

Finding source domain triggers

How corpus methodologies aid in the analysis of conceptual metaphor

Jenny Lederer

San Francisco State University

Much recent research on figurative language and conceptual metaphor theory derives from corpus examination, and analysts are increasingly focused on the development of quantificational tools to reveal co-occurrence patterns indicative of source and target domain associations. Some mappings between source and target are transparent and appear in collocation patterns in natural language data. However, other metaphors, especially those that structure abstract processes, are more complex because the target domain is lexically divorced from the source. Using economic discourse as a case study, this paper introduces new techniques directed at the quantitative evaluation of metaphorical occurrence when target and source relationships are nonobvious. Constellations of source-domain triggers are identified in the data and shown to disproportionately emerge in topic-specific discourse.

Keywords: conceptual metaphor, lexicalization patterns, collocation, economic discourse, political discourse

1. Introduction

This paper addresses emerging questions born out of the intersection of research on figurative language and corpus linguistics. With the growing accessibility of large bodies of data, the study of Conceptual Metaphor Theory (Lakoff & Johnson 1980) has moved from a discipline of introspective inquiry and detailed qualitative analysis to a field that embraces new quantitative and experimental methods (Gibbs 2010). Unlike one-off analogies or isolated figurative expressions, conceptual metaphors are systematic, structured mappings between one conceptual domain and another (Lakoff & Johnson 1980, Kövecses 2010). Linguistic expression is regularly figurative and is often lexicalized and idiomatic, as in phrases like *a*

cool cat or *a sly dog*. In these phrases, meanings broaden to encompass new senses, which are understood to be nonliteral. However, beneath most metaphorical and idiomatic language lies a system of abstract concepts, methodically structured by fixed mappings from source domain to target.

Any abstract concept which cannot be touched, seen, or felt functions as a target domain — such as love and other emotions, communication, ideas, competitions, interpersonal relationships, theories, arguments, religion and math. A smaller set of more concrete, experiential domains serves as the organizing mechanism through which target domains are understood. Common source domains are spatial and schematic and consist of elements and experiences which are frequent and universal such as journeys, objects in motion, plants, human and animal bodies, food and cooking, machines, temperature, and light (Kövecses 2010). Mapping source to target domain yields conceptual metaphors like LOVE IS A JOURNEY, in which lovers are understood to be travelers and forward motion along the path indicates progress in the romance. The conceptual metaphor LOVE IS A JOURNEY gives rise to the idiomatic language used to describe romance: *going nowhere*, *hit a dead end*, *a bumpy relationship*, *on the rocks*, and to novel elaborations of the metaphor which are just as easily comprehended: *spinning our wheels*, *reaching a crossroads*, *fast track to marriage*, etc. Complex target domains are often political in nature, prompting linguists, discourse analysts, and cognitive scientists to turn their attention from simply cataloging common source and target domains built into the grammar of English and other languages to a focus on target domains with far-reaching political implications (Lakoff 2002, 2009; Lakoff & Wehling 2012, Fausey & Matlock 2011, Matlock 2012).

One issue consistently raised in response to this genre of research is a type of methodological criticism (Gibbs 2011: 533). Analysts who carry out qualitative analyses on small samples of data can be accused of “cherry picking” examples to fit “armchair” analyses, and within the field of metaphor research there is a vocal constituency pushing investigators to pay more attention to quantitative analysis of larger samples of data (Gibbs 2010: 6–7, Deignan 2012: 447).

The goal of applied cognitive and conceptual metaphor research is commendable — to educate the public on how language embeds hidden thought patterns and damaging assumptions, often reinforcing social stereotypes. These researchers, nonetheless, often turn a blind eye toward several important questions having to do with the lexicalization patterns of metaphor. That is, what is the relationship between corpus data and metaphor activation? Can this activation be quantified? And, if so, how can semi-automated corpus technique be used to more efficiently mine source and target domains?

2. Corpus linguistics in metaphor research

Quantitative research into the system of conceptual metaphors in a variety of languages typically follows one of two methodological tracks. In critical metaphor studies, researchers manually mine topic-specific corpora to pull out metaphorical patterns (Charteris-Black 2004, 2005). Musolff (2006:24), for example, generates his own corpus of political discussion within the European Union from news media sources. He hand searches his corpus for metaphorical tokens in order to evaluate how frequently a particular source domain strain is evoked within the text. He then uses data counts to argue that certain discourse communities share specific source domain scenarios through their reliance on common folk models. There is no doubt that this manual approach yields the highest degrees of accuracy in the identification of metaphorical tokens. Line by line data coding, however, is labor-intensive, requiring researchers to read all the data in the corpus, thus constraining the size and scope of the study, with potential implications for the representativeness of the data sample. As L'Hôte (2014:28) argues, purely manual qualitative examinations of requisitely limited sample size can lead to unreliable results that may or may not apply at scale. As a result, contemporary researchers are now often working with a corpus of tokens in the millions, as opposed to the thousands.

An alternative methodological approach is to automate searching by examining only tokens which reference the target domain. Oster (2010) relies on collocation patterns in a nonspecific corpus to show which lexical units are most associated with metaphorical description of the emotion fear. Oster (2010) collects co-occurrence information — the lexical units that most frequently collocate with *fear* — to find target-specific metaphorical expressions. She uses the results of collocation searching to build a source-domain ontology, arguing that the most “relevant” metaphors are those evoked by the highest number of linked linguistic expressions (Oster 2010:742). For example FEAR IS SOMETHING INSIDE THE BODY is evoked more frequently than is FEAR IS AN ANTAGONIST. Some metaphors, however, such as FEAR IS FIRE are more creatively produced because they are evoked by a larger set of linguistic expressions. In Oster's (2010) approach, frequency information combines with lexical co-occurrence data to produce a source domain's “productivity and creativity index” (Oster 2010:748) — additional parameters by which source domains can be compared.

Following a similar semi-automated approach, MetaNet project investigators (David et al. 2014, Stickles et al. 2014) at the International Computer Science Institute have engaged in a corpus-driven, lexical approach to researching the alignment between target domain expressions, source domain frames, and the grammatical constructions that blend the two. Target and source word pairs, such

as *alleviate poverty*, in which the source domain of DISEASE is evoked to understand the target domain POVERTY, are used to quantitatively evaluate the frequency of one source domain in relation to another. Through these source-target pairings, the frequency of activation of individual frames can be compared to other frames within the same source domain. For example, in the British National Corpus, Stickles et al. (2014) show how POVERTY is more frequently discussed as a DISEASE than as a basic HARM. And, when understood as a disease, speakers are more likely to discuss the treatment of the affliction of poverty than the diagnosis of the disease of poverty. Thus, at a macro level, the corpus results lead to the conclusion that AFFLICTION and TREATMENT roles in the source domain are more salient than is the role of DIAGNOSIS (Stickles et al. 2014).

The commonality among most corpus-based approaches to metaphor research, whether through manual searching, sorting, and collection or semi-automated searching based on collocation patterns, is the focus on the target domain. In all cases described above, the researcher uses lexical items indicative of the target domain to find instances of conceptual metaphors. Oster's (2010) collocations searches are based on the word *fear* in order to find semi-fixed expressions such as *fight fear*. Likewise, the data mining approach taken by Stickles et al. (2014) is to look for common collocates of the word *poverty*, such as *spread*, *alleviate*, and *fight*.

Many target domains, however, cannot be thoroughly investigated by searching a corpus for collocates of target lexemes because not all target language occurs near or next to source domain triggers. This is the case for metaphorical concepts that are fundamentally understood as processes not as entities, and most target domains, like the economy, are built on extremely complex conceptual ecology. Thus, the ease with which metaphorical structure can be exposed has to do with the relationship between source-domain language and the structural character of the target domain. When the target domain is cognitively complex and lexically divorced from the source, target-domain directed searching limits the extraction of relevant data.

