May the FORCE be with Semantics: exploiting LLMs to Image Schematic Knowledge Enrichment

Stefano De Giorgis[1]

[1]*Institute of Cognitive Science and Technologies - National Research Council (ISTC-CNR), Italy*

Abstract

This paper addresses the underspecification of the FORCE image schema. We present a novel hybrid pipeline that combines large language model interactions, linguistic analysis, and knowledge extraction techniques to expand upon Johnson's initial categorization of FORCE types. Our methodology employs Claude 3.5 Sonnet for domain exploration, generates a dataset of 100 force-expressing verbs with contextual sentences, and integrates findings into ImageSchemaNet through AMR2FRED processing and SPARQL querying. Key contributions include: (1) a more nuanced understanding of the FORCE image schema, (2) a validated dataset of force-related linguistic expressions, and (3) an enhanced ontology with empirically derived FORCE concepts. This work bridges the gap between abstract image schema theory and specific linguistic realizations of FORCE, offering practical tools for natural language processing, knowledge representation, and cognitive computing.

1. Introduction

Image schemas, as fundamental cognitive constructs [1], have been instrumental in our understanding of embodied cognition and conceptual metaphor theory. However, while certain image schemas have been extensively investigated, others remain ambiguous, and a comprehensive, agreed-upon list of these schemas continues to elude researchers. This lack of consensus poses significant challenges for advancing the field and applying image schema theory across various domains, including knowledge representation, natural language processing, and cognitive robotics.

Large Language Models (LLMs) offer a promising avenue to address some of these challenges. As the most extensive repositories of general approximate commonsense knowledge currently available, LLMs have inadvertently internalized a degree of embodiment "by proxy" through the way language is used to describe the world [2]. This linguistic representation is inherently grounded in embodied cognition, reflecting how humans conceptualize and interact with their environment. Consequently, LLMs possess a substantial amount of knowledge that is implicitly grounded in their training data, potentially offering insights into image schemas that have yet to be fully explored or defined.

In the realm of image schemas, FORCE remains notably underspecified compared to more thoroughly explored schemas such as SOURCE_PATH_GOAL and its related families, as highlighted by [3]. While Johnson [1] provided an initial distinction of FORCE types, including COMPULSION, BLOCKAGE, COUNTERFORCE, REMOVAL_OF_RESTRAINT, ENABLEMENT, DIVERSION, ATTRACTION, and REPULSION, this categorization has remained largely static. Two significant issues persist: (a) these distinctions have not been subjected to further in-depth analysis, and (b) they lack operational applicability in practical contexts. Our research addresses this gap by employing a hybrid pipeline that combines large language model interactions, linguistic analysis, and knowledge extraction techniques. This approach encompasses initial domain exploration using Claude 3.5 Sonnet, focused data generation of force-related verbs and sentences, expert validation, and the integration of derived knowledge into existing semantic resources through AMR2FRED tool processing and SPARQL querying. By doing so, we aim to provide a more nuanced and operationally viable understanding of the FORCE image schema, bridging the gap between abstract cognitive linguistic theory and practical applications in natural language processing and knowledge representation.

This approach not only addresses the scarcity of annotated resources and comprehensive datasets in the domain of image schemas but also opens up new possibilities for understanding how humans conceptualize their experiences. By tapping into the implicit knowledge encoded in LLMs, we can potentially bridge the gap between abstract image schema concepts and their concrete manifestations in language and thought.

The paper is organised as follows: Section 2 provides useful references to IS literarture and in particular about FORCE analysis; Section 3 details the hybrid approach; Section 4 shows our results and discuss them; finally Section 5 envisions future works and concludes the paper.

The dataset, knowledge base, scripts, and full prompts will be made fully available at "camera ready" time.

2. Related Works

The concept of image schemas (IS), introduced by Lakoff and Johnson [1, 4, 5], has evolved into a foundational theory in cognitive linguistics. The process of perceptual meaning analysis (PMA) [6] in children has provided valuable insights into how knowledge can be acquired through sensorimotor interactions with the environment, and specifically through image schemas. These schemas are now understood as sensorimotor cognitive patterns that shape our perception of the world and establish semantic relations based on bodily experiences [7, 8, 9, 10, 11].

