

STYLUS: A Resource for Systematically Derived 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

Starting from an existing lexical-conceptual structure (LCS) Verb Database of 500 verb classes (containing a total of 9525 verb entries), we automatically derived a resource that supports argument identification for language understanding and argument realization for language generation. The extended resource, called STYLUS (SysTematicallY Derived Language USage), supports constraints at the syntax-semantics interface through the inclusion of components of meaning and collocations. We show that the resulting resource covers three cases of language usage patterns both for *spatially* oriented applications such as dialogue management for robot navigation and for *non-spatial* applications such as generation of cyber-related notifications.

1 Introduction

This paper presents a derivative resource, called STYLUS (SysTematicallY Derived Language USage), produced through extraction of a set of argument realizations from lexical-semantic representations for a range of different verb classes (Appendix A). Prior work (Jackendoff, 1996; Levin, 1993; Olsen, 1994; Kipper et al., 2008; Palmer et al., 2017) has suggested a close relation between underlying lexicalesemantic structures of predicates and their syntactic argument structure. Subsequent work (Dorr and Voss, 2018) argued that regular patterns of language usage can be systematically derived from lexicalesemantic representations and used in applications such as dialogue management for robot navigation. The latter investigation focused on the *spatial* dimension, e.g., motion and direction.

We adopt the view that this systematicity also holds for verbs in the *non-spatial* dimension, including those that have been metaphorically related to the spatial dimension by a range of corpus-based techniques (Dorr and Olsen, 2018). We consider one example of a non-spatial application: generation of cyber-related textual notifications. We argue that this application requires knowledge at the syntax-semantics interface that is analogous to spatial knowledge for robot navigation. A recent survey of narrative generation techniques (Kybartas and Bidarra, 2017) highlights several important components of narrative generation (including a story, plot, space, and discourse for telling the story), leaving open the means for surface-realization of arguments from an underlying lexical-semantic structure. STYLUS is designed to accommodate mechanisms for filling this gap.

STYLUS extends an existing Lexical-Conceptual Structure Verb Database (LVD) (Dorr et al., 2001) that includes 500 verb classes (9525 verb entries) with structurally-specified information about realization of arguments for both spatial and non-spatial verbs. STYLUS was produced through systematic derivation of regular patterns of language usage (Block, Overlap, Fill) without requiring manual annotation. The next section reviews related work, starting with the spatial underpinnings of the original LVD. Following this, we describe our extensions and present examples for two applications: natural language processing for robot navigation and generation of cyber-related notifications.

2 Background

2.1 Spatial Language Understanding

Spatial language understanding has made great strides in recent years, with the emergence of language resources and standards for capturing spatial information, e.g., (ISO-24617-7, 2014), which provide guidelines for annotating spatial information in English language texts (Pustejovsky and Lee, 2017;

* This work is licensed under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>.

Pustejovsky and Yocum, 2014). This work differs from the perspective adopted for STYLUS in that it provides annotation guidelines for training systems for spatial information extraction, and so it does not focus on generalized mappings at the syntax-semantics interface.

The 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). Prepositions and particles (*near*, *off*) and verbs of position and movement (*lean*, *swim*) have corresponding components of meanings and collocations 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., 2011). 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*.

2.2 Lexical-Conceptual Structure Verb Database (LVD)

Lexical Conceptual Structure (LCS) (Jackendoff, 1983; Jackendoff, 1990; Dorr, 1993; Dowty, 1979; Guerssel et al., 1985) has been used for wide-ranging applications, including interlingual machine translation (Voss and Dorr, 1995; Habash and Dorr, 2002), lexical acquisition (Habash et al., 2006), crosslanguage information retrieval (Levow et al., 2000), language generation (Traum and Habash, 2000), and intelligent language tutoring (Dorr, 1997).

LCS incorporates primitives whose combination captures syntactic generalities, i.e., actions and entities must be systematically related to a syntactic structure: GO, STAY, BE, GO-EXT, ORIENT, and also an ACT primitive developed by Dorr and Olsen (1997). LCS is grounded in the spatial domain and is naturally extended to non-spatial domains, as specified by *fields*. 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]].

