Semantic Roles in VerbNet and FrameNet: Statistical Analysis and Evaluation

Aliaksandr Huminski, Fiona Liausvia, and Arushi Goel

Institute of High Performance Computing, Singapore, 138632
{huminskia,liausviaf,arushi_goel}@ihpc.a-star.edu.sg

Abstract. Semantic role theory is a widely used approach for verb representation. Yet, there are multiple indications that semantic role paradigm is necessary but not sufficient to cover all elements of verb structure. We conducted a statistical analysis of semantic role representation in VerbNet and FrameNet to provide empirical evidence of insufficiency. The consequence of that is a hybrid role-scalar approach.

Keywords: Verb representation · Semantic role · VerbNet · FrameNet.

1 Introduction

The semantic representation of verbs has a long history in linguistics. 50 years ago the article "The case for case" [1] gave a start to semantic role theory that is widely used for verb representation. Since semantic role theory is one of the oldest constructs in linguistics, variety of resources with different sets of semantic roles has been proposed.

There are three types of resources depending on the level of role set granularity. The first level is very specific with roles like "eater" for the verb *eat* or "hitter" for the verb *hit*. The third level is very general with the range of roles from only two proto-roles [2] to nine roles. The second level is located between them and contains, to the best of our knowledge, from 10 to 50 roles approximately.

This rough classification corresponds to the largest linguistic resources: Frame Net [3], VerbNet [4] and PropBank [5] that belong to the first, second and third type of resources accordingly. All of them use semantic role representation for verbs and are combined in Unified Verb Index system[1]. They are used widely in most advanced NLP and NLU tasks, such as semantic parsing and semantic role labeling, question answering, information extraction, recognizing textual entailment, and information extraction. Knowledge of semantic representation and verb-argument structure is a key point for NLU systems and applications.

The paper is structured as follows. Section 2 briefly introduces VerbNet and FrameNet, the ideas underlying their construction and the main differences between them. Section 3 focuses on the basic statistical analysis of VerbNet and FrameNet. Section 4 describes advanced statistical analysis that shows that the

[1] http://verbs.colorado.edu/verb-index/

role paradigm itself is necessary but not sufficient for proper representation of all verbs. A hybrid role-scalar approach is presented in Section 5. The final Section 6 reports our concluding observations.

2 VerbNet and FrameNet as Linguistic Resources for Analysis

VerbNet and FrameNet are the two most well-known resources where semantic roles are used. PropBank which is considered as the third resource in the Unified Verb Index, provides a semantic role representation for every verb in the Penn TreeBank [6]. But we will not analyse it in this article since PropBank defines semantic roles on a verb-by-verb basis, not making any higher generalizations[2]. We will not use neither WordNet [7] for the analysis since this resource does not have semantic role representation for verbs.

2.1 VerbNet

VerbNet (VN) is the largest domain-independent computational verb lexicon currently available for English. In this paper we use the version 3.3[3] released in June 2018. It contains 6791 verbs and provides semantic role representation for all of them. VN 3.3 with 39 roles belongs to the second level of role set resources. In other words, the roles are not so fine-grained as in FrameNet and not so coarse-grained as in Propbank. VN was considered together with the LIRICS role set for the ISO standard 24617-4 for Semantic Role Annotation [8–10].

Idea of construction. VN is constructed on Levin's classification of verbs [11]. Verb classification is based on the idea that syntactic behavior of verbs (syntactic alternations) is to a large degree determined by its meaning. Similar syntactic behavior is taken as a method of grouping verbs into classes that are considered as semantic classes. So, verbs that fall into classes according to shared behavior would be expected to show shared meaning components. As a result of that, each verb[4] belongs to a specific class in VN. In turn, each class has a role set that equally characterizes all members (verbs) of the class.

---

[2] Each verb in PropBank has verb-specific numbered roles: Arg0, Arg1, Arg2, etc. with several more general roles that can be applied to any verb. That makes semantic role labeling too coarse-grained. Most verbs have two to four numbered roles. And although the tagging guidelines include a "descriptor" field for each role, such as "kicker" for Arg0 or "instrument" for Arg2 in the frameset of the verb *kick*, it does not have any theoretical standing [5].

