CLAVELL – A Cognitive Linguistic Annotation and Visualization Environment for Language Learning

Werner Winiwarter
University of Vienna, Faculty of Computer Science
Währingerstrasse 29, 1090 Vienna, Austria
werner.winiwarter@univie.ac.at

Abstract

In this paper, we introduce a novel sentence annotation method based on Radical Construction Grammar and Uniform Meaning Representation, covering multiple levels of linguistic analysis, ranging from interlinear morphemic glossing to PropBank rolesets, WordNet synsets, and Wikipedia page titles as concept identifiers. We visually enhance our annotation by using images to represent concepts, emojis for roles, and color-coding for constructions, the fundamental concept of Construction Grammar. The meaning representation is embedded into the syntactic parse by aligning all concepts with the surface tokens in the sentence. The main motivation for developing this type of representation was its use in second language acquisition as part of a Web-based language learning environment. By engaging in entertaining annotation tasks students assemble incrementally the representation using a bottom-up strategy. Based on language exposure while performing these exercises, we populate personal idiolectal constructicons representing the students' current status of second language comprehension. To showcase our system, we have implemented it for Japanese because of its soaring popularity in our language education program and the particular problems it poses to those trying to learn this language, especially Westerners.

Keywords: cognitive linguistics, Radical Construction Grammar, Uniform Meaning Representation, visual annotation, second language acquisition, Web-based language learning, constructicon, Japanese linguistics

1. Introduction

In the shadow of the latest LLM hype, there have been nonetheless significant developments in the field of cognitive linguistics. Two remarkable recent achievements have been Radical Construction Grammar (RCG) and Uniform Meaning Representation (UMR), which pave the way for cross-linguistic semantic annotation of documents. Building on this research work, we have extended the representation towards an interlingual annotation by linking all concepts to knowledge bases, mapping PropBank core arguments to interpretable roles, aligning concepts with words, and enhancing the display with visual elements.

Language understanding is a multi-step process where a signal is broken down into smaller units, e.g. words, morphemes, sounds, or letters, to be then interpreted in terms of meaning. Put differently, to understand, we need to recognize the categories standing for meaning or meaning relations. Categorization is a fundamental process needed both to understand the meaning of a sentence as well as to understand the regularities of the mappings between "meanings" and "forms". Thus, students have to learn two kinds of language, the target language, i.e. the one they are exposed to, and a meta-language allowing them to describe regularities.

Language learning is already difficult but it may well be even more of a challenge as one attempts to learn a typologically different language. For example, a European trying to learn Chinese has to attune his ears as Chinese has not only many unknown sounds but also specific intonations (tones) for each "word". Next to mastering the sound system, Chinese is even more intimidating when it comes to reading and writing. Regarding this aspect, things are even worse for Japanese, which has three writing systems: two for the sound (hiragana and katakana) and one mixed for meaning and sound (kanji). Drawing heavily on Chinese characters, kanji encode meaning and sound, though not always in a very regular form. To make things even more complicated, kanji most often have at least two kinds of pronunciation, one based on Chinese, the other based on Japanese. All these are difficult hurdles for the learner of Japanese. As we have already addressed many of these issues in our previous work (Wloka and Winiwarter, 2021a,b; Winiwarter and Wloka, 2022), we will focus here only on using our annotation for Japanese language learning. The goal is to alleviate the burden of converting forms to meaning and to ease the gaining of certain insights concerning the mechanics and functioning of language.

Due to the worldwide manga craze there has been an unprecedented increase in demand for Japanese language courses. As a result, there have been requests from our language education department for technological support. Our Web-based annotation exercises are supposed to im-

The Workshop on Cognitive Aspects of the Lexicon, pages 1–13

prove language learning in an appealing way. One central component of our environment are personal constructicons, which reflect each student's progress, proficiency, and individual learning path. Aggregating this knowledge is very useful as it offers invaluable insights for instructors and course designers alike.

This paper is organized as follows. In section 2, we provide some background information, citing relevant related work. In section 3, we discuss the representation of major propositional acts in RCG. In section 4, we describe implementation details, and finally, in section 5, we provide an outlook towards future work.

2. Background and Related Work

2.1. Construction Grammar

Construction Grammar (Hoffmann and Trousdale, 2013; Hilpert, 2014; Ungerer and Hartmann, 2023) is actually a broad family of theories in the area of *cognitive linguistics* (Croft and Cruse, 2004). Their common denominator is the view that *constructions* are the basic units of language, which are pairings of *form* and *meaning*. Researchers working in this paradigm/framework reject the separation of the grammar and the lexicon. Both are considered to be constructions.

Radical Construction Grammar (Croft, 2001, 2022) (RCG) was designed with *typological* applicability as main motivation. Linguistic *typology* (Croft, 2002) studies and classifies languages according to their structural features to allow their comparison. RCG considers word classes and other syntactic structures as language-specific and construction-specific (Croft, 2023). In a recent supplement to (Croft, 2022), the author provides a taxonomic/partonomic tree of constructions[1], therefore supporting an ontological view on constructions as a structured set of concepts, which, if properly combined represent the meaning underlying a sentence. For more background on this line of thought we refer to (Zock et al., 2008; Borgo et al., 2019).

One of the main consequences of Construction Grammars is that the traditional divide between lexicon and grammar is abandoned. Everything is a construction. For example, the lexicon is replaced by a network of constructions, words being constructions. Most attempts to assemble such a *constructicon* are in the context of the *FrameNet*[2] project (Lyngfelt, 2018). Whereas the original model was a *taxonomical inheritance network*, recent research on *usage-based linguistics* (Divjak, 2019) has led to a *multidimensional association network* (Diessel, 2023). This is in line with most of the work done on lexical graphs: WordNet (Fellbaum, 1998; Miller, 1990), FrameNet (Fillmore et al., 2003), BabelNet (Navigli and Ponzetto, 2012), VerbNet (Kipper-Schuler, 2005), MindNet (Richardson et al., 1998), HOWNET (Dong and Dong, 2006), ConceptNet (Liu and Singh, 2004; Speer et al., 2017), YAGO (Suchanek et al., 2007), DBnary (Sérasset, 2015), or JeuxDeMots (Lafourcade et al., 2015). Recently, related experiments were conducted to explore the role of *language exposure* on emergence and fading of constructions in the constructicon (Dunn, 2022). This novel interpretation of constructicons is also highly compatible with most modern views on the mental lexicon (see Papafragou et al., 2022; Zock and Biemann, 2020; Zock, 2022).

