35

Computational Approaches to Metaphor: The Case of MetaNet

Oana A. David

35.1 Introduction

Contrary to many speaker's intuitions about their own language, metaphor is the norm rather the exception in most spoken and written discourse, from the everyday to the poetic. Signed language and gestural communication are also not immune to figurative expression, as has been found repeatedly in the flourishing field of gesture research (Cienki and Müller 2008b). And finally, due to its basis in all forms of cognition, conceptual metaphor is pervasive not only in spoken and written media, but also in visual forms of expression, such as advertisements, works of art, and in recent years, internet memes.

Within cognitive linguistics, cognitive science, and communication studies, there is now a well-established acceptance that metaphor pervades all human thought and communication, and this acceptance is backed by ample empirical data gathered over several decades of interdisciplinary research (see Lakoff 2012 for an overview). Yet, precise formalization of metaphor, known to be a conceptual rather than solely linguistic phenomenon, is not yet widely shared across these fields, nor is an understanding of linguistic and extra-linguistic manifestations of metaphoric thought. Further, in light of recent research on the inseparable connection between grammar and metaphor (Sullivan 2007, Steen 2007, Sullivan 2013a), it has become clear that important generalizations can be stated about the metaphor interface with language that can prove useful not only in theoretical developments in the field, but also in designing Natural Language Processing implementations that are more responsive to this rich domain of human expression.

The current chapter makes a case at the same time for a structured metaphor taxonomy that is integrated with findings from other areas of studies in language and cognition, and for a computational implementation of such a taxonomy toward the goal of making the identification and analysis of metaphoric language more efficient and accurate.

Since the sound theoretical backbone of Conceptual Metaphor Theory was established by Lakoff and Johnson (1980a, 1999), Reddy (1979), Kövecses (1986), Dancygier and Sweetser (2014), and many others over several decades, we see in recent years an empirical bent in metaphor research that has brought several waves of much-needed innovation to the field. First, there is a move away from top-down analyses that rely on examples derived from introspection, whereby metaphor researchers intuit that a given metaphor should exist based on certain salient examples. Instead, we are now seeing bottom-up corpus-driven approaches to metaphor analysis (e.g. Deignan 2005, Stefanowitsch 2006b, Martin 2006), guided by the principle that metaphor discovery can only happen over large and diverse linguistic data sets, and that the theory should in turn be enhanced with newly found instances of usage.

This corpus-driven line of research has led to advances in computational methods for automated and semi-automated metaphor detection. For instance, Mason (2004) uses the selectional preference of verbs as inferred from WordNet (Fellbaum 1998) to predict metaphor via a learning process that rules out normally selected-for arguments (for instance *pour water* would be ruled out as metaphoric, so, by a process of elimination, *pour funds* would be metaphoric). There is also much important work in the domain of metaphor in discourse. Toward that end, the Pragglejaz Group put forth a metaphor identification procedure that analyzes metaphors in context, using localized meanings of lexical items in running text to identify potential metaphors arising from particular collocations (Pragglejaz Group 2007). Finally, seed-based cluster models and statistical metaphor identification methods have also produced results over unrestricted texts (Shutova, Sun, and Korhonen 2010, Shutova, Teufel, and Korhonen 2012) using algorithms and statistical methods that operate over large clusters of words that are marked in accordance with their likelihood of evoking metaphoric source and target domains (for a review of this literature, see Shutova 2010).

To complement these computational mechanisms for metaphor detection and analysis, in this chapter I introduce the MetaNet Metaphor Repository and Identification System developed by a team of linguists, computer scientists, and cognitive scientists as part of a multicampus collaboration starting in 2011, with much of the computational development occurring at the International Computer Science Institute in Berkeley, California. Relative to existing methods in metaphor computation, the system developed for MetaNet is unique in several respects. First, it combines a top-down theory-faithful CMT approach with a bottom-up data-driven approach, implementing these simultaneously as a central, defining feature of the system. This merger of methods rests on the project's adoption of the classic definition of conceptual metaphor; namely, metaphor is a deeply engrained conceptual phenomenon, whereby target domains of many kinds are *systematically* reasoned about in terms of source domains that have origins in concepts formed via embodied experience (see Sullivan this volume Ch. 24). Due to the embodied nature of the formation of source domains during early development, primary metaphors are believed to be universal and language-independent (Grady 1997a). This definition of metaphor, with the strong emphasis on the embodied nature of source domains and the unidirectional mapping from source to target domains, was put forth in Lakoff and Johnson (1980a, 1999) and was expanded on in subsequent Conceptual Metaphor Theory research.

The MetaNet automated metaphor identification system is meant to be reflective, to a certain extent, of how human minds process metaphoric language. That is, when reasoning about relatively more abstract notions, human beings use linguistic and extra-linguistic cues to evoke deep semantic structures that are grounded in more accessible embodied experience with the world, such as force-dynamic interactions with objects and other entities, with motion through space, with one's own bodily experiences and sensations, and with vertical and horizontal orientation.