Furthermore, IS constitute a finite set of relational primitives that define the uses and affordances of objects within their environments. Prominent examples include CONTAINMENT, which represents the capacity of one object to be enclosed within another, and SOURCE_PATH_GOAL, which describes the potential or actual movement of objects along specific trajectories. These schemas are not merely static representations but serve as dynamic cognitive structures that underpin more complex reasoning processes.

Indeed, image schemas are widely recognized as foundational elements in human reasoning [12] and have been demonstrated to evolve into sophisticated cognitive functions, including natural language processing and the conceptualization of abstract entities [1, 13]. This evolution occurs through the grounding of these schemas in experiential patterns, highlighting the embodied nature of cognition.

Moreover, image schemas exhibit a remarkable capacity for combinatorial complexity. Consider the concept of "transportation," which can be abstracted beyond specific objects to represent the "movement of object(s) from A to B." In image-schematic terms, this can be formally described as a combination of SOURCE_PATH_GOAL with either SUPPORT or CONTAINMENT [14]. This combinatorial property allows for the formal description of increasingly complex events through the construction of constellations and sequences of image schemas, effectively creating state spaces of conceptual structures [15, 16].

Recent advancements in image schema research have investigated the capabilities of this combinatorial capacity, leading to the development of sophisticated analytical tools and frameworks. Notable among these are the Image Schema Logic ISL^{FOL} [17], which explores the schemas' compositional nature, studies on their role in conceptual blending [18, 19], and the ImageSchemaNet ontology [20]. Additionally, a diagrammatic image schema language has been proposed for visual representation [21], further expanding the field's analytical capabilities.

While corpus-based studies [22, 23] and machine learning approaches [24, 25, 26] have made significant strides in identifying image schemas in natural language, the challenge of comprehensive image schema coverage remains an active area of research.

On the other side, the unique position of LLMs as both products and reflectors of human language use makes them valuable tools for investigating image schemas. By analyzing the patterns and structures within LLM outputs, researchers may uncover new image schemas, clarify ambiguous ones, and potentially work towards a more comprehensive list. Moreover, LLMs could be leveraged to generate synthetic data that captures the nuances of image schemas in linguistic expressions, rapidly expanding the available resources for studying these cognitive patterns.

3. Methodology

Retrieval Augmented Generation (RAG) [27] is a cross-disciplinary research topic which focuses on refining models to make them able to retrieve precise information from big compressed amount of (usually textual) data, traditionally in the form of vector embeddings. Recent advancements in knowledge extraction have made graph generation from text a relatively straightforward process. The recent emergence of Graph-RAG [28] and parallel techniques has demonstrated the feasibility of generating generic subject-predicate-object triples from textual input, even when using "smaller" language models. This capability has opened up new possibilities for automatically triplify information extracted from unstructured text.

However, the true challenge lies in aligning this extracted information with existing knowledge structures, such as ontologies or conceptual schemas in knowledge bases, and effectively leveraging existing semantic web resources. To address this challenge, our methodology, shown in Figure 1, builds upon previous work, including ImageSchemaNet [20], and enriches the formalized knowledge through two primary approaches. First, we employ a chain of thoughts prompting technique for knowledge elicitation from large language models (LLMs), as detailed in subsequent sections. This process allows us to tap into the vast knowledge encoded in LLMs while maintaining a structured approach to information extraction.

Our methodology also yields a valuable by-product: synthetic data augmentation for the Image Schema (IS) Catalogue [29], which serves as the primary resource for image schemas. We process the generated examples through AMR2FRED [30, 31], a tool capable of creating proper RDF graphs from text via Abstract Meaning Representation, and then use SPARQL queries to extract entities that can be declared as triggers for the FORCE image schema in existing repositories. This approach results in a hybrid pipeline that combines LLM-generated information (pink boxes in Figure 1) with symbolic knowledge extraction (light blue boxes). To ensure the quality and relevance of our results, the final synthetic dataset of 100 sentences undergoes manual validation by domain experts, providing a robust foundation for further research and applications in the field of image schemas and its real world application.
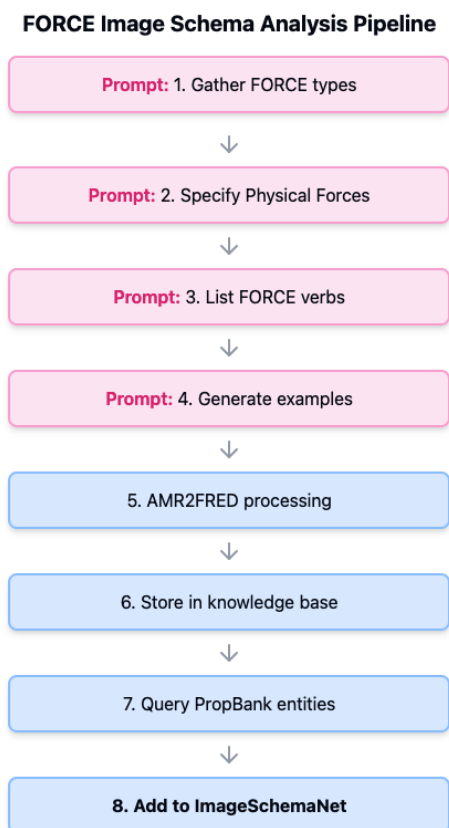

