

Xposition: An Online Multilingual Database of Adpositional Semantics

Luke Gessler
Georgetown University
lg876@georgetown.edu

Austin Blodgett
U.S. Army Research Lab
austin.j.blodgett@gmail.com

Joseph Ledford[†] Nathan Schneider
Georgetown University
{jcl240, nathan.schneider}@georgetown.edu

Abstract

We present Xposition, an online platform for documenting adpositional semantics across languages in terms of supersenses (Schneider et al., 2018). More than just a lexical database, Xposition houses *annotation guidelines*, *structured lexicographic documentation*, and *annotated corpora*. Guidelines and documentation are stored as wiki pages for ease of editing, and described elements (supersenses, adpositions, etc.) are hyperlinked for ease of browsing. We describe how the platform structures information; its current contents across several languages; and aspects of the design of the web application that supports it, with special attention to how it supports datasets and standards that evolve over time.

Keywords: adpositions, prepositions, case, lexical semantics, supersenses, interfaces, meaning representations

1. Introduction

Linguistic annotation is a challenging process that requires the coordination of annotation guidelines, primary datasets, and derived datasets such as lexical inventories. Annotation guidelines virtually always require iterative revision as the process of annotation reveals flaws and edge-cases. Datasets must subsequently be brought in line with guidelines as they change, and analytical, derived datasets like lexical inventories must also be kept synchronized.

Here we describe Xposition,¹ a linguistic database and website that has helped the SNACS project (Schneider et al., 2018) address these issues in the process of annotating several multilingual corpora for adpositional semantics. Prepositions, postpositions, and case markers across the world’s languages serve to mediate grammatical relationships while often signaling the semantics of the relation. Across languages, such elements convey broadly the same kinds of meanings (related to space, time, causality, etc.), though the conceptual boundaries of particular adposition lexemes are highly variable across languages (e.g., Bowerman and Choi, 2001). SNACS provides a unified metalanguage for characterizing the major classes of meanings typically expressed with adpositions, allowing for multilingual study of their polysemy at the lexical level as well as disambiguation in corpora.

We show how Xposition has addressed our need for an accessible presentation of SNACS corpora and guidelines for SNACS annotators and researchers across languages, supported by auxiliary documentation and lexical inventories. We focus particularly on our strategy of keeping all major parts of a multilingual annotation project—corpora, guidelines, and derived resources—together in a single web application and a single integrated database. This strategy yields synergistic ef-

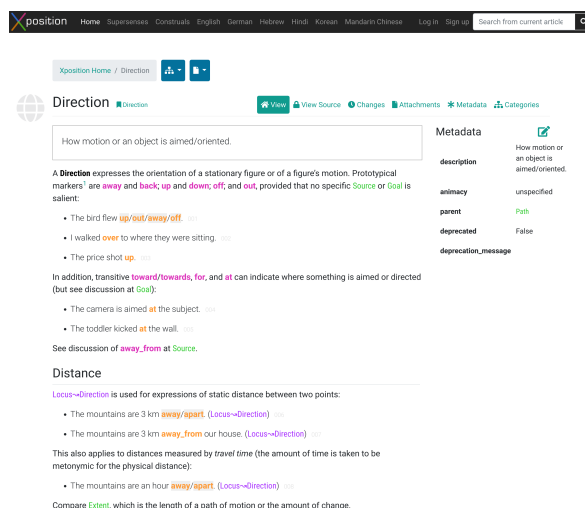


Figure 1: Xposition’s page for the **DIRECTION** supersense, one of 50 supersenses used by SNACS corpora to describe the meaning of an adposition. The page describes the meaning of the label, gives common examples which are hyperlinked to their occurrences in SNACS corpora (see Figure 2), and contrasts it with similar supersenses. Like all pages in Xposition, the page is version-controlled and has been collaboratively edited and reviewed by SNACS project members.

fects which have helped us cope with the difficulties of building crosslinguistic corpora with many different annotators over time and place, and our use of a single integrated database for all resources distinguishes us from similarly multilingual annotation projects like Universal Dependencies (McDonald et al., 2013) and PARSEME (Savary et al., 2015), which have unified websites for annotation guidelines but are not backed by an integrated database of lexical items and corpus attestations.