2.2. Meaning Representations

The annotation of sentences with *meaning representations* has established itself in the last decade as a thriving research field in computational linguistics (see Abend and Rappoport, 2017). The most influential and most actively promoted approach has been the *Abstract Meaning Representation*[3] (AMR) (Banarescu et al., 2013). There are many parsers available, the best[4] one being at the moment Lee et al. (2022). The *SPRING* parser can be tried via a Web interface[5], which also offers a nice visualization. One point of criticism concerning AMR is its reliance on numbered, hence not directly interpretable core arguments; it is addressed by the *WISeR* meaning representation (Feng et al., 2023a), which maps them to thematic roles.

AMR has been recently extended to the *Uniform Meaning Representation*[6] (UMR) (Gysel et al., 2021). It enhances AMR by adding support for other languages (in particular low-resource languages), and a document-level representation capturing intersentential coreference and temporal/modal dependencies. There is an upcoming workshop to kick-start the development of UMR parsers. According to the *UMR guidelines*[7], UMR fully embraces RCG as a theoretical foundation.

2.3. Interlinear Morphemic Glossing

Whereas this topic has been neglected for a long time by natural language processing research, it has a long tradition in linguistics and typology. An *interlinear morphemic gloss* (IMG) represents a text

[1] https://www.unm.edu/~wcroft/Papers/ConstructionRelations.pdf

[2] https://framenet.icsi.berkeley.edu/

[3] https://amr.isi.edu/

[4] https://paperswithcode.com/task/amr-parsing/latest

[5] http://nlp.uniroma1.it/spring/

[6] https://umr4nlp.github.io/web/

[7] https://github.com/umr4nlp/umr-guidelines/

as a string of elements by ideally mapping each morpheme of the source language to a morpheme of the target language or to a grammatical category.

Until the recent past (Leipzig Glossing Rules[8]), there was no common format, which resulted in a confusing variety of glossing styles. This has changed. Recently there have been some efforts by computational linguists to extend and formalize the guidelines (Mortensen et al., 2023) and to automate interlinear glossing (e.g. Samardžić et al., 2015; Zhao et al., 2020; Barriga Martínez et al., 2021). There has even been a first shared task on this topic in 2023 (Ginn et al., 2023).

2.4. Multimodal Resources

Multimodal enhancements of lexical resources have a long history but only recently they have gained momentum due to the interest on *visual question answering* (Lerner et al., 2024) or *multimodal large language models* (Bewersdorff et al., 2024). One example of an attempt towards a multimodal semantic representation is *VoxML* (Pustejovsky et al., 2016).

WordNet has been extended by *ImageNet*, which maps about 1,000 images to each synset (Deng et al., 2009). Another effort to assign cliparts to a small set of synsets was proposed by (Bond et al., 2009). Alas, this project is discontinued. A much more influential resource is *Wikipedia*, which has been increasingly enhanced with visual representations. However, the number of images provided varies widely depending on the language.

The most comprehensive effort is *BabelNet*[9] (Navigli et al., 2021) with the annotation tool *Babelfy*[10] (Moro et al., 2014) and the latest *BabelPic*[11] (Calabrese et al., 2020) dataset targeting abstract concepts.

Even though the use of pictorial illustrations has a long history in language teaching textbooks, there is a crying need for visual representations of meaning representations of sentences.

2.5. Japanese Language

Japanese is an agglutinative SOV language with topic-comment sentence structure. Both agglutinative languages and fusional languages like, for instance, German, are synthetic languages, i.e. they are statistically characterized by a higher morpheme-to-word ratio. In agglutinative languages, words contain multiple morphemes concatenated together in such a manner that individual word stems and affixes can be usually isolated and identified, whereas fusional languages combine multiple grammatical categories into one affix. Therefore, agglutinative languages tend to have more easily deducible word meanings compared to fusional languages, which allow unpredictable modifications in either or both the phonetics or spelling of one or more morphemes within a word.

In Japanese, phrases are exclusively head-final and compound sentences are strictly left-branching. The most noticeable characteristics for foreigners are the lack of articles (a/the), the absence of markers for number (sg. vs pl.) or gender (masculine/feminine), or the fact that adjectives are conjugated. On the other hand, Japanese has a complex system of honorifics, high dependency on context, hence strong likelihood of ambiguity, due to the omission of the subject or the use of zero anaphora.

There are many excellent reference grammars, e.g. Bowring and Laurie (1992); Kamermans (2010); Kaiser et al. (2013), and a lot of research done by Japanese linguists: see Hasegawa (2015) for an introduction; for a recent comprehensive overview we recommend Hasegawa (2018). There is also a wealth of typological studies of Japanese, e.g. Takezawa (1993); Washio (1997); Matsumoto (1997); Taoka (2000); Ohori (2001); Yuasa and Sadock (2002); Iwasaki (2013). Based on the *Japanese FrameNet*[12] project (Ohara et al., 2003), there have been ongoing efforts towards a Japanese constructicon (Ohara, 2014). There is also an important lexicographic work by Backhouse (2016), who organizes Japanese vocabulary according to semantic frames.

One of the main obstacles for getting proficient in Japanese is the complex writing system (see Matsumoto, 2007; Mori, 2014; Paxton, 2019). It uses a combination of logographic *kanji* and two syllabaries *hiragana* and *katakana*. Kanji are adopted Chinese characters. Since 2010 Japanese students are required to learn 2,136 so-called *jōyō kanji* in primary and secondary school.

There exist several *romanization* systems, i.e. using Latin script to write Japanese. The most widely used one is the *Hepburn romanization*, which has several variants, the most common one being the *Revised Hepburn* (see Kudo, 2011). There are many romanization tools, the most easily accessible one for the use in natural language processing software written in Python is *Pykakasi*[13] based on the *kakasi*[14] library.