Human beings are not believed to learn metaphors arbitrarily, nor piecemeal, but instead to acquire them by domain co-associations in primary experiences, which form the foundational primary metaphors upon which later, more complex abstract mappings are layered. This includes understandings of ideologies, economics, politics, complex emotions, interpersonal relationships, and sociocultural norms. An example is the expression *tax relief*, which construes the relatively more abstract domain of taxation, a financial and social activity, in terms of a concrete notion of physical relief from burdens or suffering.

The above definition of metaphor restricts the empirical domain of metaphor research. Given the design of the metaphor identification system, this also restricts the metaphoric phenomena identifiable to only those that appear in an overt linguistic format that provide at least one source-domain-evoking and at least one target-domain-evoking element. An expression such as *glass ceiling*, for instance, although highly metaphoric, would not be detected by such a system for lack of a target domain linguistic element (the target domain here remains implicit and contextually recovered). This definition of metaphor excludes other types of figurative language, such as metonymy, analogy, and conceptual blends, as well as other metaphor phenomena not grounded in primary metaphors, such as image metaphors (e.g. *hourglass figure*), which do not provide this type of asymmetric, systematic conceptual mapping grounded in primary embodied experiences (Lakoff and Turner 1989).

This theoretically driven distinction sets the MetaNet system apart from other metaphor computing systems, which tend to take a more all-encompassing view of metaphor as any type of figurative language. Although more narrow in empirical scope, the value of the MetaNet system lies precisely in its ability to link linguistic metaphors (LMs)[1] with deeper semantic domains and mappings; consequently, an LM such as *tax relief* would be found to have something in common with an LM such as *tax burden*, via their higher-level association with the primary metaphor **DIFFICULTIES ARE BURDENS** (in this case, taxation is construed as a type of financial difficulty).

The MetaNet system is designed with a set of practical applications in mind in order to serve as a useful tool for metaphor analysts. Namely, a linguist interested in the metaphoric content of a particular text can use this tool in a number of tasks – to reinforce or challenge his own intuitions about what the metaphors are in a particular domain, to discover metaphors over prohibitively large corpora, where hand-annotation would be time-consuming, and to reveal potential patterns and connections among conceptual metaphors that would otherwise take a human analyst much longer to uncover. Further, the tool helps analysts perform these tasks in a way that maintains consistency in their own work, as well as encourages consistency across analyses done by multiple analysts throughout the discipline. Finally, with each application of the system, the annotated LM database is augmented, storing data representing both the breadth and the frequency of metaphors, while furnishing metadata about the corpus, genre, register, and other potentially useful information that can help shed light on patterns in actual metaphor usage. This collected data can then be used to give feedback into the metaphor repository, in turn enhancing and diversifying existing metaphoric networks, filling in any gaps that may exist in inter-metaphor relations.

Two crucial components of the system are a well-designed semantic frame and metaphor ontology, and a connection of metaphoric language with grammatical constructions. The former feature of the design is in large part modeled on the configuration inherent in lexicographic semantic frame databases such as FrameNet (Ruppenhofer et al. 2016, Petruck 2013), which have enabled semantic frame role labeling and frame-based sentence annotation techniques to be applied automatically to large data sets. These tools are increasingly prevalent in the fields of both cognitive linguistics and NLP, being extensively applied to English as well as other languages. In addition to the popular English FrameNet resource, other examples include the FrameNet databases designed for Swedish (Borin et al. 2009), German (Boas 2009), Brazilian Portuguese (Salomão et al. 2013), and Japanese (Ohara et al. 2004). MetaNet uses much of the data structure in FrameNet, such as having numerous distinct yet interconnected frames that interact with grammatical patterns in a constructional database.

35.2 Human Cognition Meets Metaphor Computation

To observe the utility of a tool like MetaNet in action, and to demonstrate the dual implementation of a semantic and a grammatical component to the metaphor identification system, let us imagine we are faced with a common real-world text. For instance, consider the following excerpt from President Barak Obama's final State of the Union Address (January 12, 2016):[2]

> All these trends have squeezed workers, even when they have jobs; even when the economy is growing. It's made it harder for a hardworking family to pull itself out of poverty, harder for young people to start on their careers, and tougher for workers to retire when they want to.

A human reader of this paragraph effortlessly parses the linguistic metaphors, interpreting them against the background of the existing cultural models she is accustomed to and in large part takes for granted. For instance, the reader can parse *squeezed workers* metaphorically, realizing that no humans are actually, physically squeezed. Instead, there is an inference, based on the metaphor **FREEDOM OF ACTION IS FREEDOM OF**

MOTION, that worker's ability to perform a job is institutionally and systematically encumbered by limiting policies, lack of job opportunities, etc. There is a further inference that the 'squeezer' (i.e. via the metaphor INSTITUTIONS ARE PEOPLE) is an intentionally acting agent, purposely encumbering workers' actions, rather than this being an involuntary encumbrance. The fact that the verb is 'squeeze' also indicates that these limitations are also harmful to health and to life via SOCIAL HARM IS PHYSICAL HARM. The human reader proceeds equally unconsciously to parse the many other metaphoric expressions in the text, such as *pull itself out of poverty*, which evokes the metaphors NEGATIVE STATES ARE LOW LOCATIONS and IMPROVING A STATE IS MOVING UPWARD. From this we infer that the agency of making this situation-improving state-change lies with the working families themselves, rather than with the governing institutions.

