstructionMeaning cx:comePLUSnoisemannerofmovementVing_Meaning ;
  rcxn:hasMetadata cx:comePLUSnoisemannerofmovementVing_MD ;
  rcxn:hasSlots cx:comePLUSnoisemannerofmovementVing_slots ;
  rcxn:hasTitle "come + [noise/manner of movement] V-ing" .

cx:comePLUSnoisemannerofmovementVing_MD a rcxn:Metadata ;
  [...]
  rcxn:annotator membr:Patel ;
  dcterm:created "2024-11-28"^^xsd:date .

cx:comePLUSnoisemannerofmovementVing_Meaning a rcxn:ConstructionMeaning ;
  rcxn:hasMeaning "[Someone or something]1 moves towards [a refer-
ence point]4 in [a specified manner]3, often in a manner that is fast, sudden, un-
controlled, noisy, or otherwise overwhelming." .

cx:comePLUSnoisemannerofmovementVing_slots a rdf:Seq ;
  rdf:_1 cx:comePLUSnoisemannerofmovementVing_1 ;
  rdf:_2 cx:comePLUSnoisemannerofmovementVing_2 ;
  rdf:_3 cx:comePLUSnoisemannerofmovementVing_3 ;
  rdf:_4 cx:comePLUSnoisemannerofmovementVing_4 ;
  rdf:_5 cx:comePLUSnoisemannerofmovementVing_5 ;
  rdf:_6 cx:comePLUSnoisemannerofmovementVing_6 .

cx:comePLUSnoisemannerofmovementVing_2 a rcxn:SlotMandatory ;
  [...]
  rdfs:comment "perspective may imply subjective perception, enhanced expressiv-
ity, or inevitability of the action decribed by element 3" ;
  rcxn:hasSlotForm cx:comePLUSnoisemannerofmovementVing_2_Form .

cx:comePLUSnoisemannerofmovementVing_2_Form a rcxn:SlotForm ;
  rcxn:hasStem cx:eng_come ;
  rcxn:hasSyntacticForm cx:eng_Verb .

cx:comePLUSnoisemannerofmovementVing_3 a rcxn:SlotMandatory ;
  [...]
  rdfs:comment "present participle" ;
  rcxn:hasSlotForm cx:comePLUSnoisemannerofmovementVing_3_Form .