Because of the constraints imposed by manual searching and the lexical division between source and target triggers, metaphor researchers in corpus linguistics have turned to alternative approaches (Koller et al. 2008, L'Hôte 2014, Demmen et al. 2015). Demmen et al. (2015) use a semi-automated corpus-based approach to research violence metaphors active in discourse on cancer. In their method, repeated source domain verbs like *fight*, *battle*, and *struggle*, identified through manual searching, are grouped according to predetermined semantic fields such as "warfare" or "damaging and destroying" (Demmen et al. 2015:211). These fields come from an adapted version of the UCREL2 Semantic Analysis System (USAS) tagger (Rayson et al. 2004) in *Wmatrix* (Rayson 2008). Identifying relevant semantic fields, and the lexis associated with each, yields additional search tokens such as *destroy* and *shatter*, which serve as supplementary source domain triggers used to locate

additional metaphorical tokens in cancer discourse. This focus on the identification of source-domain language leads to a greater diversity of identified metaphorical lexis, allowing a systematic comparison of particular metaphorical tokens across three groups of speakers — cancer patients, family caregivers, and healthcare providers — and across two genres of data, interviews and online forum postings.

Following this same research line, my case study focuses on querying metaphorical language by concentrating on source, rather than target, language. However, the methods described below differ from previous corpus-based approaches, which rely on prespecified semantic fields in the collection of candidate metaphor triggers.

3. Data and method

As a target domain, economy, along with basic understandings of business and finance, has been well researched within metaphor analysis both in and out of academia (McCloskey 1986, Boers 1997, Boers & Demecheleer 1997, Skorczynska & Deignan 2006, Kövecses 2010, Shenker-Osorio 2012). Because the metaphors used to structure economic thinking are well understood, it serves as a good case study to investigate the ways in which the metaphors are instantiated in natural discourse. All data referenced in this paper come from a 2,084,650 token corpus built from the business and finance sections of *The Economist* magazine (2008–2015). In Section 3.1, example sentences are presented to illustrate the existing, established metaphorical ecology underpinning economic discourse. In Section 3.2, the search methodology used to catalogue lexicalization patterns specific to the *The Economist* corpus is outlined.

3.1 The economy as a metaphor

The economy, as a complex, abstract system, does not have a particularly unique metaphorical structure. Systems of all kinds, such as social organizations, governments, corporations, climate, and physical organisms are understood primarily through the same sets of metaphors. These metaphors all have one thing in common: the source domains represent different instantiations and elaborations of physical structures vis-à-vis the primary metaphor ABSTRACT SYSTEMS ARE PHYSICAL STRUCTURES. The physical structures that serve as subcases of this superordinate metaphor, however, vary: ABSTRACT COMPLEX SYSTEMS can be understood through several structure types including MACHINES, BUILDINGS, PLANTS, and HUMAN BODIES (Kövecses 2010: 156).

In some examples, the economy is understood as a building. This metaphor maps the physical characteristic of a building and the logical entailments therein over to various aspects of economic reasoning (Kövecses 2010: 136). Lexical items in *The Economist* magazine that evoke this metaphor include *build*, *foundation*, *strong*, *weak*, *crash* and *collapse* among others, as shown in Examples (1) to (6):

- (1) Bankers accept they will be forced to *build* up bigger capital buffers [...]
- (2) The bank is back on a solid financial *foundation* [...]
- (3) [...] evidence of a *stronger* economy should make bullion rise, not fall.
- (4) [...] there is no trade-off between supporting a *weak* economy
- (5) [...] after the early 1990s financial *crash* [...]
- (6) All three props have now *collapsed*. In particular, as America's housing bust [...]

The metaphorical organization of economics, finance, and business is closely tied to the concept of money. The target domain, MONEY, is structured through an understanding of the movement of liquid or water. Money *circulation* is the *flow* of capital from banks to consumers to business and back, again as illustrated in Examples (7) to (11):

- (7) Footloose capital generates *bubbles* as it *rushes in* [...]
- (8) [...] as the fall in exports was exacerbated by a sudden *drying up* in trade finance.
- (9) Best to be liquid in case the *well runs dry*.
- (10) Prices would have gone down even further had not transactions *dried to a trickle*.
- (11) Taken together, these measures have *splashed cold water* on the market.

The source domain LIQUID is elaborated and extended through rich concepts related to water including *bubbles* (soapy water) and *wells* (containers for water) along with characteristics important in the description of water like *hot* and *cold* (ACTIVITY IS HEAT; INACTIVITY IS COLD) and numerous manners of movement like *rush* or *trickle*.

Because liquid, usually instantiated as water, serves as the source domain for the understanding of money, a metaphorical understanding of the economy also relies on several source domains that have to do primarily with water. These are WEATHER EVENTS, i.e. rain, storms, clouds, BODIES OF WATER like seas, oceans, channels, rivers, and the VEHICLES that navigate water ways: boats and ships. The

concept of climate is used to discuss public sentiment on economic issues and economic performance is understood through perceptions pertaining to weather conditions: bad weather means negative performance; good weather is positive, as in Examples (12) to (20):

- (12) [...] meaning the banks have buffers to weather a medium-sized *storm*.
- (13) Edwards has dubbed these *conditions* an “ice age” which he predicts will extend [...]
- (14) If *drought* is responsible for some of China’s price pressure, a *deluge of credit* is to blame for the rest. So China-watchers were quick to welcome a *turn in the monetary weather* this week.
- (15) The euro crisis *casts a chill over a sunnier economic picture* [...]
- (16) Losses are surging as the *economic climate* worsens [...]
- (17) Plus a mass of European professionals hiding from *the economic winter* [...]
- (18) [...] returns from offices and shops *twist and turn in the economic winds* [...]
- (19) As with *monsoon rain*, so with foreign capital [...]
- (20) [...] then France will probably be part of the *hurricane*,” says a senior Italian banker.

Weather language comes in many forms including specific conditions like *stormy*, *sunny*, *wind*, and *drought*, and in the form of concepts related to weather patterns like *season*, *climate*, and *forecast*.

The economy is assumed, like a vehicle, to move forward at varying rates of speed. It can accelerate or slow down because progress, in any domain, is metaphorically understood as forward motion. Because of the water-based metaphors that structure monetary theory, the vehicle is most commonly a ship or boat. Policymakers direct the economy, just like captains steer ships. A poorly performing economy is likened to a sinking ship and economic or financial failure is talked about as if it were a shipwreck, as it can be seen in Examples (21) to (32):

- (21) What is worrying is that today’s traders are in truly *uncharted (and very cold) waters* [...]
- (22) Partly they may have been *buoyed* by robust business conditions in China.
- (23) ABC’s 320m customers [...] are reason enough to go ahead regardless of *turbulent* conditions.
- (24) Conditions are undeniably *frothy*.

- (25) Over time the economy will be weighed down by all these costs, like a barnacle-encrusted *ship*.
- (26) If we can get over the May borrowing hump, it's a relatively smooth *cruise* for the rest of the year [...]
- (27) Non-profit lenders, pawnshops and co-operatives also *swim in the microcredit sea*.
- (28) Mr. Carney would not be alone in thinking all this adds up to *rough sailing for a long time*.
- (29) The *storm-tossed* lender will reach *port* eventually [...]
- (30) [...] but Mr. Lewis's chances of being the one who *leads the crew onto dry land* are weakening fast.
- (31) True, UBS is also fortunate in not having big loan books to drag it down as the real economy *sinks*.
- (32) By contrast, IASB [...] wants to spare loans held to maturity by banks from the *vagaries* of the market [...]

The language of seafaring is rampant in *The Economist* magazine and includes many examples of words that reference water conditions, especially those that arise in inclement weather.

In a different metaphor with an unrelated source domain, the economy is understood through the frame of a (usually sick) human body, a type of elaborated personification (Shenker-Osorio 2012: 43), as illustrated in Examples (33) to (37).