Tasked with the dissection of a short paragraph such as this, a trained metaphor analyst can thus produce a synthesis of the main primary metaphors, metaphoric entailments, and necessary cultural models needed to understand how metaphor is structuring the political message expressed therein. However, what if the paragraph is only one small piece of a much larger text the analyst is interested in? What if the analyst would like to compare this State of the Union Address with previous ones by the same president, or better yet, has a corpus of all recorded State of the Union addresses and would like to do a comparative study of metaphoric discourse across presidents with respect to how not only poverty but also war, health care, and gun control are discussed over many years? For these types of data-intensive queries, an automated metaphor identification system can prove a boost to efficacy.

At this juncture, several statements must be made about how all of this metaphoric processing happens in the human mind. First, the metaphors are not in the words themselves; rather, the words evoke a more complex network of conceptual structures, which can manifest linguistically in a variety of ways. Above, *squeeze workers* could very well have been expressed in a syntactic variant, such as *led to the squeezing of workers* or as a lexicosemantic variant like *wrung workers* or *strangled workers*. Regardless of these syntactic and semantic possibilities, the same metaphoric mappings are operating unconsciously in all of these variants on the expression of the metaphor. The mind knows to parse these expressions not on an individual one-by-one basis, but against the backdrop of shared inferences.

Second, there are elements of this metaphoric network that are either conceptually universal, or high-frequency, or both, and can in fact be predicted to appear cross-linguistically. For instance, we are not surprised that poverty (as a type of state) is construed as an obstacle to motion (as *pull out of poverty* suggests[3]) and also as a low location (such as an often-found variants *raise out of poverty* and *climb out of poverty*), and we would never expect to find an expression in any language in which poverty is construed as a high location. Cross-linguistically, it has been observed that negative states, statuses, conditions, etc. are construed low on a vertical scale, and positive ones as high (Kövecses 2005). Similar cross-linguistic similarities have been observed for any number of so-called primary metaphors, of which NEGATIVE STATES ARE LOW LOCATIONS is just one. These predictions about high-level primary metaphors can be leveraged to train an automated metaphor detection system to assign potential metaphoric expressions to correct candidate conceptual metaphors underlying those expressions.

The fact that metaphor does not reside in the individual words or word collocations is good news for computational modeling and corpus identification of metaphoric expressions, because it saves us from specifying lengthy, dictionary-like lexical ontologies in which each word or collocation corresponds to one metaphor, or vice versa. A system that requires an input of every possible target domain-source domain word combination in order to identify these collocations in texts would not be economical. Instead, we only want to design a relatively small conceptual network in which the semantic frames that feed the source and target domains of metaphors are enriched with specific lexical information, and only as the latter emerges from the evidence in the data. Subsequently, when those lexical items are found in running text, they are used as pointers to some neighborhood in an existing, fully mapped semantic space. Figure 35.1 schematically illustrates the relationship among lexical items, such as *tax* and *burden*, semantic frames, such as Taxation and Physical burden, and metaphors, such as **DIFFICULTIES ARE BURDENS**, and the more specific subcase, **TAXATION IS A BURDEN.**

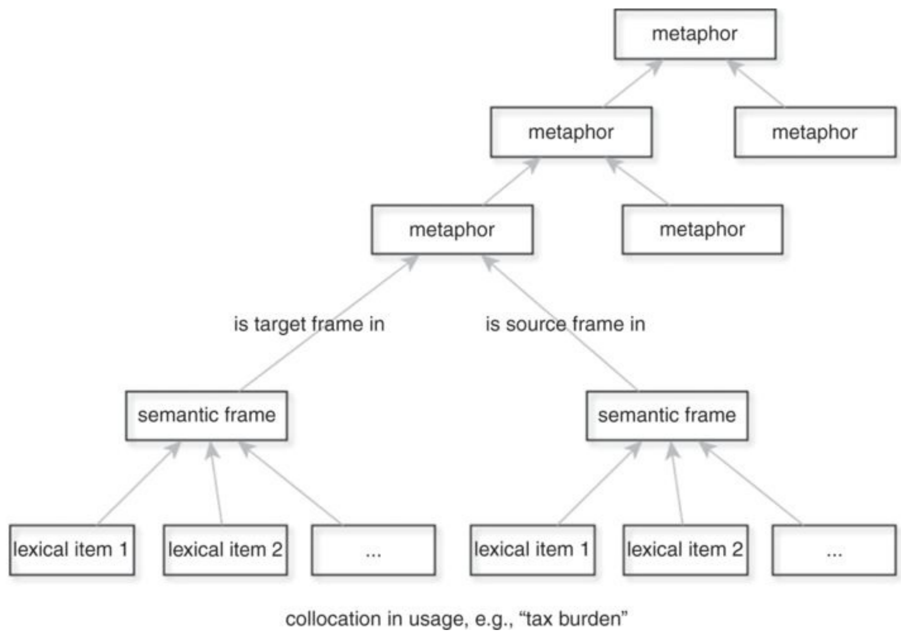

Figure 35.1 Schematic conceptual network

This is a simplified depiction of how a computational metaphor processing systems can approximate human metaphor inferencing patterns. Upon hearing a metaphoric collocation, in the mind of a human interpreter an expression such as *tax burden* would set off a cascade of inferences, leading the interpreter to not only infer the immediate metaphor that **TAXATION IS A BURDEN**, but also, by virtue of the metaphor network that metaphor is embedded in, realize connections to other metaphors. Namely, **TAXATION IS A BURDEN**, as a subcase of **DIFFICULTIES ARE BURDENS**, is not unrelated to higher-level metaphors such as **ACTION IS MOTION**, and **IMPEDIMENTS TO ACTION ARE IMPEDIMENTS TO**