The most important lexical resource for Japanese is the *Japanese Multilingual dictionary* (JMdict) (Breen, 2004), which can be searched

[8] https://www.eva.mpg.de/lingua/tools-at-lingboard/glossing_rules.php
[9] https://babelnet.org/
[10] http://babelfy.org/
[11] https://sapienzanlp.github.io/babelpic/
[12] https://jfn.st.hc.keio.ac.jp/
[13] https://pypi.org/project/pykakasi/
[14] http://kakasi.namazu.org/

online in combination with many other lexical resources via the *Online Japanese Dictionary Service* (WWWJDIC)[15]. Another very useful online service is *Honyaku Star*[16]. It references numerous dictionaries and corpora and shows translations in context. Honyaku Star includes currently over 2 million translations.

Japanese is also part of the *Open Multilingual Wordnet* (Bond and Paik, 2012)[17], which allows the mapping of Japanese words to English synsets. It is easily accessible via the *NLTK* toolkit[18].

The most prolific linguistic tool for Japanese is certainly the *CaboCha* dependency parser (Kudo and Matsumoto, 2002), which includes the *MeCab* part-of-speech and morphological analyzer (Kudo et al., 2004). More recently, trained pipelines have been added to the popular natural language toolkit *SpaCy*[19]. Another similar solution is *UniDic2UD*[20].

3. Meaning Representation in RCG

In RCG there are two central *comparative concepts* (see Haspelmath, 2010), i.e. theoretical concepts used for crosslinguistic comparison. The first one is the **construction (cxn)**, which is defined as any pairing of form and function in any language to express a particular combination of semantic content and information packaging (see Croft, 2022). The second comparative concept is the **strategy**, which further distinguishes certain characteristics of grammatical form defined in a crosslinguistically consistent way.

There are three fundamental **information packaging** functions that structure phrases and clauses: reference, modification, and predication. They are called **propositional act** functions and correspond to the prototypical **semantic classes**: objects, properties, and actions. However, any semantic class can be packaged in any information packaging function so that we end up with a 3×3 matrix. In the following subsections we provide an example for each cell of this matrix by introducing our annotation for the resulting constructions.

All theoretic concepts defined in (Croft, 2022) are emphasized in boldface. We provide a short definition for each term, for a more detailed description with examples we refer to the voluminous glossary of (Croft, 2022). As much as possible, RCG relies on terms already in use in linguistics, e.g. *reference* or *topic*, and while they try to make their definitions more precise, quite so often they depart from the traditional use.

3.1. Reference

The information packaging function **reference** indicates what the speaker is talking about. The prototypical semantic class are **object** concepts, which include persons, animals, and physical objects.

In Fig. 1, the first basic construction is an example of **object reference**. We annotate the original orthographic representation from the source text with the following information:

- morphemic representation,
- interlinear morphemic gloss,
- translation to concept(s),
- visualization of concept(s),
- construction label.

Figure 1: Three examples of reference.

We use Revised Hepburn romanization for the morphemic representation with some additional information. For example, the capitalized reading "TŌ" indicates that this is a Sino-Japanese reading. It is translated to the WordNet synset `tower.n.01`. By clicking on the image, an enlarged version can be inspected including the synset gloss as caption. The resulting construction is a **referent expression**, i.e. its prototypical use would be as **head** of a

[15] http://wwwjdic.se/
[16] http://honyakustar.com/
[17] https://omwn.org/
[18] https://www.nltk.org/
[19] https://spacy.io/models/ja
[20] https://github.com/KoichiYasuoka/UniDic2UD

referring phrase. Both semantic content and information packaging are color-coded in the annotation. The former as background color of the image, the latter as border color. As this is the prototypical combination, both are drawn in magenta.

The second construction in Fig.1 packages a **property** as reference. Properties are relational, 1-dimensional, usually scalar and stable concepts, which are drawn in green. In this example, the property "utsukushi-" ("beautiful") is translated to the PropBank roleset `beautiful-02`. The suffix "sa" acts as *nominalizer* (NR) to derive the referent expression "beauty".

By analogy, the third construction represents an **action**, i.e. relational, dynamic, and transitory concepts painted in cyan. The action "nusu-" ("steal") is translated to the PropBank roleset `thieve-01`. The continuative ending "mi" again acts as *nominalizer* (NR) resulting in the referent expression "theft".

3.2. Modification

The second propositional act provides additional information about the referent and enriches the specification of the referent for the hearer. The prototypical construction is **property modification** and is shown in Fig. 2. The property "cute" is a **modifier** for the referent expression "dog" (border color green), i.e. the **head** of an **attributive phrase**. By combining the two elements, we get a **modification cxn**, which is defined as a **referring phrase** (color magenta) consisting of a referent expression and one or several attributive phrases.

The grammatical category NPST in the gloss for "kawai-i" indicates the tense non-past, because "kawai-i" is an "i-adjective" that behaves like a verb (or, conceptually speaking, plays the same role as a verb). Therefore, it is also referred to as "verbal adjective" in many reference grammars.

In the meaning representation, the two concepts are linked by a relation with the role MOD (for more details on roles in UMR we refer to[21]). As mentioned before, we use emojis for the roles, in this case a ribbon 🎀. For an optimal alignment with the structural representation, the left-right axis conveys meaning. We do not add arrowheads to the relations because almost all relations in our annotation point from right to left due to the left-branching nature of Japanese language, therefore we define this direction as default interpretation.

Figure 3 shows an example of **object modification**. The postposition "no", indicating the modificaton relation, is not annotated. We allow to omit annotations for frequent monosyllabic postpositions in language learning scenarios because of their abundant use in Japanese and their excessive polysemy and homonymy.

[21] https://github.com/umr4nlp/umr-guidelines/blob/master/guidelines.md

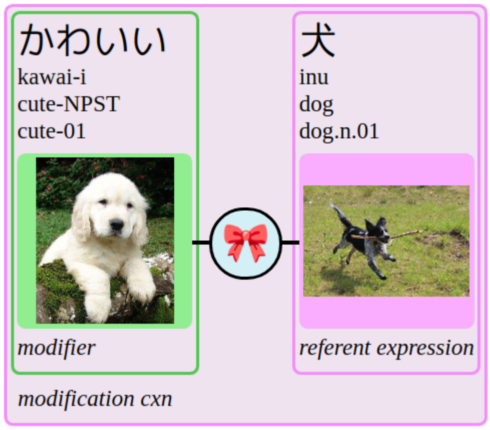

Figure 2: Example of property modification.

