

The Case for Systematically Derived Spatial Language Usage

Bonnie Dorr

Institute for Human and Machine Cognition
15 SE Osceola Ave, Ocala, FL 34471
bdorr@ihmc.us

Clare Voss

U.S. Army Research Laboratory
Adelphi, MD 20783
clare.r.voss.civ@mail.mil

Abstract

This position paper argues that, while prior work in spatial language understanding for tasks such as robot navigation focuses on mapping natural language into deep conceptual or non-linguistic representations, it is possible to systematically derive regular patterns of spatial language usage from existing lexical-semantic resources. Furthermore, even with access to such resources, effective solutions to many application areas such as robot navigation and narrative generation also require additional knowledge at the syntax-semantics interface to cover the wide range of spatial expressions observed and available to natural language speakers. We ground our insights in, and present our extensions to, an existing lexico-semantic resource, covering 500 semantic classes of verbs, of which 219 fall within a spatial subset. We demonstrate that these extensions enable systematic derivation of regular patterns of spatial language without requiring manual annotation.

1 Introduction

While prior work in spatial language understanding for tasks such as robot navigation focuses on mapping natural language into deep conceptual or non-linguistic representations—for further reasoning or embodied cognition (Perera et al., 2017; Pastra et al., 2011)—we argue that it is possible to systematically derive regular patterns of language usage from existing lexical-semantic resources (Dorr et al., 2001). Furthermore, even with access to such resources, effective solutions to many application areas such as robot navigation and narrative generation require additional knowledge at the syntax-semantics interface to capture the range of spatial expressions observed and available to natural language speakers.

The emphasis of this position paper is on the representational underpinnings of spatial expres-

sions for problems such as natural-language mediated two-way human-robot dialogue. Such communication may ultimately take place over low bandwidth networks where, for example, an autonomous robot will navigate and report back from a remote site on what it sees in cooperation with its distant human teammate who directs and responds to the robot as needed. We focus on the use and modification of existing resources to address this problem, making certain linguistically-motivated, working assumptions about:

- layers within our lexical representations,
- levels for distinct language-based modules with syntactic, semantic, and conceptual knowledge (each with primitives and operations for that level), and
- a shared computational model of an environment that includes representations of objects, agents, their relations to each other, events—thus enabling navigation information to be accessible to both robot and human.

That is, we assume first that there exist lexical-internal semantic structures with layers, and those semantic structures contain primitives that are grounded at a conceptual level (not discussed herein). We leverage Lexical Conceptual Structure (LCS) (Jackendoff, 1983; Dorr, 1993), a logical representation with compositional properties, to guide development of semantics for spatial language in language understanding and generation.¹

We note that other logical representations may also be adequate for this study, e.g., Abstract Meaning Representation (Banarescu et al., 2014), Prague Dependency Trees (Hajič et al., 2018), and descendants of such representations (Vanderwende et al., 2015). LCS has been selected due

¹We take these structures to capture language-bound meanings, that is semantic forms. In our framework, these do not, despite their name, capture language-independent, conceptual knowledge.

to its compositional, lexicon-based formalism and its potential for follow-on work in other language processing applications for which cross-lingual LCS mappings have already been devised (e.g., machine translation (Habash and Dorr, 2002)).

We assume second, that for human-robot natural-language mediated communication, a number of constraints at the syntax-semantics interface are crucial for interpreting the wide ranging flexibility of real utterances and the context of the system is central to dialogue management. We leverage previously collected dialogue data with naturally occurring spoken Bot Language (Marge et al., 2017) that provides transcripts and dialog analyses (Traum et al., 2018), but without any form of lexical semantics.

We assume third, that we will test and validate our approach by augmenting an implemented dialogue system for understanding and generation of Bot Language. The application of our foundational paradigm to this problem is a future direction outside of the scope of this position paper.

The layered lexical representations referred to in the first assumption above form the basis for this discussion. Specifically, we posit that the development of an application such as robot navigation (Bonial et al., 2018; Moolchandani et al., 2018) or generation of narrative explanations (Korpan et al., 2017; Lukin et al., 2018) requires a layered representation scheme to include a set of spatial primitives (the basis for the LCS representation) coupled with a representation of constraints at the syntax-semantics interface. Additional layers include prepositional collocates² and spatial semantics that are crucial for understanding and production of unconstrained spatial expressions.