[3] http://verbs.colorado.edu/verb-index/vn3.3/

[4] More accurate to use the term *verb sense* here because of verb polycemy.

2.2 FrameNet

FrameNet (FN) is a lexicographic project constructed on a theory of Frame Semantics, developed by Fillmore [12]. We will consider FrameNet releases 1.5 and 1.7[5]. Roles in FN are extremely fine-grained in comparison with VN. According to FN approach, situations and events should be represented through highly detailed roles.

Idea of construction. FN is based on the idea that a word's meaning can be understood only with reference to a structured background [13]. In contrast to VN, FN is first and foremost semantically driven. The same syntactic behavior is not needed to group verbs together. FN takes semantic criteria as primary criteria where roles (called *frame elements* in FN) are assigned not to a verb class, but to a frame that describes an event. Frames are empirically derived from the British National Corpus and each frame is considered as a conceptual structure that describes event and its participants. As a result of that, a frame can include not only verbs, but also nouns, multi-word expressions, adjectives, and adverbs. All of them are grouped together according to the frames. The same as in VN, each frame has a role set that equally characterizes all members of the frame. Role set is essential for understanding an event (situation) represented by a frame.

2.3 VerbNet and FrameNet in Comparison

Table 1 summarizes the differences between VN and FN[6].

Table 1. Basic differences of VN and FN.

	FrameNet	VerbNet
Basis	lexical semantics	argument syntax
Data Source	corpora	linguistic literature
Roles	fine-grained	coarse-grained
Results	frames	verb classes

3 Basic Statistical Analysis

Basic statistical analysis is considered as a necessary step for advanced analysis.

Prior to analysis of the relations across verbs, classes/frames[7] and roles, we need to extract classes/frames, those of them where at least one verb occurs, all unique roles, all verbs in classes/frames, etc.

[5] https://framenet.icsi.berkeley.edu/fndrupal/

[6] We modified the original comparison presented in [14] for our own purposes.

[7] The expression *classes/frames* is used hereinafter to emphasize that verbs are grouped into classes in VN and into frames in FN.

Table 2 summarizes the basic statistics related to VN and FN. It is necessary to provide some comments:

1. Number of classes/frames with and without verbs are different since there are classes in VN and frames in FN with no verbs. Also there are non-lexical frames in FN with no lexical units inside.
2. Calculating the number of classes in VN, we consider the main class and its subclass as 2 different classes even if they have the same role set.
3. Number of roles in FN reflects the number of unique roles that occur only in frames with verbs. We distinguish here the number of uniques roles from the number of the roles with duplicates in all frames (total 10542 for FN 1.7).
4. Number of verbs in reality is a number of verb senses that are assigned to different classes/frames. Because of polysemy the number of verb senses is larger than the number of unique verbs.

Table 2. Basic statistics of VN and FN.

Resource	Number of roles	Number of classes /frames	Number of classes/frames with verbs	Number of verbs	Av. number of verbs per class/frame
VerbNet 3.2	30	484	454	6338	14
VerbNet 3.3	39	601	574	6791	11.8
FrameNet 1.5	656	1019	605	4683	7.7
FrameNet 1.7	727	1221	699	5210	7.45

4 Advanced Statistical Analysis

We will investigate further only the latest versions of VN (3.3) and FN (1.7).

Advanced statistical analysis includes the following 2 types:

- the distribution of verbs per class;
- the distribution of roles per class.

The distribution of verbs per class is about how many verbs similar in meaning are located in one class. The distribution of roles per class is about how many verbs similar in meaning are located in different classes.

4.1 Distribution of Verbs per Class in VN and FN

Distribution of verbs can be presented in 2 mutually dependent modes. First mode consists of 3 steps:

1. calculation of the number of verbs per class;
2. sorting the classes according to the number of verbs;