Figure 1: FORCE knowledge enrichment hybrid pipeline.

4. Experiments and Discussion

This section details our experimental approach to exploring and formalizing the FORCE image schema, combining LLM interactions, linguistic analysis, and knowledge extraction techniques. Our methodology encompasses initial domain exploration, focused data generation, and the integration of derived knowledge into existing semantic web resources.

4.1. Generative Knowledge Enrichment

For the generative task it has been used a state of the art model: Claude 3.5 Sonnet, a large language model known for its comprehensive knowledge base and nuanced ability in describing even complex

concepts. The whole generation process is freely replicable since it is done via open chat interface. Our experimental approach commenced with a broad domain exploration, centered around the fundamental question: "List which kind of forces can activate the FORCE image schema."

```
List which kind of forces can activate the FORCE image schema.
```

The model provided a detailed list of generic FORCE types, which included: physical forces, psychological forces, social forces, emotional forces, gravitational forces, electromagnetic forces, nuclear forces (strong and weak), frictional forces, tensile and compressive forces, and centripetal and centrifugal forces. This initial output served as a foundation for our subsequent investigation, offering a diverse range of force categories that could potentially activate the FORCE image schema.

Building upon this initial output, we employed a chain-of-thought prompting technique to generate a controlled vocabulary for each of these FORCE types. This method involved asking the model to elaborate on each force category, providing examples and related concepts.

```
Now for each of these points generate a list of terms which can be used as controled
    vocabulary to generate a knowledge base.
```

After careful analysis of the results, we made a strategic decision to focus specifically on physical forces for several compelling reasons. Firstly, we observed that some of the other categories, such as psychological and social forces, often represented metaphorical extensions of physical forces. These metaphorical uses, while interesting, were already grounded in image-schematic concepts and would potentially introduce complexity in distinguishing between literal and figurative applications of FORCE. Additionally, certain categories, such as emotional forces, were deemed too generic and abstract for our purposes. Moreover, these emotional aspects had been previously addressed in existing literature, notably in [32], which provided a treatment of emotional forces in relation to image schemas.

Our next step involved a more focused prompt aimed at extracting a "complete list of verbs expressing FORCES." We instructed the model to concentrate solely on physical forces and provide a comprehensive list of verbs that evoke any FORCE idea.

```
Ok, now focus only on physical forces and provide a list of all verbs which evokes any
    FORCE idea.
```

The prompt was carefully crafted to elicit a wide range of verbs while maintaining relevance to physical manifestations of FORCE. This process resulted in a collection of 100 verbs expressing various types of FORCEs, ranging from common actions like "push" and "pull" to more specific verbs like "torque" and "propel." To contextualize these verbs and ensure their applicability, we then requested linguistic examples that demonstrate specific occurrences of FORCE for each item on the list.

```
Now provide a sentence as example for each item it this list, which realizes a
    situation of FORCE.
```

The prompt for this stage was designed to generate diverse, realistic sentences that clearly illustrated the FORCE concept embodied by each verb.

This iterative prompting process yielded a final output of 100 lines, each containing a lexical unit pointing to a type of FORCE, accompanied by a sentence illustrating its realization in context. For instance, one entry might include the verb "compress" along with the example sentence "The hydraulic press compressed the metal sheet into a thin disc." This comprehensive collection serves as a valuable resource for further analysis and application of the FORCE (and possibly other co-occurring) image

Lexical Trigger	Example Sentence
Push	The firefighter **pushed** against the heavy door with all his might to rescue those trapped inside.
Pull	With a strong **pull** on the rope, the sailor raised the mainsail against the wind's resistance.
Shove	In the crowded subway, an impatient commuter **shoved** his way through the mass of people.