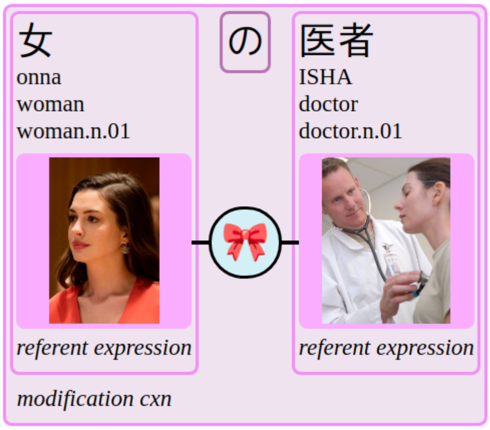

Figure 3: Example of object modification.

The third combination, **action modification**, is realized in Japanese as **relative clause**, which is modeled in RCG together with a **matrix clause** as modifier of the head of the referring phrase, the **relative clause head**, to result in a **relative clause cxn**. In the example in Fig. 4, only the relative clause and the relative clause head are shown.

The action "lose" is packaged as a **predicate**, the **head** of a **clause**. The past tense (PST) is modeled in the meaning representation as a BEFORE (⏪) relation to the special concept *DCT (document creation time)*, colored in orange, representing the present moment. This sequence of concepts is indicated in the gloss by a vertical bar. The information-packaging alternation passive voice (PASS) is not reflected in the meaning representation. Finally, the grave accent in the morphemic representation indicates that the native reading of the second kanji is unvoiced, i.e. "kami" and only changes to "gami" as second part of "tegami". This voicing is called *rendaku* in Japanese.

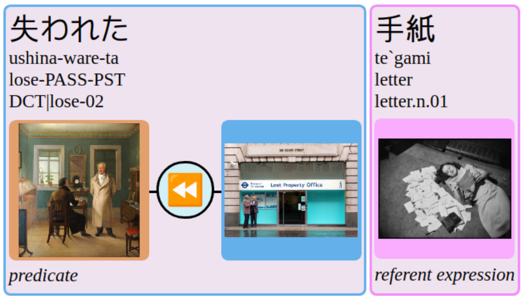

Figure 4: Example of action modification.

3.3. Predication

The third propositional act conveys what the speaker is asserting about the referents in a particular utterance. As we have already given an example of the prototypical construction **action predication** as part of Fig. 4, we focus on the two non-prototypical constructions.

An example of **object predication** can be seen in Fig. 5. The postposition "wa" indicates the **topic** of the sentence, i.e. the referent in a **topic–comment** information packaging that the **comment** is predicated about. In the meaning representation, the object predication is modeled by a special concept, which is aligned with the **copula** (COP) "da". The predication asserts what object CATEGORY (🏷️) the THEME (📜) belongs to.

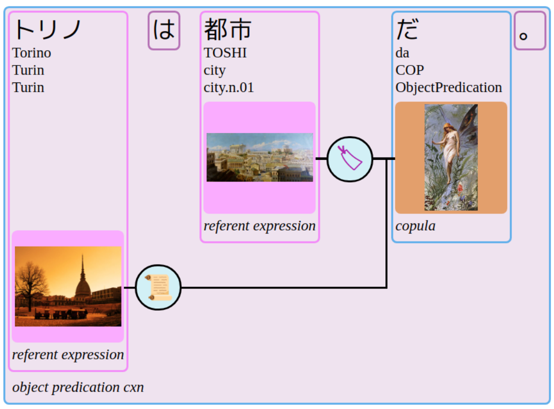

Figure 5: Example of object predication.

Finally, Fig. 6 shows an example of **property predication**. Since "kashiko-i" is also a "verbal adjective" just like "kawai-i" in Fig. 2, there is no copula and the special concept for the property predication is directly appended to the property concept in the meaning representation and linked to it by a PROPERTY (🌈) relation.

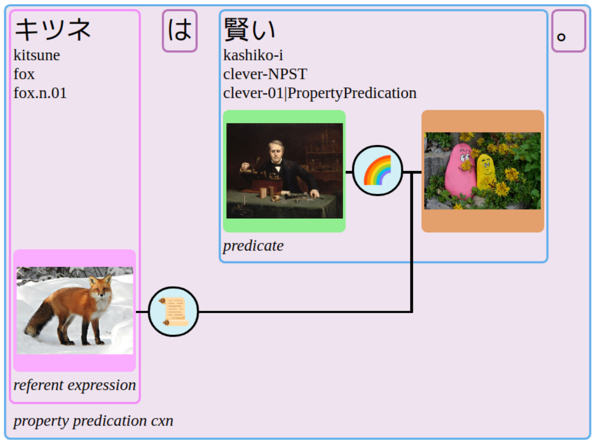

Figure 6: Example of property predication.

4. Implementation

Figure 7 highlights the main components of our system architecture. We have realized our language learning environment as Web-based solution, so that the language students can access the learning server through a Web browser by using augmented browsing enabled through *Chrome extension APIs*[22], and the *jQuery*[23] and *jQuery UI*[24] libraries. Whenever a student loads a new Japanese Web document, it is automatically analyzed and segmented into individual sentences. Each sentence is augmented with an event handler. If a student then clicks on a sentence, it is transferred to the server via *XMLHttpRequests*.

The language learning server is implemented in *SWI-Prolog*[25] (Wielemaker et al., 2012), which is not only predestined for natural language tasks but also provides a scalable Web server solution (Wielemaker et al., 2008) and libraries for efficiently handling RDF and XML files.

The server parses the sentence by using the linguistic knowledge stored in the *personal idiolectal constructicon* (reflecting the student's unique use of the learned language based on past exposure) and dynamically generates an HTML page with the annotated sentence, which is opened in a new tab in the student's browser.

The user can now add new information to the annotation, which is again sent to the server leading to an update of the constructicon, a reparsing of the sentence, and an actualization of the HTML page.