- (33) It is unreasonable to ask policymakers to worry about the long-term *side effects* of their *medicine* [...]
- (34) [...] while banks are still in *surgery*, but today's *cure* may well be the source of tomorrow's *ills*.
- (35) When an economy *sneezes*, its trading partners *catch a cold*, as demand for their exports falls. The *germs* can also spread through financial channels [...]
- (36) In Japan *contagion* risks were twice as high, despite its markets' relative lack of synchronicity. So market correlation data cannot predict *contagion*. It is like *an airborne virus that can change direction with the wind*, and *infect* countries that least expect it.
- (37) Choosing the bleakest statistic from a report issued by Kevyn Orr, Detroit's emergency manager, on his city's *financial health* is like choosing the wettest raindrop in a monsoon.

When analyzing a complex target like the economy, it is not surprising to find a concept, like temperature, playing a role in more than one source domain. For example, body temperature is one measurement used to assess health and disease in a medical setting, but air temperature is also a concept integral to our understanding of weather and machines. Thus, lexemes like *hot* and *cold* can activate alternative source domains in various contexts and can be compatible with multiple source domains simultaneously. The activation of temperature differentially applies to multiple metaphors. High temperature in a body is bad, a sign of illness in the metaphor *ECONOMY IS AN (AILING) HUMAN BODY* as illustrated in Example (38):

- (38) Banks are asked to raise a sum of money; the lead managers *take the market's temperature* and advise about the yield and maturity investors want.

But when used in reference to the *ECONOMY IS A WEATHER EVENT*, heat is good and cold is bad (as prescribed by the underlying metaphor *HEAT IS ACTIVITY*). These inferences are shown in Examples (39) and (40):

- (39) If this is right, rich-world economies may enjoy a boost after the end of the *winter freeze*.
- (40) BEIJING recently suffered its lowest temperature in 59 years, but *the economy is sweltering*.

As the above sentences demonstrate, figurative language that draws on more than one source domain is common. The intersecting nature of these related source domains helps explain the prevalence of lexical items that are compatible with more than one metaphor. Metaphorical tokens activate source domains that overlap in their conceptual structure as illustrated in Figure 1.

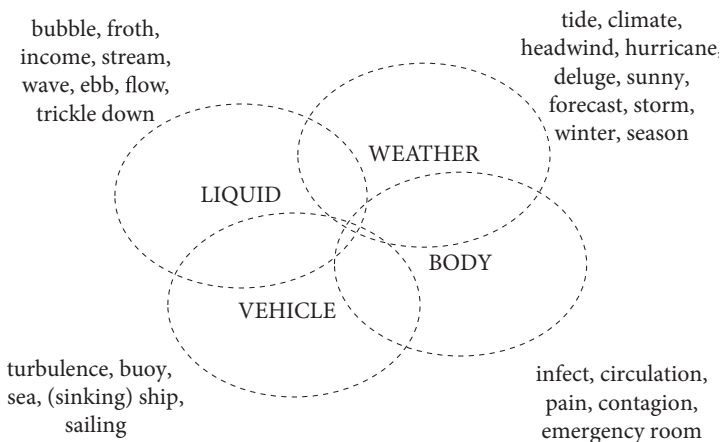


Figure 1. Illustration of overlapping source domains and associated lexical triggers

3.2 Searching for specific patterns of lexicalization

Unlike the investigation of a lexically-encoded metaphorical concept such as fear or poverty (Oster 2010, Stickles et al. 2014), a corpus approach directed at the economy cannot be comprehensive by searching for only target domain words such as *economic*, *finance*, and *business* because of the metaphorical complexity involved. Many source domain triggers in economic discourse are lexically removed from target domain language, and in this way, economic discourse mirrors the metaphorical disease language studied by Demmen et al. (2015) and Koller et al. (2008). Just as in these studies, I advocate for an approach in which metaphors for the economy are investigated through the lexicalization patterns of the source domain. In this mixed-method approach, a small portion of the restricted corpus is qualitatively scanned for frequent metaphors, and then quantitatively assessed by pulling source, not target, domain examples. However, unlike Demmen et al. (2015) and Koller et al. (2008), I adopt an approach that integrates corpus collocation patterns. Rather than using a semantic ontology to build a set of source domain triggers, I pull triggers by looking at common collocates of source domain labels. Through the *Sketch Engine* interface (Kilgarriff et al. 2014), I compiled a 2,084,650 token corpus of *The Economist* data taken from articles in the “Business” and “Finance” sections of issues published between 2008 and 2015. In order to partially automate the identification of source domain language in this specialized corpus, I used collocation searching in a baseline corpus to identify potential source domain triggers (lexical items that activate one or more source domains used to structure the target concept). In this methodology, a source domain label serves as a collocation magnet to collect a list of frequent words associated with the specified domain. The Corpus of Contemporary American English (COCA) interface and data (Davies 2008-, Davies 2009) was used in order to model a simple and accessible technique available to a wide variety of metaphor researchers, especially those with limited exposure to corpus methodology. This method was employed to investigate two of the more robust source domains listed above: the source domains in the metaphors *ECONOMY IS A SHIP* and *ECONOMY IS A WEATHER EVENT*.

In the case of *SHIP*-related language, I searched the academic section of the COCA for the frequent collocates of the word *ship* (5L:5R; MI > 3). For practical purposes the Mutual Information score was used as a measure of significance, as it is available through the COCA web interface. An MI score greater than 3.0 is interpreted as significant (Cheng 2012).¹ Although a wide collocation window introduces statistical noise to the evaluation of significance (Desagulier 2014: 155), a

1. In COCA, Mutual Information is calculated as follows: $MI = \log \left(\frac{AB * \text{sizeCorpus}}{A * B * \text{span}} \right) / \log(2)$ (Davies 2008-).

wider span reveals general associations between node and collocate (Brezina et al. 2015: 155). These broad associations capture lexis representative of the discourse surrounding seafaring in general. Of the top collocates, only those related to the semantic domain of shipping were utilized to probe the topic-specific discourse. For example, *slave* is a common collocate of *ship*, but it does not constitute lexis central to the SEAFARING/SHIPPING frame, as in the sense of Fillmore (1982), and was accordingly not used to probe *The Economist* corpus.

To find WEATHER-related language, the same collocate search was carried out with the word *weather*. This technique produced a list of the 125 top collocates for each source domain label. These candidate source domain triggers were then evaluated for their ‘metaphoricity’ (rate of metaphorical use) in relationship to economic discourse in the specialized corpus.

4. Quantitative analysis of metaphorical triggers

The COCA search method yielded a sizeable set of potential source domain triggers for two economic metaphors: ECONOMY IS A SHIP and ECONOMY IS A WEATHER EVENT. These two metaphors were selected to demonstrate the methodology because they involve two source domains with tight constellations of lexical triggers.

I have established three distinct categories to classify source domain triggers in the specialized corpus. ‘Trigger lexeme’ is the term I use to indicate any lexical item in the specialized corpus that evokes one or more relevant source domains. Many words can function as trigger lexemes. Some are words very closely tied to a source domain frame; for example, the phrase *on life support* is directly tied to our understanding of hospitals, emergency rooms, and very sick patients, and can be used to activate the metaphor ECONOMY IS (AILING) HUMAN BODY. Other words, however, activate one source domain, but that source domain structures more than one target. This would be the case for a word like *circulation*, which can be used to describe the movement of money, MONEY IS LIQUID, but can also be used as a source domain trigger for a different metaphor like IMMIGRATION IS THE FLOW OF WATER. Thus *circulation* is only counted as a trigger when used in metaphorical description of the economy.

‘Significant trigger lexemes’ are lexical items that have a significant rate of use as a source domain trigger, quantified as a frequency of three or more metaphorical uses in reference to the target. ‘Significant’ triggers must also be used in metaphorical reference to the target domain in at least 20% of the instances of total use, yielding a moderate to high rate of metaphoricity (an ‘Insignificant Trigger Lexeme’ is used either fewer than three times or less than 20% of use is metaphorical).

