Drum Up Support: Systematic Analysis of Image-Schematic Conceptual Metaphors

Lennart Wachowiak
King's College London
lennart.wachowiak@gmail.com

Dagmar Gromann
University of Vienna
dagmar.gromann@gmail.com

Chao Xu
Shanxi University
c.xu@pku.edu.cn

Abstract

Conceptual metaphors represent a cognitive mechanism to transfer knowledge structures from one onto another domain. Image-schematic conceptual metaphors (ISCMs) specialize on transferring sensorimotor experiences to abstract domains. Natural language is believed to provide evidence of such metaphors. However, approaches to verify this hypothesis largely rely on top-down methods, gathering examples by way of introspection, or on manual corpus analyses. In order to contribute towards a method that is systematic and can be replicated, we propose to bring together existing processing steps in a pipeline to detect ISCMs, exemplified for the image schema Support in the COVID-19 domain. This pipeline consists of neural metaphor detection, dependency parsing to uncover construction patterns, clustering, and BERT-based frame annotation of dependent constructions to analyze ISCMs.

1 Introduction

Building on the foundation of existing knowledge to structure and explain new experiences is a common, well-known cognitive mechanism that, if depicted as metaphorical projection, can be captured by conceptual metaphors. In the case of image-schematic conceptual metaphors (ISCMs), the structures being transferred are sensorimotor patterns. Natural language is considered a source of evidence for the existence of ISCMs, which has mostly been investigated by a top-down approach of introspectively identifying examples (e.g. Lakoff and Johnson (1999); Kovecses (2010)) or a bottom-up approach of corpus analyses (e.g. Bennett and Cialone (2014)). Automated approaches generally focus on detecting whether a given sequence is metaphoric or not (Leong et al., 2020) rather than identifying the specific type of metaphor, with few exceptions (e.g. Dodge et al. (2015)). However, effective computational tools for metaphor analysis are important as they can play a role in improving machine translation (Mao et al., 2018) and in analyzing the usage and effect of metaphors, e.g. in political discourse (Prabhakaran et al., 2021) or literature (Freeman, 2002). In this paper, we propose a pipeline, depicted in Fig. 1, to automatically detect and identify ISCMs exemplified for the image schema Support in an English COVID-19 corpus. In contrast to introspective methods, the proposed pipeline promises to be replicable, faster, less subjective and capable of uncovering novel, previously unknown metaphors.

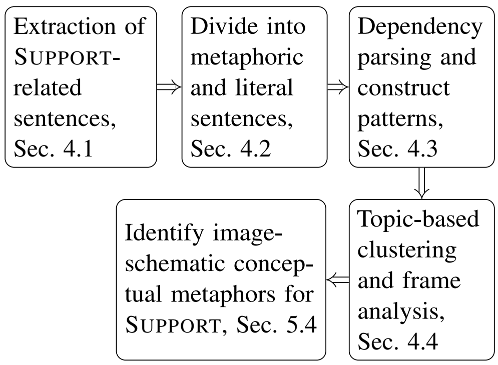

Figure 1: Overview of the proposed ISCMs analysis approach

Image schemas have been proposed by Lakoff (1987) and Johnson (1987) as cognitive building blocks to capture recurring sensorimotor interactions with the physical world. These experiential patterns are said to "reveal features of human thought and language" (Oakley, 2007), since they are mapped onto conceptual structures. ISCMs map these experiential, conceptual structures to the abstract domain. For instance, a person can physically *lean on* a concrete physical entity, e.g. a table, which entails the person pressing their body weight onto an entity that resists the push force. This physical experience can be mapped onto the abstract domain of emotional Support, such as in *He leans on his friends in these trying times.*

Our approach relies on a series of steps to semi-

Proceedings of the 3rd Workshop on Figurative Language Processing (FLP), pages 44 - 53
December 8, 2022 ©2022 Association for Computational Linguistics

automatically identify ISCMs in natural language: (a) detect whether a sequence is metaphoric or literal, (b) determine its constructional pattern, (c) identify its associated topics, and (d) identify its frames, from which we (e) derive underlying metaphoric projections. We extract sentences from the sample of The Coronavirus Corpus[1] based on seed words related to "support". In order to explore all metaphors related to SUPPORT, we decided to chose a specific, abstract domain, i.e., COVID-19 due to its ongoing relevance, abstract nature and importance to the society at large.

With this first approach to "drum up" SUPPORT for image schemas, this paper contributes a systematic method for detecting and identifying ISCMs in domain-specific natural language. To this end, constructional patterns uncover elements in a sentence that interact with metaphoric seed words, which are then frame annotated to provide evidence of the metaphor type the sentence represents. Furthermore, we contribute to the analysis of conceptual metaphors in natural language in general since the pipeline can equally be applied to other types of metaphors, image schemas and domains.

2 Preliminaries

Within the tradition of embodied cognition, physical experiences are said to shape higher-level cognition, including natural language. For instance, we learn as infants that some objects can support our weight, such as a chair, while others cannot, such as a flower. This physical support can then be transferred in *He leans on his friends in these trying times* to emotional assistance depicted by the metaphor ASSISTANCE IS SUPPORT. The proposed approach relies on theories of semantic frames and image schemas, which we briefly introduce.