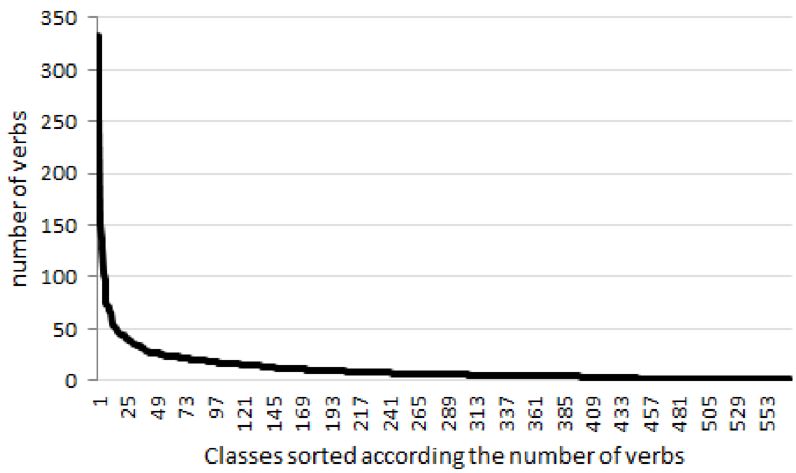

Fig. 1. Distribution of verbs per class in VN 3.3.

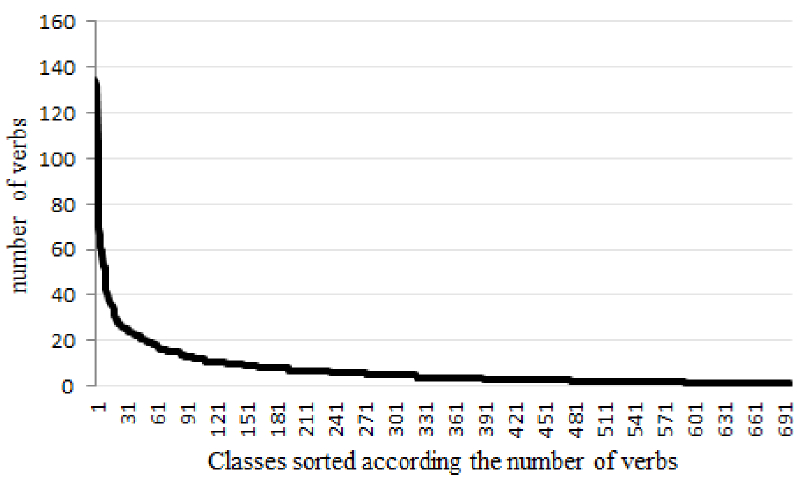

Fig. 2. Distribution of verbs per class in FN 1.7.

3. distribution of the verb number per class starting from the top class.

Fig. 1 for VN and Fig. 2 for FN illustrate the final 3rd step. Based on them one can conclude that:

- verbs are not distributed evenly across the classes/frames. There is a sharp deviation from the average value: 11.8 verbs per class in VN 3.3 and 7.45 verbs per frame in FN 1.7.
- regardless of the resource type (coarse-grained or fine-grained role set), a sharp deviation remains surprisingly the same.

Second mode has the same step#1 and step#2 but the 3rd step is different. It is a distribution of the verb coverage (from 0 to 1) starting from the top class.

Fig. 3 for VN and Fig. 4 for FN illustrate the final 3rd step. Based on them one can conclude that:

Fig. 3. Verb coverage starting from the top classes in VN 3.3.

- verb coverage is a non-linear function;
- regardless of the resource type (coarse-grained or fine-grained role set), verb coverage remains surprisingly the same non-linear function. For example, having 574 classes in VN 3.3, 50% of all verbs are covered by 123 classes, 90% are covered by 319 classes; having 699 frames in FN 1.7, 50% of all verbs are covered by 95 classes, 90% are covered by 416 classes.

4.2 Distribution of Roles per Class in VN and FN

If the distribution of verbs reflects similarity between verbs in one class/frame, the distribution of roles shows similarity between verbs located in different

Fig. 4. Verb coverage starting from the top classes in FN 1.7.

classes/frames. This similarity is unfolded through identical role sets that different classes have. We extracted all different classes that have the same role set and merged them together.

Table 3 shows the difference between total number of classes/frames and the number of merged classes/frames that have the same role set.

Table 3. Statistics of classes/frames with different role sets.

Resource	Number of classes/frames with verbs	Number of classes/frames with verbs that have different role sets
VerbNet 3.3	574	138
FrameNet 1.7	699	619

Table 4 provides some examples of role sets that different classes/frames have.