MOTION ALONG A PATH, and the inferences that EASE OF ACTION IS RELIEF FROM BURDENS. A network-based modeling of metaphor relations, such as that implemented in the MetaNet system, helps computational automated metaphor identification to better reflect how humans organize these concepts at a higher level, and how they deduce both similarity and difference among seemingly idiosyncratic forms of linguistic expressions.

In the following sections of this chapter, I detail the basic conceptually grounded components needed for designing a computational system capable of automatically identifying potential metaphoric language in any type of text, with examples of how such components follow from hypothesized human cognition patterns. Further, the system not only identifies the metaphoric language, but also provides candidate conceptual metaphors evoked by the linguistic expressions, and, by virtue of this suggestion, classifies the linguistic expression in a class of similar, previously detected linguistic expressions. This type of system has the potential for discovering patterns within and across languages.

35.3 The MetaNet Architecture

The computational system proposed is designed to perform two functions, and is thus divided into two discrete yet interdependent parts. First, the metaphor repository is organized as a structured set of ontologies, including the source and target domain frames, their frame elements, frame-to-frame and metaphor-to-metaphor relations, and the lexical units evoking particular frames and metaphors. Frames, lexical items, and metaphors represent primitives in the repository, while frame and metaphor families are built up from these primitives. An important notion is that of *cascade* – a package of pre-defined hierarchical and 'makes use of' relations among frames or metaphors (Stickles et al. 2016), specified only once at a very high, schematic level (David, Lakoff, and Stickles 2016). Subsequently, any novel or even creative linguistic occurrence of a conceptual metaphor does not require analysis from scratch, but instead makes use of frequently evoked, entrenched structures. The implementation of cascades relies crucially on well-defined and consistent frame-to-frame and metaphor-to-metaphor relations. Much of this hierarchical organization is conceptually similar to semantic hierarchies and inheritance structures proposed for other network theory-based cognitive linguistic approaches, such as Word Grammar (Hudson 2007, 2010), and is also implemented to an extent in frame-to-frame relations in FrameNet (Ruppenhofer et al. 2016, Petruck 2013). All of these entities and their interrelations are stored in the repository, alongside additional data such as illustrative example sentences and usage information.

The MetaNet repository is configured to house most of the metaphor networks encapsulating at least two hundred primary metaphors and metaphoric entailments, but also additional general and specific metaphors, resulting in a total of over eight hundred. This includes metaphors for time, events, actions, the mind, the self, morality, social interaction, emotion, and perceptual and sensorimotor representations (e.g. SIMILARITY IS CLOSENESS). It also currently includes several specialized domains of metaphoric analysis pertaining to non-primary but highly prevalent general and complex metaphors. Most notably, the repository includes those general metaphors pertaining to common social, political, and economic abstract concepts such as democracy, poverty, governance, taxation, and myriad culturally defined social issues. The repository is embedded in a larger computing pipeline, which includes several components external to but interacting with the custom-made metaphor identification system, as shown in [Figure 35.2](#).

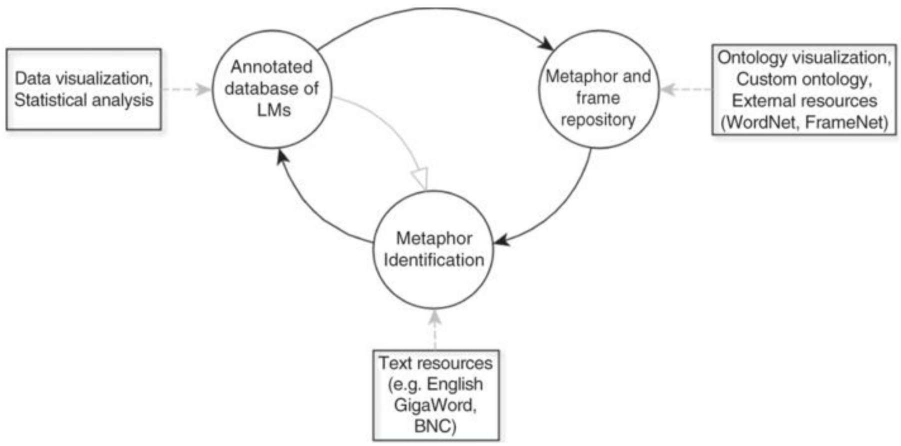

Figure 35.2 MetaNet integration of three systems – repository, identification, and annotated data

Raw texts are first processed using standard parsers, and the repository is used in the detection of metaphoric collocations. Then, any metaphoric collocations that are found are submitted to a growing database of LMs, where the data supporting existing metaphors are stored. If, for instance, the system is constrained by source domain to identify not only taxation as a burden but any other type of economic practice construed as a burden, the system will also identify sentences containing phrases such as *heavy fees, rent burden*, and *payment relief*. These would be identified by virtue of the network structure that the frames in which the LUs *fees, rent*, and *payment* share with the frame in which *tax* is located. This system has already been implemented and tested in discovering metaphors for poverty (Dodge, Hong, and Stickles 2015, Dodge 2016), gun control discourse (David, Lakoff, and Stickles 2016), and drug abuse (Stickles, David, and Sweetser 2014).