2.1 Frames Semantics and Frames

Frame semantics (Fillmore, 1982) has been highly influential in cognitive linguistics as it combines linguistic sequences with knowledge structures to describe cognitive phenomena. Words or phrases, so-called lexical units, are associated with frames based on the common scene they evoke or, as described in FrameNet, their common *situation types*. Fillmore explicitly compares frames to other notions, such as experiential gestalts (Lakoff and Johnson, 1980), stating that frames can refer to a coherent schematization of experience. Thus, widely acknowledged frames provide a theoretically well-founded and practically validated basis for detecting ISCMs in natural language sequences. In fact, an initial yet uncompleted account of image schemas on the highest level of FrameNet can be found (Gangemi and Gromann, 2019). The bottleneck in utilizing frames is the low recall and precision of most existing automated tools to identify frames in natural language, addressed in Section 4.

2.2 Image Schemas

Image schemas capture recurring sensorimotor experiences as so-called gestalts (Johnson, 1987), i.e., structure compositions of parts forming a uniform whole. Image schemas can either be static or dynamic (Lakoff and Núñez, 2000), where the former are classified as orientational (e.g. ABOVE), topological (e.g. CONTACT), or force-dynamic (e.g. SUPPORT). Image schemas are simple spatial events built from spatial primitives (Mandler, 1992). The image schema SUPPORT is built from CONTACT between two objects were one depends on the other (Mandler, 1992; Besold et al., 2017). CONTACT is defined as two objects physically touching and only with force dynamics, i.e., application or exertion of force, constitutes SUPPORT.

Herskovits (1987) proposes that an object supports another if its weight presses or pulls upon it, where the supporting object resists the push or pull force. Prototypically, one entity rests on a horizontal upward-facing SURFACE of the other. SUPPORT can also involve other topological properties (Herskovits, 1987): an object can be hanging from, adhering to or being joined by nails, screws or other devices with the supporting entity. Conceptual metaphors are not merely a linguistic phenomenon, but rather a cognitive mechanism that enables the projection of recurring experiences onto abstract domains and structures our subjective experiences (Lakoff and Johnson, 1999). They can be specialized to image-schematic metaphors (Hedblom et al., 2015), which transfer the skeletal structure of image schemas to abstract target domains.

3 Related Work

Metaphor detection is often framed as binary classification task, in which each word of a sentence is either labeled as being used metaphorically or literally. Tong et al. (2021) provide an overview of architectures used for metaphor detec-

[1] https://www.english-corpora.org/corona/

tion, datasets, and other metaphor-related tasks. Another overview (Rai and Chakraverty, 2020) takes many different approaches to computational metaphor processing into account, additionally, reflecting on the different theoretical and linguistic views on the definition of metaphors. In a recent shared task on metaphor detection, fine-tuning pre-trained language models led to the best results (Leong et al., 2020).

There is, moreover, a tradition of analyzing syntactic patterns of metaphoric language (Sullivan, 2013), e.g. verb-prep-noun in which the verb represents the source domain and the noun the target domain. Such patterns build a core assumption of various researchers with the goal of automatically identifying source-to-target domain mappings. For instance, Shutova et al. (2017) explore unsupervised methods for identifying clusters of source and target concepts as well as the connections between them, limiting their approach to verb–object/subject constructions. Dodge et al. (2015) use multiple constructional patterns to find metaphor candidates that are then further analyzed by identifying evoked frames and checking their relations in MetaNet. Rosen (2018) trains a feed-forward neural network to predict one out of 77 source domains given a target domain referent and dependencies from a contextual sentence deemed as relevant. Compared to conceptual metaphors, image schemas have received little attention from computational linguists. Existing approaches to extract image schemas include unsupervised clustering (Gromann and Hedblom, 2017) and classifying sentences with neural language models (Wachowiak and Gromann, 2022). In terms of method and domain, Wicke and Bolognesi (2020) extract sentences from a COVID-19 corpus also based on seed words and apply topic modeling to analyze the frame WAR. A broader range of COVID-19-related metaphors is considered by Semino (2021).

In contrast to previous work, we do not make any assumptions about syntactic patterns or word classes, but compute statistics on syntactic patterns after we identify metaphoric language with a language model.

4 Method

As shown in Fig. 1, in order to identify image-schematic conceptual metaphors, we first compile a list of seed words related to "support", which we use to extract sentences from an English corpus on COVID-19. Each occurrence of a seed word in the corpus is automatically annotated as literal or metaphoric. With dependency parsing the constructional pattern for each sentence with metaphoric seed words are created. These patterns are important to identify the elements directly related to metaphoric seed words, for which we then obtain frame–semantic relations. The overall topic of each sentence is analyzed by way of clustering and frames and topics serve as a basis to identify its conceptual metaphor.