2. SNACS

SNACS (Semantic Network of Adposition and Case Supersenses) is a framework for coarse supersense an-

[†]Work performed while at Georgetown University.

¹<http://xposition.org/>

notation of adpositions and case markers across languages (Schneider et al., 2018). Its most recent version, SNACS v2.5 (Schneider et al., 2020), has 50 supersenses, covering a wide range of meanings. Adpositions are functional elements (words/morphemes) that are capable of being extremely polysemous in English and many languages. Consider the following sentences:

- (1) Miriam will go to the beach **for**_{DURATION}→_{DURATION} a week.
- (2) Miriam will go to the beach **for**_{PURPOSE}→_{PURPOSE} her vacation.
- (3) Miriam will go to the beach **from**_{SOURCE}→_{SOURCE} the airport.
- (4) Miriam looked out at the sunset **from**_{LOCUS}→_{SOURCE} the beach.

The adpositions **for** and **from** can each carry a number of different meanings. As in the examples above, **for** may denote a duration or purpose, and **from** may denote a source (starting place of travel) or a location, but each can carry quite a few other meanings as well. SNACS assigns different labels to each of these meanings, allowing us to describe their semantics in detail and providing a way to disambiguate these highly polysemous words.

Note that each adposition has *two* supersense labels, drawn from the same label inventory. The first, called the **scene role**, can be thought of as its meaning in context (which may be largely determined by a predicate—e.g. **THEME**), and the second, called the **function**, is one of the meanings closely associated with the adposition. When a preposition use is analyzed as extending the lexical meaning to fit in context (e.g. via metaphor or coercion), the role and function labels will differ. Sentence (4) demonstrates this phenomenon, where **from** is used to denote a **LOCUS**, or *location*, but also evokes imagery where *the beach* is a **SOURCE** rather than a location. The pairing of supersense labels, notated **ROLE**→**FUNCTION**, is called a **construal**. (See Hwang et al. 2017 for further details of this approach.)

SNACS was originally developed for English (Schneider et al., 2014, 2018), but the guidelines have been developed to be multilingual. Currently, SNACS corpora and/or guidelines exist for English, German (Prange and Schneider, 2021), Hindi (Arora et al., 2021, 2022), Korean (Hwang et al., 2020), Mandarin Chinese (Peng et al., 2020), and Hebrew, with additional languages under development.

3. Goals

An overarching goal was to create a “one-stop shop” for accessing and editing the canonical form of all information related to the lifecycle of developing SNACS corpora: corpora, guidelines, lexical information, and more. This goal might first seem like a mere aesthetic preference, but we believe that co-locating and integrating all of these resources comes with many practical

benefits. For example, annotators uncertain about a distinction between two labels may not consult existing examples and/or guidelines if it is not as easy as possible for them to reach and digest them, or they might not realize that a PDF of the guidelines they have locally on their computer is outdated. Researchers might only infrequently derive secondary resources such as lexical inventories if the process is manual. Critically, annotation schemes (and SNACS in particular) change—labels can be added or deprecated, and guidelines for labels can change—and centralization can make the work of keeping everything in sync easier. A shared location and data representation for all of these resources facilitates good linguistic data development practices.

With this super-goal in mind, it was clear that a web application would be the best way to implement this vision. The design of the platform reflects the following goals:

Goal 1: Full Multilingual Support Each language should have room for its own lexical information, guidelines (with glossed examples), and corpora. All Unicode scripts should work without issue. A best effort should be made to accommodate variation in grammatical phenomena, such as adposition-licensed case or whether a language has prepositions or postpositions (or both).