We describe our extensions to an LCS resource covering 500 semantic classes of verbs, of which 219 fall within a spatial subset. We demonstrate that this resource is designed to systematically account for certain types of spatial expressions based on lexical-semantic constraints of spatial verbs in those expressions.

At the heart of the position presented herein is a representational framework that supports the ability to “read off” such constraints from lexical entries without requiring laborious manual annota-

²Prepositions that, when tested in collocations with otherwise non-spatial expressions, add spatial information. For example, in *The hawk screeched across the sky.*, the prepositional phrase headed by *across* introduces motion not present in the intransitive *The hawk screeched* (Talmy, 2014).

tion. Similarly, when subsequent lexicon updates occur, the ability to “read off” constraints is still available without manual annotation. This differentiates our approach from others, e.g., feature-based annotation (for a cogent review of natural language annotation approaches, see (Stubbs and Pustejovsky, 2012)). Our LCS-based approach is described next, followed by related work and concluding remarks.

2 Approach

This section introduces the notion of LCS and describes an LCS-based approach to systematic derivation of usage patterns for understanding and generation. We extend an LCS resource to include constraints (blocks, overlaps, and fills) and present the upshot of these extensions.

2.1 Lexical Conceptual Structure

Lexical Conceptual Structure (LCS) (Jackendoff, 1983, 1990; Dorr, 1993; Dowty, 1979; Guerssel et al., 1985) has been used for a range of different applications, including interlingual machine translation (Habash and Dorr, 2002), lexical acquisition (Habash et al., 2006), cross-language information retrieval (Levow et al., 2000), language generation (Traum and Habash, 2000), and intelligent language tutoring (Dorr, 1997).

The LCS representation was introduced by Jackendoff as based in the spatial domain and naturally extended to non-spatial domains, as specified by *fields*.³ For example, the spatial dimension of the LCS representation corresponds to the (*Loc*)ational field, which underlies the meaning of *John traveled from Chicago to Boston* in the LCS [John GO_{Loc} [From Chicago] [To Boston]]. This is straightforwardly extended to the (*Temp*)oral field to represent analogous meanings such as *The meeting went from 7pm to 9pm* in the LCS [Meeting GO_{Temp} [From 7pm] [To 9pm]].

An “LCS Verb Database” (LVD) developed in prior work (Dorr et al., 2001) includes a set of LCS templates classified according to an extension of (Levin, 1993)’s 192 classes, totaling 500 classes. The first 44 classes were added beyond the original set of semantic classes (Dorr and Jones, 1996). Additional classes were derived through aspectual distinctions to yield *LCS classes* that were finer-grained than the original Levin classes (Olsen

³For a more extensive, non-LCS-based analysis and accounting of the relation of spatial and temporal concepts, see (Tenbrink, 2011).

et al., 1997). Each LCS class consists of a set of verbs and, in several cases, the classes include *non-Levin words* (those not in (Levin, 1993)), derived semi-automatically (Dorr, 1997). LVD is foundational for the position adopted in this paper, as it provides a mapping from LCS-based verb classes to their surface realizations.

The representational framework provided by the LVD has many similarities with others such as FrameNet (Ruppenhofer et al., 2016) and VerbNet (Palmer et al., 2017), both of which also include classes and mappings to surface realizations. Whereas FrameNet has a richer semantics, e.g., finer grained classes than those of Levin (1993), VerbNet has a clearer mapping to surface realizations with specific mappings from thematic roles to syntactic realizations. The LVD differs from both of these in that its compositional representations support the ability to “read off” different types of lexical-semantic constraints without requiring manual annotation. For example, constraints on the mapping between semantics and syntax, e.g., blocks, overlaps, and fills, can be “read off” LVD entries, as described below.

2.2 Syntax-Semantics Interface

Prior work (Jackendoff, 1996; Levin, 1993; Dorr and Voss, 1993; Voss and Dorr, 1995; Kipper et al., 2007; Palmer et al., 2017) suggests that there is a close relation between underlying lexical-semantic structures of verbs and nominal predicates and their syntactic argument structure. The work of Voss et al. (1998) supports that the generation of a preposition (in English) as dependent on both the semantics of the predicate and structural idiosyncrasies at the syntax-semantics interface.

