5-1-2010

Figuration & Frequency: A Usage-Based Approach to Metaphor

Daniel Sanford

Daniel Sanford
Candidate

Linguistics
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Catherine Travis, PhD , Chairperson

Joan Bybee, PhD

Lynne Cameron, PhD

William Croft, PhD

Jill Morford, PhD

FIGURATION & FREQUENCY:
A USAGE-BASED APPROACH
TO METAPHOR

BY

DANIEL SANFORD

B.A., Linguistics & Classics, University of Wisconsin, 2002
M.A., Linguistics, University of New Mexico, 2005

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Linguistics

The University of New Mexico
Albuquerque, New Mexico

May 2010

DEDICATION

To my parents, Mike & Toni.

ACKNOWLEDGMENTS

Sincere thanks to all of the members of my dissertation committee: (Joan Bybee, Lynne Cameron, William Croft, and Jill Morford) for all of their insights, comments, and support: it's a far better project for their input.

In particular, I'm grateful to the chair of my dissertation committee, Catherine Travis, for her guidance throughout the process, and for the immense amount of time that she's spent reading and commenting on this dissertation.

FIGURATION & FREQUENCY:
A USAGE-BASED APPROACH TO FREQUENCY

BY

DANIEL SANFORD

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Linguistics

The University of New Mexico
Albuquerque, New Mexico

May, 2010

Figuration & Frequency: A Usage-Based Approach to Metaphor

by

Daniel Sanford

B.A., Linguistics & Classics, University of Wisconsin, 2002
M.A., Linguistics, University of New Mexico, 2005
PhD, Linguistics, University of New Mexico, 2010

ABSTRACT

Two of the major claims of the cognitivist approach to metaphor, the paradigm which has emerged as dominant over the last three decades, are 1) that metaphor is a conceptual, rather than strictly linguistic, phenomenon, and 2) that metaphor exemplifies processes which are at work in cognition more generally. This view of metaphor is here placed within the context of the functionalist approach to language, which asserts that linguistic structure is emergent in nature, the use of language directly influencing the storage and representation thereof. The dissertation argues that metaphors, as conventionalized cognitive structures, are themselves highly influenced by frequency effects, and that metaphorical cross-domain mappings exist in the mind as conceptual schemata.

Two corpus-based methods for assessing the frequency of overall metaphorical mappings are presented, both based on the use of key terms, attained using a survey method, for metaphorical source domains. These findings inform the hypotheses of a series of three experiments which test

three key predictions of the view that metaphors are affected by frequency: that frequent metaphors should be more productive, accessible, and acceptable than infrequent ones. Both the corpus and experimental approaches, as well as data from previous research on metaphor at varying levels of conventionalization, support the view that metaphors are a usage-based phenomenon. The properties of various types of metaphorical utterances (e.g., idioms and novel metaphors) are best accounted for as arising from the interaction of the conceptual schemata that license cross-domain mappings, and syntactic schemata that link meanings to syntactic templates.

TABLE OF CONTENTS

Chapter 1: Introduction & Literature review

The popularity of metaphor research among those who view language as arising out of general cognitive processes is best explained by the fact that metaphor so ideally exemplifies the cognitivist approach to language: metaphor is a linguistic phenomenon that isn't, strictly speaking, a linguistic phenomenon at all. Linguistic metaphor, rather, is simply the outermost expression of a more general system which serves, at its core, to structure our view of the world, a cognitive system allowing for the conceptualization of abstract domains of thought to benefit from the clearer, more defined structure associated with more concrete ones. Utterances such as *she blew up at me* or *the time weighed heavy on his mind* don't just describe one thing in terms of another, they evidence a system whereby the speakers of a language conceptualize nebulous domains of thought (here, emotion and time). A thesis central to the field of metaphor research has been that metaphor, while deeply important to cognition and highly worthy of study, isn't special: metaphor isn't anomalous, operating outside of the rules and principles which govern 'normal' speech, but is rather at the core of the conceptual systems which govern thought and (accordingly) language, inextricably tied up with language at the lexical, phrasal, syntactic, and discourse levels.

As many times as this assertion has been made, however, it's unclear that it has been taken all that seriously in guiding research enterprises. If it is the case that metaphor is an essential structuring feature of language, then it ought to be demonstrably true not only that metaphor has an effect on linguistic structure (as has been repeatedly and convincingly demonstrated), but that metaphors will themselves be subject to principles that govern language more generally.

Out of the functionalist tradition, meanwhile, has emerged a deeply empirical understanding of one such principle: that frequency plays a major role in the storage, use, expression and overall structure of language. “The frequency with which certain items and strings of items are used,” write Bybee & Hopper (2001: 3), “has a profound influence on the way language is broken up into chunks in memory storage, the way such chunks are related to other stored material and the ease with which they are accessed.” One of the strengths of the frequency-based approach to language lies in the fact that, as with metaphor, frequency is not a strictly linguistic phenomenon, but operates rather on cognitive structures in general. The phonological reduction of a frequent word has the same underlying causes as, for example, the streamlining, for experienced drivers, of the process of putting a key into the ignition of a vehicle (J. Bybee, personal communication, February 8, 2007).

Metaphors are cognitive entities, their psychological reality repeatedly attested in three decades of research, and as such should be subject to frequency effects. Linguistic metaphors, moreover, should be in particular operated on by those frequency effects to which language is especially subject. At the levels, then, both of linguistic expressions (e.g. ‘he churns out ideas’) and of the underlying cognitive structures which sanction them (e.g. THE MIND IS A MACHINE)—concepts that will be outlined in detail in subsequent chapters—metaphor should be affected by frequency.

The goal of this dissertation is to explore the role of frequency in affecting the storage, representation, production, and diachronic change of metaphors, and will be driven by the following research question(s):

- In what way are linguistic metaphors operated upon by frequency effects?

- What are the domain-general effects of frequency, which operate on both metaphor and on more well-researched levels of structure such as morphology?
- Specifically, how does frequency affect metaphor at the level of underlying cognitive structures, and how does it affect metaphor at the level of surface expressions?
- How do these two levels of frequency interact with one another?

It is hypothesized that metaphor is operated upon by frequency effects in a manner analogous to that which has been repeatedly observed to take place in phonological, morphological, syntactic and lexical phenomena: representations of patterns to which speakers are frequently exposed are 'strong'—easily accessed & productive—in proportion to their frequency. Specific instantiations of such patterns, as a result of the high frequency of the instantiation relative to the pattern overall, can themselves gain in strength, independently of the overall pattern.

The research questions and hypothesis are addressed using a corpus method that makes use of automated searches for similes in a large (385+ million word) corpus for the purpose of ascertaining quantitative data for underlying conceptual metaphors. The approach is predicated on the use of basic terminology, for a given source domain, as search terms. A preliminary study, based on the timed survey method used by Rosch in classic prototype theory experiments (Rosch & Mervis 1975, Rosch 1978), is used to establish basic terminology for ten separate cognitive domains which are of recurrent use as metaphorical sources: WAR, RACING, LIGHT, CHILDREN, FOOD, WRITING, PLANTS, BATTERIES, MACHINES, and THE BODY. The method provides frequency data for metaphorical mappings, which informs the experimental portion of the study. This

method, as well as a related method that makes use of a smaller corpus, are presented as exploratory towards best practices in establishing the overall frequency of metaphorical mappings. The methods and results of both methods contribute significantly to the field of corpus-based research on metaphor.

A series of three experiments directly assesses specific claims relating to how frequency affects metaphor, with sub-hypotheses informed by existing literature on frequency effects in language. For all three experiments, the variable manipulated is the overall frequency of the cross-domain mappings instantiated in the stimuli. A survey method, in which participants rate the acceptability of metaphorically predicated stimuli, is used to assess the relationship between the frequency of an underlying mapping and the acceptability of utterances that instantiate the metaphor. A computer-based reaction time experiment is used to gauge how the accessibility of metaphorical utterances follows from metaphorical frequency. A sentence completion task analyzes how the frequency of a metaphorical mapping bears on the productivity of the mapping.

The overarching goal of the dissertation is a model for understanding metaphors as emergent structures which arise out of frequency, providing a framework whereby metaphor can be approached in future research from a functionalist, as opposed to strictly cognitivist, perspective.

1.1: Review of the Literature

1.1.0 Introduction

This section presents an overview of the existing literature that is directly relevant to the research presented here, bearing either directly or indirectly on the argument that metaphors are cognitive entities which are operated on by frequency effects.

§1.1.1 presents an overview of the major theoretical approaches to metaphor which have arisen out of the tradition of cognitive linguistics, and which provide the basic framework and terminology for the approach to metaphor taken here. §1.1.2 provides an introduction to the use of corpora as a tool for researching metaphor, surveying the major authors and projects of this relatively new field.

§s 1.1.3 – 1.1.5 outline the usage-based model of language (whereby the frequency of linguistic units and patterns is seen as effecting important changes in language storage and processing), extant experimental approaches to demonstrating the effect of frequency on the processing and storage of language, and an overview of network/exemplar cognitive models (the most generally accepted approach for modeling frequency effects, and the model here applied to metaphor).

§1.1.6 provides a synthesis of these two approaches, presenting approaches to metaphor (primarily from psychology) which have attempted to account for the role of conventionalization (hypothesized here to correspond to frequency) in the processing of metaphor, and of initial forays towards applying schema-based models of language to metaphor.

1.1.1 Metaphor

The topic of metaphor has been approached from a few different disciplines, and from rather more than a few theoretical perspectives. The project outlined here, while aiming not to associate itself with any one particular extant theory of metaphor, will repeatedly draw on concepts and terminology from the existing literature. An overview of the major approaches from cognitive linguistics follows, with relevant approaches from psychology reported in §1.2.3.

It is with George Lakoff and Mark Johnson's seminal work on metaphor that the modern field of metaphor research begins, research which still frames the debate on the relationship between underlying conceptual metaphors, and their instantiations as linguistic metaphors. Within conceptual metaphor theory (CMT) (Lakoff & Johnson 1980, Lakoff 1987, Lakoff 1993), metaphor is seen as a system whereby one cognitive domain (the target domain) is structured according to our knowledge about another (the source domain), with metaphorical entailments carrying "certain patterns of reasoning… from the source domain to the target domain" (Croft & Cruse 2004: 197). A cross-domain mapping which has become conventionalized is a conceptual metaphor, which can be instantiated in any number of utterances. One of the key insights of CMT over earlier theories of metaphor is the observation and explanation of entire families of linguistic metaphors predicated on a single conceptual metaphor: for example, the utterances *our relationship has hit a dead-end street*, *look how far we've come*, and *we have to go our separate ways* all instantiate the conceptual metaphor LOVE IS A JOURNEY, while *We were made for each other* and *She is my other half* are sanctioned by the conceptual metaphor LOVE IS A UNITY (OF TWO COMPLEMENTARY PARTS). Within this or any other metaphor, every instantiation is predicated on the same mapping of source domain to target domain (with the title of the conceptual metaphor labeling the two domains involved, e.g., 'love' and 'journeys').

Set correspondences hold between the two domains, taking the form of inferences about the target based on the conceptual structure of the source. Constraining these correspondences is the Invariance Principle, according to which those aspects of the systematicity of the source domain which are mapped to the target domain maintain the

'cognitive typology'—inferences, salient aspects, thematic roles, etc.—of the source domain, provided that these don't conflict with the target domain. In LOVE IS A JOURNEY, for example, certain inferences about journeys (that they have a starting point, that they can end at any point before reaching their destination, that companions can part ways at any point, etc.) hold in their application to love and relationships. The incongruity of *our love turned back for home* would come from its incompatibility with the target domain: the metaphor maps physical distance on to time, and while distances can be retreaded, time can't be undone.

Croft & Cruse (2004) note that CMT errs in drawing a sharp line between the roles of the source domain, which provides structure, and the target domain, which receives it. Blending theory (Turner & Fauconnier 1995, Fauconnier 1997, Sweetser 1992, Fauconnier & Turner 2008), conversely, see metaphorical meaning as drawing structure out of both (or, often, many) domains. Conceptual blending is an extremely robust account of metaphor in that it is not, strictly, a theory of metaphor: it accounts for metaphor using more general cognitive tools which apply wherever a cognitive operation requires input from two or more mental spaces, these being "conceptual packet[s] built up for the purposes of local understanding and action" (Turner & Fauconnier 1995: 184). For metaphor, the theory replaces the simple source-to-target mapping of CMT with an integration network, which involves many spaces and many mappings between them. The resulting structure bears elements from all contributing domains, with systematicity and inferences having their origin not in either the source or target domain, but arising out of the integration network in staggeringly complex ways. In *three hours went by*, for example, certain aspects of the intended meaning are understandable as arising from the

application of the source domain SPACE to the target domain TIME. The inference that units of time are moving objects, however, has it source in neither domain, but rather arises out of the blend of the two (Fauconnier & Turner 2008).

The project proposed here, while not situating itself within either CMT or blending theory, will make repeated use of terminology from both theories which have become the standard for referring to metaphor: conceptual metaphor, domain, source domain, and target domain. In addition, the project will use generally recognized source and target domain labels, using the Lakoff, Espenson, & Schwartz (1991) 'Master Metaphor List' as a source (see §4.1).

1.1.2 Corpus Work on Metaphor

Any work purporting to shed light on the role of metaphor in actual discourse must, of necessity, take actual discourse (as opposed to elicited data, the metalanguage of contrived sentences, or the intuition of linguists as native speakers) as its source of information. Metaphor research from Lakoff & Johnson's seminal work onwards, however, has tended towards a reliance on sentences that 'seem' like natural metaphors, something that any speaker of English would say, but are composed by researchers rather than pulled from natural data. It's reliably the case that corpus data flies in the face of researchers' intuitions and assumptions about language (Labov 1975), and this is as true for metaphor as for any other area of linguistic inquiry (Sanford 2008a). A sound theory of metaphor (or any other aspect of language) must account for, and can only arise out of, actual, as opposed to imagined, language use.

As Chafe (1998: 96) points out, "ordinary conversational talk… occupies a special place as the kind of language that is most natural in both form and function, the kind of

language humans must be designed by evolution to produce and comprehend. It requires no special training or skill to be able to talk casually with others, and every normal person acquires this ability as a natural part of maturation. Because conversation is the form of language least influenced by acquired skills, it provides us with the most direct and uncontaminated access to natural mental processes." Metaphor, like all aspects of language, occurs in its most natural form in spontaneous, conversational speech. Corpora of spoken discourse provide ready access to spontaneous speech, and thus to the natural environment of spontaneous metaphor.

The relatively new field of corpus work on metaphor has already demonstrated convincingly how frequency-based studies can provide important insight into metaphor, both shedding light on claims made in previous theoretical approaches and leading to new ones. Such research has fleshed out how particular domains of discourse use conceptual metaphors (Boers 1999, Charteris-Black 2004, Koller 2005, Musolff 2006) as well as how individual words activate metaphorical meanings (Deignan 1999a, Tissari 2001, Musolff 2004, Shank 2007) and, perhaps most importantly, have provided empirical feedback for some of the key claims regarding metaphor which have been advanced within cognitive linguistics (Deignan 1999b, Stefanowitsch 2005, Sanford 2008a). Each of these approaches are based to some extent on frequency, indicating the variety of findings which can arise out of the application of corpus-based frequency to our understanding of metaphor.

1.1.3 The Usage-Based Theory of Language

In a usage-based account of language (Greenberg 1966, Bybee 1985, Langacker 1987, Croft & Cruse 2004, Givón 1984, Hawkins 1994, Lindblom, MacNeilage, & Studdert-

Kennedy 1983), linguistic performance isn't a byproduct of underlying abstractions, *a priori* constructs which comprise Language and of which utterances are merely indicative. Linguistic representations, rather, are seen as directly operated upon by tokens of linguistic expression, with frequency the engine whereby expression dictates representation. The traditional units of linguistic analysis (segments, syllables, morphemes, words, constructions, etc.) aren't the building blocks of language. They are, rather, entities which emerge as generalizations, abstracted away from sequences to which language users are repeatedly exposed, leading to the emergence of organizational schemata[1] and categories. An experienced linguistic event is stored as a mental representation of the event, and for any given category, those units which are most frequently experienced become strengthened, with a concomitant increase in productivity for patterns they instantiate, while infrequently experienced tokens are correspondingly weak.

Frequency can be counted in two ways: token and type. Token frequency is the raw frequency of a given unit: given a particular unit (at any level of linguistic structure), the frequency with which it occurs in a corpus (as a measure of how frequently it is

[1] In Langacker (1987) schemata are presented as abstractions over semantic, phonological, or otherwise symbolic units, which can in turn sanction specific instantiations of the schema (as, for example, the word 'tip' instantiates the phonological CVC schema). Bybee (1995), who restricts her discussion of schemata to morphological analysis, defines schemata as "emergent generalizations" over "words having similar patterns of semantic and phonological connections" (p. 430).

experienced by a language user) is its token frequency. Type frequency, on the other hand, refers to the frequency of a pattern, or more specifically, the number of items within a language that instantiate the pattern in question. Bybee (2001) gives, for example, the case of *break*. The number of times that the form *broke*, the past tense of *break*, occurs in a corpus would be the token frequency (as indicated by the corpus) of *broke*. A relevant measure of type frequency, on the other hand, would be to assess the number of verbs that form the past tense using the same vowel alternation as that used in *broke* (e.g., *spoke*, *awoke*). The type frequency of this form of the past tense would be considerably lower than that of forms which use the *–ed* ending—that is to say, the pattern applies to a much smaller set of words.

Strengthening of representations (referring to an increase in their cognitive salience) can take place at either the type- or token-frequency level. High token frequency for a particular item causes the representation of the item to be strengthened, while high type frequency causes a pattern (or schema) to be strengthened. The past tense pattern *–ed*, for example, is high in type frequency. The form *weeped* is increasing in usage because of the low token frequency of *wept* relative to the high type frequency of the *-ed* past tense. The form *kept* remains prevalent, on the other hand, because of its high token frequency (Bybee 1985). The relationship between token frequency, type frequency, and productivity is such that a schema of high type frequency is a strong, and therefore, productive, schema. The openness of a schema—the amount of specificity imposed—also plays a role in the strength of the schema, such that open schemata tend to be higher in productivity than more restrictive ones. High frequency tokens within a schema do not contribute to the frequency of the schema, because they tend to become

independent (autonomous) from it, forming their own representations rather than reinforcing the pattern (as is the case for *kept*, for example) (Bybee 1995). Thus, generally, type frequency correlates positively with productivity, but high token frequency will effect an inverse correlation between the two.