Goal 2: Approachability The graphical user interface (i.e., the website) should be clean and impose as little familiarity with the SNACS project as possible. The UI should be uncluttered and free of dense notation, and moreover should have a consistent visual language which indicates linguistic information to users. A couple dozen annotators have been involved in the SNACS project and more continue to join, some only participating for a year or two, so they need to be able to access and learn the guidelines with as little effort as possible.

Goal 3: Discoverability Conceptually related information should be bidirectionally hyperlinked in order to aid learning and maintenance. For example, guidelines pages for supersenses should be linked to sentences in which the given supersense occurs, and vice versa.

Goal 4: Easy Guideline Editability Guidelines should be easy to view and edit with as little friction as possible in the process of changing them. This is an important goal for the SNACS project because it has relied on detailed documentation of edge cases, which are constantly arising and being arbitrated by the project’s leaders, in order to achieve high consistency among annotators.

Goal 5: Versioning A full version history should be maintained for web-editable data, so that they will continue to be available for e.g. understanding decisions made in a corpus that was annotated using an older version of the guidelines. This history should be easily accessible.

Xposition Home Supersenses Construals English German Hebrew Hindi Korean Mandarin Chinese Log in Sign up Search from current article

Xposition Home / English / over / over: Direction ~> Direc...

over: Direction ~> Direction Direction-Direction View View Source Changes Attachments Metadata Categories

over: Direction ~> Direction

Metadata

- description
- adposition over
- obj_case Accusative
- construal Direction~>Direction

Data (1–9 of 9)

Ex	P	Role	Function	Adposition Part of Speech	Governor Head	Part of Speech	Governor Supersense	Object Head	Part of Speech	Object Supersense	Governor-Object Syntax	Transitive?	Typo?	Abbrev?	Sent ID
(17404) Then , the French sent their partners	over to be with them .	Direction	Direction	ADV	send	VERB					default	×	×	×	english-4d18c103-a100-ed14-fa92-cd28c4c536fb-03
(16103) I just got ta get back to my cabin then you can come	over and help me pack !	Direction	Direction	ADV	come	VERB					default	×	×	×	english-e2a01335-a830-23ab-053e-4b42cc4da021-02
(17508) He bent his head	over the drawing : " Not so small that – Look !	Direction	Direction	ADP	bend	VERB		draw	NOUN		default	✓	×	×	lpp_1943.100
(17665) He looked at me there , with my hammer in my hand , my finger black	over an object which seemed to him extremely ugly ...	Direction	Direction	ADP	bending	VERB		object	NOUN		default	✓	×	×	lpp_1943.325

Figure 2: Xposition’s page for the English preposition *over*, when it has been labeled with the **DIRECTION** supersense for both role and function. (Role and function are roughly analogous to “literal” and “contextual” meaning, and each adposition is annotated with one supersense for each in SNACS corpora.) An overview of the adposition–supersense pair is given, and all occurrences of the pair in English SNACS corpora are given in a table with the relevant occurrence of *over* highlighted in context. The page may be expanded with notes.

4. Implementation

Architecture Xposition is a fork of django-wiki,² a Python wiki system which has functionality similar to that of MediaWiki.³ Plain django-wiki provides a user system, articles with revision histories, and a plugin system for extending most aspects of its functionality. The features related to article editing and versioning are mature and stable, which enabled us to achieve goals 4 and 5, and django-wiki uses UTF-8 encoded strings throughout, a necessary part of achieving goal 1.

The plugin system was used to implement the required data model extensions required for representing corpora, annotation guidelines, and lexical inventories.

²<https://github.com/django-wiki/django-wiki>

³An initial version called PrepWiki was built using MediaWiki (the software which powers sites like Wikipedia), and specifically the Semantic MediaWiki plugin for structured data (Krötzsch et al., 2006). PrepWiki was eventually deprecated, as it became clear that fully implementing the vision for Xposition would require deep changes to MediaWiki’s data models, and making such changes would have been impractical, in part due to the size of MediaWiki’s codebase size and the scarcity of knowledge of PHP, the programming language that MediaWiki is written in, among graduate students.