Second type of role distribution is the number of role occurrences in all classes/frames (see Fig. 5 for VN and Fig. 6 for FN). Based on this distribution, one can conclude that:

- distribution of roles is a non-linear function. Top 2-3 roles occur in almost all classes.
- regardless of the resource type (coarse-grained or fine-grained role set), distribution of roles remains surprisingly the same non-linear function.
- distribution of roles correlates with the distribution of verbs (compare Fig. 5 and Fig. 6 with Fig. 1 and Fig. 2 accordingly).

Table 4. Examples of the same role sets used in different classes/frames.

Resource	Role set for representation of the class/frame	Number of classes/ frames	Number of verbs in all classes
VN 3.3	Agent:Destination:Initial Location:Theme	15	532
	Location:Theme	14	269
	Agent:Recipient:Topic	11	251
	Agent:Instrument:Patient	7	312
	Agent:Instrument:Patient:Result	6	506
FN 1.7	Entity	17	64
	Agent:Theme:Source	3	83
	Experiencer:Stimulus	2	138
	Self mover:Source:Path:Goal:Direction	2	137
	Cause:Theme:Goal or Agent:Theme:Goal	2	125

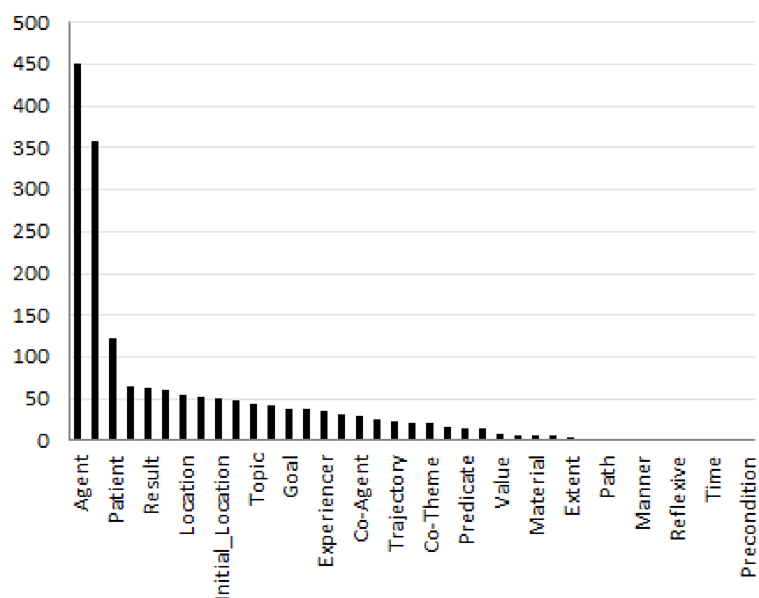

Fig. 5. Role distribution in VN 3.3.

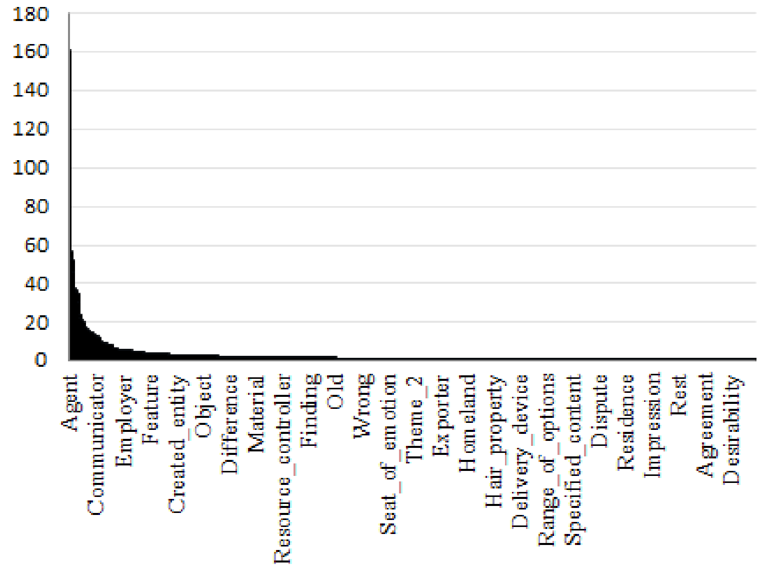

Fig. 6. Role distribution in FN 1.7.