The 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 to a total of 500 classes covering 9525 verb entries (an additional 5500+ verb entries beyond the original 4000+ verb entries). The first 44 classes were added beyond the original set of semantic classes (Dorr and Jones, 1996). The remaining classes were derived through aspectual distinctions to yield a set of LCS *classes* that were finer-grained than the original Levin classes (Olsen et al., 1997). Each LCS class consists of a set of verbs and, in several cases, the classes included *non-Levin words* (those not in Levin (1993)), derived semi-automatically (Dorr, 1997).

The original LVD provides a mapping of lexical-semantic structures to their surface realization. This mapping serves as a foundation for the enrichments that yield STYLUS. The new resource benefits from decades of prior study that led to the LVD. Specifically, Levin’s classes are based on significant corpus analysis and have been validated in numerous within-language studies (Levin and Rappaport Hovav, 1995; Rappaport Hovav and Levin, 1998) and cross-language studies (Guerssel et al., 1985; Levin, 2015). Thus, STYLUS is expected to have an important downstream impact, in both depth and breadth, for future linguistic investigations and computational applications.

2.3 Syntax-Semantics Interface

Prior work (Jackendoff, 1996; Levin, 1993; Dorr and Voss, 1993; Voss and Dorr, 1995; Kipper et al., 2004; Kipper et al., 2008; 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. VerbNet (Kipper et al., 2004) reinforces the view in this resource paper, that prepositions and their relation with verb classes serve as significant predictors of semantic content, but does not leverage an inner structure of events for compositional derivation of argument realizations.

FrameNet also sits at the syntax-semantics interface (Fillmore, 2002), with linking generalizations based on valency to map semantic frames (events and participants) to their corresponding surface structure. Osswald and Van Valin (2014) point out that such generalizations are hindered by bottom-up/datadriven frames and they argue for a richer frame representation with an inner structure of an event. This notion of “inner structure” is seen in the work of Voss et al. (1998) and Dorr and Voss

(2018), which suggests that the generation of a preposition (in English) is dependent on both the internal semantics of the predicate and structural idiosyncracies at the syntax-semantics interface. As such, the following two definitions are fundamental to the syntax-semantics mappings adopted in STYLUS¹:

- **Component of meaning:** Implicit semantic unit, such as UPWARD for the verb *elevate*
- **Collocation:** Explicit juxtaposition of a particular word, such as *up* for the verb *lift* in “*lift up*”

Leveraging these definitions, Section 3 describes the enrichments necessary to produce STYLUS without requiring training on annotated data.

3 Addition of Components of Meaning and Collocations to LVD Classes

We investigate the systematic derivation of language usage patterns for both understanding and generation of language and leverage this investigation to simplify and enrich the LVD presented in Section 2. The resulting resource, STYLUS (downloadable as described in Appendix 5), relies on lexically implicit *components of meaning* and lexically explicit *collocations* (as defined above) to cover three cases of language usage patterns applicable to language understanding and generation:

- **Block** refers to components of meaning that do not co-occur with their collocational counterparts, e.g., *elevate* and *ascend* include the UPWARD component and thus do not co-occur with the collocation *up*;
- **Overlap** refers to components of meaning that optionally co-occur with their collocational counterparts, e.g., *lift* and *raise* include the UPWARD component but optionally co-occur with the collocation *up*;
- **Fill** refers to underspecified components of meaning that fall into one of two cases:
 - Oblig: *obligatory* co-occurrence with collocations, e.g., the verb *put* does not specify a direction and thus always co-occurs with a collocation such as *up*;
 - Opt: *optional* co-occurrence with collocations, e.g., the motion verb *move* does not specify a direction but optionally co-occurs with a collocation such as *up*.

STYLUS contains simplified verb classes from the LVD, omitting the full LCS structures and thematic roles while adding prepositional *collocations* and *components of meaning*. For example, in the LVD *Verbs of inherently directed motion* (Class 51.1.a in (Levin, 1993)), the verb *leave* need not co-occur with a prepositional collocate (e.g., *leave the room*) whereas the verb *depart* can co-occur with *from* (e.g., *departed from the room*). For either case, the component of meaning is uniformly “GO TO (outside the room) FROM (inside of the room),” and the collocation *from* is optional:

LVD Class Entry:

```
(:NUMBER "51.1.a"
:NAME "Verbs Inherently Directed Motion / -from/to"
:WORDS (advance arrive ascend climb come depart descend enter escape
        exit fall flee go leave plunge recede return rise tumble)
:NON_LEVIN_WORDS (approach come! defect head)
:LCS (go loc (* thing 2)
      ((* from 3) loc (thing 2) (at loc (thing 2) (thing 4)))
      ((* to 5) loc (thing 2) ([at] loc (thing 2) (thing 6)))
      (!!+ingly 26))
:THETA_ROLES ((1 "_th,src(from),goal(towards) "
                (1 "_th,goal,src(from) "
                (1 "_th,src(from),goal(to) ")))
```