‘Super trigger lexeme’ refers to lexical items unique to the restricted corpus and exhibits a significant rate of metaphorical use in reference to the specified target domain (over 20%). Uniqueness is measured by a disproportionate use in the specialized corpus compared to a baseline corpus. To quantify uniqueness, I ranked individual lexical items through a ‘weirdness’ algorithm (Ahmad 2005), which I will call ‘keyness score’. In Ahmad’s (2005) measure, the frequency rate of the lexical item in the specialized corpus is divided by the frequency ratio of the item in a general corpus. Any item which occurs at a higher rate in the specialized corpus will measure at a keyness score greater than 1.0. This equation is illustrated below, where F is the frequency of item and N is the number of total tokens in corpus:

$$\text{Keyness (term)} = \frac{F_{\text{special}} \div N_{\text{special}}}{F_{\text{general}} \div N_{\text{general}}}$$

The keyness score quantifies the relative frequency of a particular lexical item in the restricted corpus (compared to a baseline) and allows lexical items to be both ranked by their relative frequency and numerically compared to one another. Scoring lexical items for keyness in addition to counting metaphorical uses reveals the three different types of triggers, as summarized in Table 1.

Table 1. Source domain trigger lexemes are grouped according to potency

Insignificant trigger lexeme	Significant trigger lexeme	Super trigger lexeme
Source domain trigger. Used to structure the target domain in question.	Source domain trigger. Used to structure the target domain in question.	Source domain trigger. Used to structure the target domain in question.
And	And	And
Used metaphorically in at least one instance, but fewer than three	Three or more metaphorical uses in the restricted corpus	Three or more metaphorical uses in the restricted corpus
Or	And	And
Less than 20% of use is metaphorical (in reference to the target domain in question).	At least 20% of use is metaphorical.	At least 20% of use is metaphorical.
		And
		Keyness score greater than 1.00

4.1 The SHIP source domain

Trigger lexemes reside on a scale of potency in their relationship to the source domain. Some triggers loosely connect to the source domain, while others consistently evoke it. Table 2 is comprised of a subset of the 125 top collocates of the lexical item *ship* in the academic section of COCA (5L:5R; MI > 3).

Table 2. Ship collocates listed by percent of metaphorical use in relation to economic topics*

	Metaphorical hits	Total in Econ. corpus	%Metaphorical uses	Keyness score
<i>aground</i>	5	5	1	4.962942221
<i>keel</i>	4	4	1	2.426327308
ballast	3	3	1	0.515833365
voyage	1	1	1	0.073033263
hull	1	1	1	0.091367974
<i>shore (up)</i>	43	45	0.955555556	1.380144044
<i>anchor</i>	28	31	0.903225806	2.554510637
<i>sink</i>	61	68	0.897058824	12.27200258
<i>storm</i>	36	42	0.857142857	1.020191015
<i>channel</i>	75	92	0.815217391	1.486495754
sailed	3	5	0.6	0.833471213
<i>sail</i>	7	12	0.583333333	1.313500497
<i>steer</i>	31	60	0.516666667	3.84791996
boat(s)	10	22	0.454545455	0.612771437
waters	16	43	0.372093023	0.537332571
ocean	5	22	0.227272727	0.30924545
captain	2	9	0.222222222	0.296429128
sea	12	59	0.203389831	0.322861747
cruise	1	7	0.142857143	0.583429849
crew	1	8	0.125	0.297354155
(vehicle) wreck	2	18	0.111111111	3.235103077
ship	4	127	0.031496063	2.139060635
port	2	121	0.016528926	2.293637533
board	1	332	0.003012048	1.374251918
aboard	0	1	0	0.082093781
cargo	0	22	0	1.709654117
passenger	0	18	0	1.066662209

Table 2. (continued)

	Metaphorical hits	Total in Econ. corpus	%Metaphorical uses	Keyness score
navy	0	3	0	0.084694037
coast	0	48	0	0.448416427
merchant	0	43	0	2.500635601
vessel	0	24	0	0.557539041
deck	0	3	0	0.292459095
pirate	0	22	0	5.75344679
transport	0	88	0	1.259273413
bay	0	19	0	0.226166241
sailor	0	3	0	0.359949656
harbor	0	0	0	0
fleet	0	18	0	0.79167175
dock	0	16	0	3.01199252
stern	0	0	0	0
propeller	0	0	0	0
bow	0	0	0	0
canal	0	21	0	0.473246503

* Significant Trigger Lexemes are shaded; Super Trigger Lexemes are in italics.

It is important to note that not all frequent collocates of the source domain label *ship* are used metaphorically in reference to the target domain. For example *boarding a ship* is a common fixed or semi-fixed phrase in language about ships and boats, a robust and salient aspect of our frame for ships, and an important component of our interaction with boats and ships. However, *board* is not a lexeme speakers use to discuss the economy. Apart from the word *ballast*, no language referencing the physical parts of the ship is used in the discussion of economic issues. The conceptual focus is on how the economy, as a ship, navigates its course and confronts difficult water conditions. The results of the classification schema are summarized in Table 3.

4.2 The WEATHER source domain

Table 4 is comprised of a subset of the 125 top collocates (5L:5R; MI > 3) of the lexical item *weather* in the academic section of COCA. These are all words that evoke the WEATHER frame. But, again, there is wide variation in whether or not these collocates are used as source domain triggers for the metaphor ECONOMY IS A WEATHER EVENT.

Table 3. Categorized triggers for ECONOMY IS A SHIP

Not triggers	Insignificant triggers	Significant triggers	Super triggers
Cargo	Ship	Waters	Channel
Passenger	Port	Sea	Sink
Navy	(vehicle) wreck	Boat(s)	Shore (up)
Coast	Cruise	Ocean	Storm
Merchant	Board	Sailed	Anchor
Vessel	Crew	Ballast	Sail
Deck	Voyage		Aground
Pirate	Hull		Keel
Transport	Captain		Steer
Bay			
Sailor			
Harbor			
Fleet			

Table 4. *Weather* collocates listed by percent of metaphorical use in relation to economic topics *

	Metaphorical hits	Total in Econ. corpus	%Metaphorical uses	Keyness score
stormy	2	2	1	0.77989092
vagaries	6	6	1	1.401301333
cloudy	7	7	1	2.526588767
forecast	431	433	0.995381062	61.00256464
freeze	95	99	0.95959596	10.3935463
storm(s)	36	42	0.857142857	1.020191015
channel	75	92	0.815217391	1.486495754
terrain	4	5	0.8	0.196552167
dry	81	105	0.771428571	1.327281798
flood	80	111	0.720720721	6.965232703
sunny	14	20	0.7	3.211315554
mild	31	46	0.673913043	1.423812198
severe	59	95	0.621052632	0.679832819
cool	31	56	0.553571429	1.881336866
calm	41	81	0.50617284	4.356632038
atmosphere	6	12	0.5	0.128484113
conditions	144	302	0.476821192	0.567608351

Table 4. (continued)

	Metaphorical hits	Total in Econ. corpus	%Metaphorical uses	Keyness score
<i>hot</i>	31	78	0.397435897	1.164239077
season	17	52	0.326923077	0.539055865
<i>drought</i>	14	43	0.325581395	1.716615481
<i>cloud</i>	13	46	0.282608696	1.66860383
ocean	5	22	0.227272727	0.30924545
cold	7	32	0.21875	0.215407603
wind	7	38	0.184210526	0.43230213
climate	25	139	0.179856115	0.649617006
weather	21	140	0.15	2.547643673
temperature	5	34	0.147058824	0.275238612
rain	2	18	0.111111111	0.346312796
wet	1	10	0.1	0.280680537
winter	3	31	0.096774194	0.345028195
hurricane(s)	1	14	0.071428571	0.70849882
ice	1	26	0.038461538	0.375875929
inclement	0	1	0	1.213163654
snow	0	16	0	0.431347077
tornadoes	0	1	0	0.357982718

* Significant Trigger Lexemes are shaded; Super Trigger Lexemes are in italics.

A word like *forecast* is used frequently in discourse on the economy; it has an extremely high percentage of metaphorical use and is key to *The Economist* corpus, with a keyness score of 61.00. It is a super trigger. On the other hand, *hurricanes* and *snow* are words frequently mentioned in weather discourse, but these are not used in any significant way in discussion of the economy. The results of the classification schema are summarized in Table 5.