4.3 General Analysis and Evaluation

Both verb and role distributions in VN and FN show sharp deviation from the average value.

Despite the fact that VN and FN are different in the principles of their construction (Table 1) and are significantly different in the number of roles (Table 2), we have identical picture in both VN and FN for all types of distributions. This similarity looks weird since the obvious expectation is the following: the larger is the number of roles, the more even the role/verb distribution should be and the less disproportion is expected.

The reason why it happens. Assigning a role representation to all verbs of a language assumes by default that the set of all verbs is homogeneous and because of homogeneity it can be described through one unique approach: semantic roles. We consider the statistical results as an indication that the role paradigm itself is necessary but not sufficient for proper representation of all verbs.

We argue the set of verbs in a language is not homogeneous. Instead, it is heterogeneous and requires at least 2 different approaches.

5 Hybrid Role-Scalar Approach

For the sake of getting universal semantic representation we offer a hybrid approach: role-scale.

5.1 Hypothesis: Roles Are Not Sufficient for Verb Representation

By definition, any semantic role is a function of a participant represented by NP, towards an event represented by a verb. Nevertheless, to cover all verbs, semantic role theory was extended beyond the traditional definition in such a way to represent, for example, a change of state.

In VN there are roles like Attribute, Value, Extent, Asset etc. that match abstract participants, attributes, and their changes. For example, in the sentence *Oil soared in price by 10%*, "price" is in the role of Attribute and "10%" is in the role of Extent which, according to the definition, specify the range or degree of change.

If we are going to represent a change of state through roles, we need to assign a role to state of a participant, not to a participant itself. Second, a change of state means a change in the value of state in particular direction. For example, the event "heat the water" includes values of state "temperature" for water as a participant. So, to reflect a change of state we need to introduce two new roles: initial value of state and its final value on the scale of increasing values on a dimension of temperature. These 2 new roles look like numbers, not roles, on a scale. It is unclear, what it really means: a role of value.

We argue that the attempts of semantic role theory extension contradict the nature of a semantic role. Roles are just one of the parts in event representation that does not cover an event completely. While a role is a suitable means for action verbs like "hit" or "stab", a scalar is necessary for representation of the verbs like "kill" or "heat". For instance, in semantic role theory the verb *kill* has the role set [Agent, Patient] while the meaning of *kill* contains no information about what Agent really did towards Patient. Having Agent and Patient, the verb *kill* is represented through an unknown action. Meanwhile, what is important for *kill* is not an action itself but the resulting change of state: Patient died. And this part of meaning, being hidden by roles, can be represented via a scalar change "alive-dead". Roles gives us a necessary but not a sufficient representation, since change-of-state verbs do not indicate *how* it was done but *what* was done.

The dichotomy between role and scale can be expressed in other way as the dichotomy between semantic field and semantic scale. A frame is considered as a semantic field where members of a frame are closely related with each other by their meanings, while semantic scale includes a set of values that are scattered along the scale and opposite each other.

5.2 Scale Representation

A scalar change in an entity involves a change in the value of one of its attributes in a particular direction.

Related Work The idea of using scales for verb representation was elaborated by many authors.

Dixon [15, 16] extracted 7 classes of property concepts that are consistently lexicalized across languages: dimension, age, value, color, physical, speed and human propensity .

Rappaport Hovav [17, 18] paid attention that change-of-state verbs and so-called result verbs lexicalize a change in a particular direction in the value of a scalar attribute, frequently from the domain of property concepts of Dixon.

A similar approach comes from cognitive science framework [19–21] that considers verb representation to be based on a 2-vector structure model: a force vector representing the cause of a change and a result vector representing a change in object properties. It is argued that this framework provides a unified account for the multiplicity of linguistic phenomena related to verbs.

Jackendoff [22–24] stated that result verbs representation can be derived from the physical space. Accordingly, a change in the value can be represented the same way as a movement in the physical space. For example, a change of possession can be represented as a movement in the space of possession.

Fleischhauer [25] discussed in detail the idea of verbal degree gradation and elaborated the notion of scalar change. Change-of-state verbs are considered as one of the prototypical examples of scalar verbs. There are two reasons for this: first, some of the verbs are derived from gradable adjectives, and second, the verbs express a change along a scale. It was stated that a change-of-state verb lexicalizes a scale, even if one or more of the scale parameters remain unspecified in the meaning of the verb.