A shared data model for all SNACS-related resources ensures a minimum of consistency between them. For example, a consequence of the data model is that all supersense tags used in corpora must directly refer to a supersense object in the database via a SQL foreign key. If a corpus ever tried to use a supersense that did not yet exist in the guidelines (and therefore not yet exist in the shared database as well), then an attempt to import the corpus would fail, alerting the corpus’s maintainer to the problem and ensuring that an inconsistent state would not be reached.

Django-wiki’s plugin system yielded an additional benefit, which is that it facilitates the development of self-contained modules that can be created without having to understand other plugins. Xposition has been developed over several years by multiple different developers, and the fact that a new developer did not have to understand all the Xposition-specific code that had been written before their arrival on the project was a great boon for productivity.

UI Design A visual language to consistently and unobtrusively represent aspects of SNACS-related data was designed, pursuant to goal 2. The key to achieving this goal was to extend django-wiki’s markup language

with macros that render the same HTML wherever they are used. For example, the **PURPOSE** supersense can be referred to in markup as `[ss Purpose]`, which expands into a styled HTML anchor tag `Purpose`.

The CSS class `supersense` applies consistent styles to all text representing a supersense, and moreover, this implementation also achieves goal 3 (discoverability via bidirectionally linked related data), as the anchor tag can be clicked to read more about the supersense. Examples of the UI can be seen in Figure 1 and Figure 2, and the live website may be seen at <http://xposition.org/>.

5. Annotation Guidelines

SNACS annotators need to be aware of guidelines for supersenses, role–function combinations of supersenses (“construals”), and lemmas coupled with possible supersense annotations (“usages”; these are similar to lexical senses but may be at a different level of granularity than typical dictionary senses). Language objects also exist which contain high-level typological information (such as whether the language is prepositional or postpositional) and serve as a target for usages so that usages may indicate the language that they belong to.

Figure 1 shows an example of a guidelines page for the **DIRECTION** supersense. A prose description of the supersense’s character is given, enriched by hyperlinked usages, supersenses, and construals. Note also that some metadata is available in the right hand floating column, notably including a flag that indicates whether a supersense is deprecated (deprecated supersenses should not be used, but need to stay around for old versions of datasets and documentation).

Similar pages exist for construals and usages, and additional articles may be created as needed to discuss issues which don’t neatly fit into any of the pages described so far. Supersenses have a special piece of data which tracks whether or not they are deprecated—deprecated supersenses are no longer to be used in any corpora, but it is important to still keep them around so that they can be referred to by maintainers and so that corpora that have not been transitioned off of them yet can maintain data integrity while referring to them.

6. Datasets

SNACS corpora are natively formatted in the CoNLL-U-Lex format, an extension of the widely used CoNLL-U format with 9 additional columns. A toolchain⁴ is used to validate and process CoNLL-U-Lex formatted data into equivalent JSON data, which is then used by Xposition’s import facilities to create the necessary database entries. Currently, 4 corpora from 2 different languages exist in Xposition, with more from other languages planned for addition in the future.

A SNACS corpus consists of sentences where adpositional expressions have been identified and labeled

with data. The most important labels are the supersense labels it receives for role and function, but other information is present as well, such as its part of speech and its governor and object.⁵ Data from corpora is presented in a tabular format, as seen in Figure 1, where all English corpus examples where the word **over** received **DIRECTION** as both its role and function annotation are displayed in a table. The target expression being described in the row of a table is highlighted, and its left and right context in the sentence is also shown. Similar concordances are shown for construals and supersenses.

Corpora in Xposition have an associated version number, and it is possible to have multiple versions of the same corpus in the the system. In the most highly “visible” areas, only the latest version of the corpus will be used (e.g. links to examples from a usage page), but historical versions will be kept around so that, for example, real examples of deprecated labels can be viewed.