1.1.4 Experimental Approaches to Frequency

Various experimental approaches have borne out several of the key claims of the frequency-based approach to linguistic structure—namely, that grammatical knowledge is probabilistic, that frequency increases accessibility, and that type frequency influences productivity. The great majority of such studies have focused on phonotactics, morphology, and (less commonly) constructions. The first such prediction is that grammatical judgments aren't absolute, but probabilistic, based on a speakers' previous experience with the language. In Pierrehumbert (1994), subjects were presented with pairs of nonsense words and asked questions which assessed their acceptability to a native speaker. Each pair contained one word with a low-frequency and another with a high-frequency tri-syllabic sequence (based on its frequency of occurrence in a pronouncing dictionary- a reasonable approximation, the author assumes, of a speaker's previous language experience). The author's hypothesis that the more probable a consonant cluster, the more likely it will be to be judged as acceptable by a native speaker, was supported, indicating that speakers have statistical knowledge of phonological structure which is based on their previous experience, and that they use this knowledge to evaluate novel forms, indicating that such knowledge has a place in their mental grammars. Vitevich et al. (1997) uses, in an initial experiment, a method similar to that used in Pierrehumbert (1994), but incorporating stress placement as well as

phonotactic probability and asking subjects to rate items on a scale of 1 to 10 with respect to their acceptability rather than having them pick one item out of a pair. This experiment was followed up by a second one in which processing times (as measured by recognition time, the time from the presentation of the stimuli to subjects' pressing of a button indicating that they understood it) were measured for the same stimuli. Acceptability was demonstrated to increase, and processing time to decrease, as phonotactic probability increased. In addition, then, to indicating that a word's phonotactic probability, which for any individual speaker is a function of their previous language experience, has a clear effect on their intuitions about language, the study also indicated a direct effect on processing.

The claim that frequency has a direct effect on processing speed (accessibility) is more directly assessed in Hare et al. (2001), which builds on earlier studies such as Morton (1969) and Rubenstein et al. (1970), which have indicated that access is faster for frequent than for infrequent words. Hare et al. extend these relatively uncontroversial findings to morphologically complex words in two experiments which assess processing time by having subjects write sentences in which particular verbs occur, and then by having subjects perform a lexical decision task on a past-tense verb after it has been primed by its base form (e.g., having been exposed to the sentence 'Let's run to the store', subjects were timed in deciding whether or not *ran* is an English word). They show that not only are irregular verbs subject to frequency effects which ease their access, but regular verbs are as well, corroborating the usage-based account's assertion that the units of use in language are also the units of storage: morphemes which co-occur

regularly come to be stored together, regardless of whether or not their co-occurrence can be predicted a by a regular rule (Bybee 1985, 1995).

The third of the three key claims regarding frequency with which this study will be primarily concerned is that the frequency of a pattern has a direct effect on its productivity. Studies such as Dabrowska & Szczerbinski (2006), which examined the effect of several variables on the productivity of genitive, dative, and accusative inflections of varying stages of diachronic development, Wang & Derwing (1994), which examined how speakers of English formed the past tense of nonce ablaut verbs, and others (Baayen & Lieber 1991, Moder 1992) have all lent support to the prediction that type frequency has the effect of strengthening a pattern/schema, a major factor in causing the pattern to be more likely to be applied to new items.

The experiments reported in Chapter 5 are based directly on experimental methods that have demonstrated the role of frequency in affecting how language is stored and processed, indicating that, at the level of metaphor as at other levels of linguistic structure, the frequency of a schema causes instantiating tokens to be more acceptable and accessible to speakers, and the schema itself more productive.

1.1.5 The Network/Exemplar Model

Exemplar Theory is one of a number of theories which have emerged out of psychology in the last three decades which take as central the view that categories are not discrete entities, comprising sets defined by criteria of membership, but rather are structured around a 'core' at which resides the best example(s) of the category. Exemplar Theory (Brooks 1978, Estes 1986, Hintzman 1986, Medin & Edelson 1988, Nosofsky 1986) differs most notably from prototype theory (Berlin & Kay 1969, Dirven & Taylor 1988,

Lakoff 1987b, Rosch 1973), with which it nonetheless shares many of its central assumptions, in that while prototypically defined categories are structured around a single central member which may or may not be an actual, experienced instantiation of the category, categories in the exemplar model are "cloud[s] of remembered tokens of that category" (Pierrehumbert 2001: 140). All members of a given category correspond to experienced events, and the group of remembered tokens which has the highest frequency, and is therefore strongest (with a corresponding effect on productivity), is at the core of the category.

The theory captures, then, many of the predictions of prototype theory while providing more explicitly for the role of frequency in mental storage. Graded category structure is essential to the representation of categories: language users' mental representations of phonetic units are considerably more fine-grained than would be predicted by classical or generative views of categories. Categories are structured around central members, comprising remembered tokens of use, rather an abstracted prototype. Multiple best examples exist, falling within a set of parameter values upon which context has an affect. The core of a category, functionally analogous to a prototype, consists of those items which have been most reinforced by frequency.

An exemplar model departs from a prototype model in that "all perceived tokens are categorized and stored, creating categories that directly represent the variation encountered" (Pierrehumbert 2001, p.51), The gradient nature, then of linguistic categories, such that a particular segment can have many possible phonetic expressions, and the internal structure of categories such that certain ranges of parameters correspond to what individuals deem 'best' members of a category, aren't incidental features of

language. They are, rather, natural results of the way in which language is experienced, perceived, and learned.

In the application of exemplar theory to language processing and storage (Pierrehumbert 2001, Croft 2007), categories emerge from the repetition of units in the continuous stream of linguistic data to which we are exposed in daily life—segments, morphemes, words, and constructions (as well as, it's argued here, metaphors). For any given category, those units which are most frequently experienced become strengthened, with a concomitant increase in productivity for schemata they instantiate, while infrequently experienced tokens are correspondingly weak.

Several types of evidence have emerged for exemplar representation of the phonetic shapes of words. Miller (1994) reports a series of experiments that attest to the internal complexity of phonetic categories. In contrast to studies on categorical perception, which focus on subjects' ability to make distinctions between categories, Millers' experiments focus on participants' 'goodness judgments' for specific instantiations of phonetic categories corresponding to phonemes. A core experiment in the series, for example, presented subjects with syllables containing instantiations of [p], asking subjects to rate the goodness of each instantiation as an example of /p/. She finds that not only are participants able to note differences between stimuli within a given category, they exhibit clear and consistent preferences, choosing certain stimuli as more acceptable than others based on an acoustic parameter, or on a set of acoustic parameters, relevant to the category in question (for example, subjects made clear and consistent goodness ratings for [æ] along the parameter of vowel duration). Millers' work demonstrates that graded structure is essential to the representation of phonetic categories, indicating that language

users' mental representations of phonetic units are considerably more fine-grained than would be predicted by generative theories. Her work demonstrates, as well, that categories are structured around central members, with those examples that participants label as 'best' falling within a narrow window of parameter values, and points as well to the existence of multiple 'best' examples, with the choice dependent on context—for example, rate of speech affects participants' judgments as to at what point, along the parameter of voice onset timing, the best examples of [b] and [p] are located. As the author notes, however, these findings can be explained with either prototypes or exemplars.

Bybee (2001) notes several forms of evidence for exemplar representations in the phonetic shapes of words. An exemplar model departs from a prototype model in that "all perceived tokens are categorized and stored, creating categories that directly represent the variation encountered" (2001: 51), and several experimental approaches have directly supported this claim. K. Johnson (1997), presenting the findings that individuals are able not only to identify words and utterances when they are uttered by speakers who have notably different vocalizations, but also to identify the individual voices of these speakers, argues for the storage of tokens of use, containing information at various levels, as opposed to storage only of an abstracted category or prototype. Both Palmeri et al. (1993) and Schacter & Church (1992) indicate that memories for individual tokens of words are notably strong, as well as persistent across time. All of these approaches indicate sufficient storage capacity for individual tokens of use (in particular, for these studies, for words), as well as that such tokens contain information (including phonetic detail) at various levels: they are not simply 'filed' according to an abstracted category or

prototype, but remembered in full detail. Finally, Pierrehumbert (2001), citing studies such as Bybee (2000) which point to the direct effect of word frequency on phonological variables such as deletion, presents the view that the phonetic detail associated with a word doesn't arise as a result of phonological rules in articulation. If 'rules' are being applied, with regularity, more frequently to some words than to others, even when the phonetic environment is the same, then the phonetic forms of words would seem to be stored along with the word, not derived from rules.

The exemplar model deals with higher order abstractions such as segments and features as emergent properties of the system. A generative methodology takes what might be referred to as a 'top down' approach, according to which abstractions such as segments and features are fundamental properties of the system, the organizing units which individual, experienced tokens of use merely instantiate. An exemplar model, on the other hand, adopts a 'bottom up' approach: at the core of language are tokens of use, with language as it exists in the world (and therefore, as it is experienced by individuals) directly bringing about linguistic structure. In phonology, units such as segments, syllables, and features "emerge from the inherent nature of the organization of gestures for articulation" (Bybee 2001: 85). In production, language can be reduced to an (in principle) finite number of articulatory gestures, specific movements of the articulators in speech. Ohala (1992) describes how the co-occurrence of such gestures, causing an increase in acoustic effect, can lead to salience. Places in the stream of speech where gestures co-occur in high numbers correspond to consonants, places where gestures co-occur in low numbers, vowels. The patterning of these units leads to syllable structure,

and as repeating units corresponding to specific co-occurrences of gestures emerge, so too do segments.

The exemplar model provides the mechanism by which emergence takes place: every token of use is stored, its proximity to other tokens determined by its similarity to them. The strength of an exemplar increases as frequency increases. It is, therefore, those units which tend to repeat themselves in speech which get stored most effectively, and units of different sizes (features, segments, syllables, words) emerge as a result of redundant storage: both relatively smaller and relatively larger units repeat themselves, with smaller units occurring within larger ones. The gradient nature, then of linguistic categories, such that a particular segment can have many possible phonetic expressions, and the internal structure of categories such that certain ranges of parameters correspond to what individuals deem 'best' members of a category, aren't incidental features of language. They are, rather, natural results of the way in which language is experienced, perceived, and learned.

1.1.6 Synthesis: Metaphor and frequency

One of the central assumptions of the field of contemporary metaphor research is that individual metaphors can exist anywhere along a continuum which runs between highly innovative metaphors (metaphors which set up a new, and generally 'one-shot' mapping between a source and target domain, e.g., *he's like really expensive paint: he covers everything*) at one end, and highly conventional ones (e.g., *he's a lion*, or *let the cat out of the bag*) at the other. There are two ways in which a metaphor can be conventional.

First, a metaphor can be conventional at the level of the underlying conceptual mapping. It is this type of conventionalization on which CMT has focused, elaborating

families of linguistic metaphors which all instantiate the same underlying (conceptual) metaphors. Lakoff & Johnson (1980) offer *he attacked the weakness of my argument*, *her arguments were right on target*, and *I retreated before her verbal onslaught* as examples of an underlying metaphor, ARGUMENT IS WAR. The metaphor is conventional not because any of its linguistic instantiations are especially frequent, but because, Lakoff & Johnson argue, of the conventionalization of the metaphor itself: for speakers of English, the terminology of warfare is a default mode for referring to argumentation. The authors assert that hearers are able to process such utterances with little or no effort relative to literal speech, which attests to the highly conventional nature of the metaphor.

The second way in which a metaphor can be conventional is at the level of its surface expression. This type of conventionalization has been the focus of inquiry into metaphorically motivated idioms (expressions the meaning of which are not wholly predictable based on knowledge of the individual words which comprise them) such as *let the cat out of the bag* or *spill the beans*. In the case of these examples, there is nothing especially conventional about the source-domain mappings whereby secrets are understood (respectively) as cats or as beans. The utterances, themselves, though, are highly conventionalized, belonging to the canon of set expressions in the language (as evidenced by their lexical fixedness—precluding, for example, replacing *let* with *allow* or *beans* with *peas*).

It is argued here that conventionalization at the level of the underlying metaphor is the result of type frequency, with any two linguistic metaphors which are predicated on the same cross-domain mapping being of the same type (see also Clausner & Croft 1997). Conventionalization at the level of the surface expression corresponds to token

frequency, a token being a linguistic expression which is predicated on a metaphor. This assertion opens up for review a body of literature which has, while not directly addressing frequency, made explicit claims about metaphorical conventionalization and entrenchment.

Some of the most concrete claims made to date regarding how previous experience with metaphor directly affects the way that metaphor is stored and processed have come from psychology. It has been repeatedly demonstrated that metaphorical idioms (such as 'he spilled the beans') and so-called 'formulaic metaphors' (such as 'the surgeon is a butcher', for which there is a standard interpretation) are processed more rapidly than innovative metaphors (such 'the surgeon is a woodworker', for which there is no standard interpretation), indicating that repeated exposure streamlines processing (Hoffman 1984)- a basic, domain-general frequency effect. The existence of a standard interpretation suggests repeated prior exposure. Blank (1988), in a study comparing processing speeds for literal vs. metaphorical speech, compared reaction times for words that completed sentences with either a literal or metaphorical meaning. The study found longer processing speeds for metaphorical than literal speech, but also that more familiar metaphors are processed more quickly than less familiar ones. Gentner & Wolff (1997) demonstrate that comprehension time for novel metaphors is decreased equally by the priming of either the source or target term, whereas for conventional metaphors priming for the source term speeds comprehension more than priming for the target. For the conventional metaphor *a job is a jail*, for example, prior exposure to either *jail* or *job* speeds comprehension for the entire metaphor. For *Johnny is a rocket*, on the other hand, previous exposure to *rocket*, with its highly conventionalized figurative meaning

referring to speed, primed the metaphor as a whole far more effectively than *Johnny.* They argue that these results indicate that metaphorical conventionalization arises as abstractions from source terms are repeatedly projected to target-domain terms, resulting in source-domain terms taking on easily accessible metaphorical meanings.

An ongoing debate among metaphor researchers has been whether metaphors are processed as comparisons, with speakers drawing connections between the source and target in the same way done for comparative statements such as *whales are like fish*, or whether metaphors are processed as categories, with speakers processing metaphors by revising their conception of the source as a category which incorporates the target, in the same way that is done for a statement like *Cheerleaders are athletes*. In Bowdle & Gentner (2005), the claim is advanced that, as a metaphor becomes conventionalized, it undergoes a processing shift such that while novel metaphors (to Bowdle & Gentner, metaphors in which a speaker is using a word in such a way as to innovate a new cross-domain connection) are processed as comparisons, increasing entrenchment is accompanied by a shift towards the processing of metaphors as categorization statements.

Conventionalization, for Bowdle and Gentner, refers strictly to the source term in a metaphor, which, they assert, takes on a domain-general meaning, prompting processing via categorization, over time. Their research is limited in its applicability to the current study by its focus on single metaphors rather than on families of metaphors, and on source terms as the sole carriers of conventionalized metaphorical meaning. Their experimental findings, however, that the more conventionalized the metaphorical meaning of a source term, the more subjects tend to prefer categorization-type

processing, is directly relevant to, and will be explored within the context of, the view of metaphor advanced here.

Another highly relevant debate within metaphor research has concerned whether metaphorically motivated idioms are in fact processed as metaphors, or whether the role of metaphor is purely historical in acting on idiomatic meanings, with idioms accessed on a lexical level. Gibbs and colleagues (Gibbs 1994, Gibbs et al. 1997, Gibbs & O'Brien 1990), based in large part on a methodology of analyzing experimental subjects' mental images for idioms, spearheaded the movement for understanding figurative idioms as being underpinned by metaphors which, while highly entrenched, are active in on-line processing. The traditional view to which such research is a reaction (Aitchision 1987, Cooper 1986, Cruse 1986, Strassler 1982) is that metaphorical idioms are not processed metaphorically, but are rather linked to stored meanings which are retrieved wholesale from the lexicon (so that, for example, the meaning of the string 'let the cat out of the bag' is simply stored as 'reveal a secret' rather than being motivated by an understanding of cats as secrets and bags as minds). The metaphor-based view of idioms has found experimental support more recently in work such as Keysar et al. (2000), which gauges whether or not metaphors are activated in on-line processing of idioms by seeking a priming effect from the metaphor, and by Sanford 2008b, which finds a phonological reduction effect on the main verbs in figurative idioms due to a predictability which follows from their metaphorical underpinnings. Chapter 3 addresses the metaphorical status of idioms, arguing that idioms are instances of metaphor which have become largely independent of the sanctioning schema due to their high token frequency (for example, the relatively high frequency of 'let the cat out of the bag' leads to a degree of

autonomy from the sanctioning metaphor whereby secrets are understood as entities being let out of containers).

Allbritton, McKoon, & Gerrig (1995: 614) make a critical step towards a frequency-based view of metaphor in suggesting that "conceptual metaphors can provide a schema-like structure for organizing information about a topic." Building on Gibbs and associates' research on the active role of metaphor in motivating the meaning of common idioms, they report a series of experiments in which "word and sentence recognition priming" were used to "assess the degree to which elements of a text representation were associated with one another in memory" (p. 613). Associations between items related by a metaphorical schema (and otherwise semantically unrelated) caused subjects to recognize words and sentences faster when the metaphorical schema motivating their figurative meaning had been previously activated. Thus, for example, "the sentence 'Public officials desperately searched for a cure' would be interpreted as being related to the *Crime is a disease* schema when presented after sentences about an increase in crime … but not if it followed a sentence about police officers contracting pneumonia" (p. 613). According to this line of research, pre-existing metaphorical schemata are activated by utterances instantiating the schema, and activation of the schema spreads across the semantic domains related by the schema. The operation of priming effects across domains that are semantically related only by a metaphorical mapping (for example, crime and disease) provides dramatic evidence for schemata that operate across conceptual domains.

Clausner & Croft (1997), developing the idea of metaphorical cross-domain mappings as schemata, assert that schematicity, defined by Langacker (1987) as the

extent to which precision and detail are characterized for a given schema, is an essential parameter for understanding a given metaphor, and prerequisite to assessing said metaphor's productivity. In their application of schemata to metaphor, Clausner & Croft define a metaphor's degree of schematicity as "The range of concepts characterized by a domain mapping schema" (p. 257). The proper statement, then, of the schema which sanctions a given family of linguistic metaphors will be maximally productive, while being as specific as possible. Only once schematicity has been assessed can a metaphor's productivity—"the proportion of a schema's range which can be instantiated as expressions" (p.257), and effectively the range of expressions that can be licensed by the metaphor (i.e., its type frequency, as per Bybee 1985)—be addressed, a claim in keeping with Bybee's (1995) view of a schema's productivity as dependent on its defining properties and strength. Individual metaphors can exist anywhere along a continuum of productivity, marked at one end by Lakovian conceptual metaphors and at the other by opaque idioms. Clausner & Croft (1997) offer the examples of *this argument is sound*, instantiating a conventional evaluation of arguments in terms of structural soundness, and the opaque *kick the bucket*, respectively). The gradient productivity of metaphors is offered as the primary form of evidence for conceptual schema, and for metaphorical cross-domain mappings as generalizations arising over metaphorical utterances.