Scale Representation for VerbNet and FrameNet. We just outline the approach how the verbs from the three largest frames in FN (and the classes in VN accordingly) can be additionally represented via scales. More detailed analysis of the scale representation goes beyond the limits of the article.

The benefit that such approach provides is that the large frames can be splitted by identifying within-frame semantic distinctions.

The top largest frame in FN is the frame "Self_motion". According to the definition, it "most prototypically involves individuals moving under their own power by means of their bodies"[8]. The frame contains 134 verbs and corresponds to the *run*-class (51.3.2) with 96 verbs in VN.

The necessity of scale representation for *run*-class was directly indicated by Pustejovsky [26]. To make an explicit representation of change of state he introduced the concept of opposition structure in generative lexicon (GL) as an enrichment to event structure [27]. After that he applied GL-inspired componential analysis to the *run*-class and extracted six distinct semantic dimensions, which provide clear differentiations in meaning within this class. They are: SPEED: *amble, bolt, sprint, streak, tear, chunter, flit, zoom*; PATH SHAPE: *cavort, hopscotch, meander, seesaw, slither, swerve, zigzag*; PURPOSE: *creep, pounce*; BODILY MANNER: *amble, ambulate, backpack, clump, clamber, shuf-*

[8] https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Self_motion

fle; ATTITUDE: *frolic, lumber, lurch, gallivant*; ORIENTATION: *slither, crawl, walk, backpack*.

The second largest frame in FN with 132 verbs is the frame "Stimulate_emotion". This frame is about some phenomenon (the Stimulus) that provokes a particular emotion in an Experiencer[9]. In other words, the emotion is a direct reaction towards the stimulus. It corresponds to the second largest *amuse*-class (31.1) in VN with 251 verbs.

Fellbaum and Mathieu [28] examined experiencer-subject verbs like *surprise*, *fear*, *hate*, *love* etc. where the gradation is richly lexicalized by verbs that denote different degrees of intensity of the same emotion (e.g., *surprise, strike, dumbfound, flabbergast*). The results of analysis show, first, that the chosen verbs indeed possess scalar qualities; second, they confirm the prior assignment of the verbs into broad classes based on a common underlying emotion; finally, the web-data allows to construct consistent scales with verbs ordered according to the intensity of the emotion.

The third largest frame in FN is the frame "Make_noise" (105 verbs) that corresponds to the *sound_emission*-class (129 verbs) in VN. The frame is defined as "a physical entity, construed as a source, that emits a sound"[10]. Snell-Hornby [29] suggested the following scales to characterize verbs of sound: VOLUME (*whirr* vs. *rumble*); PITCH (*squeak* vs. *rumble*); RESONANCE (*rattle* vs. *thud*); DURATION (*gurgle* vs. *beep*).

6 Conclusion

Based on statistical analysis of VerbNet and FrameNet as verb resources we showed empirical evidence of role insufficiency as a unique approach used for verb representation. It supports the hypothesis that roles as a tool for meaning representation do not cover the variety of all verbs. As a consequence of that, another paradigm – scalar approach -- is needed to fill up the gap. The hybrid role-scalar approach looks promising for verb meaning representation and will be elaborated in future.

References

1. Fillmore, Ch.: The Case for Case. In: Universals in Linguistic Theory, pp.1–88. (1968)
2. Dowty, D.: Thematic Proto-Roles and Argument Selection. Language **67**, 547–619 (1991)
3. Baker, C., Fillmore, Ch., Lowe, J.: The Berkeley FrameNet project. In: Proceedings of the International Conference on Computational Linguistics, pp. 86–90. Montreal, Quebec, Canada (1998)