35.3.1 Component 1: The Repository

The repository is a body of interconnected ontologies defining the taxonomic entities needed in conceptual metaphor analysis. These ontologies make it possible for the metaphor identification mechanism to use grammatical construction-matching patterns to identify potential LMs in unstructured texts. Figure 35.3 schematically illustrates the primitives used toward this end in the repository, showing the representation of frames and metaphors. Frames contain frame elements, and metaphors contain mappings from source domain frame elements to target domain frame elements.

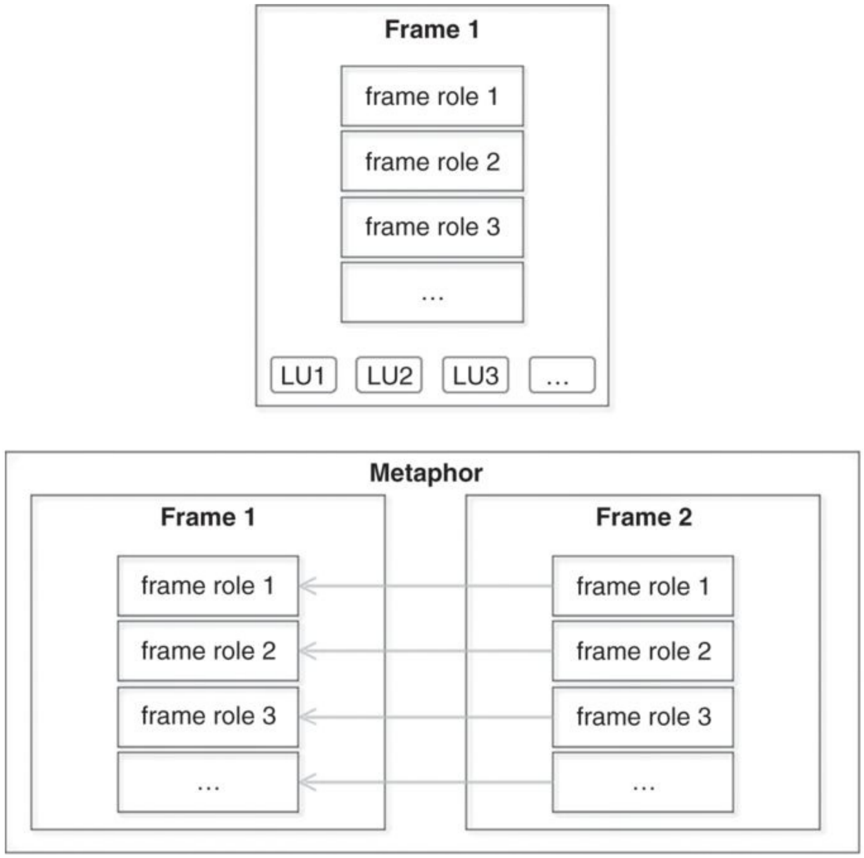

Figure 35.3 Schematic representation of frames and metaphors

The representation in Figure 35.3 illustrates the important association between lexical units (LUs) and frames, and the subsequent association between frames and metaphors. LUs are assigned to frames in much the same way as they are in FrameNet (Ruppenhofer et al. 2010), namely by likelihood that the LU will evoke the frame in any given sentential context. LUs are associated with frames in a one-to-many association, regardless of any broader polysemy structure that word may participate in. For instance, the word *break* may evoke any number of frames, such as Separation (*break off a piece of chocolate*) and Nonfunctionality (*the laptop is broken*), and is thus listed under both frames. Information about the LU-to-frame associations used in MetaNet is gathered both from existing FrameNet frames, and from custom-made frames that are explicitly designed within the context of this system for the purposes of metaphor analysis. Frames are included in the repository often with a metaphor in mind. For instance, faced with an expression such as *our relationship is broken*, the Nonfunctionality source domain frame (stemming from **RELATIONSHIPS ARE FUNCTIONAL OBJECTS**) is more pertinent than would be a Separation frame, whereas the latter frame might be better suited as the source domain frame for an expression such as *we must break ties with tradition*. Metaphors do not receive direct LU assignment, but are supplied with information about the frame elements that are mapped from the source domain frame to the target domain frame. As shown in Figure 35.4, frames are entities that have relations to one another, and this relationship is hierarchical not only in terms of how frames relate to each other, but also in terms of how the frame elements relate to each other. For instance, we can see Bodily harm as a type of Harm, which in turn is a type of

Causation, where the more schematic causer, affectee, and causal process have local, specific instantiations in the more specific frames.