Derivation of components of meaning and collocations, along with their obligatoriness or optionality, was achieved by a simple automated procedure, without manual annotation. Components of mean-

¹ Teletype font is used for components of meaning such as UPWARD. Several examples throughout this paper were purposely selected to illustrate the full range of syntactic realizations for the concept of “upwardness.” Other verbs and collocations could easily have been selected (e.g., *lower* with the collocation *down*), but a varied selection of lexical distinctions would confound the illustration of more general distinctions at the syntax-semantics interface.

ing were derived either from the LCS structure or from verb-prepositions pairs in a “Categorial Variation” database (Habash and Dorr, 2003). For example, the LCS above (from the :LCS slot) includes a sublexical component of meaning, ((** from 3*) *loc (thing 2) (at loc (thing 2) (thing 4))*), that maps optionally to a preposition *from* in the surface form (as in *exit (from) the room*). Prepositional collocations (such as *from*) were derived from the thematic roles (in the :THETA ROLES slot) of the original LVD. These prepositions are specified in parentheses, and are preceded either by a comma (,) for optional collocations or by underscore (_) for obligatory collocations. In the example above, the theme is obligatory (*_th*), whereas the source (*, src*) and goal (*, goal*) are optional. Verbs are thus paired with their corresponding prepositions by virtue of their class membership in the final STYLUS resource²:

STYLUS Class Entry:

```
(:NUMBER "51.1.a"
:NAME "Verbs Inherently Directed Motion / -from/to"
:WORDS (advance arrive ascend climb come depart descend enter escape
        exit fall flee go leave plunge recede return rise tumble)
:NON_LEVIN_WORDS (approach come! defect head)
:COLLOCATIONS ("from" "to" "towards")
:COMPONENT_OF_MEANING (FROM, AWAY_FROM, OUT_OF, OR, UP_TO, UP,
        BEFORE, INTO, TO)) ;; Fill-Opt <<SPATIAL>>
```

The derivative collocations and components of meaning obviate the need for spelling out the full LCS, thus ensuring a more compact representation of the original LVD. In addition, the Block, Overlap, Fill-Oblig, and Fill-Opt designations were easily derived automatically from the :COLLOCATIONS and :COMPONENT OF MEANING slots according to the four rules below:

- (1) Tag **Block** if :COMPONENT OF MEANING filled and :COLLOCATIONS not filled, e.g., *Face the doorway*;
- (2) Tag **Overlap** if :COMPONENT OF MEANING and :COLLOCATIONS both filled, and use of comma (,) in front of the corresponding thematic role, e.g., “*,src (behind)*” for *Follow (behind) the car*;
- (3) Tag **Fill-Oblig** if :COMPONENT OF MEANING and :COLLOCATIONS both filled, and use of underscore (_) in front of the corresponding thematic role, e.g., “*_goal (to)*” for *Put the box onto the table*;
- (4) Tag **Fill-Opt** if :COMPONENT OF MEANING not filled and either :COLLOCATIONS not filled or use of comma (,) in front of a modifier role, as in *Move the box (to the table)*.

Finally, each entry was automatically tagged <<SPATIAL>> if the LCS indicated a Loc(ational) or Poss(essional) field in association with the underlying primitive (GO, BE, etc.). All non-spatial entries were automatically tagged <<NON-SPATIAL>>.

4 STYLUS: Spatial and Non-Spatial Subset of LVD Classes

The number of classes associated with the language usage patterns introduced above is in the table below, with tallies for the full set of LVD classes, the spatial subset, and non-spatial subset.³ Representative examples of verbs in both the spatial subset (e.g., for robot navigation) and the non-spatial subset (e.g., cyber-notification generation) are provided in Table 1.