As external resources we use *PropBank Frame Files* (Pradhan et al., 2022), *WordNet* (Princeton University, 2012), *DBpedia* (Lehmann et al.,

[22] https://developer.chrome.com/docs/extensions/reference/api
[23] https://jquery.com/
[24] https://jqueryui.com/
[25] https://www.swi-prolog.org

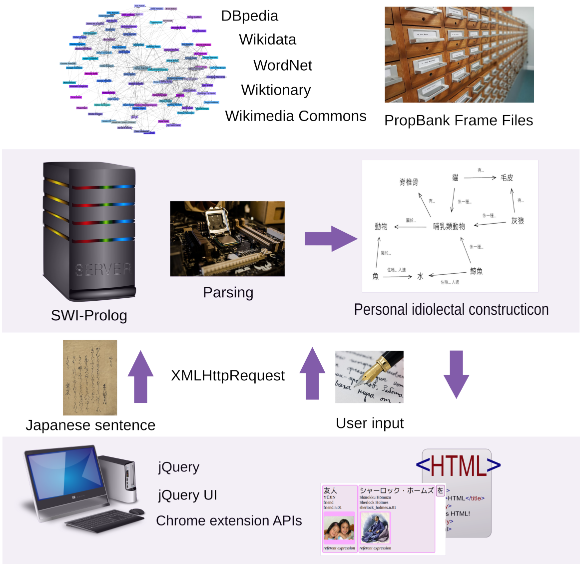

Figure 7: System architecture.

2015), *Wikidata*[26], *Wiktionary*[27], and *Wikimedia Commons*[28].

In Fig. 8, we take a closer look at the user interaction. In order to populate their personal idiolectal constructicons, the language students start at the beginning of their training from scratch with an empty knowledge base.

We have designed the annotation tasks with gaming elements in mind in that we first confront the students with a list of Japanese characters, which is the starting point to assemble the complete annotation step-by-step following a bottom-up strategy.

As Step 1 in Fig. 8, the user can form words by drawing a box with the cursor to select characters. Selected characters are highlighted in orange. As soon as the user releases the mouse button, the display changes to Step 2.

Now, the student is supposed to enter the correct morphemic representation, the interlinear morphemic gloss, and the translation to concept(s). At every step, i.e. for every input, the level of support offered to the user can be increased by displaying select menus or suggestions. This is achieved by accessing the external resources in Fig. 7 as well as the language-specific tools and lexicons mentioned in section 2.5.

As soon as the user adds some information, it is stored in the constructicon, and can be used to learn rules to apply this linguistic knowledge to new examples. Context-sensitive rules are learnt automatically and adjusted incrementally for each new item. We also store the number of times the student was exposed to this item so we can choose

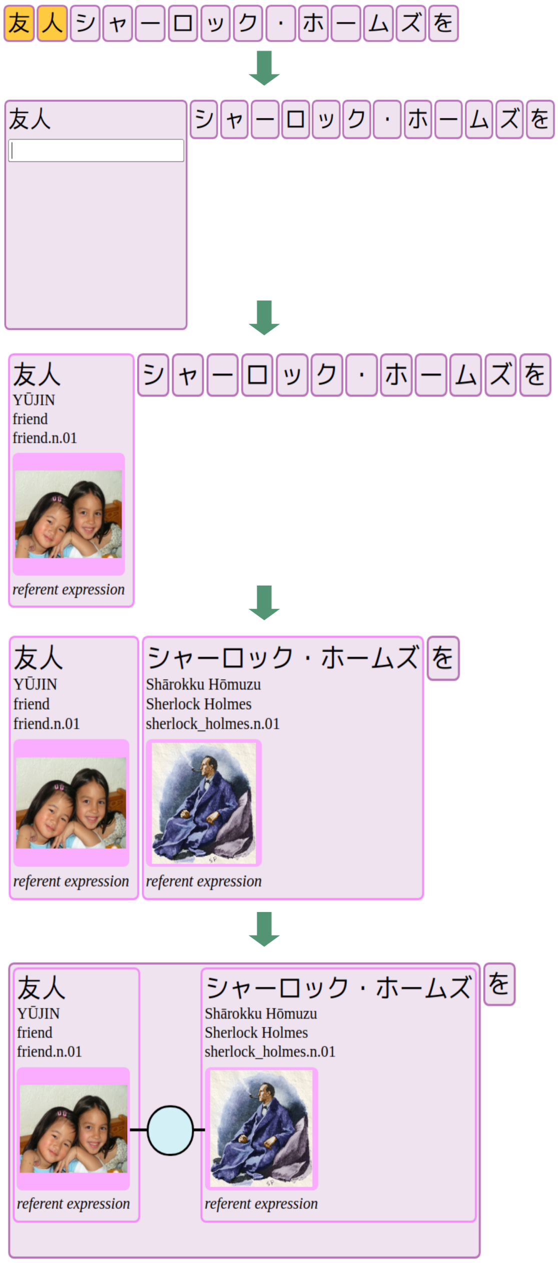

Figure 8: Example of user interaction.

the number of repetitions the student has to perform before the item is inserted automatically.

The information in the constructicon is organized as associative network, which is stored persistently using Prolog fact files. In addition, we offer routines to aggregate and focus on certain aspects like words, phrasemes, or production rules to export them in common exchange formats.

After entering all the data, the complete basic construction including the visualization of the concept(s) is displayed in Step 3. We offer default images for all concepts, which are taken from Wikimedia Commons. Whenever possible we automatically extract the image from the corresponding

[26] https://www.wikidata.org/
[27] https://en.wiktionary.org/
[28] https://commons.wikimedia.org/

Wikipedia page using DBpedia as well as Wikidata and Wiktionary for retrieval. However, since some associations between concepts and images can be culture-specific, the instructors can freely localize the images to adapt them to their target audiences. In addition, we encourage the students to choose their own images to personalize the learning experience and to collect valuable data about the observed multimodal associations.

In Step 4, the user has completed the annotation of the first two constructions. By selecting both of them with the mouse, we end up in Step 5, i.e. with a new complex construction and a new relation between the two concepts. The last thing the student has to do is to input the type of the new construction and the role of the new relation. Again, assistance can be offered for these two annotation steps. Also, the emojis for the different roles can be altered to suit the personal preferences of the student.

Figure 9 displays the complete annotation for the whole sentence. It is the first sentence from the Japanese translation of the Sherlock Holmes short story "The red-headed league" by Sir Arthur Conan Doyle. This translation is available at the Japanese digital library *Aozora Bunko*[29] under a Creative Commons license. Aozora Bunko contains over 17,000 literary works without copyright and therefore represents an invaluable resource for Japanese language education, literary studies, as well as translation studies.