Three notions introduced in this earlier work are relevant to spatial language understanding: BLOCK (where a LCS predicate preempts or blocks the composition into one of its argument positions by another LCS), OVERLAPS (where a LCS predicate allows the composition of another LCS into one of its already-occupied arguments), and FILLS (where a LCS predicate allows the composition of another correctly typed LCS into one of its empty arguments).

To investigate the systematic derivation of language usage patterns for both understanding and generation of spatial language, we first simplify and adapt the LVD to include mappings to both lexically implicit and lexically explicit directional components of meaning. We focus specifi-

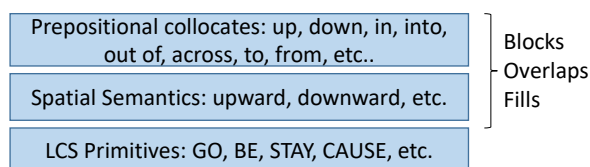


Figure 1: Layered Representation Scheme: Spatial primitives (bottom layer) are coupled with spatial semantics (middle layer) and spatial semantics (top layer) for spatial language understanding and generation

cally on directional verbs coupled with these implicit/explicit directional components of meaning.

We posit that the development of a framework for both understanding and generation of spatial language requires a layered representation scheme illustrated in Figure 1. The top two layers rely heavily on the notions of BLOCKS, OVERLAPS, and FILLS. More specifically:

- **BLOCKS** refers to *lexically implicit* directional components of meaning (such as `upward`) that cannot be lexically realized on the surface, as happens when a predicate already includes the corresponding directional component of meaning, e.g., *elevate* and *ascend* do not collocate with the preposition *up*.
- **OVERLAPS** refers to *lexically implicit and optionally explicit* directional components of meaning (such as `upward`) that may or may not be lexically realized on the surface even though the semantics of the predicate includes the corresponding directional component of meaning, e.g., *lift* and *raise* optionally collocate with *up*.
- **FILLS** refers to *lexically explicit* directional components of meanings that fall into one of two categories: (1) *obligatory* components of meaning (such as `upward`) that must be lexically realized, as the semantics of the predicate does not include the corresponding directional component of meaning, e.g., *put* always collocates with a preposition such as *up*. (2) *optional* components of meaning (such as `upward`) that may or may not be lexically realized, as the semantics of the predicate does not include directional component of meaning, e.g., *move* optionally collocates with a preposition, such as *up*.

The LVD described in Section 2.1 includes compositional structures based on primitives such as GO, BE, STAY, CAUSE. These structures, which form the foundation for the bottom layer, are outside of the scope of this paper.

2.3 Upshot of Lexico-Semantic Extensions for Spatial Language Understanding

An adapted form of the LVD has been developed for the purpose of illustrating the position taken in this paper. This derivative resource contains simplified LCS classes, omitting the full LCS structures and thematic roles from prior work, and augmenting LCS classes to include prepositional *collocations* (the top layer of Figure 1), coupled with a new *spatial component of meaning* (the middle layer of Figure 1).

The spatial component of meaning may or may not be overtly realized on the surface. For example, in the LCS Class of *Verbs of inherently directed motion* (corresponding to Class 51.1.a in (Levin, 1993)), the verb *leave* can take a NP complement (as in *leave the room*) and the verb *depart* can take a PP complement (as in *departed from the room*). For either case, the spatial component of meaning is uniformly *move to a position outside of the room*.

Whereas the collocations were derived from thematic roles in the original LVD, the spatial components of meaning were derived from verb-prepositions pairs associated with a subset of the “Categorical Variation” database (Habash and Dorr, 2003). Representative members of LCS classes were then paired with prepositions that were propagated to other members of the class.

Table 1 summarizes the number of LCS classes associated with the lexical notions introduced above (Blocks, Overlaps, Fills-Oblig, Fills-Opt).⁴ Not all LCS classes are spatial in nature; thus, the second column provides a tally for the full set of LCS classes, and the third column provides a tally for just the spatial subset. The fourth column presents the number of spatial verbs included in the corresponding spatial classes. Representative spatial examples are provided in the fifth column.

Lexical Notions	LCS Classes	Spatial Subset	#Spatial verbs	Spatial Examples
Blocks	7	7	297	elevate, face, pocket
Overlaps	17	10	84	advance, lower, lift
Fills-Oblig	310	128	2783	drive, rotate, put
Fills-Opt	87	59	1280	remove, slide
Intrans	6	3	34	float, part, squirm
N/A	73	12	162	bend, break, carry
Total	500	219	4640	

Table 1: Summary of number of classes associated with Blocks, Overlaps, Fills-Oblig, Fills-Opt, and Intrans in LCS Classes and Spatial Subset

⁴N/A refers to verb classes whose members take bare NP or S arguments. Intrans refers to Intransitive verbs.