Table 1
Lexical triggers and sentences generated by Claude 3.5 Sonnet about the *force* image schema.

schema(s) in the field. To ensure the quality and relevance of our dataset, each entry was reviewed by domain experts with previous background in image schemas. The experts evaluated the entries based on criteria such as clarity of FORCE representation, diversity of FORCE types, and linguistic naturalness of the example sentences. Any disagreements were resolved through discussion, and entries that did not meet the quality standards were replaced or refined through additional prompting sessions with the language model. This results in a synthetic dataset, manually curated, listing 100 verbs expressing FORCE.

Table ?? presents an excerpt from this curated dataset, showcasing three representative lexical units associated with different types of FORCE and their corresponding example sentences. This table not only demonstrates the diversity of FORCE-related verbs captured in our study but also illustrates how these verbs are contextualized in natural language use. The full dataset of 100 entries provides a useful extension of the IS Catalogue, with a focus on FORCE image schema and its varied manifestations in language.

4.2. Knowledge Extraction

Following the generation and initial validation of our dataset, detailed in previous section, and shown in pink boxes in Figure 1, we proceeded with a knowledge extraction process to formalize and integrate the FORCE-related information into existing semantic resources.

We employed the AMR2FRED tool to process these sentences. AMR2FRED is a sophisticated natural language processing tool that converts text into Abstract Meaning Representation (AMR) graphs, and then passes the graph to FRED [33] which performs several tasks, among others: frame extraction, entity recognition, and entity alignment to DOLCE foundational ontology [34]. This tool is particularly valuable for our purposes as it preserves the semantic richness of natural language while producing structured, machine-readable representations. Figure 2 shows the graph automatically generated from the sentence "The strong current ripped the swimmer's goggles off her face."

The RDF graphs generated by AMR2FRED for each sentence are collected and stored in a dedicated knowledge base. This knowledge base serves as a centralized repository of formalized FORCE-related semantic structures derived from our curated examples. To extract relevant entities from this knowledge base, we developed a targeted SPARQL query, shown in Figure 4.2. This query is designed to identify and extract entities from PropBank [35], a lexical resource that provides a frame like structure and semantic role labels for English lexicon. Since the AMR2FRED graph is well formed RDF graph, the verb used in the sentence is reified as an instatiation of a specific occurrence (represented as an individual), having `rdf:type` the PropBank entity on which it is disambiguated. The reasoning behind this is that an occurrence of a certain verb is the instatiation of the general concept of that verb (in our case), which is represented on the graph via subsumption relation. The extracted entities represent concepts and actions directly associated with the FORCE image schema as manifested in our dataset.

Finally, we enriched ImageSchemaNet, the existing ontology for image schemas, by adding these extracted entities as direct evocators of the FORCE image schema. This addition was implemented
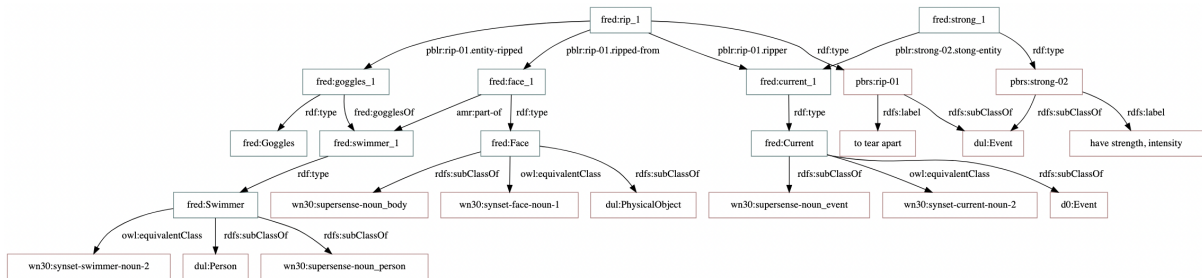
Figure 2: AMR2FRED image for the sentence "The strong current ripped the swimmer's goggles off her face."

in a separate graph within ImageSchemaNet, allowing for clear provenance and easy integration or separation of our contribution. This process not only augments ImageSchemaNet with new, empirically derived FORCE-related concepts but also establishes a concrete link between abstract image schema theory and specific linguistic realizations of force.