Another perspective on the productivity of metaphors comes from Svanlund (2007), who breaks from CMT in asserting that the conventionality of linguistic metaphors is properly attributed not to cross-domain mappings, but to words themselves. Lexical metaphors are metaphors for which a single word carries the metaphorical meaning. The conventionalization of metaphorical meanings for individual words commonly leads to

metaphorically motivated polysemy. Deignan (1999a), for example, gives the example of *deep*. Svanlund asserts that metaphors, and in particular lexical metaphors, vary in their strength, that is, their ability to evoke concepts from the source domain.

While a metaphor's strength is not the same as its degree of conventionalization, he asserts that a metaphor's strength is itself a conventionalized property, attached to individual words at the lexical level. Stressing the social nature of metaphor (as conventionalized by social behavior) rather than, as CMT does, the bodily, emergent nature of metaphor (as arising, directly or indirectly, out of embodied experience), Svanlund gives a metaphor's degree of conventionalization as a combination of how widespread it is in a community, and how deeply entrenched it is within an individual's mind within the community. Both are assessed using corpus methods, the latter using a technique of collocational analysis which assumes that a high degree of collocation between a metaphorical term and other terms from the same source domain—for example, *weigh* with *scale*—corresponds to high degree of activation for the metaphor. Adding a new layer to the notion of productivity, Svanlund acknowledges that a proper understanding of a metaphor's productivity involves both an understanding of the range of concepts from the source domain that are conventionally mapped to the target, and the number of linguistic expressions that denote these concepts. He adds, however, that since words vary in the extent to which they activate concepts from the source domain, this is itself another important factor in a given metaphor's productivity. The productivity of a mapping does not itself dictate a given word's lexical strength, because different words sanctioned by the same metaphor can differ in their metaphorical strength.

Just as metaphor evidences language's inseparability from more general cognitive processes—in this case, our ability to draw connections between disparate domains—metaphor is itself subject to more general cognitive processes such as frequency effects. The effect of frequency on shaping linguistic structure has the potential to provide fundamental insight into metaphorical processing and the careers of metaphors if it can be demonstrated that metaphorical representations are entities upon which frequency effects operate. Such an approach promises, moreover, both to corroborate and provide unity for various lines of research that have converged on the conventionality of certain metaphors, and the effect of such conventionality on processing.

1.2 Conclusion

Bybee (2001: 6) writes that

> Experience affects representation. The use of forms and patterns both in production and perception affects their representation in memory. High- [token] frequency words and phrases have stronger representations in the sense that they are more easily accessed and less likely to undergo analogical change. Low- [token] frequency words are more difficult to access and may even become so weak as to be forgotten. The lexical strength of words may change as they are used more or less in different contexts. Patterns (represented as schemata…) that apply to more items are also stronger and more accessible, and thus more productive than those applying to fewer items.

This dissertation asserts that metaphors are such patterns, and as such are subject to frequency effects. These effects operate on metaphor both at a conceptual level, pertaining to the source-target mapping itself, and at the level of linguistic realization, pertaining to individual metaphorically predicated utterances. As claims relating to

source-target mappings are stronger and have less precedence in the literature, this dissertation focuses on providing experimental evidence for frequency effects at this level. Chapter 2 outlines the usage-based approach to metaphor. While it reviews findings on the entrenchment of metaphorical meanings for idioms, lexical metaphors, and formulaic metaphors, arguing that each of these is a case of metaphorical meaning being crystallized due to high token frequency, the focus of Chapter 2 is metaphorical mappings at the conceptual level. Chapter 3 addresses the conventionalization of metaphorical meaning in idiomatic expressions, and the interaction of frequency effects at conceptual and syntactic levels. Chapter 4 addresses the issue of assessing the frequency of metaphorical mappings, presenting two corpus-based approached to doing so. Chapter 5 presents a series of experiments that assess the claims made in Chapter 2 relating to the entrenchment of metaphorical schemata overall, testing the hypothesis that the acceptability, accessibility, and productivity of metaphorical utterances are all, to at least some extent, determined by the frequency of the schema which they instantiate. An area of particular inquiry throughout will be the interaction of frequency effects at the level of conceptual mappings with effects at the level of utterances themselves. It is asserted here that a usage-based view of language, and the tools of an approach whereby language processing and storage are seen as driven by frequency effects, provide the best lens for understanding the properties of metaphor in all of its types.

Chapter 2: Outline of the Theory

2.0 Introduction

Theories of metaphor which have arisen out of both cognitive linguistics and psychology have been highly successful in describing important properties of metaphor and metaphorical systems with respect to conventionalization. It's striking, however, that they've largely failed to describe the mechanisms motivating such properties. *Why* does metaphorical meaning tend to become conventionalized over time? Why do some metaphors seem to fade over time, so that speakers of a language come to think of metaphorically motivated senses of a word such as *illuminate* merely as additional literal senses? Why do idioms seem to 'lose touch' with the metaphors that underlie them (George Orwell, for example, in the 1946 essay *Politics and the English Language*, famously pointed to expressions to such as *ring the changes on, take up the cudgel for,* and *toe the line as* expressions which once vividly evoked a mental image, but are no longer capable of 'evocative power')? Why do metaphors vary in their 'strength': their ability to evoke concepts from the relevant source domain? While many theories of metaphor have described these properties, a vanishingly small subset attempt to *account* for them.

The gap follows from the generative assumption that underlying structures, with self-evident and self-justifying properties, account for the properties of linguistic utterances. The field of metaphor research, while developing out of the cognitive tradition, has been largely constrained by a generative approach to thinking about how rules license expressions. As understood within Conceptual Metaphor Theory (CMT) in particular, the connection between underlying metaphors and specific, metaphorically predicated

utterances equates precisely to the connection between competence and performance (Chomsky 1965, 1980): underlying metaphors are 'deep' structures from which the surface-level properties of metaphor are derived. A primary goal within such research (Reddy 1979, Lakoff & Johnson 1980, Lakoff & Kovecses 1987) is to enumerate the underlying rules (metaphors) that account for metaphorical utterances that seem like natural sentences to language researchers, whether or not they are attested in natural discourse (to take one favorite example of CMT, *Her arguments were right on target*, as well as any simple variation thereof with respect to pronoun, number and tense, is unattested in the 365+-million-word Corpus of Contemporary American English {Davies 2008}). An underlying metaphor, once posited, is ascribed a sort of primacy, such that a major aspect of inquiry into metaphorical systems is head-scratching over areas where speakers' intuitions depart from the posited metaphor (why, for example, if THEORIES ARE BUILDINGS, is 'His ideas have a strong foundation' acceptable, but 'His ideas have many windows' semantically odd?). Generative approaches stress the unfathomable number of sentences and meanings that language is capable of expressing. Corpus methods, however, have given linguists insight into the fact that humans don't say anything and everything: they repeat the same things, the same chunks and phrases and constructions and sequences, over and over again (Renouf & Sinclair 1991, Renouf 1992, Erman & Warren 2000, Wray & Perkins 2000). The same is true for metaphor. When corpus methods are brought to bear, the striking feature of metaphor is not the productivity of conceptual metaphors—far from it. Rather, the same words and expressions, with the same figurative meanings, are repeated over and over (Sanford 2008a).

While the approach has yielded countless extremely valuable insights into the nature of metaphorical systems that can and should inform subsequent iterations of metaphor theory, such insights have tended towards cataloguing the properties of metaphor, rather than explaining them. The properties of metaphor and of metaphorical systems are best understood as emergent phenomena, products of fundamental aspects of human cognition. Individual metaphors, and their roles within groups of related metaphors, are best understood as entities that arise out of language in use. And the properties of metaphor can be not only enumerated, but accounted for, when the facts of language use, rather than underlying structures, are viewed as basic.

Within this chapter, the reader is presented with the outline for a usage-based view of metaphor, and the argument is presented that this theory accounts directly for key properties of metaphor with respect to conventionalization. Other properties of metaphor, including 'families' of metaphors, are not only compatible with this view, but can be viewed as emergent phenomena, arising out of a view of metaphorical structure as deriving from usage. This chapter will focus on conceptual, innovative, formulaic, and lexical metaphors; idiomaticity and idiomatic metaphors are briefly addressed here (in §2.3.4) but will be given a fuller treatment in Chapter 3. §2.1 lays out a form-based typology of metaphor, while §2.2 and §2.3 outline, respectively, the usage-based approach, and how the approach accounts for key features of metaphor.

2.1 A Typology of Metaphor

A staggeringly wide variety of utterances are sanctioned by the term *metaphor*. The typology outlined below (in which metaphors, wholly for the purpose of approaching the relevant phenomena in an orderly fashion, are divided into lexical, idiomatic, formulaic,

conceptual, and novel members of the category) reflects a useful division in terms of how the literature on metaphor has tended to divide up the phenomena at hand into approachable areas of inquiry.

The typology outlined here is a form-based typology. This is in contrast to typologies based on the nature of the relationship between the source and target domain. A prime example of the latter is Kovecses' (2002) division between structural metaphors (in which target domains are strictly structured according to source domains), ontological metaphors (which attribute to target concepts little more than status as a thing that has discrete existence), and orientational metaphors, which make use of basic, embodied concepts such as up/down to provide coherence to target domains. Typologies such as these are complementary with form-based typologies like the one offered here, such that we might discuss, for example, orientational lexical metaphors (e.g., the conventionalized use of 'higher' to mean 'numerically greater'), or structurally predicated idioms (e.g., 'time flies when you're having fun,' which depends on TIME IS MOTION for much of its meaning). The typology outlined here is also quite different from the typology offered in Sanford 2008a, which is based on the type of target to which attributes, via the metaphor, are being applied, and which again is complementary at every level with a form-based typology such as the one presented here.

Absolutely no claim is made as to the naturalness of the categories presented here; formulaic, lexical, and idiomatic metaphors are, in particular, wholly impossible to separate on any cognitive or functional criteria. This typology is presented in order to align this project with terminology used in the extant literature, to outline how different issues tend to arise with metaphorical utterances of different syntactic types, and to

structure a discussion of issues with respect to metaphor at varying levels of conventionalization and entrenchment. Issues related to the dividing line (if, indeed, any can be said to exist) between metaphorical and non-metaphorical utterances are dealt with in §s 2.3.6 and 4.2.2.

Examples used in this section are taken from either the Corpus of Contemporary American English (COCA) (Davies 2008), with the citation providing the year and file name, or from Parts I and II of the Santa Barbara Corpus of Spoken American English (SBCSAE) (DuBois 2000, DuBois et al. 2003), in which case the citation provides a transcript and line number. All examples not accompanied by a corpus citation are contrived, unless otherwise indicated.

2.1.1 Lexical Metaphor

Lexical Metaphor refers to metaphorical meanings which are carried by individual words. Typically, this metaphorical meaning is highly conventional for the word in question. The following sentences all contain examples of lexical metaphor, the lexical metaphor itself appearing in italics[2]:

(1) Whoa- that's *deep*.

COCA 2003, 'Crime Spree'

(2) She's *hot*.

COCA 2007, 'Brooklyn Bar Serves Opera on Tap'

(3) Who could doubt a *sweet* little old lady?

COCA 1996, 'The Deep End of the Ocean'

[2] In example 1, the speaker is responding to the preceding line, "Although a cartoon I feel he shares a universal theme: We are all searching for love. No?" *Deep* is in the sense of 'philosophical', or 'emotionally truthful'.

Lexical metaphor, based on the notion of the metaphorical use of individual words, is a somewhat artificial category, from a perspective on language that admits units larger than the word into the lexicon (which is done by most any model, if grudgingly, at least in the case of certain idioms): what is the special status of words, when meaning is directly attached to linguistic form at levels both above and below that of individual words? Gray areas emerge quickly in other ways, as well: are there one, or two lexical metaphors contained in the sentence, uttered of an attractive woman, *She's painfully hot*? In *shoulder the burden*, is *shoulder* a lexical metaphor, or just a portion of an idiom?

The distinction, if theoretically fuzzy, has been extremely useful for corpus work on metaphor, as individual words are so eminently searchable. Studies that look at all occurrences of a particular word in a corpus (see Deignan 1999a on *deep*, Tissari 2001 on *love*, and Musolff 2004 on *heart*, for example) are able to look at the relationship between metaphorical and literal meanings of a word, as well as at the overall distribution of metaphorical uses. Like drilling a core out of a glacier, the approach can easily miss something important, but it provides a precise and focused insight into a metaphorical system.

2.1.2 Idiom

The most commonly accepted definition of an idiom is that it's an expression with a set meaning, that doesn't follow the rules of compositionality: knowing the meaning of all the elements in it, and knowing the rules of how those elements combine, won't get you the overall meaning of the expression. Core examples of idiom include expressions such as the following (italics added):

(4) All right, so Jack had *jumped the gun* a little, buying this thing.
COCA 2008, 'Just Breathe'

(5) Each time you think you've got one thing figured out, they *throw you a curveball*.
COCA 2002, 'Rebranding the Hyena'

(6) *It's raining cats and dogs* out there, tonight.
COCA 2007, 'Guy in the Sky'

(7) But *that's neither here nor there*.
COCA 2009, 'Evie Ever After'

(8) …she was going to *blow the whistle* on him to the state medical board.
COCA 2006, 'Kill all Lawyers'

For all of the examples above, a speaker of English who had been previously exposed to all of the constituent words, but never to the overall expression, would interpret the utterance literally- and, therefore, incorrectly.

Nunberg, Sag, & Wasow (1994) make the critical observation that idioms are a prototypically defined category: there is no criterion or set of criteria by which a given utterance can be included or excluded from the category (see also Erman & Warren 2000, Wray & Perkins 2000). Idiomaticity, rather, is a continuous variable, measured along, they propose, six parameters, no single one of which need necessarily be satisfied in order for a given utterance to be considered idiomatic.

1) Conventionality:

   The overall meaning of the expression is 'set'. The gestalt meaning of the expression is not, or at least not entirely, formed by combining the constituent syntactic and semantic elements. There is, rather, a conventional interpretation for the expression as a whole. Examples (4) through (8) above all exemplify a high degree of conventionality; items low in conventionality are almost unequivocally non-idiomatic.

2) Inflexibility:

An idiomatic expression tends to bear, with its meaning intact, little syntactic alteration. Consider the alternations in the following idiomatic expressions:

(9) He *let the cat out of the bag*.
COCA 1995, 'Discussion of a politically correct comic book on The Lone Ranger and Tonto'

(9a) The cat was let out of the bag by him.

(9b) He let the cats out of the bag.

(9c) He let the dog out of the bag.

(10) …the West Germans *took the bull by the horns*, changed the currency and stabilized the economy.
COCA 1990, 'Economic Forecasts'

(10a) The West Germans will have taken the bull by the horns.

(10b) The West Germans took the elk by the antlers.

(11) People seem to be *having second thoughts*.
COCA 2008, '(NEWS BREAK) # GIGOT: We're back with more on the Illinois corruption scandal'

(11a) People seem to be owning second thoughts.

(11b) People seem to be having third thoughts.

(12) She *kicked the bucket*.
COCA 1997, 'The Barn'

(12a) She kicked the buckets.

(12b) The bucket was kicked by her.

In each set, changes such as passivization, word substitution, or number/tense alternation leave an utterance strange-sounding, non-sensical, or simply non-idiomatic.

Idiomatic expressions are, in fact, *so* fixed that they stay the same even as the language changes around them. Thus *dint*, and *chink*, despite having largely dropped out of everyday use for speakers of English, remain in the expressions

Chink in one's armor and *by dint of*. The issue plays out in interesting ways with respect to figuration- if the literal meaning of *muster*, for example, is lost to the typical speaker of English, then *muster up the courage* can't be considered metaphorical- a metaphor, after all, needs a literal basis.

3) Figuration

   There is some figurative connection between the expression's literal and figurative meanings. This connection could be metaphorical (as in *Take the bull by the horns* or *she's fishing for information*), but it need not be: *She has me pulling my hair out* and *I'm scared stiff*, for example, depend on hyperbole, rather than metaphor.

4) Proverbiality

   Idioms tend to describe some salient aspect of life, a recurrent situation of particular social interest: for example, telling a secret (*spill the beans*, *let the cat out of the bag*) or getting married (*tie the knot*, *get hitched*)[3]. Most interestingly, idioms inherently reflect and shape our conceptualizations of these situations. They also, inherently, explain these situations, either shaping or reflecting (or both) how we view them: beans, once spilled, aren't easy to get back into a

[3] For this reason, idioms are commonly used for purposes of euphemism and dysphemism, which provide agreeable and disagreeable (respectively) ways for talking around topics of recurring social interest which are also taboo: sexuality, for example, or bodily functions.

container. When one gets hitched, one loses ones power of independent movement.

5) Informality

   Generally, idioms have a high association with informal speech registers. Nunberg, Sag & Wasow assert that one is likelier to say something like *talk about beating a dead horse* in causal speech with friends than, say, in a job interview, or in speaking to one's partner's parents.

6) Affect

   Affect ties in with proverbiality, and in particular with how social acts get conceptualized via idioms: an idiom generally takes a certain evaluative stance towards the thing that it's describing. Nunberg, Sag, & Wasow assert that those things we feel relatively neutral about, we'll usually describe literally- buying tickets, or reading a book, for example. Situations towards which we have strong feelings (getting married, lying, telling secrets, exerting control, scenarios involving authority, etc.) tend to be those things that we prefer to use idiom to describe.

A certain expression might be not at all proverbial or informal, but because it has a high degree of both conventionality and inflexibility, it remains a good example of an idiom. If all of these parameters apply to a great degree, then it's a prototypical idiom. The fewer that apply, and the lesser the degree to which each applies, the more peripheral it is (for example, *tax and spend* and *right to life*, which are high in conventionality and inflexibility but low in figuration, or *render unto Caesar* and *Procrustean bed*, which are high in proverbiality and affect, but also high in formality).

Nunberg, Sag & Wasow point out that prototypical idioms also tend to involve transitive verbs, generally taking the form Verb + argument (thus, e.g., *spill the beans*, *let the cat out of the bag*, *crack the whip*, etc.). The authors note that research on idioms tends to focus on these to the exclusion of non-phrasal idioms such as *smoking gun*, *red herring*, and *pain in the neck*. These last examples point, as well, to the fact that there is no clear line between idiomatic and lexical metaphors: each of the above functions essentially as a nominal with a set metaphorical meaning. Issues pertaining to the organization of this dissertation aside, there's no reason to expect any particular utterance to fall neatly into one category or another.

2.1.3 Formulaic Metaphor

The term 'formulaic metaphor' is generally used to refer to metaphors that have a highly conventional interpretation, and almost invariably to metaphors of the form '*x* is *y*'. Their overall meaning draws not from the wide range of possibilities for meaning that arise when a source is applied to a target, but rather a single, narrow interpretation. Take the following examples:

(13) Men are dogs.
COCA 2009, 'Aussie Rules'

(14) For a boy she's kind of cute, but for a girl *she's a dog*.
COCA 2000, 'Fiction Crushed'

(15) She's a fox.
COCA 2009, 'Star Tracks'

(16) That surgeon is a butcher.
(Turner & Fauconnier 2002)

Of all of the possible attributes of dogs that the first *could* be being attributed to men in example (13)—loyalty, a keen sense of smell, hairiness—it's strictly a dog's sense of

selfish opportunism that makes the transfer across domains. We cannot attribute this strictly to a single conventionalized metaphorical meaning associated with *dog*: in example (14), an entirely different aspect of dogs—their lack, by human standards, of physical attractiveness—is highlighted. (14) and (15) provide another interesting and surprising contrast: the set of *possible* interpretations for the two utterances overlap to high degree, and yet in practice, the interpretations to which any native speaker of English will leap are precisely opposite. It appears to be the case, for such expressions, that their conventionalized meaning isn't linked strictly to the source domain term, as in lexical metaphors, but rather to the overall mapping.