4.1 Extraction of SUPPORT-Related Sentences

As a first step, we compile a list of seed words related to SUPPORT by taking the top 100 words related to "support" from `relatedwords.org`, which bases its results on combined similarity metrics from resources such as ConceptNet and word embeddings. Moreover, we add words related to physical senses of "support" in WordNet synsets, FrameNet frames, and MetaNet frames. Based on these seed words, we extract sentences related to the image schema SUPPORT from the publicly available sample of The Coronavirus Corpus[2] consisting of 3.2 million words.

Seed words that entirely resulted in sentences unrelated to senses of SUPPORT as defined in Section 2.2 were excluded, e.g. "stomach" only related to the body part and not the related verb or "brook" could only be found in named entities, such as *Brook Park*. The resulting list of seed words with its count of sentences is provided in Section 5.1.

4.2 Automatic Metaphor Detection

Given the list of SUPPORT-related sentences, we automatically labeled each word of a sentence as literal or metaphoric. For this sub-task, we trained a metaphor-detection model on the VU Amsterdam Metaphor Corpus (Steen, 2010), which was annotated at word-level according to the metaphor identification protocol presented in the same paper. Based on the success of large pre-trained language models in a recent shared task on metaphor detection using the same corpus (Leong et al., 2020), we used the multilingual pre-trained language model XLM-RoBERTa (Conneau et al., 2020).

We trained the model with a learning rate of 2e-5 for eight epochs and loaded the model with the best validation performance at the end. We used the same train–test split as in the shared task and used

[2] `https://www.corpusdata.org/formats.asp`

randomly allocated 10% of the training data for validation. Code and model are publicly available[3].

4.3 Dependency Parsing for Comparison of Syntactic Structure

For each seed word, we investigated its syntactic function and relation to other words in the sentence by using the part-of-speech tagger and dependency parser from Stanford's neural NLP library Stanza (Qi et al., 2020). We provide statistics on incoming and outgoing relations to and from the seed words in Table 1. We first identify all dependency relations to and from the seed words, and then remove the relations with the following tags: cc, conj, fixed, flat, list, parataxis, orphan, goeswith, reparandum, punct, root, dep, aux, mark, det. They are considered as having no direct relevance for our purposes, for instance, only indicating function or coordination words. Some seed words used as nouns only have compound relations, with most of the syntactic information being stored in the relations of the compound word. Thus, we also extract the incoming and outgoing relations of the words constituting a compound together with the seed word. All elements identified in this step are then annotated with frames to identify the conceptual metaphor.

4.4 Identifying Topics and Conceptual Metaphors

We clustered the extracted sentences, allowing us to group similar sentences semantically. With this procedure, we quickly explore how SUPPORT is used in a literal and metaphorical sense. We created the clusters by using the BERTopic-library (Grootendorst, 2022). BERTopic represents each sentence using BERT-based sentence embeddings (Reimers and Gurevych, 2019). In a second step, it reduces the dimensionality using UMAP (McInnes et al., 2018), before clustering the resulting data points using the density-based hierarchical clustering algorithm HDBSCAN (McInnes et al., 2017).

Each sentence is automatically annotated with semantic frames by utilizing BERT-for-FrameNet (Minnema and Nissim, 2021) in its configuration of only predicting frames and not jointly predicting also semantic roles, relying on BERT layer 12. Frames related to each seed word and its dependent words or compounds are then manually analyzed and compared. While most frame parsers experience relatively low recall and precision, the BERT-for-FrameNet model returned a considerably higher number of frames than previous approaches. Nevertheless, specific seed words were almost never annotated, which could potentially be alleviated by querying other resources, such as Wikidata. However, for this case study, we opted for analyzing the frame-annotated sentences. The code and data for our approach are publicly available[4].

5 Results and Analysis

5.1 Extraction of SUPPORT-Related Sentences

Our final list of SUPPORT-related seed words and their frequencies is:

> advocacy (54), affirm (19), aid (315), assist (242), assistance (331), back (2154), back up (41), backbone (16), backing (18), backup (17), base (906), bear (142), bear out (2), bolster (37), boost (223), brace (37), bracket (15), buttress (2), commitment (191), corroborate (2), defend (102), endorse (41), endorsement (10), establish (250), financial backing (2), financial support (51), foot (271), help (2985), hold (1169), hold up (40), lifeline (22), livelihood (92), maintain (511), maintenance (95), patronage (5), prop (23), prop up (16), reinforcement (4), resource (563), sponsorship (10), stand (508), subscribe (119), substantiate (4), support (2317), supporter (104), supportive (40), sustain (92), sustenance (11), undercarriage (1), underpin (15), unsupported (10), uphold (34).

5.2 Automatic Metaphor Detection

Our metaphor-detection model achieves an accuracy of 95% on the test set. For the label *literal*, it achieves an F1 score of 0.97 with a precision of 0.96 and a recall of 0.98; and an F1 score of 0.76 for the label *metaphoric* with a precision of 0.82 of and a recall of 0.71. Its performance is, thus, comparable with the best-performing model of the 2020 metaphor-detection task (Leong et al., 2020).