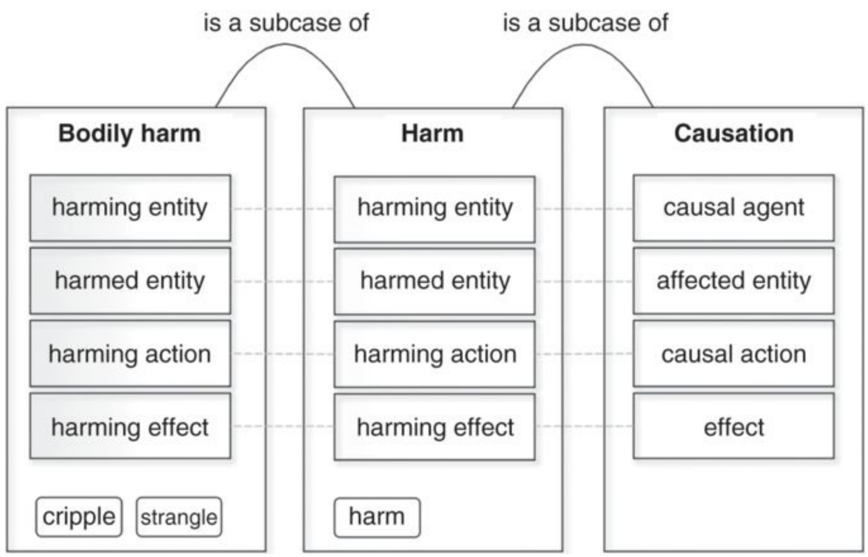

Figure 35.4 Relations among frames

Another important effect of this type of semantic layering is that lexical items are categorized by degrees of semantic specificity. In the example in Figure 35.4, words like *cripple* and *strangle* specifically evoke bodily harm, while verbs like *harm* do not refer exclusively to bodily harm. More general LUs can evoke more specific frames, but the reverse is not true.

At the same time, metaphors (Figure 35.5) are also entities structured relationally to each other, as well as to the frames that populate their source and target domains. Subcase relations among frames that feed the source and target domains of the metaphor are reflected as subcase relations among metaphors. As an example, Figure 35.5 can be posited as the metaphoric cascade for an expression such as *crippling poverty*.

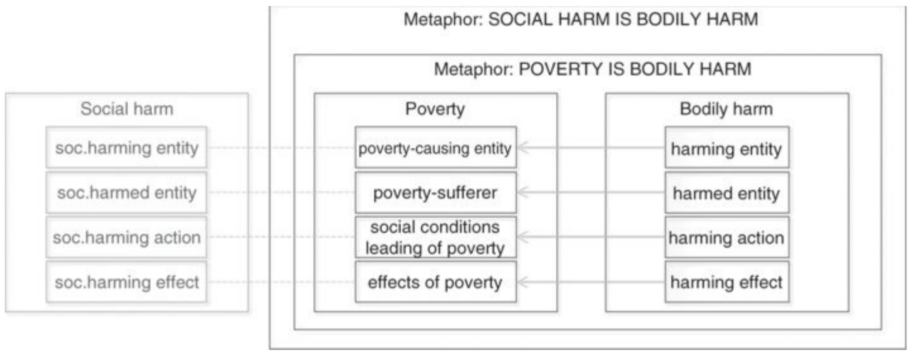

Figure 35.5 Relations among metaphors, and relationship of frames to metaphors

The latter captures two generalizations about the ontology. First, frames that are the source and target domains of metaphors are themselves simultaneously embedded in their own relational networks with other frames. The frame Poverty, for instance, is simultaneously the target domain of the metaphor **POVERTY IS BODILY HARM** as it is a subcase of the Social harm frame. In parallel, the more specific metaphor **POVERTY IS BODILY HARM** is a subcase of a more general **SOCIAL HARM IS BODILY HARM** (as represented by the nested boxes), which itself is potentially embedded in even more general metaphors.

The subcase relationships among metaphors are in place when there are equivalent subcase relationships among frames, but not all subcase relationships among frames necessarily receive an equivalent metaphor entry.

35.3.2 Component 2: The Metaphor Extractor

The metaphor identification mechanism is separate from the repository, yet crucially incorporates it as part of its functioning. While the repository is a place to store and organize information about conceptual, grammatical, and lexical structures involved in metaphor decoding, the metaphor identification mechanism uses custom scripts to implement information from the repository in seeking new instances of metaphors in the wild. Its purpose is to detect automatically potential candidates for linguistic metaphors (LMs) from large bodies of texts, such as the English Gigaword Corpus (Graff and Cieri 2003). It comprises several interdependent components, whose sequence is diagrammed in Figure 35.6.