4.3 Limitations

Even though source domain collocations yield a large quantity of metaphorical language, there are limits. Not all trigger lexemes can be found through this automated technique. Manually tagging a subset of *The Economist* corpus reveals that there are robust source domain triggers that are not frequent collocates of their source frame labels. That is, there are trigger lexemes that come from an understanding of ships and weather, which do not frequently co-occur with the word *ship* and *weather*, and the fact that these triggers exist shows the limitations in

Table 5. Categorized triggers for ECONOMY IS A WEATHER EVENT

Not triggers	Insignificant triggers	Significant triggers	Super triggers
Inclement	Stormy	Terrain	Vagaries
Snow	Wind	Severe	Cloudy
Tornadoes	Climate	Atmosphere	Forecast
	Weather	Conditions	Freeze
	Temperature	Season	Storm(s)
	Rain	Ocean	Channel
	Wet		Dry
	Winter		Flood
	Hurricane(s)		Sunny
	Ice		Mild
			Cool
			Calm
			Hot
			Drought
			Cloud

using a methodology that exclusively relies on source frame collocation magnets. Many of these source domain triggers are not salient concepts in our understanding of the source domain frames, but are frequently used jargon in economic discourse. Several are listed in Table 6.

Table 6. Lexical triggers for LIQUID, WEATHER, and SHIP source domains*

	Metaphorical hits	Total in Econ. corpus	%Metaphorical uses	Keyness score
<i>ebb</i>	22	22	1	6.623206413
<i>froth</i>	7	7	1	10.18568865
<i>frothy</i>	29	29	1	50.63742358
<i>muddy</i>	5	5	1	1.054419115
<i>buoy</i>	35	35	1	7.41676358
<i>fizzle</i>	10	10	1	11.19306445
<i>cleanse</i>	4	4	1	0.220469451
<i>turbulent</i>	18	18	1	1.693433458
<i>choppy/ier</i>	8	8	1	11.26527777
<i>headwind(s)</i>	25	25	1	136.415473
<i>bubble(s)</i>	458	459	0.997821351	29.86096077
<i>flow</i>	466	471	0.989384289	1.655971335
<i>circulation</i>	29	32	0.906250000	1.02037578

Table 6. (continued)

	Metaphorical hits	Total in Econ. corpus	%Metaphorical uses	Keyness score
<i>wave(s)</i>	124	139	0.899280576	2.41186767
navigate	8	9	0.888888889	0.5069375
<i>float</i>	78	89	0.876404494	2.584905303
<i>circulate</i>	29	35	0.828571429	1.051516378
stream	35	46	0.760869565	0.456994939
<i>turbulence</i>	9	13	0.692307692	1.576356577
swim	4	6	0.666666667	0.150874256
<i>tide</i>	14	32	0.4375	1.242788651
wash	8	21	0.380952381	0.427970111
dam(s)	2	6	0.333333333	0.138288125
breeze	1	3	0.333333333	0.550247286
tsunami	3	21	0.142857143	4.303812106

* Significant Trigger Lexemes are shaded, Super Trigger Lexemes are in italics.

The words in this list exhibit high percentages of metaphorical use, which indicate that they are used figuratively most of the time in *The Economist* corpus — their metaphoricality, as measured by percentage of metaphorical use, is high. A subset of these triggers are ‘super triggers’ as well because they not only exhibit a significant percentage of metaphorical use, but they are also key to discussion of economic issues. These categorical results are listed in Table 7.

4.4 Fixed phrases

Researching source domain triggers reveals nuanced patterns of figurative use and fixed collocation — synonymous verbs do not follow the same patterns. For example, the words *flow* and *stream* are near synonyms in everyday language, and both words activate the LIQUID source domain in the metaphorical understanding of money and the movement of money. However, these synonyms do not equally participate in the activation of the metaphor. While *cash flow** (20 tokens in *The Economist* corpus) is a very frequent lexical cluster used in economic language, *cash stream* is not. In fact there are no tokens of *cash stream* in *The Economist* corpus at all. *Stream*, on the other hand, tends to co-occur with the word *income*, as in *income stream* (12 tokens in *The Economist* corpus).

A comparison of these two source domain triggers also reveals a sizeable difference in their overall frequency in economic language. The super trigger lexeme *flow* is used 466 times out of 471 total instances to metaphorically reference the movement of money and assets, an extremely high rate of metaphorical usage (99%).

Table 7. Categorized triggers for WEATHER, LIQUID, and SHIP source domains

Insignificant triggers	Significant triggers	Super triggers
Tsunami	Cleanse	Ebb
	Navigate	Froth
	Stream	Frothy
	Swim	Muddy
	Wash	Buoy
	Dam(s)	Fizzle
		Turbulent
		Choppy/ier
		Headwind(s)
		Bubble(s)
		Flow
		Circulation
		Float
		Circulate
		Turbulence
		Tide
		Wave

Likewise *flow* has a significant keyness score of over 1.66. *Stream* also has a high percentage of metaphorical uses in relation to the economy (76%), but its overall frequency is much lower (35 metaphorical uses out of 46 total tokens). And, *stream* is not a lexical item unique to the restricted corpus, as it does not have a significant keyness value (.46). These quantitative differences suggest that, should one want to find metaphorical language evoking the MONEY IS LIQUID metaphor in discussion of the economy, the lexeme *flow* is a more robust trigger and a trigger that can be used to find a larger quantity of metaphorical examples. The differing rates of use also suggest nuanced patterns in how money is understood. *Stream* entails a unidirectional flow of water, from water source to water outlet. *Flow*, on the other hand, does not incorporate such a salient source-path-goal asymmetry. Thus, *stream* appears most apt in the description of money movement from business venture to investor; whereas *flow* is more basic and can be used in many instances of money transfer.

4.5 Frequency and trigger strength

Some source domain triggers are highly figurative, meaning that their literal sense is infrequently, if ever, evoked. These are words used rarely in unrestricted English

discourse but used consistently with a metaphorical meaning in economic language, often a characteristic of professional jargon in general. The words in Table 3 with high keyness scores, for the most part, fall into this category. *Froth*, *frothy*, *buoy*, *fizzle*, and *bubble* are good examples. All of these words have a much higher likelihood to occur in *The Economist* corpus in a metaphorical capacity than either metaphorically or literally in the academic section of COCA. For example, there are 29 instances of *frothy* in *The Economist* corpus and only 25 total instances in the much larger academic section of COCA, resulting in an extremely high keyness score of 50.64. *Froth* and *frothy*, in their literal sense are used in reference to high concentrations of bubbly water, but in economic discourse, *froth* refers to markets in which the price of assets begins to rise above their “real” value because of high demand as illustrated in Example (41).

- (41) As a result of the crash the industry faces four big obstacles to recovery.
Thanks to *frothy* equity markets, the industry is closest to overcoming the first barrier — exiting current investments.

Froth is a metaphorical extension of a market *bubble*, a prolonged period of asset overvaluation, shown in Example (42).

- (42) But the list also includes half a dozen rich-country crashes, from Japan’s slump after its property *bubble* burst in the late 1980s, to the Nordic bank crises in the early 1990s.

What is important to remember about these lexemes is that, although they evoke fixed senses unique to economic terminology, they are still source domain triggers. They are part of a *system* of source domains that structure our understanding of money, markets, finance, and the economy. In fact, these particular word uses arise because of an organized and integrated combination of source domains. In this case, water, as a form of liquid, when infused with air, creates bubbles and froth. Because quantity is understood as verticality (MORE IS UP), visual representation in a line graph or histogram of the rapid increase and decrease of values looks like a downward parabola. Bubbles raise the surface of the water, resulting in a similar configuration. Thus, the concept of a bubble evokes the source domains of VERTICALITY and LIQUID. As an extension of these two source domains, the idea of a literal bubble offers several important entailments: bubbles are ephemeral, fragile, and can easily burst. An *economic bubble* is a temporary, fragile, and precarious valuation of assets.