The frequency of seed words in each sentence classified as metaphoric or literal is depicted in Fig. 2, which reveals that some seed words are more regularly used in a metaphoric sense than others. While words like "boost", "maintain", and "hold"

[3] https://github.com/lwachowiak/Multilingual-Metaphor-Detection

[4] https://github.com/lwachowiak/ISCMs/

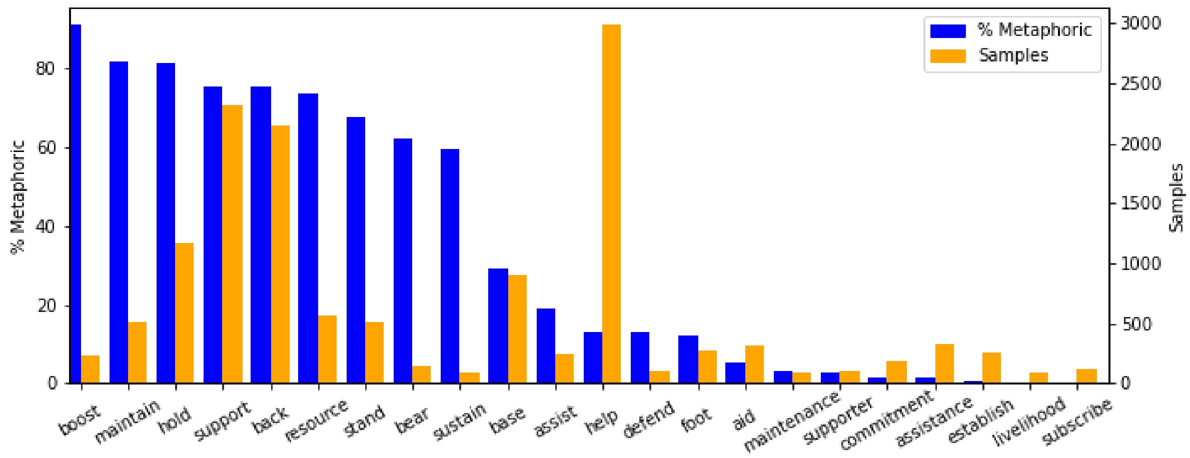

Figure 2: Seed words with over 75 samples ordered by how often they were used metaphorically

are used more than 70% of the time metaphorically, "establish", "livelihood", or "subscribe" are used less than 5% of the time metaphorically. These labels and statistics give us a good indication of which sentences to explore further in order to identify conceptual metaphors based on SUPPORT and which sentences' syntactic structure to investigate.

5.3 Dependency Parsing for Comparison of Syntactic Structure

For each sentence, we compute a constructional pattern centered on the seed word using dependency parsing as described in Section 4.3. For each pattern, we count how many seed words are used metaphorically in that syntactic constellation. Thus, the highest possible count for any pattern is 52 — the number of seed words. Counting all sentences per pattern would dip the statistics towards frequent patterns for a specific seed word that, however, is not necessarily an overall frequent pattern. The resulting most frequent constructional patterns grouped by word class of the seed word and examples are shown in Table 1. Word classes and dependency relations are presented in word order and concatenated by an underscore. If the dependency tag stands after the word class, it is an incoming relation to the seed word, if after the word class, it is an outgoing relation from the seed word, e.g. verb_obj ⟪noun⟫ indicates the relation `obj` going from the verb to the seed noun.

A variety of common patterns was detected for both, verb and noun seed words, where the seed words represent the source domain. The noun seed words appear most frequently as the object of a verb, with only one of the ten most common patterns having the seed word as the subject. Moreover, five of the patterns contain a nominal modifier relation. For verb seed words, the target domain noun frequently occurs as object, frequently co-occurring with a preceding noun or verb. Patterns for adjective and adverb seed words are much rarer, and we did not include those only occurring once.

5.4 Identifying Topics and Conceptual Metaphors

For an exploration of the senses and themes of SUPPORT-related words used in the Coronavirus discourse, we conducted a cluster analysis of different subsets of sentences. To obtain clusters of mostly metaphorical sentences, we clustered all 2,322 samples based on the seed word "support" (76% labeled as metaphoric); to obtain clusters of mostly literal sentences, we clustered all 2,988 samples based on the seed word "help" (13% metaphoric). BERTopic successfully clustered 1,288 sentences with the seed word "support" and 1,442 sentences with the seed word "help", grouping the rest of the sentences in a cluster of outliers. Fig. 3 and 4 show the two resulting cluster hierarchies, with more similar clusters being iteratively grouped together. Each cluster can be identified by the three words representing it best according to their Term Frequency-Inverse Document Frequency (TF-IDF). The TF-IDF value assumes each cluster to be a document and offsets the frequency of a word by the number of clusters containing the same word.