In many cases, the meaning of such expressions is highly contrary to what might be expected. Turner & Fauconnier (2002) note, of example (16), that while *butcher* is a specialized profession and that a particular butcher (just like a particular surgeon) might fall anywhere on a scale from incompetent to highly skilled, the surgeon in the example is clearly being called incompetent. Turner & Fauconnier argue that such examples show the weakness of Conceptual Metaphor Theory in treating all metaphorical mappings as applying the systematicity of a given source domain to target domain: here, the blend of two domains gives rise to an idiosyncratic interpretation without clear precedent in either of the two contributing domains.

2.1.4 Conceptual Metaphor

Conceptual Metaphor, outlined in §1.1.1, was originally advanced in Lakoff & Johnson 1980. Included here for the sake of terminological clarity, conceptual metaphor does *not* refer to a type of utterance, but rather to an underlying metaphor that sanctions any number of actual utterances. LOVE IS A JOURNEY, for example, is a conceptual metaphor.

Each of the following examples (drawn from Lakoff & Johnson 1980) instantiates the metaphor.

(17) We're just starting out together.

(18) We've decided to go our separate ways.

(19) We've come so far- we can't turn back now.

2.1.5 Novel Metaphor

Metaphors can be innovative in one of two ways: they can explore a new aspect of an existing source-target mapping, or they can posit an altogether novel connection between domains. Both forms are highly valued in the domains of wit and wordplay, as well as in literature (where, indeed, formulaic, lexical, and idiomatic metaphor tend to be denigrated as cliché). Consider the following examples:

(20) He was like a little Australian sheepdog, running around.
SBCSAE Text 6, lines 794-799

(21) Those tables are museums, could you please, chill out in the uh, art work here.
SBCSAE Text 6, lines 523-527

(22) Human nature does not possess free will. It is like a horse. Ridden by God or the Devil. The rider possesses the will. The horse obeys.
SBCSAE Text 25, lines 175-183

Example (20), spoken by a woman who is making fun of her male partner for flirting with other women, is a novel instantiation of a well-established conceptual metaphor: PEOPLE ARE ANIMALS (which, as Kovecses 2002 points out, tends to constrain interpretations towards negative characteristics of the animal used as a source). Here, the specific equation of a man to a 'little Australian sheepdog, running around' suits perfectly the particular, current discourse needs of the speaker: to portray her partner's flirting in a

rather undignified light. Its participation in a well-established metaphor, however, means that it's easily interpretable by other participants in the conversation, needing little set-up.

Examples (21) and (22), on the other hand, are wholly one-shot mappings, constructed on the fly for the purposes for the current conversational topic. (21), spoken by a parent to an overly rambunctious child, represents a clever bit of wordplay on the part of a speaker who uses the first clause to set up the second. (22) is a more extended metaphor, a simile which is then elaborated in several ways.

In some cases, of course, one-shot mappings catch on and become established metaphors. The Oxford English Dictionary gives the first citation of 'virus' in the computing sense, from a 1984 article on computer security, as following (*virus* n. 2d. In *Oxford English Dictionary.* Retrieved November 9, 2009, from http://dictionary.oed.com):

(23)We define a computer 'virus' as a program that can 'infect' other programs by modifying them to include a possibly evolved copy of itself.

From this point (or from whatever the initial coinage was), the metaphor propagated across speech registers and discourse communities to become, eventually, the default mode for speaking of malicious software. Not only are the quotation marks no longer necessary, but the metaphor has in fact become so pervasive and automatic that it's generally used with little if any awareness of its metaphorical status—an extremely common process in language change.

From the perspective of on-line processing, little (if anything) separates a lexical metaphor from an idiom, no boundary separates either from formulaic metaphors, and the difference between novel and conventionalized metaphors is one of degree, not kind. There is an association, however, between each of the categories above, and certain

properties of metaphor, such that a complete theory of metaphor must, in order to account for the phenomena at hand, account for each type of metaphor. The sections that follow will refer repeatedly to the typology provided here, as various aspects of metaphor in language are explained as emerging from a usage-based model.

2.2 Conventionalization, Entrenchment, and Metaphorical Systems

Schemata are defined by Langacker (1987, 1988) as abstractions over semantic, phonological, or otherwise symbolic units, which can in turn sanction specific instantiations of the schema (as, for example, the word *tip* instantiates the phonological CVC schema). Bybee (1995) defines schemata as "emergent generalizations" over "[forms] having similar patterns of semantic and phonological connections" (p. 430). Schemata form over units that tend to both co-occur and re-occur, with items that consistently occur alongside one another emerging, via repetition, as salient units of speech and in turn as the stored units of language.

In the application of a schema-based view to metaphor, the units which co-occur are domains of thought corresponding to Lakovian conceptual domains, which emerge from categories of experience. These units are semantic, as opposed to phonological, morphological, or syntactic, and the co-occurrence is simultaneous rather than sequential (as in the common co-occurrence of *do not* or *going to*). They are no less governed than schemata at other levels, however, by principles whereby frequency affects storage.

Speakers, as they engage in language in any mode, encounter linguistic metaphors. As metaphors are encountered, each individual token of use encountered is stored. Figure 2.2.1 is a representation of the sum of stored linguistic metaphor tokens for a hypothetical

speaker of English who, prior to reading §2.1 of this dissertation, had never before been exposed to metaphor[4].

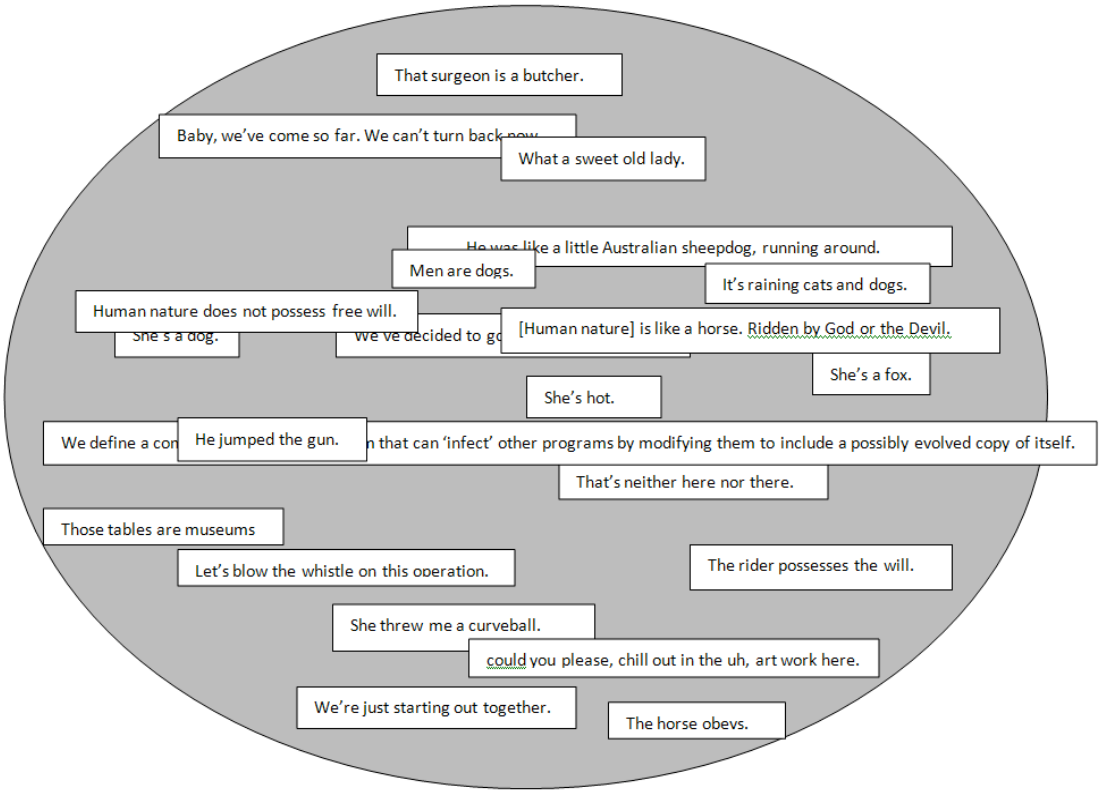

Figure 2.2.1
Linguistic Metaphor Tokens

As tokens accumulate, they are stored on the basis of similarities with other items. Based on similarities between lexical metaphors, then, 'clouds' of stored tokens begin to

[4] Note that Figure 2.2.1 presents each item as being stored separately, with no connections to other items. This is a convenience in demonstrating the process, and should not be taken to imply a processing stage in which token of use are stored, but not yet connected to other items. In practice, each token of use is attached to (and affects) representations at multiple levels as it is encountered.

develop, as similar items are stored in close proximity to one another within the conceptual space. Utterances can be similar to one another in many ways, and it's unclear that any possible parameter can be excluded here. The most relevant types of similarity for metaphor, however, seem to be semantic (based on categorizations of the source domain and categorizations of the target domain) and syntactic (based on the surface form, in terms of constituent elements, of the expression itself). Similarity judgments forming along one parameter in no way exclude categorizations forming along another with the resulting categorizations (and, eventually, schemata) being highly redundant. Figure 2.2.2 indicates categorizations, over the tokens of use encountered in Figure 2.2.1, along semantic similarities. At the same time, as we shall see, categorizations are also made over syntactic similarities (see Figure 2.3.4, below).

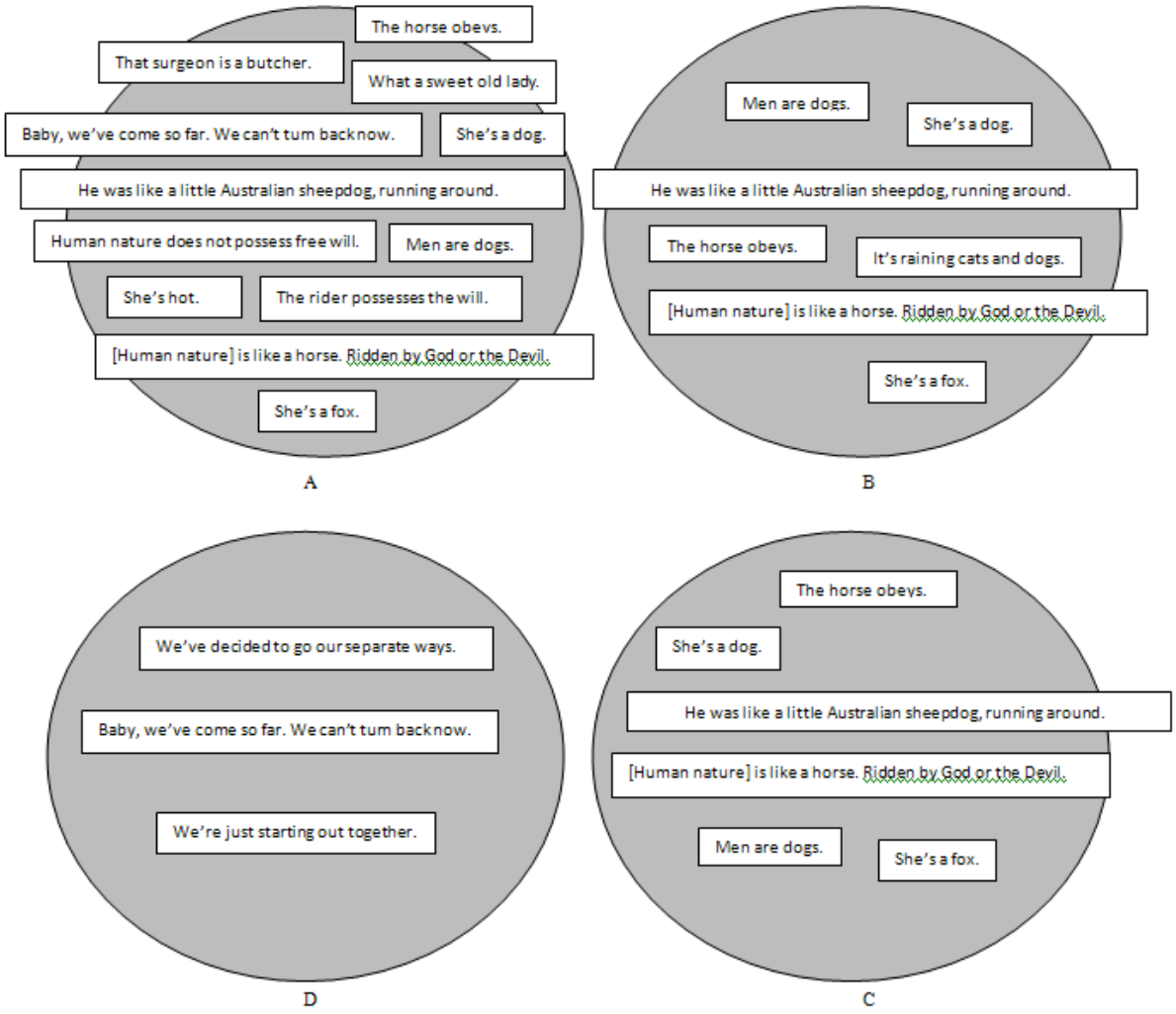

Figure 2.2.2
Categorizations based on Semantic Similarity
A. utterances referring to people
B. utterances based on animals
C. utterances based on animals and referring to people
D. utterances referring to relationships

C, it should be noted, is here effectively the set of items that are common to A and B. This should not be taken to imply a hierarchical organization, such that C is a special case of either A or B, or that in on-line processing C is a function of logical operations upon A and B. Rather, items are added to each category independently (and redundantly).

Schemata form as speakers, within categorizations, make connections over individual tokens of use[5].

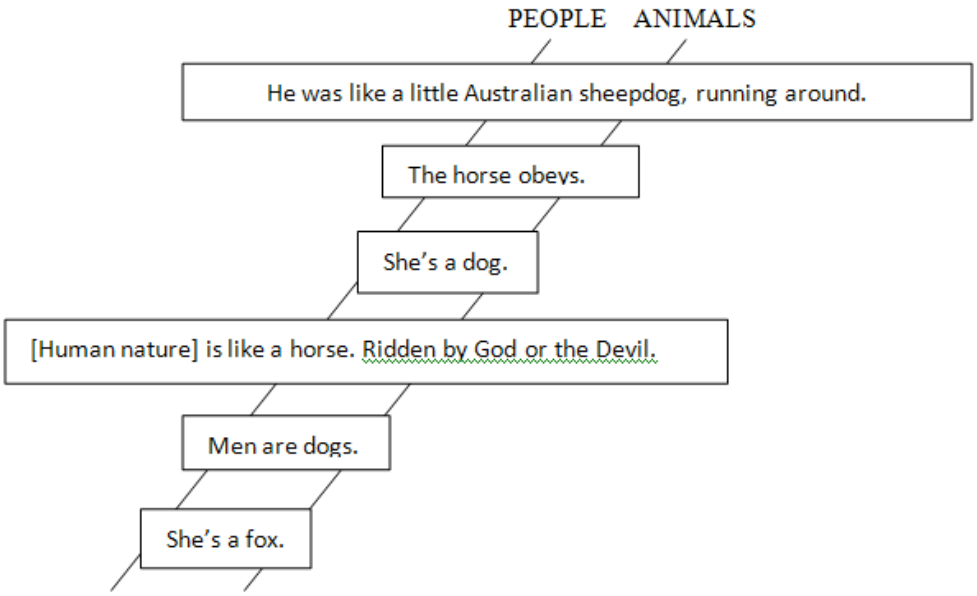

Figure 2.2.3
Conceptual schema: PEOPLE ARE ANIMALS

Figure 2.2.3 shows a schema emerging over the categorization shown in Figure 2.2.2(C), in which humans are being equated to animals. Based on semantic properties in common to all items within the cluster of tokens, the speaker generalizes a pattern[6].

[5] The system for diagramming schemata as connections over tokens of use is borrowed from Bybee (1985, 2001).

[6] The number of distinct tokens necessary for a schema to form is an open empirical question, and the presentation of §2.2 should not be taken as implying that (for any given *x*) *x* tokens accumulate within a category before a schema can be formed. Once a schema *does* emerge, subsequent tokens are linked to the schema.

Tokens are analyzed based on the schema, with specific elements in the clause being equated to one of the two semantic elements[7].

At this stage, a cross-domain domain mapping has effectively formed. Further tokens of utterances equating people to animals will be analyzed according to the schema, each iteration strengthening the schema as its type frequency increases. In proportion to the strength of the schema, new tokens' semantic acceptability will be judged according to their semantic proximity to, as well as the strength of, the prototype, and processed more quickly than items not sanctioned by a schema or sanctioned by a less entrenched schema (Pierrehumbert 1994, Vitevich et al. 1997, Hare et al. 2001, Bybee & Eddington 2006, Wilson 2009). Most critically, the schema becomes a template for creating new utterances. Speakers referring to given topics metaphorically are likely to choose source terms based on the relative strengths of attested metaphorical schemata that involve the target. Within a speech community, other language users will have encountered the pattern at similar levels of frequency, and a metaphorical utterance produced based on a schema deeply entrenched for a speaker is likely to be deemed semantically acceptable, and interpreted with ease, by a listener for whom the schema is similarly entrenched. At the point that such a schema has become entrenched across many speakers, and pervasive within a speech community, a Lakovian Conceptual Metaphor has essentially formed (here, PEOPLE ARE ANIMALS). Critically, however, the direction of causation is precisely

[7] An explicit comparison need not be made on order to serve as a basis for similarity comparison, as it's not words that generalizations form over, but domains. In the *horse obeys*, from example (22), people are described as horses implicitly. A comparison of horses to people is made nonetheless.

the reverse of what is proposed in Conceptual Metaphor Theory: in the account outlined here, conceptual metaphors are accounted for, rather than merely proposed[8]. Instantiations don't just follow from underlying conceptual metaphors, they are, rather, integral to the process by which such mappings arise.

[8] The model outlined here takes the perspective of an individual language user, with a history of use of the language developed through exposure to tokens of use. Schemata at all levels, including conceptual, form as language users make generalizations over recurring patterns. A metaphor can't be property said to exist until a schema such as the one in Figure 2.2.3 has emerged, but in that case, what metaphor motivated *she's a dog* or *the horse obeys* in the act of language use to which our speaker was exposed? Another schema, in another language user's mind. Every speaker recreates the language anew, based on exposure to the language: the only option, unless language is to be treated as an abstract system rather than as a being rooted in the mind of an individual.