cx:comePLUSnoisemannerofmovementVing_3_Form a rcxn:SlotForm ;
  rcxn:hasSyntacticForm cx:eng_VerbING .

```

Figure 3: Fragments of the A-box of the RCxn: RDF-Turtle code for the “come + [noise/manner of movement] V-ing” Construction. Some triples have been omitted to enhance readability.

- What research questions motivated the investigation of this construction?
- What empirical findings or theoretical insights have researchers reported about the construction, and how do these contribute to broader debates in Construction Grammar?

Together, this modular ontology ensure that the RCxn not only documents the *what* of linguistic constructions but also the *how* and *why* of the scientific research.

Rooted in the theoretical framework of Construction Grammar, we did our best to make this ontology inclusive to all currents in Construction Grammar. While their immediate application is tailored to the needs of our RCxn, their development reflects a broader ambition: to provide a shared resource for the the community. By adopting open standards and modular design principles, we aim to facilitate interoperability with other constructicography projects, enabling a community-driven ecosystem.

5.2. Technical Implementation and Standards

The development of our ontology builds upon established semantic web standards. We used existing ontologies such as RDFS (Resource Description Framework Schema) for foundational modeling, FOAF (Friend of a Friend) for describing researchers and their contributions, SKOS (Simple Knowledge Organization System) for conceptual organization, and especially OLiA (Ontologies of Linguistic Annotation) for linguistic features in describing constructions.

The ontology is formalized in OWL (Web Ontology Language), a standard for defining and instantiating ontologies on the Semantic Web. We used the ontology editor Protégé to write the different modules (except the `comcon` module, as explained below). In the OWL standard, classes are defined as `owl:Class` and properties as `owl:ObjectProperty`. Classes represent the categories of linguistic and research entities—such as constructions (`rcxn:Construction`), construction elements (`rcxn:Slot`), research questions (`rsrch:Project`), or language variety (`lg:variety`), while properties describe the relationships between these entities, such as a construction belonging to a specific language (`lg:partOfLanguage`), having specific construction elements (`rcxn:hasSlots`), or a researcher working on a specific research question (`rsrch:hasResearchQuestion`).

Each class and property is assigned a unique IRI (Internationalized Resource Identifier) for global identification, along with human-readable labels (`rdfs:label`) and, where applicable, alternative labels (`skos:altLabel`). Definitions are provided using `rdfs:comment`. Hierarchical relationships between classes are expressed using `rdfs:subClassOf`. Some OWL features can be used to enrich the description, such as `owl:disjointWith`, which states that two classes are not mutually compatible (a single IRI cannot belong to both classes simultaneously). For example, the class `rcxn:SlotMandatory` is defined as a subclass of `rcxn:Slot`, disjoint from the class `rcxn:SlotOptional` (i.e., a construction element is either optional or non-optional); its label is “Non-optional element”, and its definition (`rdfs:comment`) is “Non-optional construction elements that need to be realized.”

Hierarchical relationships between properties are expressed using `rdfs:subPropertyOf`. For properties, we employ OWL features such as `owl:SymmetricProperty`, `owl:TransitiveProperty` or `owl:equivalentProperty` where relevant.

For illustrative purpose, Figure 4 displays the code for the property

`rcxn:hasSemanticContribution`, which links a construction element to the semantic meaning it brings to the construction. Our contributors can use OLiA’s rich vocabulary for semantic roles to describe this semantic contribution, in which case the subproperty `rcxn:hasSemanticRole` is used. This property ranges over an object of the class `olia:SemanticRole` and is indicated as being an equivalent to `olia:hasSemanticRole`, thus ensuring as many interoperability with OLiA as possible.⁴ However, contributors who are not willing to use the terminology of semantic role can use the other subproperty `rcxn:hasOtherSemanticContribution` whose object is a literal object.

5.3. Brief Overview of the Modules

The RCxn is supported by a suite of ontology modules. A comprehensive technical documentation of the ontology and its modules is available on the project’s GitHub repository.⁵ Below, we provide a concise overview of each module and its role within the RCxn.

5.3.1. The Constructicography Module (`rcxn.rdf`)

The Constructicography module, stored in `rcxn.rdf` and accessible via the prefix `rcxn:` (for <https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#>), forms the core of our framework for representing constructions in a Cxn. The prominent class of this module is `rcxn:Construction`, which serves to identify an IRI as a construction. Each construction is linked to a human-readable title (as a literal object) and associated metadata (an object of class `rcxn:Metadata`).

In line with the principles of Construction Grammar, the construction is also linked to a form and a meaning. Meaning is represented as an object of the class `rcxn:Meaning`, and can be further described with the help of associated classes and properties. Form, however, is not treated as a monolithic entity, but rather decomposed into a sequence of construction elements (objects of type `rcxn:Slot`). This design applies uniformly, whether the construction consists of a single element (e.g., morphemes or monomorphemic lexemes) or multiple elements. Each slot can then be further characterized by a series of features,

⁴It is technically not possible to use the property `olia:hasSemanticRole` directly because it links a linguistic annotation to a semantic role, whereas we need a property that takes a construction element as subject.

⁵https://github.com/ElodieWinckel/RCxn/blob/master/ontologies/ontology_doc.md