The results show that financial support is one of the most common contexts in which the seed word "support" is being used. A sentence from the cluster 23_billion_package_

Table 1: The most common constructional patterns of metaphorical ⟨⟨seed words⟩⟩. Count indicates how many unique seed words labeled as metaphoric appeared in such a pattern. Abbreviations: prep=preposition, adj=adjective, noun=noun or noun phrase, ppr=personal pronoun, adv=adverb; acl=clausal modifier of noun, advcl=adverbial clause modifier, amod=adjectival modifier, nmod=nominal modifier, nmod:poss=possessive nominal modifier, nsubj=nominal subject, obj=object, obl=oblique nominal, xcmop=open clausal complement

Noun Seed Words		
Dependency Pattern	Language Example (order as in sentence)	Count
verb_obj ⟨⟨noun⟩⟩	give a ⟨⟨lifeline⟩⟩	12
noun_nmod case_prep ⟨⟨noun⟩⟩ nmod_noun	supply on the ⟨⟨back⟩⟩ of demand	11
verb_obj ⟨⟨noun⟩⟩ nmod_noun	form ⟨⟨backbone⟩⟩ (of) speech	10
verb_obj amod_adj ⟨⟨noun⟩⟩	(COVID-19 restrictions) won broad ⟨⟨support⟩⟩	10
verb_obj nmod:poss_ppr ⟨⟨noun⟩⟩	(citizens) strengthen their (politicians) ⟨⟨backbones⟩⟩	9
verb_obl case_prep ⟨⟨noun⟩⟩	put (the industry) on ⟨⟨hold⟩⟩	9
⟨⟨noun⟩⟩ nmod_noun verb_nsubj	⟨⟨backing⟩⟩ (of a) brand becomes (invaluable)	8
verb_obl prep_case amod_adj ⟨⟨noun⟩⟩	(government needs to) get on the “front ⟨⟨foot⟩⟩”	7
⟨⟨noun⟩⟩ nmod_noun	⟨⟨boost⟩⟩ (to) economy	6
verb_obl case_prep ⟨⟨noun⟩⟩ nmod_noun	go (ahead) on ⟨⟨foot⟩⟩ (of) advice	6
Verb Seed Words		
acl_noun ⟨⟨verb⟩⟩ obl_noun	team ⟨⟨standing⟩⟩ (on) the front lines (of the outbreak)	14
verb_xcomp ⟨⟨verb⟩⟩ obj_noun	(war on corruption) continues to ⟨⟨bear⟩⟩ fruits	12
verb_advcl ⟨⟨verb⟩⟩ obj_noun	cut (down on expenses) to ⟨⟨sustain⟩⟩ (these difficult) times	11
nsubj_noun ⟨⟨verb⟩⟩ obj_noun	righteousness ⟨⟨upholds⟩⟩ (the) nation	10
acl_noun ⟨⟨verb⟩⟩ obj_noun	evidence to ⟨⟨back⟩⟩ (this) fear	9
nsubj_noun ⟨⟨verb⟩⟩ obj_noun obl_noun	businesses ⟨⟨bearing⟩⟩ the brunt (for) months	9
verb_ccomp nsubj_noun ⟨⟨verb⟩⟩ obj_noun	ensure everyone ⟨⟨maintains⟩⟩ (stable) housing	9
nsubj_noun ⟨⟨verb⟩⟩ obj_noun	authority ⟨⟨boosts⟩⟩ measures	8
njsub_noun ⟨⟨verb⟩⟩ obj_noun advcl_verb	unit ⟨⟨held⟩⟩ a protest to reiterate (their demands)	8
xcomp_verb ⟨⟨verb⟩⟩ obl_noun	to rebuild (our economy) ⟨⟨based⟩⟩ (on a green energy) future	8
Adjective and Adverb Seed Words		
⟨⟨adj⟩⟩ amod_noun	⟨⟨unsupported⟩⟩ market	3

spending is for example *6bn of new funding to support NHS*. Some clusters revolve around financial, political, and other forms of support for specific groups: artists, football clubs, farmers, businesses, students, children, or journalists. A sample from cluster 20, identified by the keywords “music”, “artists”, and “great”, is simply the phrase *Support for Artists*. Another interesting example from the same cluster shows that support can also go the other way around and music can take the role of the support-giver: *... the songs they ’re turning to right now for support, peace, hope, and inspiration*. Another type of support is life support given in the context of COVID, such as in *To leave the ICU, Dr Monika said Mr Efendi must first be taken off breathing support*. All these examples are covered by the already existing metaphors ASSISTANCE IS SUPPORT and HELP IS SUPPORT in MetaNet. However, the clusters give a more fine-grained overview of what types of assistance and help can be given, as well as who the supporting and the supported entity are.

In comparison to “support”, the seed word “help” is used in more diverse contexts in this corpus, resulting in a larger number of clusters. As before, different groups can be identified as giver and receiver of help, e.g. journalists as in *If you can help us, please click the button to ensure we can continue to provide quality independent journalism you can trust*. “Help” is used in a literal way and does not evoke the physical SUPPORT frame. However, in many sentences “help” could be replaced with “support” without changing the meaning of the sentence other than adding a metaphoric sense.

To more closely investigate the types of metaphors, all elements dependency-related to metaphoric seed words were automatically annotated with frames utilizing BERT-for-FrameNet. These frames provide insights into the potential type of metaphor of SUPPORT-related words and were counted once per seed word. Fig. 5 shows the top 14 most frequent frames. From the metaphorically labelled seed words, only 45% were provided with a frame, where the 1,429 examples of “back” and a surprisingly large 1,345 variations of “support” (including supportive, unsupported, etc.) remained without a frame. Nevertheless, an overall picture of types of frames related to SUPPORT can be obtained as shown in Fig. 5.