Interestingly, the Spatial Subset of classes is sizeable (44% of the entire set). However, STYLUS development has yielded generalizations beyond prior LCS-inspired work that focused strictly on spatial verbs (Voss et al., 1998). Most notably, the Block, Overlap and Fill patterns are generalizable to a high number of LCS classes that are non-spatial as well. The obvious cases are those with non-zero values in Table 1 (Overlap and Fill): e.g., *tread on* (Overlap), *cram for* (Fill-Oblig), and *blend together* (Fill-Opt). However, although the number of non-spatial Block cases appears to be 0, the corresponding 7 spatial usages also extend to metaphorical extensions, e.g., in the phrase *elevated her spirits*, the use of “up” is blocked. Although STYLUS encodes *elevate* as a spatial verb, this same lexico-semantic

² In the original LVD, the :SPATIAL COMPONENT OF MEANING field name was used. More recently, such components have been determined to be useful for non-spatial verbs as well. For example, *admire* (in class 31.2.c) has a *toward* component of meaning, whereas *abhor* has an *away-from* component of meaning. Thus, the more general :COMPONENT OF MEANING field name has been subsequently adopted.

³ N/A refers to verb classes whose members take bare NP or S arguments. Intrans refers to Intransitive verbs.

property applies for its non-spatial (metaphorical) usage. Analogously, STYLUS encodes the contrasting verb *lift* as a spatial verb, designated as Overlap, to enable *lifted up her hand*. A metaphorical (non-spatial) usage would be *lift up her spirits*. Such cases indicate that Block, Overlap, and Fill apply even more broadly than was conceived in the original LVD—which means that the numbers in the Non-Spatial Subset column above are understated and would benefit from additional analysis of metaphorical extensions.

Lexical Ops	LVD Classes	Spatial Subset	Spatial Examples	Non-Spatial Subset	Non-Spatial Examples
Block	7	7	elevate, face, pocket	0	infect, archive
Overlap	17	10	advance, lower, lift	7	follow, precede
Fill-Oblig	310	128	drive, rotate, put	182	mount, install
Fill-Opt	87	49	remove, slide	38	move, add
Intrans	6	3	float, part, squirm	3	choke, drown, snap
N/A	73	12	—	61	—
Tot. Classes	500	219		281	
Tot. Verbs	9525	4640		4885	

Table 1: Language Usage Patterns for Spatial and Non-Spatial Verbs with Representative Examples

To explore the broader applicability of the Block, Overlap, and Fill patterns, we first examined verbs in the Spatial Subset and determined that many of them are among those relevant to robot navigation, e.g., *move*, *go*, *advance*, *drive*, *return*, *rotate*, and *turn*. Others are easily accommodated by extending classes—without modification to the spatial language usage patterns described above. For example, *back up* matches the class template containing *advance*, and *pivot* matches the class template containing *rotate*. We then considered the Non-Spatial Subset for generation of cyber-related notifications, a sub-task of an ongoing cyber-attack prediction project (Dalton et al., 2017). We determined that the Fill, Overlap, and Block patterns apply to members of the same classes that were relevant to robot navigation. Specifically, we examined the four classes shown in Table 2.

Class [Pattern]	Robot Navigation		Cyber Notification (Generation Only)
	Understanding	Generation	
9.8-Fill [Block]	<i>Face</i> the doorway <i>Face</i> to the doorway	Where do I <i>face</i> ? *Where do I <i>face</i> to?	Risk that trojan horse virus will <i>infect</i> (*to) System A increased 5%
47.8.1-Contig [Overlap]	<i>Follow</i> the car Follow behind the car	Which car do I <i>follow</i> (behind)?	Risk increased 25% that VPN failure will <i>follow</i> (behind) malware attack
9.1-Put [FillOblig]	<i>Put</i> it on the table * <i>Put</i> the box	Where do I <i>put</i> it? *Do I <i>put</i> the box?	Risk of malware increases 5% if disk A <i>mounted</i> on file server B
11.2-Slide [FillOpt]	<i>Move</i> it to the table <i>Move</i> it	Where do I <i>move</i> it? Do I <i>move</i> it?	10% increased malware risk if System A files are <i>moved</i> (to System B)

Table 2: Class-Based Patterns for Robot Navigation and Cyber Notification

Note that Block, Overlap, and Fill can be applied as validity constraints on language usage patterns in the *spatial* domain of human-robot commands (for understanding and generation).⁴ Correspondingly, for the *non-spatial* domain of cyber notification generation, these same four classes included other class members (e.g., for 9.8 and 9.1) or meaning extensions of the same class members (e.g., for 47.8.1 and 11.2) that exhibited the same language usage patterns as their robot navigation counterparts.