It has to be emphasized that the visualization of the annotation is realized purely by using HTML and CSS without any additional libraries apart from basic jQuery UI widgets. This way we can gain flexibility towards alternative application scenarios which would favor a representation which is directly embedded into the original Web page, by avoiding conflicts with original libraries and control flow. One consequence of this design choice is that we only include concepts and roles in the Web document generated at the language server and then add the connecting lines for the relations dynamically at the client after loading the page for an exact rendering of start and end points.

The only new constructions in Fig. 9 not covered so far are the **auxiliary** "koto ga atta", an expression, which indicates the perfect (PRF) ASPECT (💅), the corresponding **auxiliary cxn**, and **clause** for the complete sentence.

Additional roles are TEMPORAL (⏰), SEASON (🌻), and UNDERGOER (🚑). The background color green for the role 🌻 indicates an *inverse relation*, i.e. "autumn" is a SEASON-OF "last year". Inverse relations are mainly used for focusing in UMR, in this case on "autumn". A similar situation is the relation between the perfect aspect and the predicate "visited". However, in this case the predicate is the head of the auxiliary cxn, therefore we change the direction of the relation from left to right, indicated by the line color violet. Finally, a hopefully redundant amendment is that the original text passage reads as "I had called upon my friend, Mr. Sherlock Holmes, one day in the autumn of last year, …".

5. Conclusion

We have presented a Web-based Japanese language learning environment, which offers engaging annotation exercises through a visually enhanced sentence representation. The current user interface design is the result of several iterative development cycles, which included feedback rounds with volunteer language students.

In the future we are going to continue to improve the user experience. For that purpose we intend to have our system soon ready for more widespread experimental use in language classrooms to obtain further feedback, which is also essential for issue tracking and system stabilization. Once we have reached the desired level of maturity, we plan to make the environment available on GitLab.

Apart from the application to other languages, a more ambitious and long-term research target will be the extension of our annotation to incorporate the document-level representation of UMR to be able to model intersentential dependencies.

We will also experiment with different user interaction modalities with varying degrees of automatic linguistic analysis and annotation. In addition, we will consider other application scenarios for additional target user groups. The quite unique setting of annotation tasks for language learning certainly requires additional skills including metalinguistic knowledge that have to be taught to the students. This restricts the applicability of our methodology to certain user groups like, for instance, university students. On the positive side, this also significantly widens the potential user base to students of translation studies, literature studies, linguistics, etc. For example, conducting psycholinguistic experiments represents a fascinating challenge for future work.

We are also very curious about the results of analyzing the constructicon data, which we will collect from the students. Future work in this subfield will address the research question of an optimal interaction with LLMs (Feng et al., 2023b) to create a neuro-symbolic AI system (Wan et al., 2024).

We see a strong potential of personal idiolectal constructicons to become a foundation for the next generation of AI to reach the desired faculties of conceptualization (Singer, 2021), generalization (Hupkes et al., 2023), reasoning (Arkoudas, 2023), and self-reflection (Whitten, 2023) on the long road to self-awareness (Chandha, 2021).

[29] https://www.aozora.gr.jp/

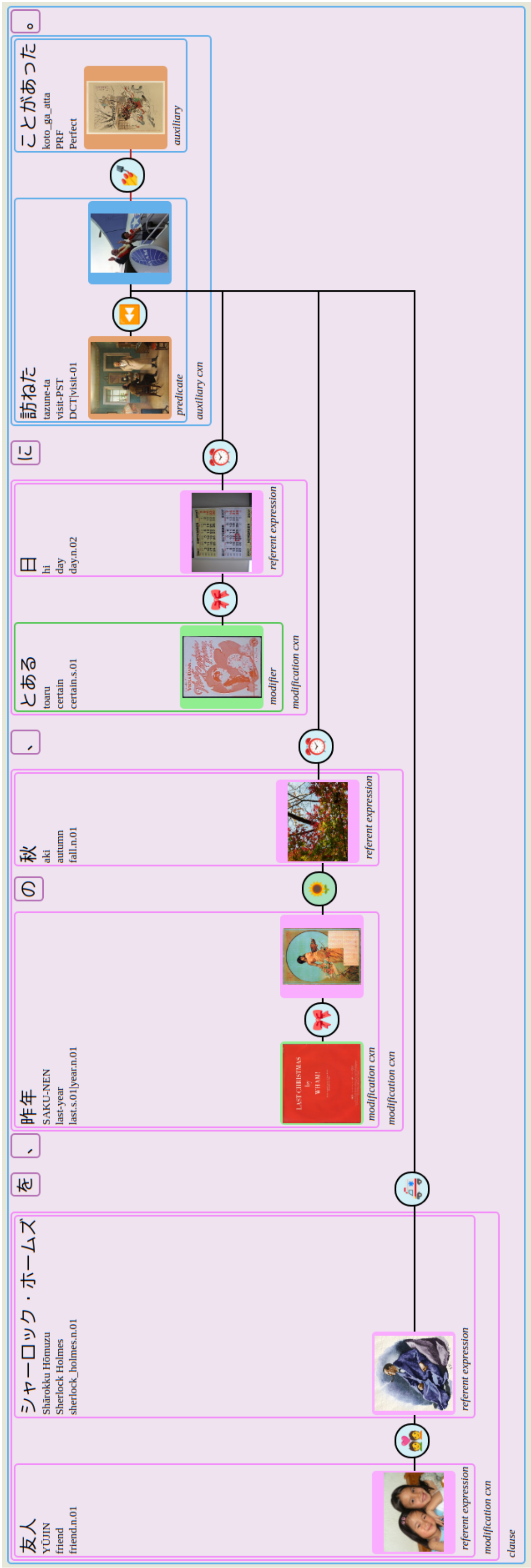

Figure 9: Complete annotation example.

6. Bibliographical References

Omri Abend and Ari Rappoport. 2017. The state of the art in semantic representation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 77–89, Vancouver, Canada.

Konstantine Arkoudas. 2023. GPT-4 can't reason. arXiv:2308.03762 [cs.CL].

A. E. Backhouse. 2016. *Using Japanese Synonyms*. Cambridge University Press.

Laura Banarescu et al. 2013. Abstract Meaning Representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 178–186, Sofia, Bulgaria.

Diego Barriga Martínez, Victor Mijangos, and Ximena Gutierrez-Vasques. 2021. Automatic interlinear glossing for Otomi language. In *Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas*, pages 34–43, Online.