```

<owl:ObjectProperty rdf:about="https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#hasSemanticContribution">
  <rdfs:domain rdf:resource="https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#Slot"/>
  <rdfs:comment xml:lang="en">Describes the meaning (e.g., semantic role) of the construction element in the construction.</rdfs:comment>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#hasSemanticRole">
  <rdfs:subPropertyOf rdf:resource="https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#hasSemanticContribution"/>
  <rdfs:range rdf:resource="http://purl.org/olia/olia-top.owl#SemanticRole"/>
  <rdfs:comment xml:lang="en">Describes the semantic role of the construction element in the construction. The semantic roles are de-
  fined by the OLiA ontology.</rdfs:comment>
  <rdfs:label xml:lang="en">Semantic role</rdfs:label>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://purl.org/olia/olia.owl#hasSemanticRole">
  <owl:equivalentProperty rdf:resource="https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#hasSemanticRole"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#hasOtherSemanticContribution">
  <rdfs:subPropertyOf rdf:resource="https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rcxn#hasSemanticContribution"/>
  <rdfs:range>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://www.w3.org/2002/07/owl#topDataProperty"/>
      <owl:someValuesFrom rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal"/>
    </owl:Restriction>
  </rdfs:range>
  <rdfs:comment xml:lang="en">Describes the meaning of the construction element in the construction that cannot be captured by the seman-
  tic roles of the OLiA ontology.</rdfs:comment>
  <rdfs:label xml:lang="en">Semantic contribution</rdfs:label>
</owl:ObjectProperty>

```

Figure 4: Fragment of the ontology definition of rcxn: RDF-XML code for the property rcxn:hasSemanticContribution and its two supproperties rcxn:hasSemanticRole and rcxn:hasOtherSemanticContribution

many of which are drawn from the OLiA ontology (e.g., `olia:hasAnimacy`, with possible values like `olia:Inanimate`). The module defines however a series of features, for example `rcxn:hasPhonology` (which takes as object a literal object).

Some classes and properties are furthermore dedicated to the bibliographical documentation of the construction, and to other aspects of its metadata (for example `rcxn:annotator`).

5.3.2. The Research Module (`rsrch.rdf`)

The Research module, stored in `rsrch.rdf` and accessible via the prefix `rsrch`: (<https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/rsrch#>), defines classes and properties that document the research process itself, including research questions, findings, and the researchers involved. For example, the class `rsrch:Project` represents individual research initiatives, and `rsrch:Finding` encapsulates empirical results, theoretical insights, or methodological contributions derived from these projects and from the study of individual constructions. To represent the researchers and their contributions, the ontology uses the FOAF (Friend of a Friend) vocabulary, enabling detailed descriptions of individuals.

This module is essential to the RCxn project that aims not only at documenting constructions but also the evolving scientific discourse surrounding them. Constructicography projects that do not concentrate on this aspect might have no need of this module.

5.3.3. The Links Module (`links-1.0.rdf` and beyond)

The Links module is dedicated to modeling the relationships between constructions. It is currently available in two versions (`links-1.0.rdf` and `links-1.1.rdf`) under the prefix `links:`. The initial version (`links-1.0.rdf`) was the topic of a previous publication (Winckel, 2025) aimed at a non-computational audience, focusing on lexicographic and theoretical issues in Construction Grammar. Evolutions in the project led to the development of `links-1.1.rdf`, which remains fully backward-compatible with the 1.0 version. This is a commitment to compatibility we intend to maintain in all future updates.

This ontology is grounded in the theoretical framework proposed by Diessel (2019, 2023), capturing both vertical relationships, that is inheritance links (e.g., a specific construction inheriting properties from a more general schema), and horizontal relationships, also called sister constructions, that share functional or formal similarities. Additionally, the ontology accommodates a cross-linguistic dimension, enabling comparative analyses of constructions across languages for projects in contrastive linguistics. Beyond these core relationships, the ontology also supports other types of links, such as metaphorical relationships, which highlight semantic extensions or conceptual mappings between constructions.

5.3.4. The Comparative Concepts Module (`compcon.ttl`)

The Comparative Concepts module, stored in `compcon.ttl` and accessible via the prefix `compcon`: (<https://bdlweb.phil.uni-erlangen.de/RCxn/ontologies/compcon#>), is dedicated to facilitating

cross-linguistic comparisons of constructions. Croft (2022) proposed a series of Comparative Concepts, which are abstract, typologically grounded categories (e.g., the construction category “free relative clause construction”, or the semantic category “agent”) that serve to compare constructions across Cxn, and therefore languages. This module is part of the broader MoCCA initiative (Lorenzi et al., 2024), a collaborative effort among Cxn projects to leverage the typological power of Comparative Concepts in order to align construction descriptions.

While the RCxn actively participates in the MoCCA enterprise, the `compcon.ttl` ontology itself is not an original development of our project. Instead, it represents an RDF translation of a pre-existing database of comparative concepts, originally developed by the MoCCA team and based on Croft (2022). This database is publicly available on their GitHub repository as a YAML database⁶, and the aim of this module is to ensure interoperability of our RCxn with their framework.

5.3.5. Additional Modules: Language and Project-Specific Resources

In addition to the core modules, the RCxn includes a few auxiliary modules designed for specific purposes. The Language module (`lg.rdf`) serves as a repository for language instances used within the RCxn, providing definitions for the languages documented in our RCxn. While functional for our internal workflows, this module is highly project-specific. The CASA module (`casa.rdf`) is tied to a research project in collaboration with the English CASA Constructicon. Given its project-specific focus, this module does not need be detailed further here. It is available for reference in the project’s GitHub repository.