7. Lexical Inventories

It is often helpful for annotators to refer to a concordance of all expressions which have received at least one annotation for a given language. Xposition provides this list for each language (see Figure 3 for Hindi’s page). Expressions are separated by whether they are single-word or multi-word adpositional expressions, and adpositional idioms are also listed separately. Examples of the former two in English are **from** and **instead_of**, which both have syntactic behavior typical of adpositions, and an example of the latter is **on_time**, which does not as a whole function as an adposition but is PP idiom that is syntactically headed by an adposition.

8. Related Work

Other corpus development projects with detailed linguistic annotations have partially relied on websites for dissemination of project data. FrameNet (Baker and Sato, 2003) has online interfaces⁶ for browsing annotated data and lexical resources (i.e., frames, lexical units, etc.), though these interfaces do not share a single integrated multilingual database like Xposition, and separate software is used to edit the lexicographic descriptions and corpus annotations offline. Modeled off of FrameNet, PrepNet (Saint-Dizier, 2006) was an attempt to provide similar online resources focused on prepositions, but it was never fully implemented. The Preposition Project and extensions thereof (e.g., Litkowski and Hargraves, 2005; Litkowski, 2014) is the main computational lexicographic research effort focused just on English prepositions; the web infrastructure for that project includes a sense dictionary linked to annotated corpora, with a detailed UI designed for a small number of users who work with the tool on a regular basis.

⁵Governor and object can roughly be thought of as syntactic head and syntactic complement, respectively.

⁶https://framenet.icsi.berkeley.edu/fndrupal/framenet_data

⁴<https://github.com/nert-nlp/conllullex>

Figure 3: Xposition’s page for the Hindi language. Structured metadata about the language is included in the right-hand column, including information about whether adpositions are prepositional or adpositional, and the grammatical cases that are found in the language. An inventory of Hindi’s single-word and multiword adpositional expressions is given, both in the Devanagari script and in a romanized transcription. The number next to each expression indicates how many constructions are attested for each expression.

More recently, the large-scale, massively-multilingual annotation projects Universal Dependencies (McDonald et al., 2013) and PARSEME (Savary et al., 2015) have also moved to centralize their guidelines online with examples from many languages, though their interfaces are not backed by a single, central database, and examples must always be manually sourced. Universal Dependencies’ guidelines/documentation setup in particular served as inspiration for Xposition: their guidelines are developed as Markdown files in a Git repository (Pyysalo and Ginter, 2014), which motivated initial interest in a wiki system for Xposition.

9. Conclusion

We have presented Xposition, a unified web application for adpositional semantics that has guided the SNACS project. Xposition has been instrumental in facilitating the flow and maintenance of all data in SNACS projects. Users report that the tidy and convenient UI encourages them to refer to the guidelines and existing corpora whenever a question arises, and constraints at the database level guard against many types of potential inconsistencies, as discussed in §4. We have been pleased with our ability to meet our five goals, and have found django-wiki to be a solid foundation for our application. Xposition is also able to receive upstream updates from django-wiki, as it is a fork of django-wiki at the level of the version control system (Git). While

we do not expect that Xposition’s code could be easily applied to many other similar linguistic annotation projects, we hope that developers of other linguistic resources might be inspired to consider whether the goals discussed in §3 might also be worth pursuing for their projects, whether it is with django-wiki or some other solution.

9.1. Future Work

There are some functionality-related extensions which we wish to make to Xposition to continue to serve the needs of SNACS corpora:

1. Parallel corpus support: many SNACS corpora are parallel corpora (translations of *Le Petit Prince*), and we would like to extend Xposition so that alignments between translations and the original could be viewable.
2. User-facing export: there is currently no convenient way for users to export parts of the data in Xposition, e.g. all sentences used in guidelines, or all language-specific guidelines, and we anticipate this would be useful.
3. Tutorial: it would be nice to have a tutorial to train new annotators to annotate new corpus data.
4. Tagger demo: a live demo of the SNACS tagger⁷ would be helpful.
5. Analytic comparisons: statistical comparisons across

⁷<https://github.com/nelson-liu/lexical-semantic-recognition>

corpora (e.g. for relative frequencies of supersenses across corpora) would be useful for annotators and maintainers.