Arne Bewersdorff et al. 2024. Taking the next step with generative artificial intelligence: The transformative role of multimodal large language models in science education. arXiv:2401.00832 [cs.AI].

Francis Bond and Kyonghee Paik. 2012. A survey of wordnets and their licenses. In *Proceedings of the 6th Global WordNet Conference (GWC 2012)*, pages 64–71.

Francis Bond et al. 2009. Enhancing the Japanese WordNet. In *Proceedings of the 7th Workshop on Asian Language Resources (ALR7)*, pages 1–8, Suntec, Singapore.

Stefano Borgo et al., editors. 2019. *Ontology Makes Sense: Essays in Honor of Nicola Guarino*, volume 316 of *Frontiers in Artificial Intelligence and Applications*. IOS Press.

Richard John Bowring and Haruko Uryu Laurie. 1992. *An Introduction to Modern Japanese*. Cambridge University Press, Cambridge, UK.

James Breen. 2004. JMdict: A Japanese-Multilingual dictionary. In *Proceedings of the Workshop on Multilingual Linguistic Resources*, pages 71–79.

Agostina Calabrese, Michele Bevilacqua, and Roberto Navigli. 2020. Fatality killed the cat or: BabelPic, a multimodal dataset for non-concrete concepts. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4680–4686, Online.

Saikiran Chandha. 2021. What if AI becomes self-aware. Express Computer.

William Croft. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford University Press.

William Croft. 2002. *Typology and Universals*, 2nd edition. Cambridge Textbooks in Linguistics. Cambridge University Press.

William Croft. 2022. *Morphosyntax: Constructions of the World's Languages*. Cambridge Textbooks in Linguistics. Cambridge University Press.

William Croft. 2023. Word classes in Radical Construction Grammar. In *The Oxford Handbook of Word Classes*. Oxford University Press.

William Croft and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge Textbooks in Linguistics. Cambridge University Press.

Jia Deng et al. 2009. ImageNet: A Large-Scale Hierarchical Image Database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL.

Holger Diessel. 2023. *The Constructicon: Taxonomies and Networks*. Elements in Construction Grammar. Cambridge University Press.

Dagmar Divjak. 2019. *Frequency in Language: Memory, Attention and Learning*. Cambridge University Press.

Zhendong Dong and Qiang Dong. 2006. *HOWNET and the computation of meaning*. World Scientific.

Jonathan Dunn. 2022. Replication data for: Exposure and emergence in usage-based grammar: Computational experiments in 35 languages. DataverseNO.

Christiane Fellbaum, editor. 1998. *WordNet: An electronic lexical database and some of its applications*. MIT Press.

Lydia Feng et al. 2023a. Widely interpretable semantic representation: Frameless meaning representation for broader applicability. arXiv:2309.06460 [cs.CL].

Zhangyin Feng et al. 2023b. Trends in integration of knowledge and large language models: A survey and taxonomy of methods, benchmarks, and applications. arXiv:2311.05876 [cs.CL].

Charles J. Fillmore, Christopher R. Johnson, and Miriam R.L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography*, 16.3:235–250.

Michael Ginn et al. 2023. Findings of the SIGMORPHON 2023 shared task on interlinear glossing. In *Proceedings of the 20th SIGMORPHON workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 186–201, Toronto, Canada.

Jens E. L. Van Gysel et al. 2021. Designing a Uniform Meaning Representation for natural language processing. *KI – Künstliche Intelligenz*, 35:343–360.

Yoko Hasegawa. 2015. *Japanese: A Linguistic Introduction*. Cambridge University Press, Cambridge, UK.

Yoko Hasegawa, editor. 2018. *The Cambridge Handbook of Japanese Linguistics*. Cambridge Handbooks in Language and Linguistics. Cambridge University Press.

Martin Haspelmath. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language*, 86:663–687.

Martin Hilpert. 2014. *Construction Grammar and its Application to English*. Edinburgh University Press.

Thomas Hoffmann and Graeme Trousdale, editors. 2013. *The Oxford Handbook of Construction Grammar*. Oxford University Press.

Dieuwke Hupkes et al. 2023. A taxonomy and review of generalization research in NLP. *Nature Machine Intelligence*, 5(10):1161–1174.

Shoichi Iwasaki. 2013. *Japanese*, revised edition. John Benjamins, Amsterdam, The Netherlands.

Stefan Kaiser et al. 2013. *Japanese: A Comprehensive Grammar*, 2nd edition. Routledge, London and New York.

Michiel Kamermans. 2010. *An Introduction to Japanese – Syntax, Grammar & Language*. SJGR Publishing, Rotterdam, The Netherlands.

Karin Kipper-Schuler. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. University of Pennsylvania.

Taku Kudo and Yuji Matsumoto. 2002. Japanese dependency analysis using cascaded chunking. In *CoNLL 2002: Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)*, pages 63–69.

Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. 2004. Applying conditional random fields to Japanese morphological analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pages 230–237, Barcelona, Spain.

Yoko Kudo. 2011. Modified Hepburn romanization system in Japanese language cataloging: Where to look, what to follow. *Cataloging & Classification Quarterly*, 49(2):97–120.

Mathieu Lafourcade, Alain Joubert, and Nathalie Le Brun. 2015. *Games With a Purpose (GWAPSs)*, chapter The JeuxDeMots Project – GWAPs and Words. Wiley.

Young-Suk Lee et al. 2022. Maximum Bayes Smatch ensemble distillation for AMR parsing. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5379–5392, Seattle, United States.

Jens Lehmann et al. 2015. DBpedia – a large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web*, 6(2):167–195.

Paul Lerner, Olivier Ferret, and Camille Guinaudeau. 2024. Cross-modal retrieval for knowledge-based visual question answering. arXiv:2401.05736 [cs.CL].

Hugo Liu and Push Singh. 2004. ConceptNet: A practical commonsense reasoning tool-kit. *BT Technology Journal*, 22(4):211–226.

Benjamin Lyngfelt, editor. 2018. *Constructicography: Constructicon development across languages*, volume 22 of *Constructional Approaches to Language*. John Benjamins.

Hiroshi Matsumoto. 2007. Peak learning experiences and language learning: A study of American learners of Japanese. *Language Culture and Curriculum - LANG CULT CURRIC*, 20:195–208.