Interestingly, the spatial subset of classes is sizeable (44% of the entire set of 500 classes). The percentage of verb entries in the spatial subset is also quite high (42% of the 11K total number of verb entries). Several verbs in the Spatial Subset are relevant to those used in robot navigation, e.g., *move*, *go*, *advance*, *drive*, *return*, *rotate*, and *turn*. Others are easily accommodated by extending classes—without modification to the spatial notions described above. For example, *back up* matches the class containing *advance*, and *pivot* matches the class containing *rotate*.

Note that the BLOCKS, OVERLAPS, AND FILLS notions are generalizable to a high number of LCS classes that are non-spatial as well. These typically correspond to metaphorical extensions of spatial components of meaning to other domains, e.g., *lifted her spirits up*, *elevated her spirits*. Thus, these notions are more broadly applicable than just to the spatial dimension.

Ultimately, surface realizations of verbs with collocations include lexically explicit prepositions as in *lift up*, whereas no such collocates are available when spatial components of meaning are internally conveyed as in *elevate* and thus are lexically implicit. Adding this information to the derivative resource supports a refined formulation of BLOCKS, OVERLAPS, and FILLS notions—which are central to a range of important problems, e.g., dialogue management in robot navigation (Bonial et al., 2017) and generation of narrative explanations (Korpan et al., 2017).

3 Related Work

The ever-growing number of interdisciplinary research programs that now involve natural language processing but are published outside of computational linguistics, provides both challenges and opportunities to all communities seeking to leverage emerging insights from beyond their own areas of expertise. In this short position paper, we highlight but two areas pertinent to our work, while acknowledging there exists much other research in situated dialogue for robots (e.g., (Mavridis and Roy, 2006; Kruiff et al., 2007)) and spatial cognition (e.g., publications of the Spatial Cognition collaborative research center in Germany) that is not as central to our focus.

3.1 Spatial Language Understanding

Spatial language understanding has made great strides in recent years, with the emergence of lan-

guage resources and standards for capturing spatial information. For example, the ISO 24617 standard provides guidelines for annotating spatial information in English language texts (24617-7, 2014) that continues to evolve (Pustejovsky and Lee, 2017). This Semantic Annotation Framework (semAF) identifies places, paths, spatial entities, and spatial relations that can be used to associate sequences of processes and events in news articles (Pustejovsky et al., 2011). Spatial prepositions and particles (such as *near*, *off*) and verbs of position and movement (such as *lean*, *swim*) in text have corresponding spatial components of meanings, collocations, and classes of spatial verbs in the perspective adopted in this paper.

Spatial role labeling using holistic spatial semantics (i.e., analysis at the level of the full utterance) has been used for identifying spatial relations between objects (Kordjamshidi et al., 2010). The association between thematic roles and their corresponding surface realizations has been investigated previously, including in the LCS formalism (described next), but Kordjamshidi et al’s approach also ties into deeper notions such as *region of space* and *frame of reference*. Their work differs from the perspective adopted in this paper in that they provide annotation guidelines for training systems that do spatial information extraction, and so do not focus on generalized mappings at the syntax-semantics interface to predict possible linguistic constructs for spatial relations.

3.2 Embodied Cognition

Another research area relevant to the position adopted herein is that of embodied cognition for the development of language processing tools (Pastra et al., 2011). A European-funded project (POETICON) has resulted in a suite of embodied language processing tools relating symbolic and sensorimotor representation spaces. This work sheds light on the nature of the relationship between language and action, enabling exploration of a range of different projects concerning language learning and human-robot interaction.

Other researchers have focused on natural language grounding for embodied interaction (Al-Omari et al., 2017) to learn components of language and the meanings of each word. The acquired knowledge that emerges from this approach is used to parse commands involving previously unseen objects. Thus, that work assumes no prior knowledge of the structure of language; rather,

word meanings are learned from scratch. In contrast, the perspective put forward in this paper is one in which this knowledge already exists and can be leveraged for support of both language understanding and generation.