6. Conclusion

6.1. Summary

This paper has presented the theoretical motivations, technical implementation, and modular design of the ontology developed for our RCxn project. We began by outlining the need for a dynamic, community-driven resource that documents linguistic research. We showed that a network-based approach is particularly suited to model descriptions rooted in Construction Grammar. Semantic Web Technologies have the further advantage of being flexible and the interoperability with other linguistic ontologies is an important asset.

We presented a modular ontology, with each module tailored to address specific aspects of con-

structionist research. The Constructicography module captures the core structure of constructions, modeling their form and meaning. The Research module documents the scientific process, linking research questions, findings, and contributors to the constructions they investigate. The Links module formalizes the relationships between constructions, while the Comparative Concepts module aligns with the MoCCA initiative to enable cross-linguistic comparisons.

6.2. Future of the Project

In the near future, we plan to identify and implement additional classes and properties to further refine the ontology. We are also exploring the development of a dedicated module to document empirical evidence for constructions, as well as an additional module to model examples of constructions, which require specific features such as glosses or translations. However, we remain open to using existing ontologies for this purpose if they meet our requirements.

Beyond its immediate goals, the RCxn project offers several broader benefits for the field of Construction Grammar. It contributes indeed to linking resources in this field. This makes future connections with other Cxn projects possible and enable a more holistic perspective on construction knowledge. The RCxn serves as a model for future projects, demonstrating how constructicographic work can be integrated with research documentation. The project also provides a robust software infrastructure and a well-designed ontology that can be adopted by other projects in Construction Grammar.

The project will remain accessible on the FAU server, ensuring that all current IRIs and resources stay available. The ontology will continue to be open for use and expansion by the community, even if no further development occurs. We aim, however, to secure additional funding to further enhance the project.

This work is very much a work in progress, and we welcome feedback and collaboration from the Construction Grammar community. With this an open and adaptive resource, we hope the RCxn will serve as a valuable tool in constructicography.

7. Acknowledgements

This research was funded by the German Research Foundation (DFG) as part of the Research Training Group *Dimensions of Constructional Space* (project no. 468527017).

We are deeply grateful to the members of the Research Training Group for their invaluable contributions throughout this project. We also extend

⁶<https://github.com/comparative-concepts/cc-database>

our sincere thanks to three anonymous reviewers for their insightful feedback.

8. Bibliographical References

- Christian Chiarcos and Maria Sukhareva. 2015. *OLiA – Ontologies of Linguistic Annotation*. *Semantic Web*, 6(4):379–386.
- William Croft. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford University Press, Oxford.
- William Croft. 2022. *Morphosyntax: Constructions of the World's Languages*. Cambridge University Press, Cambridge.
- Holger Diessel. 2019. *The Grammar Network: How Linguistic Structure Is Shaped by Language Use*. Cambridge University Press, Cambridge.
- Holger Diessel. 2023. *The Constructicon: Taxonomies and Networks*. Elements in Construction Grammar.
- Adele E. Goldberg. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.
- Thomas Herbst and Thomas Hoffmann. 2024. *A Construction Grammar of the English Language: CASA – a Constructionist Approach to Syntactic Analysis*. John Benjamins Publishing Company.
- Yassine Iabdounane. 2025. *Multimodal Constructional Space*. Ph.D. Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.
- Arthur Lorenzi, Peter Ljunglöf, Ben Lyngfelt, Tiago Timponi Torrent, William Croft, Alexander Ziem, Nina Böbel, Linnéa Bäckström, Peter Uhrig, and Ely E Matos. 2024. *MoCCA: A Model of Comparative Concepts for Aligning Constructions*. In *Proceedings of the 20th Joint ACL - ISO Workshop on Interoperable Semantic Annotation @ LREC-COLING 2024*, pages 93–98, Torino, Italia. ELRA and ICCL.
- Malin Patel. in prep. *Corpus Evidence for Delineating Constructions*. Ph.D. Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.
- Aria Rastegar. 2026. *Understanding idioms across languages*. Ph.D. Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.
- Elodie Winckel. 2025. *Defining relationships in the constructional network: A Semantic Web ontology for Construction Grammar*. *Lexicographica*, 41(1):299–317.

9. Language Resource References

- Kardelen Babal. 2026. *‘Konstruktion: Intensivierung Attribut nachgestellt:NP_pur’ in FrameNet-Konstruktikon des Deutschen*. <https://framenet-constructicon.hhu.de/constructicon/construction?id=1323> (accessed 19.02.2026).
- Chiarcos, Christian and Adam, Angelika and Hellmann, Sebastian and Sukhareva, Maria and Fäth, Christian and Abromeit, Frank and Ionov, Maxim and Dimitrova, Vanya. *Ontologies of Linguistic Annotation (OLiA)*. Applied Computational Linguistics (ACoLi) Lab at the Goethe University Frankfurt, Germany.
- Chiarcos, Christian and Fäth, Christian and Ionov, Maxim. 2020. *The ACoLi Dictionary Graph*. European Language Resources Association.
- DELPH-IN Consortium. DELPH-IN: Deep Linguistic Processing with HPSG.
- Scott Farrar and D. Terence Langendoen. 2010. *General Ontology for Linguistic Description*. Institute for Language Information and Technology.
- FKD. *FrameNet-Konstruktikon des Deutschen*. <https://framenet-constructicon.hhu.de> (accessed 19.02.2026).
- Herbst, Thomas and Hoffmann, Thomas and Uhrig, Peter and Garibyan, Armine and Evert, Stephanie. 2023–. *CASA ConstructiCon of the English Language*.
- W3C OntoLex Community Group. *lemon - The Lexicon Model for Ontologies*.