Economic jargon of this type exhibit high keyness scores and high relative frequency in the restricted corpus, suggesting that keyness can be used as a numeric evaluation to quantify and compare idiomatic usage patterns. In usage-based theories, idioms — nonliteral, unanalyzable word senses tied to a particular discourse

context — are understood as graded on a scale of idiomaticity, not categorical (Gibbs & Nayak 1989, Wulf 2010). The keyness scores in Table 6 hint that one indirect measure of idiomaticity is keyness. By ranking the lexemes in Table 7 according to keyness score, a division is revealed. Items with high keyness scores are the least likely metaphorical triggers to be understood via context, whereas words with lower keyness scores include, although still figurative, meanings perhaps more easily deciphered through context. This division is summarized in Table 8.

Table 8. Keyness as a measure of idiomaticity

Most idiomatic	Least idiomatic
Headwinds	Flow
Froth/y	Turbulent/ce
Bubble	Tide
Choppy/ier	Muddy
Fizzle	Circulate/tion
Buoy	Navigate
Ebb	Stream

4.6 Variation in source domain activation

There is great variability in how dense and connected the constellation of source domain triggers is within a given source domain. The source domain of an unhealthy human body is a lexically diffuse conceptual domain. That is, there is no fruitful source domain label to use as a collocate magnet in COCA. Thus, the automated technique of using collocation patterns to find source domain seed language does not work here. In this case, source domain triggers have to be directly collected from *The Economist* corpus in order to investigate the metaphor by searching for likely source domain triggers given the identified source frame. Directing searching reveals the set of source triggers for the HUMAN BODY domain as shown in Table 9.

Table 9. Lexical triggers for (AILING) BODY source domain*

	Metaphorical hits	Total in Econ. corpus	%Metaphorical uses	Keyness score
<i>ailing</i>	9	9	1	1.418326939
<i>life support</i>	8	8	1	3.01054837
code blue	2	2	1	87.30590272
<i>contagion</i>	63	63	1	11.50684492
<i>ailment</i>	5	5	1	3.410386825
autopsy	3	3	1	0.834132828
chest pains	2	2	1	7.936900247
epidemiology	2	2	1	0.088815771
transfusion	1	1	1	0.196634916
<i>recovery</i>	550	552	0.996376812	4.491412703
cure	32	35	0.914285714	0.843651738
pain(s)	127	152	0.835526316	0.818863212
emergency room	2	3	0.666666667	0.723529581
<i>healthy</i>	106	159	0.666666667	1.35140562
medicine(s)	24	46	0.52173913	0.29521255
disease(s)	29	58	0.5	0.132164283
pulse	3	6	0.5	0.229350007
sick	9	19	0.473684211	0.52065667
health	107	300	0.356666667	0.23115553
ill	3	33	0.090909091	0.425442231

* Significant Trigger Lexemes are shaded; Super Trigger Lexemes are in italics.

This source domain is notable in several ways. Apart from the use of frequent triggers, writers incorporate particularly literary elaborations. For example, phrases like *code blue*, *emergency room*, and *chest pains*, are all expressions tightly linked to our understanding of a very sick patient in the hospital. These extensions work in economic discourse because they simultaneously evoke the metaphor ECONOMY IS AN (AILING) BODY and import a set of unique entailments, which convey detailed inferences. For example, when central banks are understood as emergency rooms for very sick economies, they are seen as agents of brief treatment, not prolonged intervention, as in Example (43).

- (43) But this support was supposed to be short-term, not continuous: a central bank should be an *emergency room*, not a *hospice*.

Standout source domain triggers among this set are the words *healthy*, *recovery*, *contagion*, and *ailing*. These words, which occur frequently in the data, have high percentages of metaphorical use in reference to economic concepts and significant keyness scores, meaning they are more frequently represented in economic discourse than in academic English as summarized in Table 10.

Table 10. Categorized Triggers for ECONOMY IS AN (AILING) BODY

Insignificant triggers	Significant triggers	Super triggers
Code blue	Autopsy	Ailing
Chest pains	Cure	Life support
Epidemiology	Pain(s)	Contagion
Transfusion	Medicine(s)	Ailment
Emergency room	Disease(s)	Recovery
Ill	Pulse	Healthy
	Sick	
	Health	

The ailing patient frame is activated through the concept of illness. The economy and business finance in general are referred to as *ill* three times, as *sick* in nine instances, and as *ailing* also nine times in *The Economist* corpus. These words have similar meanings and are used in similar ways to address the economy and business as in Examples (44) to (46).

- (44) Policymakers should keep an eye on this growing body of research for guidance on how to marshal health-care resources when economies fall *ill*.
- (45) For all Mr Putin's apparent self-confidence, a *sick* economy weakens his hand and makes Russia more vulnerable to sanctions.
- (46) They decided, among other things, to allow the new permanent bail-out fund to recapitalise banks in *ailing* economies directly rather than via their governments.

However, unlike *ill* and *sick*, there are no instances of *ailing* used either literally or metaphorically in reference to a target domain apart from the economy. This statistic exemplifies how words, like *ailing*, which are not particularly frequent in basic English discourse, can still serve as robust, super trigger lexemes and exhibit a high degree of frequency as economic jargon.

4.7 Collocates of source triggers

Conceptual metaphor interacts with constructional grammar resulting in asymmetries in how conceptual metaphors are lexically encoded, and some of these configurations are revealed in collocation patterns (Deignan 2005, Stefanowitsch 2005, Sullivan 2009). The process of cataloging trigger lexemes can expose these collocational details and further refine the mining process. The significant trigger word *health*, for instance, evokes the metaphor ECONOMY IS HUMAN BODY in some (36%) tokens in *The Economist* corpus. But, the collocation *health of*, on the other hand, is used exclusively (100%) as a metaphorical reference to financial institutions and economies, never used in reference to human health, as in Example (47).

- (47) The crisis between the capitalisation and *health of* America's banks and those in Europe.

In comparison, the collocations *health or* and *health and* are never used in metaphorical reference to finance and neither are common lexical clusters like *health problems* or *health warnings*. When *health* appears directly before *and* in *The Economist* corpus, it is used to refer to human health and the healthcare system.

4.8 Derivational asymmetries

Just as specific bigrams can be used to further refine mining for source domain language, so too can derivational asymmetries. Shown in Table 4, *health*, used as a noun, qualifies as a significant trigger lexeme (36% metaphorical use applied to economy target), but is not a super trigger (keyness = .23). In adjectival form, however, the word is a super trigger: *healthy* is used in 67% of examples to metaphorically reference the economy and has a significant keyness score (1.35).

*Ail** is another example lexeme to examine closely. Unlike *health* and *healthy*, *ailing* and *ailment* both have the same rate of metaphoricity — that is, they are both used exclusively to metaphorically reference finance. However, *ailing* never appears as a verbal gerund; it is always used as a modifier. In fact, *ail*, is never used as a verb in *The Economist* corpus.

4.9 Explaining super triggers

We can, to some extent, explain super triggers through the overlapping of source domains, as was illustrated in Figure 1. The list in Table 11 contains all the super triggers from the various metaphors investigated above. The words in italics in the list, numbering about half of the total super triggers, are those which are compatible with more than one source domain.

nor is there a way to foresee what source domain language will be adopted into the speech community and used consistently as idiomatic jargon.

5. Implications for metaphor research

There are several implications and applications of the quantification and categorization of source domain triggers. Painting a clearer picture of how source domain concepts are lexically encoded in target domain discourse allows researchers to probe important questions related to metaphorical reasoning. For example, the finding that language related to inclement weather dominates the activation of the WEATHER source domain suggests that economists view market behavior and economic performance as events that can be anticipated, but not controlled, as in Examples (48) and (49).

- (48) Last year Indonesia was struck by the financial *storm* that pummeled emerging markets, earning itself a place among the so-called “fragile five” of the developing world.
- (49) By international standards, the loans extended to customers are backed by plenty of equity, meaning the banks have *buffers to weather a medium-sized storm*.