The work of Spranger et al. (2016) is the closest to our perspective, particularly in its use of spatial relations such as *across* and *in front of*, both for hearing and for producing utterances for robot-robot communication. However, the position adopted here is one in which generalizations about language structure are assumed and available in natural language generation for both *use* (“lift up”) and *suppression* (“elevate”) of spatial prepositions in phrases containing motion and direction verbs, depending on the context.

4 Conclusions and Future Work

We have made a case for the systematic derivation of regular patterns of spatial language usage from an existing lexical semantic resource (LCS Verb Lexicon). We have focused on a refined formulation of BLOCKS, OVERLAPS, and FILLS, lexical-semantic notions that are central to problems dialogue management in robot navigation and generation of narrative explanations. We demonstrated that these extensions enable systematic derivation of regular patterns of spatial language without requiring manual annotation.

Future work motivated by the position set forth in this paper is investigation of systematic derivation of mappings at the syntax-semantics interface for other parts of speech involving access to a “Categorical Variation” database (CatVar) (Habash and Dorr, 2003) to map verbs in the LCS classes to their nominalized and adjectivalized forms. For example, the CatVar entry for *depart* includes the nominalized form *departure*, which takes a prepositional-phrase complement (e.g., *from the room*)—analogous to the verbal counterpart specified in the simplified LCS classes.

Another future direction is one where these generalized mappings are used in conjunction with data collected within an ongoing Bot Language project (Marge et al., 2017) to enable spatial language understanding in robot navigation. That project has heretofore focused on dialogue annotation (Traum et al., 2018) and has not yet incorporated deeper semantics necessary for automatically detecting incomplete, vague, or implicit navigation commands within dialogues in the spatial domain—issues addressed by our extensions.

Acknowledgments

We would like to acknowledge and thank three anonymous reviewers for their careful reading of our manuscript and their many insightful comments and constructive suggestions. This research is supported, in part by the Institute for Human and Machine Cognition, in part by the U.S. Army Research Laboratory, and in part by the Office of the Director of National Intelligence (ODNI) and the Intelligence Advanced Research Projects Activity (IARPA) via the Air Force Research Laboratory (AFRL) contract number FA875016C0114. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of ODNI, IARPA, AFRL, ARL, or the U.S. Government.

References

- ISO 24617-7. 2014. [Language Resource management Semantic Annotation Framework Part 7: Spatial information \(ISOspace\)](https://www.iso.org/standard/60779.html). <https://www.iso.org/standard/60779.html>.
- Muhannad Al-Omari, Paul Duckworth, David C. Hogg, and Anthony G. Cohn. 2017. Natural Language Acquisition and Grounding for Embodied Robotic Systems. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA.* pages 4349–4356.
- Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2014. [Abstract Meaning Representation \(AMR\) 1.2.1 Specification](https://github.com/amrisi/amrguidelines/blob/master/amr.md). <https://github.com/amrisi/amrguidelines/blob/master/amr.md>.
- Claire Bonial, Stephanie Lukin, Ashley Fouts, Cassidy Henry, Matt Marge, Ron Artstein, David Traum, and Clare Voss. 2018. Human-robot dialogue and collaboration in search and navigation. In *LREC 2018 AREA Workshop (Annotation, Recognition and Evaluation of Actions)*.
- Claire Bonial, Matthew Marge, Ron Artstein, Ashley Fouts, Felix Gervits, Cory J. Hayes, Cassidy Henry, Susan G. Hill, Anton Leuski, Stephanie M. Lukin, Pooja Moolchandani, Kimberly A. Pollard, David R. Traum, and Clare R. Voss. 2017. Laying Down the Yellow Brick Road: Development of a Wizard-of-Oz Interface for Collecting Human-Robot Dialogue. In *AAAI Fall Symposium on Natural Communication for Human-Robot Collaboration*.
- Bonnie Dorr and Doug Jones. 1996. Acquisition of Semantic Lexicons: Using Word Sense Disambiguation to Improve Precision. In *In Proceedings of the Workshop on Breadth and Depth of Semantic Lexicons, 34th Annual Conference of the Association for Computational Linguistics*. Kluwer Academic Publishers, pages 42–50.
- Bonnie J. Dorr. 1993. *Machine Translation: A View from the Lexicon*. MIT Press, Cambridge, MA.
- Bonnie J. Dorr. 1997. Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation. *Machine Translation* 12:271–322.
- Bonnie J. Dorr, Mari Olsen, Nizar Habash, and Scott Thomas. 2001. [LCS Verb Database Documentation](http://www.umiacs.umd.edu/~bonnie/Demos/LCS_Database_Documentation.html). http://www.umiacs.umd.edu/~bonnie/Demos/LCS_Database_Documentation.html.
- Bonnie J. Dorr and Clare R. Voss. 1993. Machine Translation of Spatial Expressions: Defining the Relation between an Interlingua and a Knowledge Representation System. In *Proceedings of the Twelfth Conference of the American Association for Artificial Intelligence*. pages 374–379.
- David Dowty. 1979. *Word Meaning and Montague Grammar*. Reidel, Dordrecht.
- M. Guerssel, K. Hale, M. Laughren, B. Levin, and J. White Eagle. 1985. A Cross-linguistic Study of Transitivity Alternations. In W. H. Eilfort and P. D. Kroeber and K. L. Peterson, editor, *Papers from the Parasession in Causatives and Agentivity at the Twenty-first Regional meeting of the Chicago Linguistic Society*. pages 48–63.
- Nizar Habash and Bonnie Dorr. 2003. A categorial variation database for english. In *In NAACL/HLT 2003, Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference*. pages 96–102.
- Nizar Habash and Bonnie J. Dorr. 2002. Handling Translation Divergences: Combining Statistical and Symbolic Techniques in Generation-Heavy Machine Translation. In *Proceedings of the Fifth Conference of the Association for Machine Translation in the Americas*. Tiburon, CA, pages 84–93.
- Nizar Habash, Bonnie J. Dorr, and Christof Monz. 2006. Challenges in Building an Arabic GHMT system with SMT Components. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas*. Boston, MA, pages 56–65.
- Jan Hajič, Eduard Bejček, Alevtina Bémová, Eva Buráňová, Eva Hajičová, Jiří Havelka, Petr Homola, Jiří Kárník, Václava Kettnerová, Natalia Klyueva, Veronika Kolářová, Lucie Kučová,