5 Conclusion and Future Work

STYLUS provides the basis for both understanding and generation in the spatial robot navigation domain and for generation in the *non-spatial* cyber notification domain. A larger scale application and evaluation of the effectiveness of STYLUS for understanding and generation in these two domains is a

⁴ An asterisk at the start of a sentence indicates an invalid generated form. Verbs in the designated class are italicized.

future area of study. A starting point is an ongoing Bot Language project (Marge et al., 2017) that has heretofore focused on dialogue annotation (Traum et al., 2018) and has not yet incorporated lexicon-based knowledge necessary for automatically detecting incomplete, vague, or implicit navigation commands.

Another avenue for exploration is the enhancement of cyber notifications through systematic derivation of mappings to surface realizations for other parts of speech. This work will involve access to a “Categorial 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 *infect* includes the nominalized form *infection*, which provides additional options that may be more fluent in cyber-related notifications, e.g., *viral infection of system* might be considered less stilted than *virus will infect system*.

Although STYLUS is strictly for English, the lexical-semantic foundation from which it is derived has been investigated for a range of multilingual applications (Habash et al., 2006; Cabezas et al., 2001; Levow et al., 2000). Future work will examine the derivation of this resource for non-English languages.

Acknowledgements

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

- Clara Cabezas, Bonnie J. Dorr, and Philip Resnik. 2001. Spanish Language Processing at University of Maryland: Building Infrastructure for Multilingual Applications. In *Proceedings of the Second International Workshop on Spanish Language Processing and Language Technologies (SLPLT-2)*, Jaen, Spain, <http://www.umiacs.umd.edu/users/bonnie/Publications/slplt-01.htm>.
- Adam Dalton, Bonnie Dorr, Leon Liang, and Kristy Hollingshead. 2017. Improving cyber-attack predictions via unconventional sensors discovered through information foraging. In *Proceedings of the 2017 International Workshop on Big Data Analytics for Cyber Intelligence and Defense, BDA4CID '17*, pages 4560–4565. IEEE.
- Bonnie Dorr and Doug Jones. 1996. Acquisition of Semantic Lexicons: Using Word Sense Disambiguation to Improve Precision. In *Proceedings of the Workshop on Breadth and Depth of Semantic Lexicons, 34th Annual Conference of the Association for Computational Linguistics*, pages 42–50. Kluwer Academic Publishers.
- Bonnie J. Dorr and Mari Broman Olsen. 1997. Deriving Verbal and Compositional Lexical Aspect for NLP Application. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, pages 151–158.
- Bonnie J. Dorr and Mari Broman Olsen. 2018. Lexical Conceptual Structure of Literal and Metaphorical Spatial Language: A Case Study of Push. In *Proceedings of the NAACL Workshop on Spatial Language Understanding*, pages 31–40.
- 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.
- Bonnie J. Dorr and Clare Voss. 2018. The Case for Systematically Derived Spatial Language Usage. In *Proceedings of the NAACL Workshop on Spatial Language Understanding*, pages 63–70.
- 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.