As noted by Shenker-Osorio (2012), the repeated use of this metaphor highlights the belief that the economy is uncontrollable, a natural phenomenon separate from humans and ignores the role that governments, institutions, and corporations play in economic performance. The fact that the most frequent lexical triggers of the WEATHER source domain are words like *storm* and *cloud* serves as more evidence for the robustness of the metaphor itself and shows how economists think about economic players. Financial institutions are entities that need protection from the storm as opposed to understanding them as entities that *cause* the storm. Poor weather, after all, is never understood to have a cause nor understood as the result of human decision-making and bad behavior.

When it comes to the metaphorical analysis of other issue areas, it is not uncommon for researchers to speculate on the dominance of one model over another with no real quantitative analysis. An impressionistic assessment of active metaphors has its place in applied conceptual metaphor research. Yet, corpus assessment is increasingly available in the evaluation of discourse in many applied areas of linguistic research (Cheng 2012). It is still an open research question as to whether corpus frequency statistics translate into conceptual potency as suggested by Musolff (2006) and Oster (2010). This proposal rests on the assumption that the

model most frequently evoked within a speech community is also the model most conceptually salient in the minds of the speakers — a kind of metaphor theory version of the entrenchment hypothesis described for grammatical representation (Schmid 2000: 39) and conceptual blending (Fauconnier & Turner 2002: 49, Coulson & Oakley 2005: 1530).

This connection could be probed further through the use of corpus metrics, including metaphorical usage percentages and keyness scores. If a source domain is actively structuring an abstract idea, language from the source frame should be frequent and robust in topic specific discourse. The “over-use” of model-evoking lexical units points to conceptual reliance on a holistic domain. Therefore, one way to argue a given model is dominant among a particular group of speakers is to compare the frequency of lexical triggers in and out of the speech community. If a particular model is consistently evoked to address a metaphorical concept, then some of the lexemes associated with that model should appear in topic-specific data more frequently than in the language overall.

However, in order to evaluate the robustness or potency of a complete metaphorical model, it is insufficient to probe the frequency of just one lexical item. A better representation of model prominence is to establish the comparative frequency of a collection of lexical items all related to the same source domain. Importantly, though, the evaluation of a constellation of source domain triggers must be carried out with caution. Language related to the source domain is not language exclusive to the source domain. Direct frequency count comparisons of individual words will not necessarily yield useful information. Trigger lexemes need to be evaluated as a group for both metaphoricality in relation to the target concept and keyness in relation to the restricted corpus.

The abovementioned factors are at the heart of the difference between the approach suggested here and the methodology implemented in other corpus-based approaches to culling source domain language (e.g. Koller et al. 2008, L'Hôte 2014, Demmen et al. 2015). Three important questions need to be addressed when probing topic specific-data for conceptual metaphors, especially when the focus is on finding source domain language. Firstly, the source domain seed language, whether culled from a semantic ontology, as in Demmen et al. (2015), or culled from source domain collocation magnets in a general corpus, needs to be evaluated for how representative it is of the source frame. How germane are the lexical items to the frame of knowledge that structures the source? Quantity of related lexis — multiple synonyms for the same action or entity — needs to be balanced with breadth of coverage, e.g. are essential frame elements represented in the lexis? Secondly, the potential source triggers in use should be closely evaluated in regard to which specific target domain they activate. If metaphor is in fact conceptual, not just lexical, discourse context should not be confused with target concept. Many

target domains share source domain frames, thus it is easy to falsely group source domain hits as meaningfully framing one target domain, when, in fact, unique instances are activating separate metaphors.

Lastly, it is of note to mention that the methodology presented in this paper allows the researcher to quantify the frequency of use of metaphorical language in genre-specific discourse in relationship to other contexts. This additional parameter may be of interest in the examination of metaphorical language in target domains of importance to corpus linguists. For example, how unique is violence-based language in the discussion of cancer as opposed to the discussion of illness in general? Is *devastate* used more frequently by caregivers discussing loved ones with cancer than the general population discussing sad circumstances of any kind? Is opponent language used more frequently to talk about fear, as opposed to other negative emotions? Is disease language especially unique to poverty or is it an equally common framing of other social ills like hunger, crime, and addiction?

6. Conclusion

Collections of naturally occurring language data serve as repositories of metaphor and can be used to investigate lexical patterns indicative of specific source domains. Since the introduction of big data, in the form of sizable, computationally searchable corpora, metaphor analysts in academia have begun to probe questions of quantitative validity (Deignan 2005). This movement toward quantifiable indicators of metaphorical salience has allowed for increasing focus on the quantification of conceptual metaphors, which is seen as a mechanism to shield conceptual metaphor research from ongoing methodological criticism. For some, the validity of the theory rests on movement away from a fine-grained qualitative analysis of exemplary data to more robust experimental and quantitative measures designed to gauge salience and nuanced details of how conceptual metaphors are lexicalized and expressed in natural discourse (Gibbs 2011, Deignan 2012).

Most analyses, however, are based on metaphorical data that are easy to mine. That is, probing a corpus for metaphorical data in which both source and target domain language is paired and collocated is a straightforward process. But this method is not possible for many metaphorical concepts due to the nature of how target domains are represented. When a target domain like the economy is understood as a complex system and is based on multiple conceptual metaphors, direct lexical searches using target domain language will not recover all pertinent structural information about active source domains.

In this study, I have modeled a different mining technique centered on source, rather than target, language. The conceptual metaphors used to understand the

economy are overwhelmingly activated by certain metaphorical tokens, which I categorize as ‘super triggers’. That is, specific frames used to understand the economy are linked to a subset of lexemes, which are consistently used metaphorically and occur more frequently in economic discourse than in the general corpus. This methodology involved probing central metaphors like *ECONOMY IS A SHIP* and *ECONOMY IS A WEATHER EVENT* by pulling common collocates of source domain labels such as *weather* and *ship* from a general corpus. The resulting constellation of collocates was then evaluated to see if they are in fact source domain triggers by searching the specialized corpus for instances of use. In the case of the economy, a subset of source-domain triggers are habitually used metaphorically and occur more frequently in *The Economist* corpus than in nonspecific English discourse. This relative frequency differential was assessed through a ‘keyness score’, which is a numerical measure of how frequent, on average, metaphorical triggers occur in the restricted corpus compared to the baseline corpus. Words that consistently activated the metaphorical source domain fit into a category of ‘super triggers’ and were then used to mine more metaphorical data.

In addition to providing metaphor researchers supplementary strategies to validate impressionistic conclusions of model dominance, the methods used in this study suggest several future lines of inquiry concerning the automated extraction of metaphorical data. The programmed extraction of metaphorical tokens is not only of interest to corpus linguists but has also been a focus of much research in natural language process and computer science (Babarczy et al. 2010, Tang et al. 2010). A subgroup of these researchers is focused on the automatic detection of metaphorical tokens in relationship to the system of conceptual metaphors that structure a given language (Shutova et al. 2013, Dodge et al. 2015). These efforts can be furthered through the identification of significant and super source domain triggers. The effort to build ontologies of conceptual metaphor, necessary for any adequate computational model of semantic processing, should occur alongside the identification of salient trigger language, and the construction of such systems will undoubtedly benefit from the methodologies outlined in this paper.

References

- Ahmad, K. (2005, June). Terminology in text. Paper presented at the *Tuscan Word Centre Workshop*, Siena, Italy.
- Babarczy, A., Bencze, I., Fekete, I., & Simon, E. (2010). The automatic identification of conceptual metaphors in Hungarian texts: A corpus-based analysis. In N. Bel, B. Daille, & A. Vasiljevs (Eds.), *LREC 2010 Workshop on Methods for the Automatic Acquisition of Language Resources: Proceedings* (pp. 30–37). Malta.