It does seem worthwhile, however, to point out that from a diachronic perspective, there is a chicken-or-the-egg question to be addressed: there must have been some original emergence of PEOPLE ARE ANIMALS (or of any other schema). But how it could it arise, when there were no utterances for it to form over? There are two options: the first is that a schema can emerge due to reanalysis of tokens of other schemata. The other is that analogical reasoning can and does create wholly novel metaphorical utterances (although existing schemata will some play a role in determining its acceptability), over which schemata then form.

2.3 Properties of Metaphor

§2.3 looks at key properties of metaphor and of metaphorical systems, accounting for how these properties are accounted for in a usage-based approach.

2.3.1 Idiosyncratic Interpretations, Autonomy

A metaphorical schema, once formed, is not a static entity. Frequency effects are on-going, and metaphorical schemata differ from one another in strength (as the experiments outlined in Chapter 5 will demonstrate) by virtue of speakers' continued exposure to language, and to tokens that instantiate different schemata. Moreover, frequency effects have an on-going effect on the internal complexity of a schema. The frequency of instances of a schema can lead to their entrenching away from the sanctioning schema overall, taking on properties not associated with the more general schema—as Bybee (2001: 125) notes, the "frequency of a form weakens associations with other forms." A particular linguistic form can, by virtue of its high token frequency, become entrenched in its own right, losing connections to other forms sanctioned by a particular schema and acquiring a degree of autonomy (Bybee 1995, Hay & Baayen 2002).

Autonomy goes hand-in-hand with the reassignment of constituent structure that accompanies high-frequency items. Repetition conditions chunking: elements that consistently co-occur develop, over time, constituent structure (Haiman 1994) such that the reoccurring string becomes a unit of use and storage. In a network model, schemata form over connections (semantic, phonetic and syntactic) between utterances, on the basis of shared or similar elements. As a form develops constituency due to the effect of chunking, internal elements contribute less to the overall meaning of the expression, and become less salient. As items internal to a frequent collocation come to participate less

and less in the overall meaning of an expression (accompanied, in many cases, by phonological reduction), high-frequency connections lose the basis upon which connections to similar forms are made (Bybee & Scheibman 1999, Beckner & Bybee 2009). Connections to other instances of the schema fade, a direct processing route forms, and the form takes on properties not consistent with the schema in which it once participated (or once participated in to a greater extent). Bybee & Scheibman (1999) note the effect with the high-frequency collocation *be supposed to*, which has taken on a function and meaning increasingly distinct from other forms containing *supposed.*

We see the effect in an example such as *men are dogs. Dog*, used as a source term applied to men, has a conventionalized meaning of sexual promiscuity and lack of loyalty. This departs notably from what might be expected as an interpretation for the metaphor. For a speaker of American English who participates in the common culturally shared conception of dogs as loyal companions, and who has a fully formed schema for PEOPLE ARE ANIMALS whereby salient traits of animals are applied to humans, the most intuitive interpretation of *men are dogs* would be that men are loyal, steadfast companions: the inverse of the meaning intended here. On the other hand, *she's a dog* has the specific meaning that the woman referred to is physically unattractive. *She's a fox* means the opposite. Even within the tendency for PEOPLE ARE ANIMALS to profile negative aspects of animals (and therefore of people) —a feature of the schema, emerging over tokens such as those cited here —the interpretation for each example is unpredictable. In each case, the frequency of the specific form (*dog* and *fox*) as used in reference to a particular target has caused the form's connection to other, related forms to

weaken, and its particular representation to become strengthened, such that the interpretation for the form in question is largely idiosyncratic.

Note that it makes little difference whether these examples are described as lexical metaphors or formulaic metaphors. All three examples look, on the surface, like textbook lexical metaphors, with a single word carrying a set figurative meaning. The contrast between *men are dogs* and *she's a dog*, on the other hand, suggests that both are more akin to *the surgeon is a butcher*, with a set interpretation arising out of the application of a particular source domain term to a particular target term. All cases are analyzed as particular forms abstracting away from the sanctioning conceptual schema, such that they take on a degree of autonomy.

2.3.2 Lexical Strength

Here, in the notion of autonomy, we find accounted for the idea of a metaphor's lexical strength. Svanlund (2007), as outlined in §1.1.6, notes that the ability of particular words to evoke concepts from a source domain is not wholly dependent on the strength of overall cross-domain mappings. The conventional figurative meanings of lexical metaphors, Svanlund asserts, tend to be associated more with individual words than with the underlying Conceptual Metaphors. Metaphor *is* conventionalized at the level of broad cross-domain mappings, in the form of schemata that emerge as abstractions over similar utterances. Individual words, however, can, by virtue of the frequency with which they are used to refer to given targets, entrench away from schemata in which they participate. As a result, the conventionalized metaphorical meaning of an individual word increases inversely to its connection with the schema overall, such that when a given word has a

fixed figurative meaning it may or may not be predictable based on the cross-domain mapping overall.

Over time, as a word is used with a certain figurative meaning (or meanings), the particular mapping evoked by the word takes on a degree of autonomy from the sanctioning schema. At this point, the figurative interpretation for a word can become highly idiosyncratic, and difficult to predict based on the sanctioning schema. In the *dog* examples above, for example, two specific mappings that make use of the word *dog* as a source term are entrenched independently of the overall schema PEOPLE ARE ANIMALS, as well as of the narrower schema PEOPLE ARE DOGS.

On the other hand, however, even extremely entrenched figurative meanings for individual words can remain highly consistent with schematic metaphorical mappings in which they participate. Sweetser (1990), for example, points out an overall pattern in semantic shift for words with a base meaning rooted in sight and vision, motivated by a metaphor whereby understanding a thing is conceptualized as being able to see it clearly (corresponding, in CMT, to IDEAS ARE PERCEPTIONS): *point of view, crystal clear, illuminate, transparent*, *opaque, muddy, clear-sighted*, *bright, brilliant*: all have highly entrenched figurative meanings motivated by the above conceptual metaphor. In the case of each of these, entrenchment doesn't seem to have led the particular aspect of the mapping instantiated by the word too far astray for the metaphorical mapping overall. On the other hand, the high degree of entrenchment for the representations whereby these words are equated to set metaphorical meanings is itself indicated by the fact that these are 'go-to' words for evoking the overall mapping: while other words and phrases from the source domain (e.g., 'shine a light on' or 'invisible') are likely to be interpreted,

based on the strength of the overall schema, in a way consistent with the IDEAS ARE PERCEPTIONS mapping, the frequency with which terms like *illuminate* are selected leads to their being entrenched as the default terminology for evoking sight to describe thought.

Thus, as the metaphorical use of a word becomes conventionalized, the particular figurative meaning of the word gains in strength relative to the metaphor that licenses the overall mapping[9]. This idea has a clear correlate elsewhere as well.

2.3.3 The Career of Metaphor

The Career of Metaphor Hypothesis (Bowdle & Gentner 2005) asserts that when speakers are exposed to novel metaphorical utterances, they process them as comparisons: the listener draws analogical connections between the source and target domain, drawing inferences based on a comparison of the structure of the source to the target, and interprets the source term in the context of the target by mapping the relevant structure from the source onto the target. As the metaphorical use of a word becomes conventionalized, however, processing shifts away from comparison, and towards

[9] The issue invites the question of whether, for some highly entrenched figurative meanings for words and constructions, there is any connection whatsoever to the broader metaphorical schema that originally motivated the figurative meaning of the word. The question of whether autonomy can be absolute (even to the point that metaphorical motivations can't be reconstructed), and if so at what point this can be said to occur, is an important one, and a major direction for future research. The finding that metaphorical autonomy can indeed be complete would mirror findings at other levels of linguistic structure.

categorization. The concept that is drawn from the base is generalized to form a new category, based on traits common to both the literal and figurative meanings of the term. At this stage, the term in question has taken on a set figurative meaning, at the core of the newly constructed category.

Thus, an utterance of *men are dogs* would prompt, for a listener who had never been previously exposed to it, an attempt to use analogical reasoning to map the relevant feature of *dogs* onto *men*. A listener who had a sufficiently high amount of previous exposure to the metaphor, however, would have constructed a category over the subordinate categories MEN and DOGS, characterized by an abstract structure common to both (OPPORTUNISTS?). This figurative meaning of *dog* is the prototype of the abstract category.

Bowdle & Genter (2005) test this hypothesis in the context of the assumption that the grammatical form of similes (*x* is like *y*) and metaphors (*x* is *y*) is intimately connected with how each is processed: similes, which look like literal comparison statements, bias processing towards a comparison of attributes, metaphors towards categorization. The first of two experiments they conduct demonstrates that, for sentences containing figurative uses of a word, participants tend to prefer phrasing the sentence as a simile when the figurative use is novel, and as a metaphor when the figurative is conventional. The second indicates that novel figurative uses of a word are processed more rapidly in simile form, conventional uses in metaphor form.

The predictions made by the Career of Metaphor Hypothesis, with respect to processing for novel and conventional figurative uses of a word, align precisely with those of the schema-based approach, and the two theories are wholly consistent with one

another. A listener exposed to a novel figurative use of a word will align (based on the target being referenced and the source being drawn upon) the use of the word with the most appropriate metaphorical schema, based on semantic proximity to relevant schemata. The word is then interpreted within the context of the cross-domain mapping licensed by the schema selected, with multiple competing interpretations at hand. A speaker exposed to, for example, a comparison of a woman to an otter would find PEOPLE ARE ANIMALS the most proximate schema, but within the schema would still be faced with any number of possible interpretations (She's clever? She's a strong swimmer? She has a sleek coat?).

For a conventionalized usage, however, the listener doesn't need to refer to the overall mapping to interpret the word, as a figurative meaning is entrenched for the word itself. The overall schema is activated, and interpretation of the word may or may not be wholly consistent with the schema overall. The word can be reanalyzed on the basis of the schema if context suggests that such is necessary. But the most rapid route to processing is via the representation for the figurative meaning attached to the word itself. Exposed to example 15, *she's a fox*, interpretation according to the conventionalized figurative interpretation involving physical attractiveness is the default. Context, however, can prompt reanalysis based on PEOPLE ARE ANIMALS (e.g., *I can't imagine how she got out of that situation- she's a fox*).

Critical support for the role of frequency in affecting how metaphorical uses of words are processed was garnered in a follow-up experiment also reported in Bowdle & Gentner (2005). Repeating a design similar to the first experiment, the follow-up preceded the sentence-form preference portion of the experiment with an initial phase that exposed

subjects to figurative uses of the relevant source terms, such that the initial phase produced, artificially, conventionalization for the figurative uses of the source terms. The results of the first experiment were reproduced, with conventionalization produced by repeated exposure. Conventionalization is here aligned with repeated exposure to—which is to say, the frequency of— the form being used with a given metaphorical meaning.

2.3.4 Idiomaticity

At the same time that categorizations form based on semantic similarities, they also form over syntactic, form-based ones—in the figure that follows, based on the occurrence of the copula, *like*, or a transitive verb:

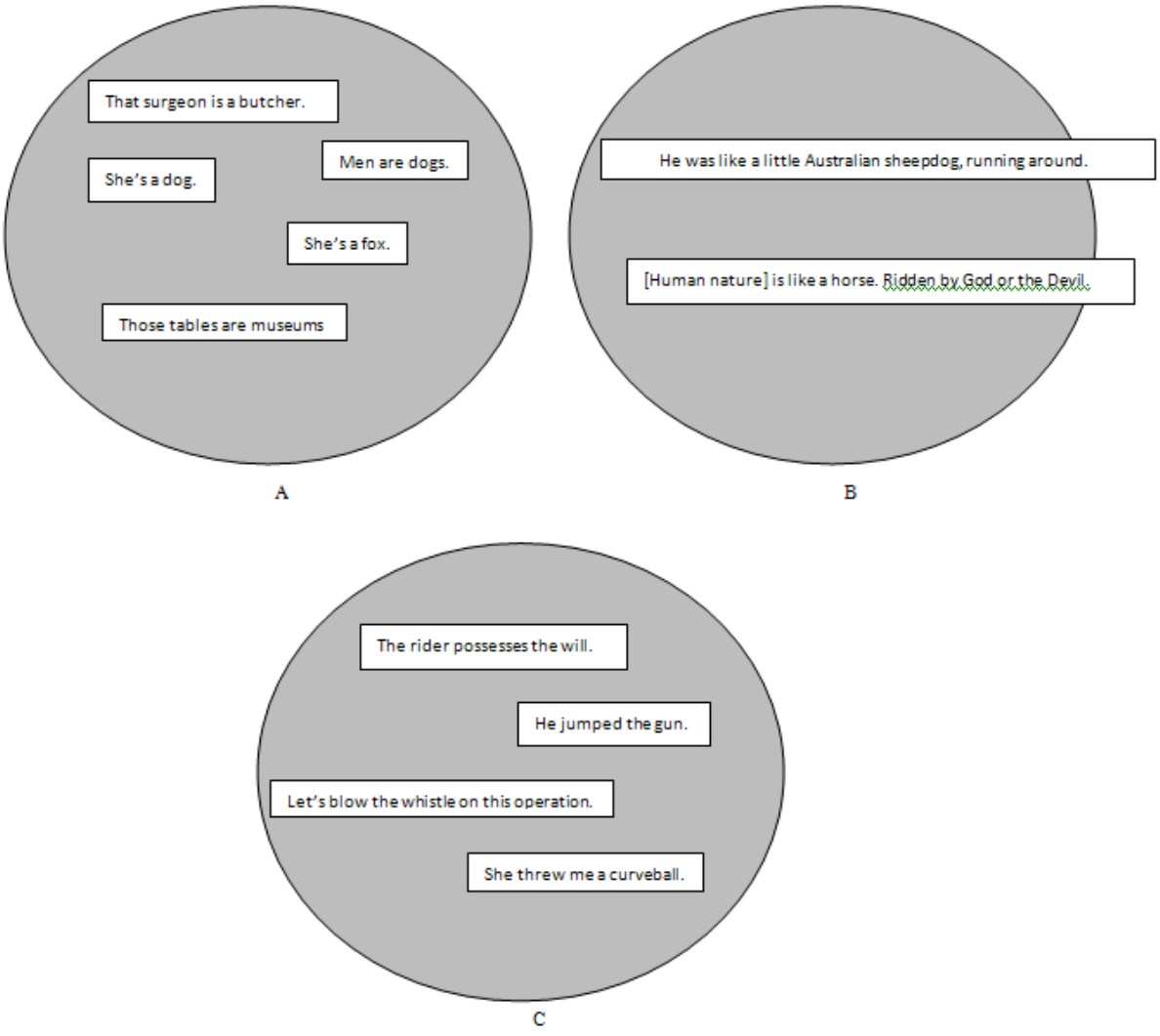

Figure 2.3.4
Categorizations based on Syntactic Similarity
A. [N] is [N]
B. [N] is like [N]
C. [N] [transitive V] [N]

Schemata arising over such categories are fundamentally objects of syntactic representation: they specify a template for a construction in the form of positions that are filled with a set word, or a member of a set class of words. Syntactic schemata (Barlow & Kemmer 1994, Goldberg 1995, 2006, Taylor 1998, Croft 2001) also, however, contain

semantic and pragmatic information. Constructions bear an overall meaning that puts constituent elements into set semantic roles, and are accompanied by constraints on usage. As §2.3.3 outlines, the schema arising over both items in A and over literal categorization statements has a strong association with conventionalized mappings, the schema over B and literal comparison statements with innovative, novel ones.

Where a specific mapping becomes partially or wholly autonomous, taking on conventionalized properties that are not predictable based on the overall conceptual schema that forms over the mapping, it often does so in the context of a set expression. *Blow the whistle on x,* for example, participates in a broad conceptual schema whereby perceptual saliency is conceptualized in terms of audibility (*that's a loud tie, that outfit screams 'available'*), but has a specific meaning of revealing a wrongdoer that isn't wholly derivable from the broader schema. Alongside the semantic idiosyncracy comes syntactic inflexibility: while limited operations on *blow the whistle* in terms of tense are allowed, the words must occur in a fixed order with no intervening elements, and there are strict restrictions on the preceding and following word[10]. The effect reflects what has been described by Company Company (2006) as the "cancellation of syntax" (p. 97), whereby, due to the role of frequency in shaping syntactic structure, subjective expressions can lose, over time, their normal syntactic capacities (see also Travis 2006).

All of the items in 2.3.4(C) are participants in a broad schema with the abstract form [N] [transitive V] [N]. *Jump the gun* and *blow the whistle*, while each representing in itself a

[10] The alternate form *whistle blower* is most certainly related to the *blow the whistle on x*—but is itself another entrenched, relatively inflexible expression.

fixed construction, also exemplifies a sub-schema that has close associations with idiom: [V] the (adj) [N] (e.g., *spill the beans*, *hit the ceiling*, *blow the lid off,* etc.).

In a connectionist model, schemata (at least until they become autonomous) operate as activation networks (McClelland & Elman 1984, Carr & Thompson 1996). Any instantiation of a schema, once formed, triggers activation of the schema overall. The intersection of metaphorical and syntactic schema means that a single utterance (such as *blow the whistle*) can activate both a cross-domain mapping (PERCEPTUAL SALIENCE IS AUDIBILITY) and constructional pattern ([V] the (adj) [N]). Metaphorically motivated idiom is essentially the class of utterances that activates both a narrow, highly autonomous cross-domain mapping and a highly fixed syntactic construction—an argument elaborated in Chapter 3.

2.3.5 Families of Metaphor, Internal Structure of Mappings

A principal feature of CMT is that conceptual metaphors form families of related metaphors, with the structure of a source domain providing a coherent way of conceptualizing the target domain. A cross-domain mapping sanctions the use of concepts and terminology from one domain to describe parallel, if more abstract, ideas in the target. Critically, such mappings are uneven in how they draw on the source: in practice, linguistic metaphors don't generally sample terminology and concepts evenly from the domain, but rather draw repeatedly on particular items[11].