Besides the typical frames related to ASSISTANCE IS SUPPORT, Fig. 5 shows the interesting case of BODY PARTS as in *get back on their feet*

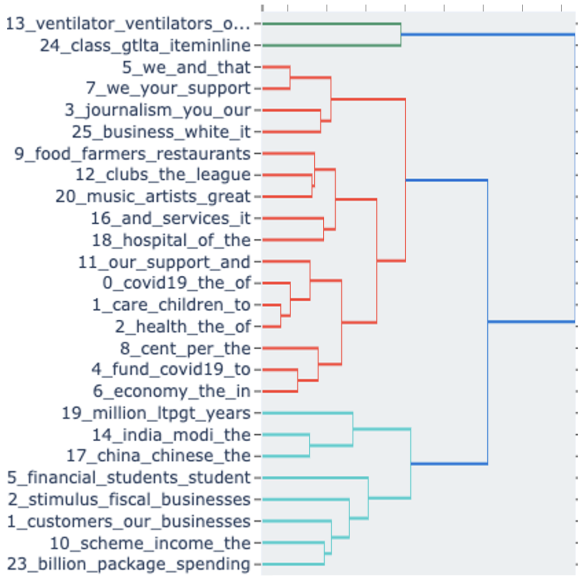

Figure 3: Clusters of sentences containing the word "support'. Clusters have a unique ID, followed by the words representing the cluster based on TF-IDF values.

and *be on the front foot* in the sense of being at an advantage. Adapting MetaNet metaphors, this could be interpreted as RECOVERY IS BODILY SUPPORT since *get back on their feet* means recovery, while *nimble on their feet* indicates endangerment. The orientation here is important since the *front foot* and *best foot forward* represent an advantage and the *back foot* puts one at a disadvantage, which collocates this metaphor with PROGRESS IS FORWARD MOTION. The expression *dragging their feet*, annotated with the frame MANIPULATION, relates it to a lack of support by body parts, i.e., MANIPULATION IS LACK OF BODILY SUPPORT.

The frame TAKING SIDES is mostly related to "support" and "back" as in *backing the campaign* and requires one person to metaphorically push or pull the weight of one side, so TAKING SIDES IS SUPPORT. For SELF MOTION the most frequent contender is "step", where similar to "foot" forward is progress, e.g. *people have stepped forward for this*, and backwards or away is withdrawal of support, e.g. *he backed away from calling for a quarantine*. Thus, a specialization of the PROGRESS IS FORWARD MOTION could be PROGRESS IS SUPPORT BY SELF MOTION, which can be backed by examples of the frames SELF MOTION as well as BODY PARTS, e.g. *put our best foot forward*. The frame COMPLIANCE mostly relates to abide, but provides interesting cases for

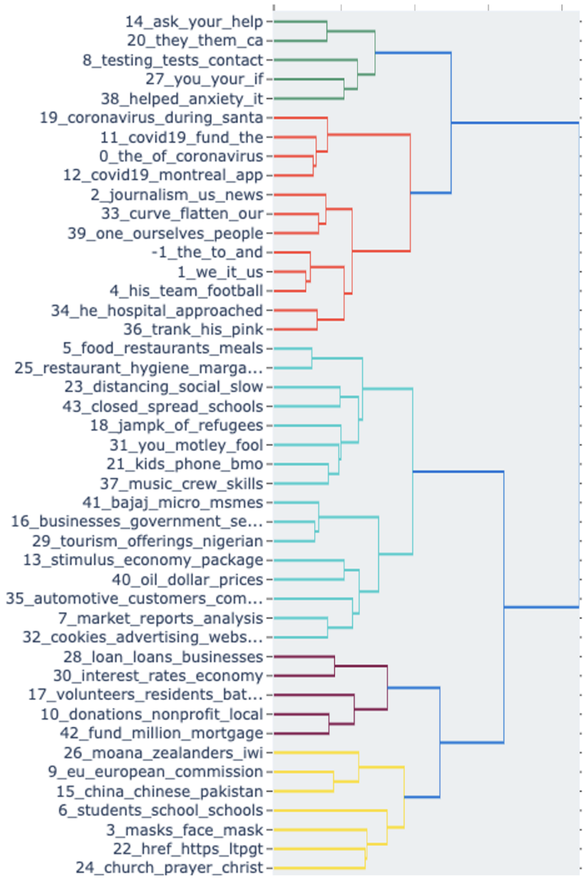

Figure 4: Clusters of sentences containing the seed word "help". Clusters have a unique ID, followed by the words representing the cluster based on TF-IDF values.