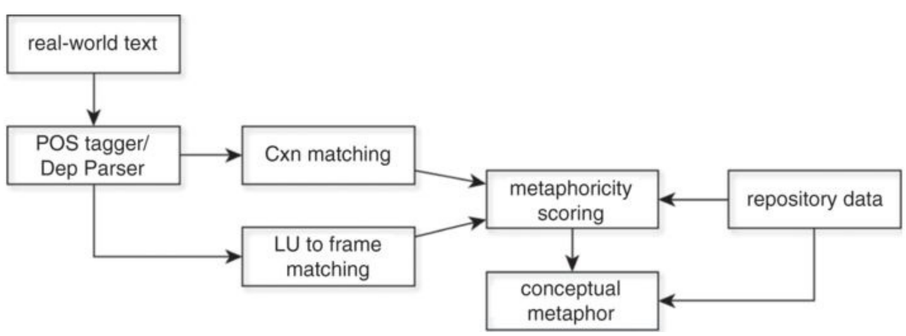

Figure 35.6 The metaphor identification pipeline

First, corpus texts are pre-processed for uniformity and compatibility with the part-of-speech and dependency tagging conventions adopted (Hong 2016). Then, texts are subjected to part-of-speech and dependency parsing, and that information is subsequently interpreted by a patented constructional pattern-matching system. The system consists of several pre-determined schematic grammatical constructions covering many, but not all, possible metaphoric collocations in which at least one constituent evokes the source domain and one evokes the target domain of a metaphor.[4] These constructional patterns are rough generalizations over actual grammatical constructions, and they are designed so that the metaphor identification system is able to assign source and target domain frames appropriately, depending on where the lexical unit appears in the sentence. Frames in the repository are not themselves tagged as being candidates for source or target domain slots in any given metaphor; it is thus the job of the construction-matching pattern to make this assignment by virtue of having slots designated for target and for source domain frames. The frames are subsequently recognized via the lexical items populating the constructional slots, which in turn are associated with frames in the repository frame entries.

Figure 35.7 shows an illustration of the pipeline for one particular linguistic metaphor that could easily be found in a large corpus, particularly one focusing on political and news texts (as is the case with the English Gigaword Corpus).

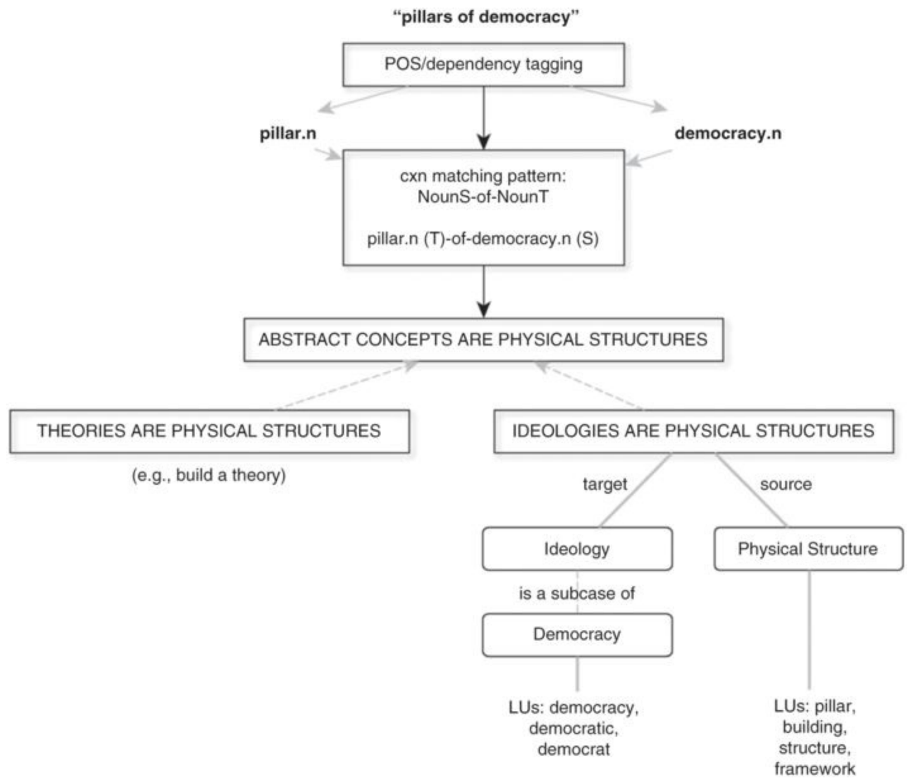

Figure 35.7 Sample metaphor identification run identifies *pillars of democracy* as metaphoric

In the automated metaphor identification run illustrated in Figure 35.7, the system is tuned to search the corpus based on a target domain of interest, in this case Democracy, which already contains several LUs that can potentially evoke this frame. The system can be constrained to either search for a particular target frame, a hand-picked subset of target frames, a known target frame family (e.g. all frames pertaining to Ideas/Thinking) or by target frames in general. For instance, an analyst can have a set of blog entries that they suspect ahead of time to contain many metaphors on issues of democracy and/or freedom and/or rights. Alternatively, the analyst might not know anything about the metaphors in that set of blog entries and may want the system to crawl for all possible metaphors that the system is pre-equipped with. It can also be tuned to search by source domains, for instance, to show all metaphors for which the target domain is construed as object manipulation versus motion along a path. Finally, all constraints can be removed, and the system can be told to find all possible metaphors, fitting all available grammatical construction matching patterns. The 'wild card' by target or source domain search method is possible precisely due to the specification of grammatical construction matching patterns. Thanks to insights from recent writings on the interdependence between metaphoric structure and grammatical structure (Croft 1993, Sullivan 2013a, Stickles, David, and Sweetser 2014), we know to expect certain constructional slots to always be target-domain or source-domain evoking.