6. Annotation integration: it could be useful to have interfaces for directly annotating new corpora, whether it is raw text or data that has already received silver-quality annotations from a tagger.

7. Enhanced query tools: SNACS users currently rely on other tools for more sophisticated queries, such as ones that rely on contextualized word embeddings (Gessler and Schneider, 2021) or ones that leverage other annotations like dependency trees and POS tags (via export to ANNIS; Krause and Zeldes, 2016), and it would be nice to have this integrated into Xposition itself. Additionally, users have requested that the django-wiki search functionality be extended to corpus sentences (it currently only searches wiki articles, i.e. the guidelines).

At a higher level, the guidelines currently have a hierarchical structure: guidelines are primarily written for English, and other languages develop their own separate guidelines to the extent that the English guidelines are inadequate. We feel this may result in particulars of English grammar having undue sway on SNACS guidelines decisions, and may consider integrating more examples from other languages into the core guidelines.

Acknowledgments

We are grateful to Jena Hwang, Vivek Srikumar, and members of the CARMLS and NERT teams who have contributed to the SNACS guidelines/datasets or provided user feedback on Xposition. This research was supported in part by NSF award IIS-1812778 and grant 2016375 from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel.

10. Bibliographical References

References

- Arora, Aryaman, Venkateswaran, Nitin, and Schneider, Nathan (2021). Hindi-Urdu Adposition and Case Supersenses v1.0. *arXiv:2103.01399 [cs]*.
- Arora, Aryaman, Venkateswaran, Nitin, and Schneider, Nathan (2022). MASALA: Modelling and analysing the semantics of adpositions in linguistic annotation of Hindi. In *Proc. of LREC*. Marseille, France.
- Baker, Collin F. and Sato, Hiroaki (2003). The FrameNet Data and Software. In *The Companion Volume to the Proceedings of 41st Annual Meeting of the Association for Computational Linguistics*, pages 161–164. Association for Computational Linguistics, Sapporo, Japan. doi:10.3115/1075178.1075206.
- Bowerman, Melissa and Choi, Soonja (2001). Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In Bowerman, Melissa and Levinson, Stephen, editors, *Language Acquisition and Conceptual Development*, number 3 in Language, Culture & Cognition, pages 475–511. Cambridge University Press, Cambridge, UK.
- Gessler, Luke and Schneider, Nathan (2021). BERT Has Uncommon Sense: Similarity Ranking for Word Sense BERTology. In *Proceedings of the Fourth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, pages 539–547. Association for Computational Linguistics, Punta Cana, Dominican Republic. doi:10.18653/v1/2021.blackboxnlp-1.43.
- Hwang, Jena D., Bhatia, Archana, Han, Na-Rae, O’Gorman, Tim, Srikumar, Vivek, and Schneider, Nathan (2017). Double Trouble: The Problem of Construal in Semantic Annotation of Adpositions. In *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (*SEM 2017)*, pages 178–188. Association for Computational Linguistics, Vancouver, Canada. doi:10.18653/v1/S17-1022.
- Hwang, Jena D., Choe, Hanwool, Han, Na-Rae, and Schneider, Nathan (2020). K-SNACS: Annotating Korean adposition semantics. In *Proc. of the Second International Workshop on Designing Meaning Representations*, pages 53–66. Barcelona, Spain (Online).
- Krause, Thomas and Zeldes, Amir (2016). ANNIS3: A new architecture for generic corpus query and visualization. *Digital Scholarship in the Humanities*, 31(1):118–139.
- Kröttsch, Markus, Vrandečić, Denny, and Völkel, Max (2006). Semantic MediaWiki. In Cruz, Isabel, Decker, Stefan, Allemang, Dean, Preist, Chris, Schwabe, Daniel, Mika, Peter, Uschold, Mike, and Aroyo, Lora M., editors, *The Semantic Web - ISWC 2006*, Lecture Notes in Computer Science, pages 935–942. Berlin, Heidelberg. doi:10.1007/11926078_68.
- Litkowski, Ken (2014). Pattern Dictionary of English Prepositions. In *Proc. of ACL*, pages 1274–1283. Baltimore, Maryland, USA.
- Litkowski, Ken and Hargraves, Orin (2005). The Preposition Project. In *Proc. of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, pages 171–179. Colchester, Essex, UK.
- McDonald, Ryan, Nivre, Joakim, Quirnbach-Brundage, Yvonne, Goldberg, Yoav, Das, Dipanjan, Ganchev, Kuzman, Hall, Keith, Petrov, Slav, Zhang, Hao, Täckström, Oscar, Bedini, Claudia, Bertomeu Castelló, Núria, and Lee, Jungmee (2013). Universal Dependency Annotation for Multilingual Parsing. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short*