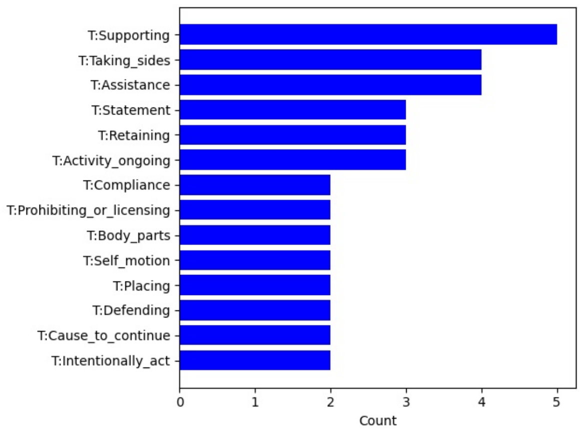

Figure 5: Frequent frames associated with seed words

"upholding" as in *upholding the rule of law* which indicates COMPLIANCE IS SUPPORT, since upholding in its literal sense to keep elevated requires the pulling of weight. One highly frequent frame in terms of occurrence across sentences that, however, only occurs with the two seed words "step" and "boost" is CAUSE CHANGE OF POSITION ON A SCALE. For instance, in the sentence of *He will* *step down as CEO* it is collocated with ACTIVITY STOP leading to ACTIVITY STOP CAUSES CHANGE OF POSITION ON A SCALE. The seed word "bear" is frequently annotated with TOLERATING as in *patients bear the pain*, indicating that TOLERATING IS SUPPORT.

This frame annotation step provides an excellent method for analyzing the (lack of) semantic richness of seed words, e.g. "aid" always relates to ASSISTANCE and "abide" always to COMPLIANCE. In contrast, the seed word "hold" relates to 13 different frames. While not all meanings of all seed words directly relate to ISCMs of SUPPORT, the above examples show that this method can still facilitate their exploration. Nevertheless, with a representative amount of human-curated data, a more rigorous evaluation can be foreseen, also taking other sources of metaphoric and image-schematic information into account. In any case, the final formulation of ISCMs will most likely always benefit from human refinement.

6 Discussion and Conclusion

In this paper, we presented a method to semi-automatically explore image-schematic conceptual metaphors, their related topics and constructional patterns in natural language. A pipeline returns syntactic patterns, thematic clusters, and frames for seed words related to a specific image schema. This approach enables the analysis of how the image schema SUPPORT is used within the context of COVID-19 in a more systematic and comprehensive way than possible with introspective methods. Besides detecting examples of well-known metaphors, it allowed us to uncover new metaphors, e.g. RECOVERY IS BODILY SUPPORT. To this end, building constructional patterns in a bottom-up manner without prior assumptions was important. In terms of topics, a wide variety of supporters and support recipients could be detected.

To apply the same method to other image schemas, a set of related seed words would need to be compiled as input to the method. For instance, a seed word list for the image schema CONTAINMENT could include words such as "inside", "boundary", or "vessel". One drawback of this seed word approach is that polysemy in the sense of multiple literal or even metaphoric meanings of a seed words is not explicitly considered. Nevertheless, given that no repositories of ISCMs exist and repositories on conceptual metaphors, such as MetaNet, contain a limited number of natural language examples or ISCMs, this semi-automated approach is an important step forward to drum up support of ISCMs.

As a knowledge extraction approach rooted in cognitive science, a natural next step would be to explore the taxonomic structures of frames provided by MetaNet, FrameNet or similar resources to query interdependencies between and relations among ISCMs. Furthermore, existing semantic resources, such as DBpedia and Wikidata, should be utilized to increase the number of annotated frames.

Currently, this approach heavily relies on recent advances of methods in computational linguistics brought together in a pipeline. Errors of one step are then propagated to the next. From the set of analysis steps, the metaphor detection performed best. The part-of-speech tags assigned in the process of dependency parsing are highly problematic for specific seed words, such as "back" that is frequently mistagged as noun or adverb when used as verb, negatively affecting the obtained dependency relations and constructional patterns. Another shortcoming is that the clustering method groups many samples into a cluster of outliers. The number of identified outliers, however, is so large that valuable information is inevitably lost, and only a subpart of the semantic topics is represented in the results. For the frame parsing, the return of frames per sentence was considerably higher than with similar approaches (as also reported in (Minnema and Nissim, 2021)), however, the number of frames for seed words was less than half of the overall count of sentences. Very short, heading-like sequences that lack context were generally not frame-annotated at all, e.g. *standing in line for essentials*. This reinforces the need to supplement frame parsing with other processes and resources.

To improve the pipeline and reduce the amount of manual labor required, it would be beneficial to be able to automatically label the target domains for which a specific image schema is used — a step that currently is mostly done manually with the resulting frames and clusters, due to the lack of frame coverage. In order to automatically identify the target domain, we plan to train a sequence-to-sequence model, e.g. T5 (Raffel et al., 2019), to predict the target domain given a source domain and a contextualizing sentence. For instance, the sample SUPPORT*: He leans on his friends in these trying times* should be labeled as ASSISTANCE.

References

B. Bennett and C. Cialone. 2014. Corpus guided sense cluster analysis: a methodology for ontology development (with examples from the spatial domain). In *The Proceedings of 8th International Conference on Formal Ontology in Information Systems (FOIS)*, pages 213–226. IOS Press.

Tarek R Besold, Maria M Hedblom, and Oliver Kutz. 2017. A narrative in three acts: Using combinations of image schemas to model events. *Biologically Inspired Cognitive Architectures*, 19:10–20.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451. Association for Computational Linguistics.