[9] https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Stimulate_emotion

[10] https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Make_noise

4. Kipper-Schuler, K.: VerbNet: A broad-coverage, comprehensive verb lexicon. Ph.D. thesis. Computer and Information Science Dept., University of Pennsylvania, Philadelphia. PA. (2005)
5. Palmer, M., Gildea, D., Kingsbury, P.: The proposition bank: An annotated corpus of semantic roles. Computational linguistics **31**(1), 71–106 (2005)
6. Marcus, M., Kim, G., Marcinkiewicz, M., MacIntyre, R., Bies, A., Ferguson, M., Katz, K., Schasberger, B.: The Penn Treebank: Annotating predicate argument structure. In: ARPA Human Language Technology Workshop, (1994)
7. Fellbaum, Ch., Miller, G.: Folk psychology or semantic entailment? a reply to Rips and Conrad. The Psychological Review **97**, 565–570 (1990)
8. Petukhova, V., Bunt, H.: LIRICS semantic role annotation: Design and evaluation of a set of data categories. In: Proceedings of the Third International Conference on Language Resources and Evaluation (LREC), Marrakech, Morocco (2008)
9. Claire, B., Corvey, W., Palmer, M., Bunt, H.: A hierarchical unication of LIRICS and VerbNet semantic roles. In: Proceedings of the 5th IEEE International Conference on Semantic Computing (ICSC 2011), Palo Alto, CA, USA (2011)
10. Bunt, H., Palmer, M.: Conceptual and representational choices in defining an ISO standard for semantic role annotation. In: Proceedings of the Ninth Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-9), Potsdam, Germany (2013)
11. Levin, B.: English Verb Classes and Alternations: A preliminary investigation. University of Chicago Press, Chicago, IL (1993)
12. Fillmore, Ch.: Frame semantics. In: Linguistics in the Morning Calm, pp.111–137. Soeul: Hanshin (1982)
13. Fillmore, Ch., Atkins, T.: Toward a frame-based lexicon: The semantics of RISK and its neighbors. In: Frames, Fields, and Contrasts: New essays in semantic and lexical organization, pp.75–102. Hillsdale, NJ: Erlbaum (1992)
14. Baker, C., Ruppenhofer, J.: FrameNet's Frames vs. Levin's Verb Classes. In: Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society, pp. 27–38. Berkeley, California, USA (2002)
15. Dixon, R.: Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax. Mouton, Berlin (1982)
16. Dixon, R.: A Semantic Approach to English Grammar. Oxford Univ. Press, UK (2005)
17. Rappaport Hovav, M.: Lexicalized Meaning and the Internal Temporal Structure of Events. In: Crosslinguistic and Theoretical Approaches to the Semantics of Aspect, pp. 13–42. John Benjamins, Amsterdam (2008)
18. Rappaport Hovav, M.: Scalar Roots and their Results. In: Workshop on Roots: Word Formation from the Perspective of "Core Lexical Elements", Universitat Stuttgart, Stuttgart (2009)
19. Gardenfors, P.: The geometry of meaning: Semantics based on conceptual spaces. MIT Press, Cambridge, Massachusetts, USA (2017)
20. Gardenfors, P., Warglien, M.: Using conceptual spaces to model actions and events. Journal of Semantics **29**(4), 487–519 (2012)
21. Warglien, M., Gardenfors, P., Westera, M.: Event structure, conceptual spaces and the semantics of verbs. Theoretical Linguistics **38**(3-4), 159-193 (2012)
22. Jackendoff, R.: Semantics and Cognition. Massachusetts: MIT Press, Cambridge (1983)
23. Jackendoff, R.: Semantic structures. Massachusetts: MIT Press, Cambridge (1990)
24. Jackendoff, R.: Foundations of language. Oxford University Press, Oxford (2002)

25. Fleischhauer, J.: Degree gradation of verbs. Dusseldorf University Press, Dusseldorf (2016)
26. Pustejovsky, J., Palmer, M., Zaenen, A., Brown, S.: Verb meaning in context: Integrating verbnet and gl predicative structures. In: Proceedings of the LREC 2016 Workshop: ISA-12. Potoroz, Slovenia (2016)
27. Pustejovsky, J.: Events and the semantics of opposition. In: Events as Grammatical Objects, pp. 445-482. Center for the Study of Language and Information (CSLI), Stanford, CA, USA (2010)
28. Fellbaum, Ch., Mathieu, Y.: A Corpus-based Construction of Emotion Verb Scales. In: Linguistic Approaches to Emotions in Context. John Benjamins, Amsterdam (2014)
29. Snell-Hornby, M.: Verb-descriptivity in German and English. A Contrastive Study in Semantic Fields. Winter, Heidelberg (1983)