- Papers*), pages 92–97. Association for Computational Linguistics, Sofia, Bulgaria.
- Peng, Siyao, Liu, Yang, Zhu, Yilun, Blodgett, Austin, Zhao, Yushi, and Schneider, Nathan (2020). A corpus of adpositional supersenses for Mandarin Chinese. In *Proc. of LREC*, pages 5988–5996. Marseille, France.
- Prange, Jakob and Schneider, Nathan (2021). Draw mir a sheep: a supersense-based analysis of German case and adposition semantics. *KI - Künstliche Intelligenz*, 35(3):291–306.
- Pyysalo, Sampo and Ginter, Filip (2014). Collaborative development of annotation guidelines with application to universal dependencies. Unpublished technical report.
- Saint-Dizier, Patrick (2006). PrepNet: a Multilingual Lexical Description of Prepositions. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*. European Language Resources Association (ELRA), Genoa, Italy.
- Savary, Agata, Sailer, Manfred, Parmentier, Yannick, Rosner, Michael, Rosén, Victoria, Przepiórkowski, Adam, Krstev, Cvetana, Vincze, Veronika, Wójtowicz, Beata, Losnegaard, Gyri, Parra Escartín, Carla, Waszczuk, Jakub, Constant, Matthieu, Osenova, Petya, and Sangati, Federico (2015). PARSEME – PARSing and Multiword Expressions within a European multilingual network. In *Proceedings of the 7th Language & Technology Conference*. Poznań, Poland.
- Schneider, Nathan, Hwang, Jena D., Bhatia, Archana, Srikumar, Vivek, Han, Na-Rae, O’Gorman, Tim, Moeller, Sarah R., Abend, Omri, Shalev, Adi, Blodgett, Austin, and Prange, Jakob (2020). Adposition and Case Supersenses v2.5: Guidelines for English. *arXiv:1704.02134 [cs]*. ArXiv: 1704.02134.
- Schneider, Nathan, Hwang, Jena D., Srikumar, Vivek, Prange, Jakob, Blodgett, Austin, Moeller, Sarah R., Stern, Aviram, Bitan, Adi, and Abend, Omri (2018). Comprehensive Supersense Disambiguation of English Prepositions and Possessives. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 185–196. Association for Computational Linguistics, Melbourne, Australia. doi:10.18653/v1/P18-1018.
- Schneider, Nathan, Onuffer, Spencer, Kazour, Nora, Danchik, Emily, Mordowanec, Michael T., Conrad, Henrietta, and Smith, Noah A. (2014). Comprehensive Annotation of Multiword Expressions in a Social Web Corpus. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 455–461. European Language Resources Association (ELRA), Reykjavik, Iceland.