- Markéta Lopatková, Marie Mikulová, Jiří Mírovsk, Anna Nedoluzhko, Petr Pajas, Jarmila Panevová, Lucie Poláková, Magdaléna Rysová, Petr Sgall, Johanka Spoustová, Pavel Straňák, Pavlína Synková, Magda evčíková, Jan tpánek, Zdeňka Ureová, Barbora Vidová Hladká, Daniel Zeman, Šárka Zikánová, and Zdeněk Žabokrtský. 2018. [Prague dependency treebank 3.5](#). LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. <http://hdl.handle.net/11234/1-2621>.
- Ray Jackendoff. 1983. *Semantics and Cognition*. MIT Press, Cambridge, MA.
- Ray Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge, MA.
- Ray Jackendoff. 1996. The Proper Treatment of Measuring Out, Telicity, and Perhaps Even Quantification in English. *Natural Language and Linguistic Theory* 14:305–354.
- Karin Kipper, Anna Korhonen, Neville Ryant, and Martha Palmer. 2007. A Large-scale Classification of English Verbs. In *Language Resources and Evaluation*.
- P. Kordjamshidi, M. Van Otterlo, and Marie-Francine Moens. 2010. Spatial Role Labeling: Task Definition and Annotation Scheme. In *Proceedings of Language Resources and Evaluation Conference*.
- Raj Korpan, Susan L. Epstein, Anoop Aror, and Gil Dekel. 2017. WHY: Natural Explanations from a Robot Navigator. In *AAAI 2017 Fall Symposium on Natural Communication for Human-Robot Collaboration*.
- Geert-Jan Kruiff, Hendrik Zender, Patric Jensfelt, and Henrik Christensen. 2007. Situated dialogue and spatial organization: What, where ... and why? In *International Journal of Advanced Robotic Systems*.
- Beth Levin. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. The University of Chicago Press.
- Gina Levow, Bonnie J. Dorr, and Dekang Lin. 2000. Construction of Chinese-English Semantic Hierarchy for Cross-language Retrieval. In *ICCLC'2000 International Conference on Chinese Language Computing*.
- Stephanie Lukin, Reginald Hobbs, and Clare Voss. 2018. A pipeline for creative visual storytelling. In *NAACL 2018 StoryNLP*.
- M Marge, C Bonial, A Fouts, C Hayes, C Henry, K Pollard, R Artstein, C Voss, and D Traum. 2017. Exploring Variation of Natural Human Commands to a Robot in a Collaborative Navigation Task. In *ACL2017 RoboNLP workshop*.
- Nikolaos Mavridis and Deb Roy. 2006. Grounded situation models for robots: Where words and percepts meet. In *Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems*.
- Pooja Moolchandani, Cory Hayes, and Matthew Marge. 2018. Evaluating robot behavior in response to natural language. In *HRI '18 Companion: ACM/IEEE International Conference on Human-Robot Interaction Companion*.
- Mari Broman Olsen, Bonnie J. Dorr, and Scott Thomas. 1997. Toward Compact Monotonically Compositional Interlingua Using Lexical Aspect. In *Proceedings of the Workshop on Interlinguas in MT*. San Diego, CA, pages 33–44.
- Martha Palmer, Claire Bonial, and Jena D. Hwang. 2017. VerbNet: Capturing English Verb behavior, Meaning and Usage. In Susan Chipman, editor, *The Oxford Handbook of Cognitive Science*, Oxford University Press.
- Katerina Pastra, Eirini Balta, Panagiotis Dimitrakis, and Giorgos Karakatsiotis. 2011. Embodied Language Processing: A New Generation of Language Technology. In *Language-Action Tools for Cognitive Artificial Agents: Papers from the 2011 AAI Workshop (WS-11-14)*.
- Ian E. Perera, James F. Allen, Lucian Galescu, Choh Man Teng, Mark H. Burstein, Scott E. Friedman, David D. McDonald, and Jeffrey M. Rye. 2017. Natural Language Dialogue for Building and Learning Models and Structures. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA.*, pages 5103–5104.
- J. Pustejovsky, J. L. Moszkowicz, and M. Verhagen. 2011. [Using ISO-Space for Annotating Spatial Information](#). <http://www2.denizyuret.com/bib/pustejovsky/pustejovsky2011cosit/COSIT-ISO-Space.final.pdf>.
- James Pustejovsky and Kiyong Lee. 2017. Enriching the Notion of Path in ISOspace. In *Proceedings of the 13th Joint ISO-ACL Workshop on Interoperable Semantic Annotation (ISA-13)*.
- Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, Collin F. Bakerand, and Jan Scheffczyk. 2016. [Framenet ii: Extended theory and practice](#). https://framenet.icsi.berkeley.edu/fndrupal/the_book.
- Michael Spranger, Jakob Suchan, and Mehul Bhatt. 2016. Robust Natural Language Processing Combining Reasoning, Cognitive Semantics, and Construction Grammar for Spatial Language. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*.