- Boers, F. (1997). "No pain, no gain" in a free market rhetoric: A test for cognitive semantics? *Metaphor & Symbol*, 12(4), 231–241. doi: 10.1207/s15327868ms1204_2
- Boers, F., & Demecheleer, M. (1997). A few metaphorical models in (western) economic discourse. In W. A. Liebert, G. Redeker & L. Waugh (Eds.), *Discourse and Perspective in Cognitive Linguistics* (pp. 115–129). Amsterdam: John Benjamins. doi: 10.1075/cilt.151.10boe
- Brezina, V., McEnery, T., & Wattam, S. (2015) Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, 20(2), 139–173.
- Charteris-Black, J. (2004). *Corpus Approaches to Critical Metaphor Analysis*. London: Palgrave-Macmillan. doi: 10.1057/9780230000612
- Charteris-Black, J. (2005). *Politicians and Rhetoric: The Persuasive Power of Metaphor*. London: Palgrave-Macmillan. doi: 10.1057/9780230501706
- Cheng, W. (2012) *Exploring Corpus Linguistics*. England: Routledge.
- Coulson, S., & Oakley, T. (2005). Blended and coded meaning: Literal and figurative meaning in cognitive semantics. *Journal of Pragmatics*, 37(10), 1510–1536. doi: 10.1016/j.pragma.2004.09.010
- David, O., Dodge, E., Hong, J., Stickles, E., & Sweetser, E. (2014, September). Building the MetaNet metaphor repository: The natural symbiosis of metaphor analysis and construction grammar. Paper presented at the 8th International Conference on Construction Grammar (ICCG 8), Osnabrück, Germany.
- Davies, M. (2008-). The Corpus of Contemporary American English: 450 million words, 1990-present [Online Corpus].
- Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*. 14(2), 159–190. doi: 10.1075/ijcl.14.2.02dav
- Deignan, A. (2005). *Metaphor and Corpus Linguistics*. Amsterdam: John Benjamins. doi: 10.1075/celcr.6
- Deignan, A. (2012). Figurative language in discourse. In H. J. Schmid (Ed.), *Cognitive Pragmatics* (pp. 437–462). Berlin: Mouton de Gruyter.
- Demmen, J., Seminoi, E., Demjén, Z., Koller, V., Hardie, A., Rayson, P., & Payne, S. (2015). A computer-assisted study of the use of violence metaphors for cancer and end of life by patients, family carers and health professionals. *International Journal of Corpus Linguistics*, 20(20), 205–231. doi: 10.1075/ijcl.20.2.03dem
- Desagulier, G. (2014). Visualizing distances in a set of near synonyms: *Rather, quite, fairly, and pretty*. In D. Glynn & J. Robinson (Eds.), *Polysemy and Synonymy: Corpus Methods and Applications in Cognitive Linguistics* (pp.145–178). Amsterdam: John Benjamins. doi: 10.1075/hcp.43.06des
- Dodge, E., Hong, J., & Stickles, E. (2015, June). Deep semantic automatic metaphor analysis. Paper presented at The Third Workshop on Metaphor in NLP, Denver, Colorado.
- Fauconnier, G., & Turner, M. (2002). *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York, NY: Basic Books.
- Fausey, C. M., & Matlock, T. (2011). Can grammar win elections? *Political Psychology*, 32(4), 563–574. doi: 10.1111/j.1467-9221.2010.00802.x
- Fillmore, C. J. (1982). Frame semantics. In Linguistic Society of Korea (Ed.), *Linguistics in the Morning Calm* (pp. 111–137). Seoul: Hanshin Publishing Co.

- Gibbs, R. W. (2010). The wonderful, chaotic, creative, heroic, challenging world of researching and applying metaphor. In G. Low, A. Deignan, L. Cameron, & Z. Todd (Eds.) *Researching and Applying Metaphor in the Real World* (pp.1–18). Amsterdam: John Benjamins. doi: 10.1075/hcp.26.01gib
- Gibbs, R. W. (2011). Evaluating Conceptual Metaphor Theory. *Discourse Processes*, 48(8), 529–562. doi: 10.1080/0163853X.2011.606103
- Gibbs, R., & Nayak, N. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, 21(1), 100–138. doi: 10.1016/0010-0285(89)90004-2
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovár, V., Michelfeit, J., Rychly, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 1–30. doi: 10.1007/s40607-014-0009-9
- Koller, V., Hardie, A., Rayson, P., & Semino, E. (2008) Using a semantic annotation tool for the analysis of metaphor in discourse. *Metaphorik.de*, 15(1), 141–160.
- Kövecses, Z. (2010). *Metaphor: A Practical Introduction*. Oxford: Oxford University Press.
- Lakoff, G. (2002). *Moral Politics: How Liberals and Conservatives Think*. Chicago, IL: University of Chicago Press. doi: 10.7208/chicago/9780226471006.001.0001
- Lakoff, G. (2009). *The Political Mind: A Cognitive Scientist's Guide to your Brain and its Politics*. New York, NY: Penguin Books.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: University of Chicago Press.
- Lakoff, G., & Wehling, E. (2012). *The Little Blue Book: The Essential Guide to Thinking and Talking Democratic*. New York, NY: Free Press.
- L'Hôte, E. (2014). *Identity, Narrative and Metaphor: A Corpus-based Cognitive Analysis of New Labour Discourse*. London: Palgrave Macmillan. doi: 10.1057/9781137427397
- Matlock, T. (2012). Framing political messages with grammar and metaphor. *American Scientist*, 100(6), 478–483. doi: 10.1511/2012.99.478
- McCloskey, D. N. (1986). *The Rhetoric of Economics*. Brighton: Wheatsheaf Books and Harvester Press.
- Musolf, A. (2006). Metaphor scenarios in public discourse. *Metaphor and Symbol*, 21(1), 23–38. doi: 10.1207/s15327868ms2101_2
- Oster, U. (2010). Using corpus methodology for semantic and pragmatic analyses: What can corpora tell us about the linguistic expression of emotions? *Cognitive Linguistics*, 21(4), 727–763. doi: 10.1515/cogl.2010.023
- Rayson, P., Archer, D., Piao, S., & McEnery, T. (2004). The UCREL semantic analysis system. In *Proceedings of the workshop on Beyond Named Entity Recognition Semantic labeling for NLP tasks in association with the 4th International Conference on Language Resources and Evaluation (LREC 2004)* (pp. 7–12). Paris, France: European Language Resources Association.
- Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4), 519–549. doi: 10.1075/ijcl.13.4.06ray
- Schmid, H. (2000). *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin: Mouton de Gruyter. doi: 10.1515/9783110808704
- Shenker-Osorio, A. (2012). *Don't Buy It: The Trouble with Talking Nonsense about the Economy*. Philadelphia, PA: Public Affairs.

- Shutova, E., Teufel, S., & Korhonen, A. (2013). *Computational Linguistics*, 39(2), 301–353.
doi: 10.1162/COLI_a_00124
- Skorczynska, H., & Deignan, A. (2006). Readership and purpose in the choice of economics metaphors. *Metaphor and Symbol*, 21(2), 87–104. doi: 10.1207/s15327868ms2102_2
- Stefanowitsch, A. (2005). The function of metaphor: Developing a corpus-based perspective. *International Journal of Corpus Linguistics*, 10(2), 161–198. doi: 10.1075/ijcl.10.2.03ste
- Stickles, E., Dodge, E., & Hong, J. (2014, November). A construction-driven, MetaNet-based approach to metaphor extraction and corpus analysis. Paper presented at the *12th meeting of Conceptual Structure, Discourse, and Language (CSDL 12)*, Santa Barbara, California.
- Sullivan, K. (2009). Grammatical constructions in metaphoric language. In B. Lewandowska-Tomaszczyk & K. Dziwirek (Eds.) *Studies in Cognitive Corpus Linguistics* (pp. 57–80). Frankfurt: Peter Lang Publishers.
- Tang, X., Qu, W., Chen, X., & Yu, S. (2010, December). Automatic metaphor recognition based on semantic relation patterns. Paper presented at the *International Conference on Asian Language Processing*, Harbin, China.
- Wulf, S. (2010). *Rethinking Idiomaticity: A Usage-based Approach*. London: Continuum International Publishing.

Author's address

Jenny Lederer
Department of English Language and Literature
San Francisco State University
1600 Holloway Ave.
San Francisco CA 94132
United States
lederer@sfsu.edu

Copyright of International Journal of Corpus Linguistics is the property of John Benjamins Publishing Co. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.