The efficacy of the system comes in the form of its expandability from the compendium of primary metaphors (a relatively fixed set) and of more general metaphors; the latter can be a language-specific or a commonly occurring set of non-primary metaphors. However, we need not encode every possible specific metaphor into the repository in order to ensure a metaphor identification result for a specific metaphoric collocation. In the example in Figure 35.7, the metaphor DEMOCRACY ARE PHYSICAL STRUCTURES happens not to be a metaphor in the repository, and yet the expression *pillars of democracy* is nevertheless recovered as a potential LM with a high metaphoricity score. This is by virtue of the subcase relation between the frame Democracy and the frame Ideology, where it is the latter frame, and not the former, that acts as the source domain of a general metaphor IDEOLOGIES ARE PHYSICAL STRUCTURES. This more general metaphor has many more LMs than just those pertaining to democracy, such as *tear down fascism, dismantle totalitarianism brick by brick*, and *communism crumbled at the end of the Cold War*. We can then capitalize on a network that already provides us with 'heavy traffic nodes' in the conceptual structure, where most of the inferences are packed for a particular metaphor, while minimally adding frame subcase relations when interested in a new abstract domain for which existing general metaphors can potentially be informative.

Upon performing a crawl, the metaphor identification system then provides all candidate LMs found in the text, each accompanied by their metaphoricity scores. Metaphoricity is scored in a gradient manner, such that an LM can be evaluated as more-or-less likely to be metaphoric, rather than metaphoric or non-metaphoric. The metaphoricity score is an index of the certainty with which the system can make the suggestion that any given LM be considered metaphoric, based on pre-specified parameters. The metaphoricity score is in part calculated based on the distance between the immediate frame (e.g. Democracy) and the higher-level frame for a candidate metaphor, measured in numbers of subcase relations between the two (Hong 2016). It is also measured in terms of the relationship between the source and target domain frames. A collocation will pick out two LUs (for illustrative purposes, A and B), each of which connects to a frame (e.g. A evokes frame C, and B evokes frame D). The collocation will receive a low metaphoricity score under two circumstances: 1) if A evokes a frame that is itself a subcase (or a subcase of a subcase) of D, and 2) if A and B both evoke D. For instance, *cure* *cancer* is ruled out as metaphoric because *cancer* and *cure* both evoke the Disease frame; however, *cure poverty* would receive a high metaphoricity score since the Disease source domain frame and the Poverty target domain frame do not find themselves in the same frame network (Dodge 2016).

Finally, to expand on the ability of the extractor to identify more and more LMs of a similar kind as *pillars of democracy*, additional LUs can be added to both the source and target frames in the lowest-level metaphor in the cascade. In the case illustrated in Figure 35.7, adding *brick* and *mortar* in the Physical structure frame can ensure that the expression *the brick and mortar of democracy* is extracted with a high metaphoricity score on a subsequent run.

35.4 Conclusions

Adapting standard NLP tools for corpus text processing to accommodate metaphor identification is an important component in CMT theory development, in empirical enrichment of the study of metaphor, and in improving computational linguistic techniques better to reflect diverse figurative forms of expression. Here, the MetaNet metaphor repository and identification system was presented as one possible solution to accurately and automatically detect metaphoric collocations. Researchers in any field working with linguistic data evaluate texts not only in terms of the percentages and distributions of linguistic collocations within a domain of interest, but also in terms of how those linguistic collocations relate to bigger conceptual structures that are informative in the understanding of human decision-making and reasoning processes. For this reason, the proposed textual analysis architecture – consisting of both a hand-built frame and metaphor ontology and an automated metaphor identification component – can prove fruitful. Most importantly, by implementing a structured conceptual network model to both conceptual and computational analyses of metaphor, diverse linguistic metaphors can be organized and understood against the backdrop of complex cultural models and shared schemas grounded in embodied cognition.

---

[1] Here, linguistic metaphor is defined simplistically as any linguistic manifestation of a conceptual metaphor, and not as a separate type of metaphor parallel to conceptual metaphor. For instance, *tax relief, tax burden*, and *heavy taxes* are three LMs that evoke the same conceptual metaphor, TAXATION IS A BURDEN with one entailment, RELIEVING TAXATION IS RELIEVING A BURDEN.

---

[2] www.whitehouse.gov/the-press-office/2016/01/12/remarks-president-barack-obama-%E2%80%93-prepared-delivery-state-union-address.

---

[3] Although *pull out of poverty* could be argued to be evoking a scene involving someone getting help in overcoming an obstacle on a horizontal, and not necessarily a vertical path, the corpus data shows that poverty is most often construed as a low location – e.g. *pull out of poverty, raise out of poverty, leap/climb out of poverty*. On the other hand, few examples arise in which poverty is unambiguously an obstacle on a horizontal path (e.g. *the wall of poverty, get over the hurdle of poverty*, etc.). This suggests that *pull* is evoking an aid to vertical motion from a lower to a higher location.

[4] Some examples of the grammatical patterns include: verb-object (*cure poverty*), noun-noun (*poverty trap*), and noun-of-noun (*jaws of poverty*), among others. See Dodge, Hong and Stickles (2015) and David, Lakoff, and Stickles (2016) for a detailed discussion and exhaustive list of construction-matching patterns currently implemented in the system, as well as the details of the semantic libraries and ontologies used.