- 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.
- David Dowty. 1979. *Word Meaning and Montague Grammar*. Reidel, Dordrecht.
- Charles J. Fillmore. 2002. Linking sense to syntax in FrameNet (keynote speech). In *19th International Conference on Computational Linguistics*, Taipei. COLING, <http://www.lirmm.fr/~lafourca/ML-enseign/DEA/DEA03-TALN-articles/fillmore.pdf>.
- M. Masten Guerssel, Kenneth Hale, Mary Laughren, Beth Levin, and Josie White Eagle. 1985. A Cross-linguistic Study of Transitivity Alternations. In W. H. Eilfort and P. D. Kroeber nad 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 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*, pages 84–93, Tiburon, CA.
- Nizar Habash and Bonnie Dorr. 2003. A categorial variation database for English. In *NAACL/HLT 2003, Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference*, pages 96–102.
- 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*, pages 56–65, Boston, MA, August.
- ISO-24617-7. 2014. Language Resource management Semantic Annotation Framework Part 7: Spatial information (ISOspace), <http://www.iso.org/standard/60779.html>.
- 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, Benjamin Snyder, and Martha Palmer. 2004. Using Prepositions to Extend a Verb Lexicon. In *Proceedings of the HLT/NAACL Workshop on Computational Lexical Semantics*, pages 23–29.
- Karin Kipper, Anna Korhonen, Neville Ryant, and Martha Palmer. 2008. A Large-scale Classification of English Verbs. *Language Resources and Evaluation*, 42(1):21–40.
- Parisa Kordjamshidi, Martijn Van Otterlo, and Marie-Francine Moens. 2011. Spatial Role Labeling: Task Definition and Annotation Scheme. In *Proceedings of the Seventh Conference on International Language Resources and Evaluation*, pages 413–420.
- Ben Kybartas and Rafael Bidarra. 2017. A survey on story generation techniques for authoring computational narratives. *IEEE Transactions on Computational Intelligence and AI in Games*, 9(3):239–253.
- Beth Levin and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface, Linguistic Inquiry Monograph 26*. MIT Press, Cambridge, MA.
- Beth Levin. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. The University of Chicago Press.
- Beth Levin. 2015. Verb Classes Within and Across Languages. In B. Comrie and A. Malchukov, editors, *Valency Classes: A Comparative Handbook*, pages 1627–1670. De Gruyter, Berlin.
- Gina-Anne Levow, Bonnie J. Dorr, and Dekang Lin. 2000. Construction of Chinese-English Semantic Hierarchy for Cross-Language Retrieval. In *Workshop on English-Chinese Cross Language Information Retrieval, International Conference on Chinese Language Computing*, pages 187–194, Chicago, IL.
- Matthew Marge, Claire Bonial, Ashley Fouts, Cassidy Henry Cory Hayes, Kimberly Pollard, Ron Artstein, Clare Voss, and David Traum. 2017. Exploring Variation of Natural Human Commands to a Robot in a Collaborative Navigation Task. In *ACL2017 RoboNLP workshop*, pages 58–66.

- Malka Rappaport Hovav and Beth Levin. 1998. Building Verb Meanings. In M. Butt and W. Geuder, editors, *The Projection of Arguments: Lexical and Compositional Factors*, pages 97–134. CSLI Publications, Stanford, CA.
- 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*, pages 33–44, San Diego, CA, October.
- Mari Broman Olsen. 1994. The Semantics and Pragmatics of Lexical and Grammatical Aspect. *Studies in the Linguistic Sciences*, 24(1–2):361–375.
- Rainer Osswald and Robert D. Van Valin. 2014. Framenet, Frame Structure, and the Syntax-semantics Interface. In T. Gamerschlag, D. Gerland, R. Osswald, and W. Petersen, editors, *Frames and Concept Types: Applications in language and philosophy (Studies in Linguistics and Philosophy 94)*, pages 125–56. Heidelberg: Springer.
- 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.
- 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)*, <http://aclweb.org/anthology/W17-7415>.
- James Pustejovsky and Zachary Yocum. 2014. Image annotation with iso-space: Distinguishing content from structure. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 426–431.
- James Pustejovsky, Jessica Moszkowicz, and Marc Verhagen. 2011. Using ISO-Space for Annotating Spatial Information, <http://www2.denizyuret.com/bib/pustejovsky/pustejovsky2011cosit/COSIT-ISO-Space.final.pdf>.
- 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.
- David Traum, Cassidy Henry, Stephanie Lukin, Ron Artstein, Felix Gervits, Kimberly Pollard, Claire Bonial, Su Lei, Clare Voss, Matthew Marge, Cory Hayes, and Susan Hill. 2018. Dialogue Structure Annotation for Multi-Floor Interaction. In Nicoletta Calzolari (Conference chair), Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hlne Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, and Takenobu Tokunaga, editors, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 7-12, 2018. European Language Resources Association (ELRA).
- Clare R. Voss and Bonnie J. Dorr. 1995. Toward a Lexicalized Grammar for Interlinguas. *J. of Machine Translation*, 10:143–184.
- 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, editor, *Representation and Processing of Spatial Expressions*, pages 133–148. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

Appendices

Appendix A. Supplemental Material: STYLUS

STYLUS (SysTematically Derived Language USAge) is a simplified and extended version of Lexical Conceptual Structure Verb Database (LVD) (Dorr et al., 2001) with 500 classes of verbs that incorporate components of meaning and prepositional collocations (in ascii .txt format). The classes include over 9500 verb entries—about 5500 entries beyond the original 4000+ entries in Levin 1993. This resource can be downloaded from the link below:

<https://www.dropbox.com/s/3fwz1g1wrirjhv1/LCS-Bare-Verb-Classes-Final.txt?dl=0>