- Amber Stubbs and James Pustejovsky. 2012. *Natural Language Annotation for Machine Learning*. O'Reilly Media.
- Leonard Talmy. 2014. Foreward: Past, present, and future of motion research. In Iraide Ibarretxe-Antugano, editor, *Motion and Space across Languages: Theory and Applications. HCP (Human Cognitive Processing) Series*, John Benjamins.
- Thora Tenbrink. 2011. Reference frames of space and time in language. In *Journal of Pragmatics*. volume 43, pages 704–722.
- D Traum, C Henry, S Lukin, R Artstein, F Gervitz, K Pollard, C Bonial, S Lei, C Voss, M Marge, C Hayes, and S Hill. 2018. Dialogue Structure Annotation for Multi-Floor Interaction. In *LREC*.
- David Traum and Nizar Habash. 2000. Generation from Lexical Conceptual Structures. In *Proceedings of the Workshop on Applied Interlinguas, North American Association for Computational Linguistics / Applied NLP Conference*. pages 34–41.
- Lucy Vanderwende, Arul Menezes, and Chris Quirk. 2015. An amr parser for english, french, german, spanish and japanese and a new amr-annotated corpus. In *HLT-NAACL*.
- Clare Voss and Bonnie J. Dorr. 1995. Toward a Lexicalized Grammar for Interlinguas. *J. of Machine Translation* 10:14–3.
- Clare R. Voss, Bonnie J. Dorr, and M. U. Şencan. 1998. Lexical Allocation in Interlingua-based Machine Translation of Spatial Expressions. In Patrick Oliver and Klaus-Peter Gapp, editors, *Representation and Processing of Spatial Expressions*, L. Erlbaum Associates Inc., Hillsdale, NJ, USA, pages 133–148.