SPARQL Query

```
1  SELECT DISTINCT ?e
2  WHERE {
3     ?e rdf:type ?pb_entity .
4  }
```

Box 4.2 SPARQL query to retrieve entities generated out of the original lexical unit in the graph, and the PropBank entity on which it is disambiguated.

The resulting enhanced ontology provides a valuable resource for researchers and practitioners working at the intersection of cognitive linguistics, natural language processing, and knowledge representation.

5. Conclusions and Future Works

In this work, we have addressed the long-standing issue of underspecification in the FORCE image schema. Our research has made significant strides in bridging the gap between abstract theoretical constructs and practical, operational applications. By employing a novel hybrid pipeline that combines large language model interactions, domain experts linguistic analysis, and knowledge extraction techniques, we have expanded upon Johnson's initial categorization of FORCE types. Our methodology, which included domain exploration using Claude 3.5 Sonnet, generation of FORCE-related verbs and contextual sentences, expert validation, and integration with existing semantic resources, has yielded several key achievements. First, we have developed a more nuanced and comprehensive understanding of the FORCE image schema, expanding beyond the original eight categories to include a wider range of force manifestations in language. Second, our approach has resulted in a validated dataset of 100 force-expressing verbs and their contextual uses, providing a valuable resource for future research in this area. Third, through the use of AMR2FRED tool processing and SPARQL querying, we have successfully integrated our findings into ImageSchemaNet, enhancing this ontology with empirically derived FORCE-related concepts. This integration establishes a concrete link between image schema theory and specific linguistic realizations of FORCE, opening new avenues for applications in natural language processing, knowledge representation, and cognitive computing.

Acknowledgment

References

[1] M. Johnson, The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason, The University of Chicago Press, Chicago and London, 1987.

[2] S. Nolfi, On the unexpected abilities of large language models, Adaptive Behavior (2023) 10597123241256754.

[3] M. M. Hedblom, O. Kutz, F. Neuhaus, Choosing the right path: image schema theory as a foundation for concept invention, Journal of Artificial General Intelligence 6 (2015) 21–54.

[4] G. Lakoff, M. Johnson, Metaphors we live by, University of Chicago press, 1980.

[5] G. Lakoff, M. Johnson, et al., Philosophy in the flesh: The embodied mind and its challenge to western thought, volume 640, Basic books New York, 1999.

[6] J. M. Mandler, The foundations of mind: Origins of conceptual thought, Oxford University Press, 2004.

[7] R. W. Langacker, Foundations of cognitive grammar: Theoretical prerequisites, volume 1, Stanford university press, 1987.

[8] R. W. Langacker, Cognitive grammar, Basic Readings 29 (2008).

[9] E. Dodge, G. Lakoff, Image schemas: From linguistic analysis to neural grounding, From perception to meaning: Image schemas in cognitive linguistics (2005) 57–91.

[10] B. Bennett, C. Cialone, Corpus guided sense cluster analysis: a methodology for ontology development (with examples from the spatial domain)., in: FOIS, 2014, pp. 213–226.

[11] A. Cienki, Image schemas and gesture, From perception to meaning: Image schemas in cognitive linguistics 29 (2005) 421–442.

[12] G. Lakoff, The invariance hypothesis: Is abstract reason based on image-schemas? (1990).

[13] R. Núñez, G. Lakoff, The cognitive foundations of mathematics: The role of conceptual metaphor, in: The handbook of mathematical cognition, Psychology Press, 2005, pp. 109–124.

[14] W. Kuhn, An image-schematic account of spatial categories, in: International Conference on Spatial Information Theory, Springer, 2007, pp. 152–168.

[15] M. M. Hedblom, O. Kutz, R. Peñaloza, G. Guizzardi, Image schema combinations and complex events, KI-Künstliche Intelligenz 33 (2019) 279–291.

[16] R. St. Amant, C. T. Morrison, Y.-H. Chang, P. R. Cohen, C. Beal, An image schema language, in: Proc. of the 7th Int. Conf. on Cognitive Modeling (ICCM), 2006, pp. 292–297.