[11] The impression of unevenness in how a metaphor samples a source domain can also arise as a result of how broadly a source-target mapping is interpreted, or directly from the misstatement of a metaphor's range. Clausner & Croft (1997), for example, argue that THE CONVINCINGNESS OF AN ARGUMENT IS THE STRUCTURAL INTEGRITY OF A

One reason for this is the entrenchment of set metaphorical meanings in particular words and phrases. The repeated use of a word or construction to evoke a particular aspect of a source domain causes the form in question to take on a degree of autonomy from the sanctioning schema, in the sense that the entrenchment of the form itself is accompanied by a concomitant weakening of the form's connection to the metaphorical schema governing the mapping. The entrenched form takes on a relatively fixed figurative interpretation. The stronger the representation for the entrenched form, the more accessible it becomes, and the likelier a speaker is to choose it when selecting a vehicle term from the source domain. We see the end result in a word like *illuminate*, upon which frequency effects have operated to create a default term for evoking the source domain PERCEPTION in reference to IDEAS. We can contrast *illuminate* with *brighten*, which has a roughly equivalent literal meaning, but doesn't carry the same automatic metaphorical meaning of 'cause to understand'. *Illuminate* participates in the schema IDEAS ARE PERCEPTIONS, but has a highly fixed and idiosyncratic interpretation within that schema (hypothesized here to be a function of the high token frequency of its use in reference to perceptions). While the schema is activated when the word is used

---

BUILDING is a more accurate statement of the scope of the metaphor licensing expressions such as *is that the foundation of your theory* and *the argument is shaky* than the form in which it was posited by Lakoff & Johnson (1980), THEORIES AND ARGUMENTS ARE BUILDINGS. Grady (1997), on the other hand, asserts that THEORIES AND ARGUMENTS ARE BUILDINGS is not a basic-level metaphor, but rather arises at the intersection of other, core-level mappings: ORGANIZATION IS PHYSICAL STRUCTURE and PERSISTING IS REMAINING ERECT.

figuratively, speakers don't need to *use* the overall schema to reconstruct anew its metaphorical meaning every time that the word is used or uttered in a figurative sense: the word invokes a direct connection between a particular concept within the source (increasing lighting making something more visible) and a particular concept within the target (making an idea more easy to understand). While a metaphorical use of *brighten* will evoke the same source domain (PERCEPTIONS), and even the same concept within the source domain (increasing lighting), the word does not have an entrenched metaphorical meaning, and therefore, for a listener, its metaphorical meaning must be interpreted based on the schema. The example here pertains to individual words, but the same effect applies to longer units—the effect noted for *illuminate* happens also, for example, with *shed light on x*.

Not only isolated words and constructions, however, are given preferential treatment within a schema. Within the overall cross-domain mapping, some aspects of the source and target domain are consistently invoked, while others go unexplored. The issue of unexplored aspects of a mapping, described by Grady (1997: 270) as "the poverty of the mapping", is one that has proven difficult within metaphor theory. If there is a cross-domain mapping whereby IDEAS ARE FOOD, such than an idea can be half-baked, conjectures ruminated upon, and information digested, then why are 'my ideas are completely boiled' and 'the information tasted terrible' strange? If STRONG EMOTIONS ARE MADNESS (such that one can be 'mad with hate' or 'out of one's mind with grief'), then why aren't interventions for insanity used to refer to the calming of emotions?

Questions such as these put the cart before the horse, getting the direction of causality precisely wrong. Metaphorical mappings don't pre-exist utterances; they're not static

structures that reside deep in one's linguistic competence, allowing and disallowing various utterances. It's true that such a schema, once formed, can be used to interpret and coin novel metaphors—interpretation will be more rapid, and coinages more frequent, in proportion to the strength of the schema, which is a function of its type frequency. The schema, however, does not as a rule pre-exist metaphorical utterances, it forms as an abstraction over them. Novel metaphor can explore new aspects of a mapping, but by its nature, a schema forms over entrenched forms and conventionalized schemata, which will invariably 'outcompete' non-entrenched forms as speakers cast about for words and constructions of which to make metaphorical use. The experiments outlined in Chapter 5 point to this effect, but are geared towards addressing frequency effects as they pertain to overall mappings, rather than to the entrenchment of forms themselves. This prediction of the usage-based account with respect to competition among forms sanctioned by the same schema presents a clear direction for future research.

Models of metaphor that treat metaphors as *a priori* constructs treat special cases of a metaphor as entailments: language users, having applied the systematicity of one domain to another, make logical inferences about the target based on the source (for example, that if IDEAS ARE PERCEPTIONS, then to increase perception must be to increase a person's understanding). As noted above, a usage based-model is bottom-up, rather than top-down: while the idea of entailment may apply in some cases of novel metaphor, 'special cases' are in most cases more appropriately treated as schemata in themselves. More general schemata form over the tokens comprising the special case, as well as other tokens instantiating the more general schema.

Lakoff, Espenson & Schwartz (1991) posit the metaphor SUBJECTS ARE AREAS, whereby areas of research are conceptualized as areas of space. Within this overall mapping is the special case that RESEARCH IS EXPLORATION (see Figure 2.3.5 for examples). Within RESEARCH IS EXPLORATION, moreover, are two further subcategories: metaphors according to which making discoveries is seeing new land or objects, and metaphors according to which a field, once surveyed, can then be used for agriculture.

Just as we might ask why *brighten* isn't used instead of *illuminate* to describe making something more understandable, we can pose the question here as to why, if SUBJECTS ARE AREAS, all sorts of properties of areas of land— whether they're hilly or flat, for example, or whether they have access to fresh water—don't apply. RESEARCH may be EXPLORATION, but the explorers in question don't seem to be paying attention to many of things that explorers of new areas of land pay attention to— native flora and fauna, for example, go unnoticed in this system of metaphors.

Again, the resolution is that the cross-domain mapping doesn't pre-exist tokens of use, it emerges as a schema over them. This applies at the level of a schema emerging over a set of tokens, but also to a larger schema, encompassing several sub-schemata, that emerges as an abstraction over all constituents. Figure 2.3.4 gives a schematic representation for SUBJECTS ARE AREAS; all examples are from Lakoff, Espenson, & Schwartz (1991).

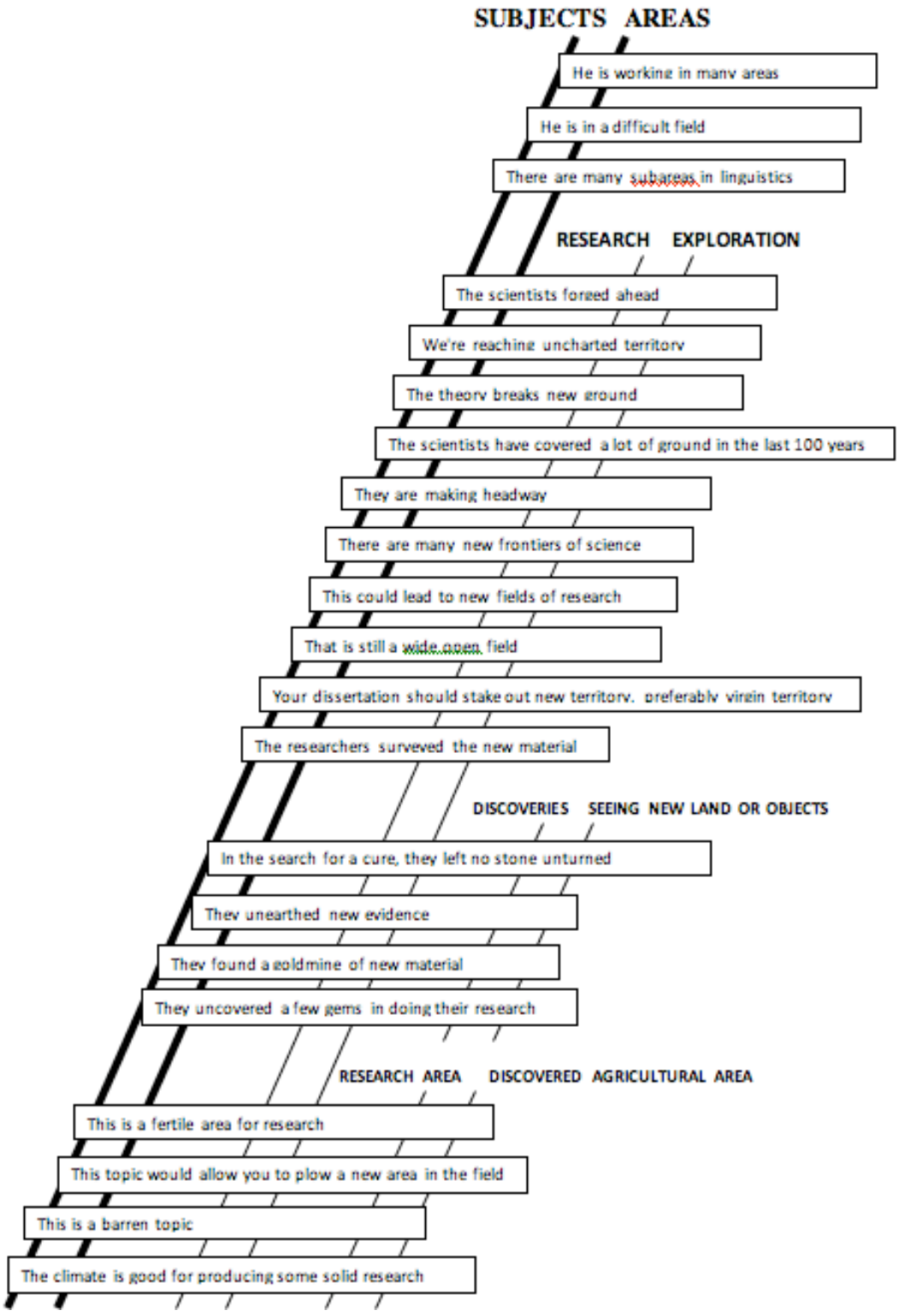

Figure 2.3.5
Family of Related Metaphors

A single utterance can activate schemata at multiple levels. 'They unearthed new evidence' activates DISCOVERIES ARE SEEING NEW LAND OR OBJECTS, but also RESEARCH IS EXPLORATION and SUBJECTS ARE AREAS. 'He is working in many areas', however, activates only the master schema, SUBJECTS ARE AREAS.

2.3.6 Metaphor as a Continuous Variable

"Some metaphors", notes Deignan (1999a: 107), "are more metaphorical than others." Metaphors tend to lose strength over time as they become conventionalized, such that any particular metaphorical utterance can be placed along a scale from more metaphorical to less metaphorical. It's a difficult aspect of metaphor to deal with for researchers looking at metaphor use in natural language, as identifying instances of metaphor in discourse necessitates placing a dividing line between metaphorical and non-metaphorical utterances—choices forced at the periphery are difficult, because 'metaphorical' and 'non-metaphorical' aren't categories, they're endpoints on a continuum. In such cases, as Cameron (1999) points out, explicitness of criteria for an operational definition of metaphor used in a study must serve the function that a clear and objective dividing line between metaphorical and literal language cannot[12].

Even within specific categories of metaphor, utterances can be placed along a continuum which ranges from literal to metaphorical. Shank (2007), in a study analyzing semantic extension on tactile verbs with figurative meanings relating to perception,

[12] Deignan (1999a) and (Sanford 2009a) outline many of the issues associated with corpus work on metaphor, and specifically separating metaphorical utterances from non-metaphorical ones for the purpose of a corpus study.

isolates all tokens of the four verbs *touch*, *handle*, *hold*, and *feel* from a corpus of spoken English, placing them along a continuum from literal uses, referring to physical manipulation, to more abstract and figurative uses, referring to intellectual understanding. Sanford (2008b) reports a study in which one group of subjects was given the task of rating a series of idioms on a scale from 1 to 5 relating to how metaphorical they perceived the idiom to be. A second group of subjects had their spoken performance of the same idioms monitored, and the length of the main verb in each idiom (eg., *spill* in *spill the beans*, and also in the control utterance *spill the peas*) monitored. The correlation between the two sets of results, following a control for the reduction effect from the frequency of the phrases themselves, indicates a verification of the hypothesis that an idiom's degree of metaphoricity directly effects the phonological reduction of internal elements. It also, implicitly, validates the gradedness of metaphor for idioms, as corroborated by both methods of measuring metaphoricity. A host of other studies on idiom (Nunberg 1978, Cacciari & Glucksberg 1995, Bosman 1999) have placed idioms along a continuum relating to their degree of novelty, and analyzability—both highly related to the extent to which they activate underlying cross-domain mappings.

The gradedness of metaphor is not a feature well handled by most theories of metaphor: whatever the proposed cognitive underpinnings of metaphor are—cross-domain mapping, integration network, statement of categorization—the process either does or does not take place. Here, a form's degree of metaphoricity is taken as a function of its degree of autonomy from its sanctioning schema. Lexical metaphors, formulaic metaphors, metaphorically predicated idioms, and other metaphorical utterances at any level of conventionalization: in each of these cases, an utterance is metaphorical to the

extent that it activates an underlying mapping. Such activation depends on the strength of the form's connection to the metaphorical schema governing the mapping, and the strength of the form's connection to the schema varies alongside its degree of autonomy. The more frequent a form, the weaker its connection to the sanctioning schema, and the less it activates the underlying schema.

2.3.7 Metaphor Processing and Age

Within a language user's mind, schemata form over tokens of use, and gain strength in proportion to the number of exemplars. Within a given speech community, speakers have been exposed to the frequencies of metaphorical tokens of use at similar levels, and accordingly have roughly congruent schematic structures. The correlate, of course, is that the less time that a speaker has had to be exposed to metaphorical tokens of use, the less their cross-domain mappings will correspond to those of a language user who has had more time to accumulate exemplars. For a language user whose experience with language was sufficiently brief, few enough exemplars would have been encountered, for many clouds of tokens, to allow the formation of schemata. And language experience is most brief, naturally, for those who joined the community of language users most recently: children.

Metaphor comprehension in children has been a fairly well-documented area, numerous studies confirming the early emergence of metaphorical understanding in children, but also a strong correspondence between age and the ability to comprehend metaphor. Such studies have generally focused on factors which lead to erroneous metaphor interpretation in children. Billow (1975) draws a distinction between similarity metaphors, in which the referent is equated to something else on the basis of a shared

quality (e.g., 'hair is spaghetti') and proportional metaphors, which establish an analogical relationship among four items, one of which is left implied (e.g.'my head is an apple without any core', in which 'head' is to 'apple' as 'brain' is to 'core'). While understanding of both types of metaphor improve with age, the understanding of similarity metaphors emerges first, and is already well-developed by the ages of 5-7. Proportional metaphors, on the other hand, emerge considerably later, at ages 9-13. Billow concludes that the comprehension of metaphor emerges alongside higher cognitive structures and the systems of classification necessary to understanding analogical relationships between categories. Similarly, Nippold and Sullivan (1987) suggest that the capacity for understanding metaphor is tied directly to the emergence of analogical reasoning, the ability both to solve proportional analogy problems and to comprehend proportional metaphors emerging as early as age 5 and progressing parallel to one another thereafter. Broderick (1990) notes that while there is undeniably an improvement in metaphoric comprehension between early childhood and adulthood, these are "related to general improvements in overall comprehension ability rather than to the emergence of specific metaphoric capacities such as relating psychological and physical domains" (p. 65).

Vosniadou et al. (1984) address not only the early emergence of metaphorical understanding in children, but the extent to which children draw on an array of cues, both linguistic and situational, in order to draw inferences regarding the meaning of metaphors, accessing a matrix of contextual information in order to test alternate hypotheses as to the meaning of a figurative phrase. Comprehension difficulties arise in response to a lack of predictability (in relation to linguistic context) of the metaphor, as

well as overall difficulty of the metaphor itself. Older children are better able to cope with both sets of confounding factors than are preschoolers. Siltanen (1989) notes that the more difficult a metaphor, the more comprehension depends on context: while older children depend less on context than younger children, difficult metaphors are better understood at all ages when a greater amount of context (in this case, a longer story) is provided. A child who hadn't been previously exposed to a metaphor needed more context to figure it out.

Waggoner and Palermo (1989) call attention both to the importance of context to children's understanding of metaphors, and the prerequisite to metaphor comprehension of familiarity with the domain used as a referent by the metaphor: very young children understood best those metaphors which pertained to emotions of which they had the most understanding (the negative emotions fear, sorrow, and anger). In the authors' study, while competence at comprehension of psychological metaphors increased with age at least as far as the college level, above-chance levels of performance were observed in the youngest of the study's participants (5 years old). Seitz (1997) points to the emergence, by the age of 6, of the use of linguistic knowledge to apply already present metaphorical capacity (seen in younger children mostly through visual perception of metaphorical relations) to psychological concepts.

Evans and Gamble (1988) found that most errors in metaphor comprehension were tied to attribute saliency: where children fail to explain a metaphor in a way similar to how an adult would, it is often because they single out different aspects of what is important about concepts. The authors assert that "what comes readily to mind for young children regarding certain words is quite different than for older children and adults. For

example while older children and adults listed 'fight in wars' and 'carries weapons' as important characteristics of soldiers, young children mentioned 'marches', 'wears a black and red uniform' and 'stands straight'" (p. 435). In the partial mapping from source to target domain which takes place in innovative metaphors, such variability can make a great deal of difference on interpretation: whereas an adult might interpret a metaphor involving a soldier as highlighting the soldier as one who fights, a child would tend to focus on the physical attributes of a soldier.

Metaphorical ability, in short, emerges early on, but children don't generally interpret metaphors in the same way that adults do until considerably later. What's notable, in the case of errors, is that children do interpret metaphors, they just don't necessarily do so in the same way that adults from the same speech community do. Whereas such findings are explained in the research outlined above in terms of the development of various cognitive capacities that precede various types and levels of complexity of metaphor use, it would seem just as reasonable to explain such errors in terms a lack of exposure to the metaphorical systems motivating the target interpretations. Palermo (1986:132), noting the issue, sums it up as follows:

> There is little question that a metaphor such as "Bees are the buccaneers of buzz" would not have the same meaning for a child of 5 as for Emily Dickinson who created it. The child, however, may fail to understand the metaphor, not because he or she does not have the cognitive capacity, structure, or cognitive processes to comprehend metaphors, but because the child does not have the knowledge of buccaneers that is the key to comprehending the message about buzzing bees that Dickinson was trying to convey.

Not only individual words, but also metaphorical systems and families of metaphorically motivated idioms, are themselves culturally specific, and thus acquired (Lakoff &

Johnson 1980, Lakoff 1987a). If metaphorical systems develop, within a language user, as abstractions over stored tokens of use, then it's wholly expected that while young children might have the capacity for cross-domain mappings, they wouldn't have accrued sufficient language experience to have developed the specific schemata to facilitate the mappings that allow an adult speaker to use and interpret a given metaphor in roughly the same way as other speakers within the community. Children enter the metaphorical world inhabited by adults slowly, as they form, through exposure to individual metaphorical forms that comprise tokens of use, schemata governing the cross-domain mappings that comprise a culture's shared system of metaphors.

2.4 Conclusion

My proposal is that metaphorical systems, contrary to what is popularly believed, don't underlie and license metaphorical tokens of use. Rather, they emerge over them in the form of schemata that link cognitive domains. Many essential features of metaphorical systems, especially as they pertain to conventionalization, have been noted elsewhere, but not followed to the conclusion that metaphorical systems are emergent in nature, arising over the fact of language in use. Such features of metaphor follow naturally from this conceptualization of metaphor.

Even for highly productive metaphors, a large proportion of instantiating tokens are accounted for by a handful of highly entrenched forms. This phenomenon, as well as the existence of subcases/inferences for overall mappings, speaks to the way in which schemata form over actual tokens of use: metaphors aren't entities that govern an evenly distributed range of possibilities, some of which are mysteriously absent from language in use. Metaphors are abstractions formed over uneven input. Such mappings correspond

to conceptual metaphors, and entrenched forms to lexical, formulaic, and idiomatic metaphors.