Ellen Dodge, Jisup Hong, and Elise Stickles. 2015. MetaNet: Deep semantic automatic metaphor analysis. In *Proceedings of the Third Workshop on Metaphor in NLP*, pages 40–49, Denver, Colorado. Association for Computational Linguistics.

Charles J Fillmore. 1982. Frame semantics. In *Linguistics in the Morning Calm*, pages 111–138. Seoul: Hanshin.

Margaret H Freeman. 2002. Momentary stays, exploding forces: A cognitive linguistic approach to the poetics of emily dickinson and robert frost. *Journal of English Linguistics*, 30(1):73–90.

Aldo Gangemi and Dagmar Gromann. 2019. Analyzing the imagistic foundation of framality via prepositions. In *Proceedings of the Joint Ontology Workshops (JOWO)*.

Dagmar Gromann and Maria M Hedblom. 2017. Kinesthetic mind reader: A method to identify image schemas in natural language. *Advances in Cognitive Systems*, 5:14.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Maria M Hedblom, Oliver Kutz, and Fabian Neuhaus. 2015. Choosing the right path: Image schema theory as a foundation for concept invention. *Journal of Artificial General Intelligence*, 6(1):21–54.

Annette Herskovits. 1987. *Language and Spatial Cognition*. Cambridge University Press.

Mark Johnson. 1987. *The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Reasoning*. University of Chicago Press.

Zoltan Kovecses. 2010. *Metaphor: A Practical Introduction*. Oxford University Press.

George Lakoff. 1987. *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. The University of Chicago Press.

George Lakoff and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press.

George Lakoff and Mark Johnson. 1999. *Philosophy in the Flesh: the Embodied Mind & its Challenge to Western Thought*. Basic Books.

George Lakoff and Rafael Núñez. 2000. *Where Mathematics Come From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books.

Chee Wee (Ben) Leong, Beata Beigman Klebanov, Chris Hamill, Egon Stemle, Rutuja Ubale, and Xianyang Chen. 2020. A report on the 2020 VUA and TOEFL metaphor detection shared task. In *Proceedings of the Second Workshop on Figurative Language Processing*, pages 18–29. Association for Computational Linguistics.

Jean M. Mandler. 1992. How to build a baby: II. conceptual primitives. *Psychological Review*, 99(4):587–604.

Rui Mao, Chenghua Lin, and Frank Guerin. 2018. Word embedding and wordnet based metaphor identification and interpretation. In *Proceedings of the 56th annual meeting of the association for computational linguistics*. Association for Computational Linguistics (ACL).

L. McInnes, J. Healy, and J. Melville. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv preprint arXiv:1802.03426*.

Leland McInnes, John Healy, and Steve Astels. 2017. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11):205.

Gosse Minnema and Malvina Nissim. 2021. Breeding fillmore's chickens and hatching the eggs: Recombining frames and roles in frame-semantic parsing. In *Proceedings of the 14th International Conference on Computational Semantics (IWCS)*, pages 155–165.

Todd Oakley. 2007. Image schemas. In *The Oxford Handbook of Cognitive Linguistics*, pages 214–235. Oxford: Oxford University Press.

Vinodkumar Prabhakaran, Marek Rei, and Ekaterina Shutova. 2021. How metaphors impact political discourse: A large-scale topic-agnostic study using neural metaphor detection. *arXiv preprint arXiv:2104.03928*.

Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2019. Exploring the limits of transfer learning with a unified text-to-text transformer. *arXiv preprint arXiv:1910.10683*.

Sunny Rai and Shampa Chakraverty. 2020. A survey on computational metaphor processing. *ACM Computing Surveys (CSUR)*, 53(2):1–37.

Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Zachary Rosen. 2018. Computationally constructed concepts: A machine learning approach to metaphor interpretation using usage-based construction grammatical cues. In *Proceedings of the Workshop on Figurative Language Processing*, pages 102–109.

Elena Semino. 2021. "Not soldiers but fire-fighters"–metaphors and covid-19. *Health Communication*, 36(1):50–58.

Ekaterina Shutova, Lin Sun, Elkin Darío Gutiérrez, Patricia Lichtenstein, and Srini Narayanan. 2017. Multilingual metaphor processing: Experiments with semi-supervised and unsupervised learning. *Computational Linguistics*, 43(1):71–123.

Gerard Steen. 2010. *A Method for Linguistic Metaphor Identification: From MIP to MIPVU*. John Benjamins Publishing.

Karen Sullivan. 2013. *Frames and Constructions in Metaphoric Language*. John Benjamins Publishing.

Xiaoyu Tong, Ekaterina Shutova, and Martha Lewis. 2021. Recent advances in neural metaphor processing: A linguistic, cognitive and social perspective. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4673–4686. Association for Computational Linguistics.

Lennart Wachowiak and Dagmar Gromann. 2022. Systematic analysis of image schemas in natural language through explainable multilingual neural language processing. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 5571–5581.

Philipp Wicke and Marianna M Bolognesi. 2020. Framing covid-19: How we conceptualize and discuss the pandemic on twitter. *PLoS ONE*, 15(9):e0240010.