[17] M. M. Hedblom, O. Kutz, T. Mossakowski, F. Neuhaus, Between contact and support: Introducing a logic for image schemas and directed movement, in: Conference of the Italian Association for Artificial Intelligence, Springer, 2017, pp. 256–268.

[18] G. Righetti, D. Porello, N. Troquard, O. Kutz, M. M. Hedblom, P. Galliani, Asymmetric hybrids: Dialogues for computational concept combination, in: Formal Ontology in Information Systems, IOS Press, 2021, pp. 81–96.

[19] G. Righetti, O. Kutz, The moving apple: An image-schematic investigation into the leuven concept database, in: Proceedings of The Seventh Image Schema Day co-located with The 20th International Conference on Principles of Knowledge Representation and Reasoning (KR 2023), Rhodes, Greece, September 2nd, 2023, CEUR-WS, 2023.

[20] S. De Giorgis, A. Gangemi, D. Gromann, Imageschemanet: Formalizing embodied commonsense knowledge providing an image-schematic layer to framester, Semantic Web Journal forthcoming (2022).

[21] M. M. Hedblom, F. Neuhaus, T. Mossakowski, The diagrammatic image schema language (disl), Spatial Cognition & Computation (2024) 1–38.

[22] A. Papafragou, C. Massey, L. Gleitman, When English proposes what Greek presupposes: The cross-linguistic encoding of motion events, Cognition 98 (2006) B75–B87.

[23] J. A. Prieto Velasco, M. Tercedor Sánchez, The embodied nature of medical concepts: image schemas and language for pain., Cognitive processing (2014). doi:10.1007/s10339-013-0594-9.

[24] D. Gromann, M. M. Hedblom, Body-mind-language: Multilingual knowledge extraction based on embodied cognition, in: AIC, 2017, pp. 20–33.

[25] D. Gromann, M. M. Hedblom, Kinesthetic mind reader: A method to identify image schemas in natural language, in: Proceedings of Advancements in Cogntivie Systems, 2017.

[26] L. Wachowiak, D. Gromann, Systematic analysis of image schemas in natural language through explainable multilingual neural language processing, in: Proceedings of the 29th International Conference on Computational Linguistics, 2022, pp. 5571–5581.

[27] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, et al., Retrieval-augmented generation for knowledge-intensive nlp tasks, Advances in Neural Information Processing Systems 33 (2020) 9459–9474.

[28] D. Edge, H. Trinh, N. Cheng, J. Bradley, A. Chao, A. Mody, S. Truitt, J. Larson, From local to global: A graph rag approach to query-focused summarization, arXiv preprint arXiv:2404.16130 (2024).

[29] J. Hurtienne, S. Huber, C. Baur, Supporting user interface design with image schemas: The iscat database as a research tool., in: ISD, 2022.

[30] A. GANGEMI, et al., Amr2fred, a tool for translating abstract meaning representation to motif-based linguistic knowledge graph, in: Proceedings of the Extended Semantic Web Conference (ESWC2017), DEU, 2017, pp. 43–47.

[31] A. Meloni, D. Reforgiato Recupero, A. Gangemi, Amr2fred, a tool for translating abstract meaning representation to motif-based linguistic knowledge graphs, in: The Semantic Web: ESWC 2017 Satellite Events: ESWC 2017 Satellite Events, Portorož, Slovenia, May 28–June 1, 2017, Revised Selected Papers 14, Springer, 2017, pp. 43–47.

[32] S. De Giorgis, Ethics in the flesh: formalizing moral values in embodied cognition (2023).

[33] A. Gangemi, V. Presutti, D. Reforgiato Recupero, A. G. Nuzzolese, F. Draicchio, M. Mongiovì, Semantic web machine reading with fred, Semantic Web 8 (2017) 873–893.

[34] S. Borgo, R. Ferrario, A. Gangemi, N. Guarino, C. Masolo, D. Porello, E. M. Sanfilippo, L. Vieu, Dolce: A descriptive ontology for linguistic and cognitive engineering, Applied ontology 17 (2022) 45–69.

[35] S. Pradhan, J. Bonn, S. Myers, K. Conger, T. O'gorman, J. Gung, K. Wright-Bettner, M. Palmer, Propbank comes of age—larger, smarter, and more diverse, in: Proceedings of the 11th joint conference on lexical and computational semantics, 2022, pp. 278–288.