Idiosyncratic, conventionalized meanings for lexical, formulaic, and idiomatic metaphors are a result of the autonomy of forms with a high degree of entrenchment relative to an overall schema, resulting in a weakening of the connection between the entrenched form, and other forms sanctioned by the schema. A metaphorical form's degree of autonomy corresponds inversely to the extent to which it will be perceived as metaphorical by the speakers of a language.

The metaphorical systems that are shared across individuals within a culture are a result of shared membership within a speech community, and roughly approximate levels of exposure to metaphorical tokens of use over which emerge conceptual schema that govern cross-domain mappings. The ability to use and process metaphor according to cultural norms for adult speakers results from children not having been exposed to sufficient tokens of use for many metaphors to be firmly entrenched.

There are, notably, many features of metaphor that a usage-based approach does not account for. Metaphorical aptness, in particular, is only partially explained. The strength of a schema, and of metaphorical forms within it, are important aspects of why a speaker, in casting about for a metaphor, settles on one mapping over another, and within the mapping, one form over another. There are other important factors, however, as well. Sanford 2008a combined a corpus-based study in which all instances of metaphor were isolated from a corpus of approximately 40,000 words and coded as to their target domain, with a survey instrument in which participants were asked to rate the concreteness (for the purposes of the study, being easily understood, clearly defined, and

easily conceptualized) of each of the categories used for coding the target domains in the corpus. The study found a significant positive correlation between concreteness and the frequency of a given category being referred to metaphorically, indicating that the concreteness of a referent is a clear factor in speakers' deciding whether or not to refer to it metaphorically. The more concrete a referent is, the less likely it to be referred to metaphorically. Furthermore, the concreteness of the source relative to the target is a clear factor in choosing a source domain, such that targets are almost invariably less concrete than sources. Such constraints aren't accounted for by the usage-based theory, but neither are they outside of its purview: they feed, and are in turn reinforced by frequency effects, such that a mapping or form with an advantage in aptness will be rapidly conventionalized or entrenched.

Chapter 3: Idiom: Syntactic Schemata, Graded metaphor

3.0 Introduction

Chapter 2 of this dissertation presented the argument that metaphors are conceptual schemata that form over cross-domain mappings, as language users encounter metaphorical tokens of use. The conventionalization of metaphorical cross-domain mappings is asserted to result from the entrenchment of such schemata, and Lakovian conceptual metaphors (of the form X IS Y), it was argued, describe entrenchment at this level. The frequency effects described in Chapter 2 pertain to such mappings overall, with individual instantiations of metaphors processed through alignment with metaphorical schemata. Each experienced token of use, once aligned with a schema, strengthens the mapping overall. Chapter 5 will present a series of experiments that tested the claims related to metaphorical entrenchment which were made in Chapter 2, demonstrating that a language user's previous exposure to metaphorical tokens of use (assessed for the purposes of the experiment using large corpora of English which are assumed to be valid reflections of individual English speakers' history of the language) has a demonstrable effect on how novel token are processed. This chapter elaborates on entrenchment at the level of individual instantiations of a metaphor, a level of conventionalization that has far greater precedent in the literature on metaphor. Idiomatic metaphors are described here as arising from the association of high-autonomy metaphorical mappings with syntactic constructions.

Idiom exists, by its nature, at the intersection of the study of figurative language and of syntax effects, and has proven a singularly problematic issue in both areas of inquiry. For a generative model of language, in which linguistic performance emerges as the

operation of rules upon the lexicon, idioms are oddballs, in that they are lexical units larger than individual words. At the same time, however, they've generally been viewed as the exception that proves the rule: idioms may be strange bedfellows for conventional lexical units such as words and morphemes, but aside from containing multiple words, they're otherwise unsurprising members of the lexicon (Chomsky 1965, 1980, Jackendoff 1992). According to this view, when speakers interpret an utterance such as *let the cat out of the bag* or *jumped the gun*, they retrieve them wholesale from the lexicon, where the expressions are stored alongside their stock figurative meanings ('reveal a secret' or 'begin too early').

As many researchers (Barlow & Kemmer 1994, Nunberg, Sag, Wasow 1994, Goldberg 1995, Croft 2001, Wray 2002) have pointed out, however, this concept of idiom falls short of accounting for important aspects of idiom: First, there is no clear distinction between idiom and non-idiom, such that 'normal' language operates according to compositional principles, and the set of idioms in a language operate according to another set of rules. Second, many idioms do bear some internal structure, such that the overall meaning of the expression is distributed over semantic units smaller than the expression as a whole. Once relegated to a dusty annex of the lexicon, idiom has come to serve as a core around which functionalists have built an understanding of grammar. Within syntax, idiom has demonstrated the impossibility of drawing a clear distinction between lexical items and rules which operate upon them. For metaphor, idiom demonstrates the interaction of conceptual schemata and constructional templates, as well as the impossibility of treating metaphoricity as a binary category.

3.1 Syntactic Schema

Generative theories of syntax have tended to draw a line between the 'normal' language and idiomatic speech.

(24)The cat is asleep on the television.

(25)I let the cat out of the bag.

The meaning of a sentence such as example (24) can be modeled as a function of the predictable operation of grammatical rules and semantic principles. On the other hand, a sentence such as (25) seems to demand a different, and by definition abnormal, route to interpretation in that the gestalt meaning of the expression is not what one would expect based on the application of morphosyntactic rules to lexical units.

As the tools of corpus linguistics have been brought to bear on language, and as corpus data have reliably demonstrated both the paucity of 'normal' sentences and the preponderance of repetition in discourse, theories of syntax have reoriented around accounting for phenomena previously considered peripheral. Within the constellation of research that analyzes sentence-level patterns as syntactic schema, or constructions (Fillmore, Kay, & O'Connor 1988, Barlow & Kemmer 1994, Nunberg, Sag, & Wasow 1994, Goldberg 1995, Croft 2001), analysis of clause- and sentence-level structure embraces idiomaticity and idiosyncracy as fundamental features of language. As the tools developed to account for idiomatic speech are applied to more canonically compositional sentences, a syntactic model emerges that accounts for both types of sentences using the same apparatus.

Syntactic constructions, as defined by Fillmore, Kay, & O'Connor (1988), are objects of syntactic representation that also contain semantic information. Constructions are

lexical, in that they directly relate a form to a semantic content: Goldberg (2006) asserts that the essential feature of a construction is to have an overall meaning or function not entirely derivable from internal elements[13]. They're also, however, complex in that they contain discrete, identifiable internal elements which are themselves lexical units.

Critically, constructions are variable in their degrees of fixedness.

(26) kith and kin

(27) That's neither here nor there.

(28) pain in the butt

(29) I don't want to play baseball at all, let alone play in the rain.

(30) He sneezed the napkin off the table.

At one end of the continuum are highly fixed expressions such as (26) or (27), all elements of which must occur, and which must furthermore occur in a set order, in order for the string to carry its constructional meaning. The slightly less fixed construction seen in example (28) (*pain in the X*) allows for variations only within a small set of words (*neck*, *butt*, and informal synonyms for *butt*) to fill the final slot. Towards the other end of the continuum one finds constructions such as Fillmore, Kay, & O'Connor's (1988) *let alone* construction, licensing expressions such as the one seen in example (29), and relatively open sentence templates such as that seen in example (30) (from Goldberg 1995). The former, the *let alone* construction, imposes semantic roles on the elements preceding and following *let alone*, demonstrating as well the extent to which grammar can be highly idiosyncratic for a given construction. The latter demonstrates the caused

[13] Bybee (2006), in an account of constructions more explicitly consistent with a usage-based account, defines constructions as automated sequences or processing units.

motion construction. The construction licensing this last example is highly productive, allowing for the creation of a potentially limitless number of new sentences, and highly schematic, in that it allows for variation, within a class defined by a category or set, as to elements filling particular roles within the construction. It is the template itself, however, rather than the verb, contributing the idea of caused motion.

With *any* change in syntax, there is *some* difference in meaning: Goldberg (1995) demonstrates the effect with examples such as (31) and (32):

(31) I loaded the truck with hay.

(32) I loaded hay on the truck.

From whence, if not the construction itself, comes the idea that the truck is fully loaded in the former, but not necessarily so in the latter? From the word *with*? *on*? The difference between basic sentence patterns and canonical idioms is only the degree of idiosyncraticity. In every case—from the most fixed expressions, to general sentence patterns, to the thousand shades of gray that fall between—grammatical patterns are associated directly with semantic content. There's no point on the continuum, in short, at which the realm of the idiomatic ceases and 'normal' language begins.

Syntactic constructions are grammatical schema (Barlow & Kemmer 1994, Bybee & Thompson 1997, Taylor 1998, Dabrowska 2001). Constructions form as abstractions generalized over multiple instances of utterances, forming prototypically defined categories around clusters of similar tokens of use, with schematic patterns strengthened alongside the frequency of exposure to tokens of use instantiating the schema. The acceptability of incoming forms is assessed based on similarity to entrenched schemata, and stronger schemata are more easily accessed and more productive.

With all syntactic patterns, from the most restricted to the most general, viewed as lexical pairing of form to meaning, the domain of the idiomatic is extended to cover all of a speaker's syntactic knowledge, as shaped by the speaker's previous exposure to multiple-word patterns. Canonically idiomatic phrases, however, are clustered far towards the fixed end of the continuum, where the entrenchment of highly restricted patterns and of relatively fixed strings leads to constructions the overall meaning of which is far narrower than (and, in many cases, highly divergent from) the set of possible meanings, based on their constituent elements.

3.2 Idiom: syntactic & metaphorical schemata

The essence of idiom is the tension between metaphorical and syntactic schema. Within the theory of metaphor outlined in Chapter 2, metaphorical autonomy occurs where some instances or aspects of metaphorical mapping become partially or wholly autonomous from an overall sanctioning metaphorical schema, taking on conventionalized properties that are not predictable based on the overall conceptual schema that forms over the mapping. Where this happens in association with a relatively fixed construction, we find prototypical idioms: utterances with a figurative basis (supplied by a metaphorical schema), a conventionalized meaning (a function of autonomy for some aspect of the mapping from the overall mapping), and a form that bears little lexical or syntactic alteration without its meaning being changed (as a function of the fixedness of the construction underlying the expression).

(33) keep it under your hat

(34) zip your lips

(35) spill the beans

(36) let the cat out of the bag

The family of idioms represented in examples (33) through (36) revolves around the idea of revealing or keeping information through verbalization, with information conceived of as a physical entity within a container. A considerable hierarchy of metaphorical schema are here involved: the highly general IDEAS ARE OBJECTS (Lakoff, Espenson & Schwartz 1991) is emergent over idioms such as examples (33) through (36), as well as a wide host of other utterances at varying levels of entrenchment (such as examples (37) through (40)).

(37) I grasped her arguments.

(38) This idea is stuck in my head.

(39) I gave her the idea.

(40) I get it.

Within this more general schema, (33) through (36) all exemplify THE MIND IS A CONTAINER FOR OBJECTS (Lakoff, Espenson, & Schwartz 1991): information is an object in a container (the mind), and verbalizing the information is to have the information escape the container. Over expressions such as (33) and (34) emerges the generalization that KEEPING A SECRET IS PREVENTING AN OBJECT FROM ESCAPING A CONTAINER, over expressions akin to (35) and (36) that REVEALING INFORMATION IS ALLOWING AN OBJECT TO ESCAPE FROM A CONTAINER (both sub-schemata that would be treated, in Conceptual Metaphor Theory, as entailments of IDEAS ARE OBJECTS). Each example individually illustrates a specific mapping which has become partially autonomous, with a conventionalized interpretation not wholly predictable based on either the most general or most narrow of the schema over them.

At the same time, however, examples (33) through (36) exemplify syntactic constructions. Each is a narrow sentence template that bears some, but not much, lexical and syntactic alternation with its meaning intact. (34) admits 'button' in the role of 'zip', and variations in tense and person are possible for all four, but passivization is questionable in all cases, and lexical substitutions quickly erode the meaning of the expression for all examples. (35) and (36) are prime examples of the transitive '*x* the *y*' construction, closely associated with idiomaticity.

As has been argued throughout, the strength of a schema, whether semantic or syntactic, is dependent on the sum of a language user's history of exposure to tokens of use. Activation itself, however, is tied to short-term factors: once activated by an utterance instantiating a given schema, an utterance sanctioned by the same schema will be more rapidly processed (Bock 1986, Branigan, Pickering, & Cleland 2000 , Poplack 1980, Scherre & Naro, 1991, Allbritton, McKoon, & Gerrig 1995, Gries 2005, Szmrecsanyi 2005). Based on the immediate preceding context of a given idiomatic utterance, a language user might be expected to be 'biased' towards an interpretation of idiomatic meaning based either on the utterance's affinity to tokens instantiating the same syntactic schema, or towards tokens instantiating the same conceptual schema.

Longer-term factors certainly come into play as well: the strong entrenchment of a specific mapping, or of a relatively fixed string of words, relative to an overall schema, leads to some degree of autonomy. Schemata form as connections—essentially routes of activation—are made between tokens on the basis of shared properties. With autonomy, connections between a given mapping/string and other tokens sanctioned by the overall

schema weaken, such that the extent to which the schema is activated depends on the degree of autonomy for the string/mapping in question.

A metaphorical idiom represents both an entrenched mapping, and an entrenched string: in both regards, it has some degree of autonomy from its immediate sanctioning schema. For both syntactic and conceptual schemata, autonomy, schema strength, and priming together determine the role of sanctioning schemata in on-line processing. Conceptual schemata enable the distribution of idiomatic meaning out over constituent elements, while syntactic schemata are associated with the reanalysis of idiomatic meaning based on constructional similarity to other idioms.

3.2.1 Conceptual schemata and compositionality

The issue of the compositionality of idioms weighs two alternatives as to how language users process idiomatic utterances: when people use idioms, do constituent elements contribute to the overall meaning, or are the overall meanings of idioms retrieved intact from the lexicon? The nature of idioms as utterances with unpredictable meanings has steered researchers in idiom understandably towards the latter. In an expression such as *he jumped the gun*, how could the individual words *he, jump, the,* and *gun*, possibly combine to form the meaning 'begin too early'? What word in the expression *he kept me in the dark* contributes the notion that knowledge is being willfully kept from someone? Issues such as these have led to the canonical view of idiom that while expressions such as these may once have been formed compositionally, they are now essentially irreducible units. From a synchronic perspective, idioms may look on the surface like normal sentences, but in fact words internal to the idiom play no role in contributing to the overall meaning of the expression.

Nunberg, Sag & Wasow (1994) argue that a great deal of misunderstanding has arisen out of the misguided conflation of conventionality and conventionality (see §2.1.2). The conventionalization of meaning which accompanies idiomatic expressions needn't lead to an assumption of noncompositionality because this conventionalization of meaning may well be attached to individual words within an idiom, rather than to the idiom as a whole. While acknowledging the existence of a large number of idioms which are not, in fact, analyzable (being low on each of the above three scales), the authors assert that the bulk of phrasal idioms are, in fact, idiomatically combining expressions (ICEs): "idioms whose parts carry identifiable parts of their idiomatic meanings" (p. 496). The compositionality of such expressions is evidenced by the fact that individual words within them can be modified (*leave no legal stone unturned*), quantified (*touch a couple of nerves*), emphasized via topicalization (*those strings, he wouldn't pull for you*), omitted in elliptical constructions (*my goose is cooked, but yours isn't*), or referenced anaphorically (*we worried that Pat might spill the beans, but it was Chris who spilled them*), each of these operations indicating that speakers assign identifiable meanings to the word involved (all examples from Nunberg, Sag, & Wasow 1994: 500-502).

The view that an idiom's meaning can be distributed over its parts is supported by several lines of evidence. McGlone, Glucksberg, & Cacciari (1994) assert that constituent words in an idiom take on phrase-specific meanings. In a series of experiments which assessed subjects' comprehension of idioms, it was ascertained that while subjects understand familiar idioms more rapidly in their canonical form than in a variant form, variant idioms are understood as quickly as their literal paraphrases (e.g., *he bit off less than he could chew* as quickly as *he did not challenge himself*). The authors interpret the

results as indicating that variant idioms are processed in essentially the same manner as literal speech: compositionally. They suggest that "the words that form familiar idioms, by repeated usage, come to incorporate at least part of the figurative sense that they have when embedded in idioms" (p. 181), a process aided by "transparent conceptual relations between an idiom's constituents and that idiom's meaning" (p. 182): thus, provided the right context, a word's figurative/idiomatic sense is activated, its meaning contributed to the expression in which it occurs[14]. Billig & MacMillan (2005) provide a corpus perspective on the compositionality of idioms, their study on the use of the expression 'smoking gun' in political discourse providing evidence that the use of a linguistic metaphor musn't necessarily become any more automatic as it becomes conventionalized, but rather that the metaphor underlying the idiom can continue to be 'negotiated' in discourse—exploring new aspects of the systematicity of the source domain as it applies to the target domain—in the same way that novel metaphors do. They point, as one example, to the following rhetorical move which takes place around *smoking gun* (p. 473):

> A Downing Street spokesman insisted the so-called smoking gun, evidence of a continuing concealed weapons programme, was not the only justification for war set out in Resolution 1441.
> (*Guardian*, 27 January 2003)

[14] Cameron & Deignan's (2006) concept of a metaphoreme, "non-literal expressions with a relatively fixed form and highly specific semantics and pragmatics" (p. 1), is useful in accounting for such contextual and social effects. Such aspect of metaphor being accounted for in a schema-based model of supported by Goldberg's (1995) inclusion of pragmatic information in syntactic constructions.

> One might note the addition of 'so-called', as the government spokesman distances the government from the idiom, casting suspicion over the legitimacy of its usage… The shift is from the smoking gun as an object, which only has to be seen or heard to be believed, to a rhetorical claim that was to be dismissed and whose upholders were to be construed as holding discreditable and irrational motives. The use of quotation marks, or the addition of 'so-called', helps the shift back from apparent objectivity to contestable rhetoric.

The extent to which an overall (however conventionalized) idiomatic meaning can be attributed to the individual parts of a metaphorical idiom is attributable to the role of an active metaphorical schema in motivating the meaning of the idiom, the mapping providing a basis for analyzing out meaning over constituent elements. The autonomy of the specific mapping instantiated in an idiom nonetheless means that overall idiomatic meaning is not wholly derivable from these internal elements.

3.2.2 Syntactic schema and reanalysis

The fact that idioms instantiate not only semantic, but also syntactic schemata, means that a given metaphorically motivated idiom is connected to a network of semantically similar items, and also to a network of syntactically similar items. Examples (41) and (42), for example, both represent highly entrenched instances of more general schemata, both semantically and syntactically.

(41) You're pulling my leg.

(42) You're yanking my chain.