Yoshiko Matsumoto. 1997. *Noun-Modifying Constructions in Japanese: A frame semantic approach*. John Benjamins, Amsterdam, The Netherlands.

George A. Miller. 1990. WordNet: An on-line lexical database. *International Journal of Lexicography*, 3(4):235–244.

Yoshiko Mori. 2014. Review of recent research on kanji processing, learning, and instruction. *Japanese Language and Literature*, 48(2):403–439.

Andrea Moro, Alessandro Raganato, and Roberto Navigli. 2014. Entity Linking meets Word Sense Disambiguation: a Unified Approach. *Transactions of the Association for Computational Linguistics (TACL)*, 2:231–244.

David R. Mortensen et al. 2023. Generalized glossing guidelines: An explicit, human- and machine-readable, item-and-process convention for morphological annotation. In *Proceedings of the 20th SIGMORPHON workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 58–67, Toronto, Canada.

Roberto Navigli and Simone Paolo Ponzetto. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250.

Roberto Navigli et al. 2021. Ten years of BabelNet: A survey. In *Proceedings of IJCAI 2021*, pages 4559–4567.

Kyoko Ohara. 2014. Relating frames and constructions in Japanese FrameNet. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 2474–2477, Reykjavik, Iceland.

Kyoko Ohara et al. 2003. The Japanese FrameNet project: A preliminary report. In *Proceedings of Pacific Association for Computational Linguistics (PACLING'03)*, pages 249–254.

Toshio Ohori. 2001. *Clause integration as grammaticalization: a case from Japanese Tokoro-complements*, pages 279–301. Kuroshio, Tokyo, Japan.

Anna Papafragou, John C. Trueswell, and Lila R. Gleitman, editors. 2022. *The Oxford Handbook of the Mental Lexicon*. Oxford University Press.

Simon Paxton. 2019. Kanji matters in a multilingual Japan. *The Journal of Rikkyo University Language Center*, 42:29–41.

Sameer Pradhan et al. 2022. PropBank comes of Age—Larger, smarter, and more diverse. In *Proceedings of the 11th Joint Conference on Lexical and Computational Semantics*, pages 278–288, Seattle, Washington.

James Pustejovsky et al. 2016. The development of multimodal lexical resources. In *Proceedings of the Workshop on Grammar and Lexicon: interactions and interfaces (GramLex)*, pages 41–47, Osaka, Japan.

Stephen D. Richardson, William B. Dolan, and Lucy Vanderwende. 1998. MindNet: Acquiring and structuring semantic information from text. In *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 2*, pages 1098–1102, Montreal, Quebec, Canada.

Tanja Samardžić, Robert Schikowski, and Sabine Stoll. 2015. Automatic interlinear glossing as two-level sequence classification. In *Proceedings of the 9th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH)*, pages 68–72, Beijing, China.

Gadi Singer. 2021. Conceptualization as a basis for cognition – human and machine: A missing link to machine understanding and cognitive AI. Towards Data Science.

Robyn Speer, Joshua Chin, and Catherine Havasi. 2017. ConceptNet 5.5: An open multilingual graph of general knowledge. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. YAGO: a core of semantic knowledge. In *Proceedings of the 16th International Conference on World Wide Web*, pages 697–706.

Gilles Sérasset. 2015. DBnary: Wiktionary as a Lemon-based multilingual lexical resource in RDF. *Semantic Web Journal*, 6(4):355–361.

Koichi Takezawa. 1993. Secondary predication and locative/goal phrases. In Nobuko Hasegawa, editor, *Japanese Syntax and Comparative Grammar*, pages 45–77. Kurosio, Tokyo, Japan.

Chiaki Taoka. 2000. *Aspect and argument structure in Japanese*. Ph.D. thesis, University of Manchester.

Tobias Ungerer and Stefan Hartmann. 2023. *Constructionist Approaches: Past, Present, Future*. Elements in Construction Grammar. Cambridge University Press.

Zishen Wan et al. 2024. Towards cognitive AI systems: a survey and prospective on neuro-symbolic AI. arXiv:2401.01040 [cs.AI].

Ryuichi Washio. 1997. Resultatives, compositionality and language variation. *Journal of East Asian Linguistics*, 6:1–49.

Allison Whitten. 2023. AI agents that "self-reflect" perform better in changing environments. Human-Centered Artificial Intelligence, Stanford University.

Jan Wielemaker, Zhisheng Huang, and Lourens Van Der Meij. 2008. SWI-Prolog and the Web. *Theory and Practice of Logic Programming*, 8(3):363–392.

Jan Wielemaker et al. 2012. SWI-Prolog. *Theory and Practice of Logic Programming*, 12(1-2):67–96.

Werner Winiwarter and Bartholomäus Wloka. 2022. VISCOSE – a kanji dictionary enriched with VISual, COmpositional, and SEmantic information. In *7th Workshop on Cognitive Aspects of the Lexicon (CogALex-VII)*.

Bartholomäus Wloka and Werner Winiwarter. 2021a. AAA4LLL – Acquisition, Annotation, Augmentation for Lively Language Learning. In *3rd Conf. on Language, Data and Knowledge (LDK2021)*.

Bartholomäus Wloka and Werner Winiwarter. 2021b. DARE – a comprehensive methodology for mastering kanji. In *23rd Intl. Conf. on Information Integration and Web Intelligence (iiWAS2021)*, pages 427–435.

Etsuyo Yuasa and Jerrold M. Sadock. 2002. Pseudo-subordination: a mismatch between syntax and semantics. *Journal of Linguistics*, 38:87–111.

Xingyuan Zhao et al. 2020. Automatic interlinear glossing for under-resourced languages leveraging translations. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 5397–5408, Barcelona, Spain (Online).

Michael Zock. 2022. The mental lexicon: A blueprint for the dictionaries of tomorrow? *Frontiers in Psychology and Artificial Intelligence*, 5.

Michael Zock and Chris Biemann. 2020. Comparison of different lexical resources with respect to the tip-of-the-tongue problem. *Journal of Cognitive Science*, 21(2):193–252.

Michael Zock, Paul Sabatier, and Line Jakubiec. 2008. Message composition based on concepts and goals. *International Journal of Speech Technology*, 11(3-4):181–193.

7. Language Resource References

Princeton University. 2012. *WordNet 3.1*. Princeton University, ISLRN 379-473-059-273-1.