The two utterances also share a more general schema: again, both semantic (HAVING A JOKE AT SOMEONE'S EXPENSE IS TO PULL SOMETHING OF THEIRS would seem to be as accurate a statement as any of this relatively unproductive schema) and syntactic (where

'You're pulling/ yanking my leg/chain' is prototypical', but a great deal of lexical variation remains possible with meaning intact). [15]

Where a strong (relative to the strength of the specific mapping) metaphorical schema, or priming of the metaphorical schema, causes activation of an overarching conceptual schema, the idiom in question is analyzed on the basis of qualities shared with semantically similar utterances, and the figurative meaning of the idiom is distributed out over internal elements. When, on the other hand, the same factors come into play for a relatively fixed syntactic schema instantiated by an idiom, an idiom is analyzed primarily on the basis of its connection to syntactically similar utterances, and conceptual schemata are eclipsed by syntactic ones, leading in many cases towards a shift in meaning.

(43) Shit or get off the pot.

(44) Fish or cut bait.

The figurative meaning of example (43) (with its decidedly un-mysterious etymology) means something along the lines of 'do something, or get out of the way.' (44), on the basis of its constructional similarity to (43), is generally analyzed as meaning the same thing, despite saying something rather different: 'be helpful in some way' (i.e., if you're not going to fish, could you cut bait for those of us who are?). This is a result of

[15] In this case, (42) seems in fact be a coinage on the basis of the pattern established by much older (1888 is first usage in its current sense cited by the OED) example (41). The OED cites the 1975 usage, in the Washington Post, "If he told you his elephant story anywhere but his own home, you might think, as he puts it, that he was 'pulling your chain'". By 1986, we have today's form intact: "Looks like somebody out there's sure trying to jerk our chain, doesn't it?"

a syntactic schema eclipsing a semantic one, in this case attributable in large part to the weakness of the connection between the utterance and the image schema that motivates (or, it may well be more accurate to say in this case, once motivated) its meaning.[16]

Something similar happens with (45) and (46):

(45) Don't jump the gun.

(46) He jumped the shark.

Example (45) dates back to the beginning of the 20th century, a variation on 'beating the pistol', an expression referring to a false start in athletic events. (46), on the other hand, is far more recent: the first instances in the Corpus of Contemporary American English (Davies 2008) are from 2001. The origin of the phrase lies with the 1950s television series 'Happy Days', in the autumn of which 'The Fonz' jumps a shark on waterskis. 'Happy Days' remains popular in syndication, and the expression was coined more recently in reference to a ridiculous, last-ditch effort to increase ratings for a series. Not soon afterward, however, instances of 'jump the shark' synonymous with 'jump the gun' begin to emerge, as the semantics of the expression were reanalyzed on the basis of constructions similarity to 'jump the gun'. Idiomatic meaning changes, over time, for similar reasons in exx. (47) through (48):

(47) We've got this in the hole.

(48) It's in the bag.

(49) He's in the tank for Obama.

In example (49), 'in the tank' is shown in a usage meaning 'firmly in the camp of'. The expression dates back, again, to early 20th century athletics, when 'to go in the tank'

[16] Thanks to Gabriel Waters of the University of New Mexico for examples (50) and (51).

meant to intentionally lose a boxing match. The expression was extended to sacrificing oneself for some greater good, usually in a political arena, but later came to be reanalyzed on the basis of features shared with expressions such as those seen in examples such as (47) and (48), where the *y* in '*the x is in the y'* refers to something that it's well, right, and good for *x* to be in.

The effect of syntactic priming can be found in the highly productive [Verb + ing] the (adj) [N] schema, corresponding to dysphemism for masturbation.

(50) beat the bishop

(51) jerk the chicken

(52) stroke the salami

On the basis of priming from expressions such as (50) to (52), nearly any instantiation of the schema that doesn't have an otherwise entrenched meaning (although tactile verbs are preferred in the verb slot) is interpreted as referring to masturbation: a joke driving, for example, the on-line 'Euphemism Generator', a site which returns *Xing the Y* expressions in response to the click of a button.

3.3 Are Idioms metaphorical?

The traditional view of the relationship between metaphor and idiom is that metaphors play only a historical role in motivating idiomatic meanings: that expressions such as *he spilled the beans* or *they went behind my back* are merely the fossils of dead metaphors, and metaphor plays no role in the on-line processing of idioms (Aitchison 1987, Cooper 1986, Cruse 1986, Strassler 1982, Glucksberg, Brown & McGlone 1993). At the opposite extreme, Conceptual Metaphor Theory makes the assertion that

metaphorical mappings are no less active in structuring idiomatic expressions than they are in structuring less conventionalized expressions (Lakoff 1993).

Nunberg, Sag & Wasow note that "idiomatic meanings are generally derived from literal meanings in conventionalized, but not arbitrary, ways" (p. 503). Metaphor, for many idioms, provides the motivation for idiomatic meaning, with the literal meaning of the idiom sharing event structure and thematic roles with the figurative meaning. In (53), the meaning of the idiom is not the same as the literal meaning of the string, and yet the situation to which it refers figuratively "preserves certain properties of pulling, and an affected object that participates in the idiomatic activity in a way that is similar in certain key respects to the way strings are pulled" (p. 504).

(53) pull strings

(54) kick the bucket

Hamblin & Gibbs (1999) note that even within relatively opaque idioms such as (54), the figurative meaning often preserves key aspects of the event structure associated with the literal meaning of the verb (preventing, in this case, the possibility of interpreting the idiom as referring to a slow death).

With respect to example (53), it's the invariance principle (by which those aspects of the systematicity of the source domain which are mapped to the target domain maintain the 'cognitive typology'—inferences, salient aspects, thematic roles, etc.—of the source domain, provided that these don't conflict with the target domain) which, according to CMT, provides the apparatus by which aspects of the literal meaning of the expression are mapped to the figurative meaning (Lakoff 1992). At the same time, within a family of idioms referring to the same target domain, the relationships which define the family are

structured by the application of the invariance principle to a given target domain and the central metaphor which defines it (Lakoff 1987).

The claim that metaphor motivates idiomatic meanings has been made most directly by Gibbs (1990, 1994, Gibbs & O'Brien 1990, Gibbs, Bogdanovich, Sykes, & Barr 1997), who, working within the framework of CMT, has offered several forms of experimental evidence to the hypothesis that underlying metaphors constrain the meaning of idioms: most notably, Gibbs & O'Brien 1990, which reports an experiment in which the high degree of consistency in different subjects' mental images for the same idiomatic expressions, as well as the consistency of subjects' mental images across different idioms which are hypothesized to be predicated on the same metaphor, are interpreted as evidencing the metaphors which underlie and motivate the meaning of common idiomatic expressions. Gibbs (1994) describes the analyzability of a given idiom as being closely related to its metaphorical underpinnings. The extent to which an idiom is analyzable directly follows from the degree of the saliency of elements internal to the idiom: for example, "many speakers view the phrase *fall off the wagon* as being less decomposable than *pop the question* because the meaning that *fall* contributes to *fall off the wagon* is not as salient as the meaning that *pop* contributes to *pop the question*" (pp. 278-279). The metaphors which underlie such idioms both 1) increase the analyzability of idioms by contributing to the salience of words within the idioms, because metaphorical mappings facilitate speakers' ability to find the connection between an idiom's literal and figurative meaning (p. 279), and 2) are facilitated by an idiom's analyzability, because the easier it is for speakers to separate out individual components, the easier it is for the them to

identify such components with corresponding elements from the source domain (pp. 285, 291).

Sanford (2008b) corroborates Gibbs' key claims regarding the metaphorical nature of idiom in a study that finds a phonological reduction effect on words internal to idiomatic utterances, positing such reductions as evidence of underlying metaphoricity. The study monitors subjects' performance of the stimuli used in Gibbs & O'Brien (1990) to find a shorter duration for main verbs in metaphorically motivated idioms (for example, *kick* in *kick the bucket*) relative to literal counterparts (*kick* in *kick the dog*). These findings were interpreted as evidence of metaphors underlying the idioms being active in on-line processing, with the metaphor licensing high predictability for the main verb (Sanford 2008b, p. 5):

> According to the predictability-based account of word-level reduction, words which are highly predictable based on their context carry a low semantic load, which is reflected in production by shortening (Bolinger 1991, Fowler & Housum 1987). Elaborating on this view, Gregory et al. (1999) assert that frequency-based and probability-based effects are in fact different facets of the same phenomenon: speakers' probabilistic knowledge of their own language reflected in their production of it, such that highly probable—whether said probability be due to high frequency or predictability—words are reduced in duration. Metaphors, by their nature, use the systematicity of highly cognitively structured domains to structure the conceptualization of less concrete domains. The image schema which forms the basis for an idiom such as *let the cat out of the bag* is highly structured to the extent that *let* is highly predictable based on the remainder of the string. The same conceptual metaphors, then, that Gibbs & O'Brien (1990) view as governing the high degree of consistency in speakers' mental images for idioms, should also be expected to provide a high degree of 'contextual probability' which, Gregory et al. assert, will license word shortening. Such a finding [is] consistent with research which has linked the predictability of words in idiomatic expressions to the same idioms' ease of processing (Cacciari and Tabossi 1988), an effect described by Cronk, Lima & Schweigert (1993: 69) as "the biasing context of the phrase itself."

Gibbs & O'Brien's (1990) study demonstrates that speakers *can* recover underlying metaphorical motivations for idioms, and the consistency of speakers' mental images supports the positioning of such motivations within a conceptual system shared across speakers of a language. Sanford 2008b demonstrates that underlying metaphorical schemata *are* automatically activated during on-line processing, with a corresponding effect on surface expression. Moreover, the study found a meaningful relationship between the degree of phonological reduction and its perceived metaphoricity, such that the main verbs of metaphorically idioms were reduced in proportion to their degree of metaphoricity. In a preliminary survey instrument, a separate group of subjects was asked to rate, on a numerical scale, the extent to which the utterances used in the experiment are metaphorical. These responses were averaged, across subjects, to arrive at a mean 'perceived metaphoricity' rate for each of the utterances used in the performance task. The study reports a meaningful relationship between metaphoricity and phonological reduction, such that in a regressive model, the two factors of metaphoricity and frequency (both treated as continuous variables) account for the observed reduction in the length of the main verb more accurately than frequency alone. Two separate metrics, in short, of the continuous variable metaphoricity were applied (perceived metaphoricity, and phonological reduction on the main verb). Both indicated that idiomatic utterances range along a scale from more to less metaphorical, with the relationship between the two sets of measurements indicating the viability of treating metaphoricity as a continuous variable and the viability of both approaches for measuring it.

The notion of metaphoricity as a continuous variable has direct repercussions on the debate to as the metaphorical status of idioms, pointing to a resolution shy of Lakoff's claim that conceptual metaphors are as active in idioms as in novel metaphors, but well beyond the traditional view that idioms as a class are non-metaphorical, their meanings retrieved as an irreducible whole from the lexicon: idioms can be *a little bit* metaphorical. A particular idiom's degree of metaphoricity is a function of the extent of its autonomy from the sanctioning schema, which in turn depends on the entrenchment of the specific aspect of the mapping instantiated in the idiom relative to the entrenchment of the overall metaphorical schema.

3.4 Conclusion

Any model of language seeking to account for language as it exists in the world, rather than as it exists in minds of language researchers, must account for the recurring patterns that comprise much of language. Many essential insights into syntax have emerged from the study of idiom, which exemplifies the way in which form can be attached directly to meaning at any level: not only individual morphemes and words, but also strings. Idiom can and should play a similar role for the study of metaphor, demonstrating how metaphorical schemata can be entrenched at a variety of levels, and how metaphorical schemata interact with syntactic schemata.

A usage-based approach to metaphor accounts for key features of metaphor. Bybee (2001: 7) says of morphologically complex words that "New forms can be produced by reference to existing forms, but most multimorphemic words are stored whole in the lexicon". Idioms, similarly, are ready-made metaphors: their meaning can, in many cases, be analyzed out on the basis of 'reference to existing forms', but the idiom itself, with a

set metaphorical interpretation, is entrenched discretely from an overall metaphorical mapping. The extent to which an idiom activates such a mapping is a function of the degree of autonomy for the sub-schema instantiated in the idiom itself. Idioms as a class are neither metaphorical nor non-metaphorical, but rather run along a broad continuum from highly metaphorical to non-metaphorical, based on autonomy from the overall metaphorical schema. Highly idiosyncratic interpretations for metaphorical idioms are associated with a high degree of entrenchment for the specific mapping profiled in the idioms relative to the entrenchment of the overall mapping, while idiomatic analyzability is associated with the inverse. Finally, metaphorical idioms cannot be wholly understood as highly entrenched instances of metaphorical mappings, nor can they be analyzed entirely as syntactic constructions: it is out of the interaction of these two types of schemata that the rich properties of idioms emerge, and a complete understanding of figurative idioms is possible only when this dual nature is embraced.

Chapter 4: Metaphorical Frequency

4.0 Introduction

This chapter outlines the corpus methodology used for the purposes of determining the frequency of metaphorical mappings, which is prerequisite to analysis of the experiments reported in Chapter 5.

Two methods for assessing the frequency of metaphorical mappings in a corpus are presented here, both based on the use of 'key terms', elicited in a survey instrument, as search items: a small-corpus method that codes instantiations of key terminology at multiple levels, and a large-corpus method that uses a similar method but restricts coding to a single level, and searches only for simile forms. The small-corpus method uses a corpus of 3,749,000 words—large by many standards, but the high frequency of metaphor in discourse notwithstanding[17], the relatively low frequency of individual metaphors in discourse makes a corpus of this size necessary for making meaningful assessments of metaphorical frequency. This corpus is small, at any rate, relative to the

[17] Sanford 2008a asserts, based on a study that hand-coded ten conversations from the Santa Barbara Corpus of Spoken American English (Du Bois, Chafe, Meyer, & Thompson 2000, Du Bois, Chafe, Meyer, Thompson, & Martey 2003, Du Bois & Englebretson 2004, Du Bois & Englebretson 2005) for all instances of metaphor, that one out of every 11.9 intonation units is or contains a metaphorical utterance. An intonation unit is defined by DuBois et. al (1993: 17) as "a stretch of speech uttered under a single, coherent intonation contour", and appears (generally) as a single line of text within the corpus.

much larger (385+ million-words) corpus used in the large-corpus method. The large-corpus method is used as the basis for analysis of the experiments presented in Chapter 5; both methods are presented here as exploratory contributions to best practices for determining the corpus frequency of metaphorical mappings.

4.1 Metaphors Used in the Study

For the purposes of the experiment outlined below, it is desirable to use a group of underlying metaphors that fit the following three criteria:

1) Each metaphor is generally accepted as a productive cognitive entity that sanctions instantiating metaphorical utterances.

To this end, all metaphors used are taken from the Master Metaphor List (MML) (Lakoff, Espenson, & Schwartz 1991). The MML is described by its compilers as an "attempt to compile in one place the results of metaphor research since the publication of Reddy's The Conduit Metaphor and Lakoff and Johnson's Metaphors We Live By. [The] list is a compilation taken from published books and papers, student papers at Berkeley and elsewhere, and research seminars" (p. 1). While the list is far from complete (comprising, by the best estimates of its compilers, about 20% of conceptual metaphors currently reflected in the English language), those entries which are included represent conceptual metaphors that are widely accepted as such within the community of metaphor researchers, and which are relatively uncontroversial with respect to their formulation[18].

[18] 'Relatively' is the key term. The formulation of labels for cross-domain mappings is notoriously contentious; it's unlikely that there is *any* wholly undisputed mapping in the literature (see footnote 8 in chapter 2 for one of many possible examples). Many readers

A sample entry from the MML follows:

BELIEFS ARE BEINGS WITH A LIFE CYCLE
The belief lives on.
That belief died out years ago.
That belief was born of the early philosophers.
Source Domain: beings, life cycle
Target Domain: beliefs
Special case 1: BELIEFS ARE PLANTS
1. Development of a Belief is Growth of a Plant
This is just the seed of a belief.
This belief stems from my basic morality.
This belief is an offshoot of my faith.
This belief has been growing in me for years.
This belief has taken root in my mind.
This is a flourishing belief in this culture.
I planted the belief in his mind.
2. Basis of a Belief is Rooting of a Plant
Related metaphors: MAINTAINING EXISTENCE IS SUPPORTING
This is a deeply rooted belief.
This belief is rooted in fact.
3. Encouraging a Belief is Cultivating a Plant
I cultivated a belief in my infallibility among my subordinates.

The entry is typical in that it provides a statement of a metaphoric at a general level, then moves on to providing 'special cases': sub-metaphors, essentially, of the overall metaphor. Examples are provided at each level.

2) All of the metaphors exist at approximately the same level with respect to metaphorical hierarchies.

Families of metaphors are characterized by hierarchical relationships, such that many metaphors have sub-metaphors. The basic-level metaphor EMOTIONS ARE ENTITIES WITHIN A PERSON ('I was filled with rage'), for example, encompasses another metaphor whereby emotions are conceptualized as liquids within a person ('he poured out his

---

will disagree with the formulation of individual metaphors posited by the MML. The use of mappings proposed by the MML for the purposes of this study is advanced as the least of many evils, rather than as a perfect solution.

hate'). EMOTIONS ARE LIQUIDS itself has the special case of emotions being viewed as liquids in the eyes ('his eyes were full of love'). Thus:

Level 1: EMOTIONS ARE ENTITIES WITHIN A PERSON

Level 2: EMOTIONS ARE LIQUIDS

Level 3: EMOTIONS ARE LIQUIDS IN THE EYES

This property of metaphorical systems, also demonstrated by the **BELIEFS ARE BEINGS WITH A LIFE CYCLE** metaphor (above), is consistent with schemata more generally. Describing phonological schemata, Bybee (2001: 32) writes that

> schemata may be formed at many different levels of generality. The representation of particular word, such as *send*, would be a very specific or local schema. A schema for the rhyme *–end$* is at a more general level of representation. Then there could be a more general schema for *–Vnd$,* and a still more general – *vowel-nasal-voiced stop$*, or even more general *–vowel-sonorant-stop$*, and so on. The presence of any of these levels of generality for a schema does not preclude the existence of others.

A specific metaphorical utterance can instantiate any number of schemata, with more local schemata themselves instantiating more general ones. For a meaningful comparison across all of the metaphors used in the experiments, it would seem to be desirable for all of the metaphors used to exist at a similar schematic level. All metaphors participate, to some extent, in more general metaphorical systems. To as great an extent as possible, however, all of the metaphors selected for use in the experiments are relatively isolated, neither instantiating more general metaphors, nor having clearly patterned special cases.

3) The metaphors come in matched pairs that can be meaningfully compared.

The independent variable is, in all three experiments, frequency. A categorical analysis of the experiments demands that the metaphors used as the basis for the stimuli come in

matched pairs, as close to one another as possible, with each pair comprising a more and less frequent metaphor (see below for pairs used; the methods used for determining frequency, as well as the frequencies for each of the metaphors used in the study, are presented in §4.2).

On this basis, pairs of metaphors were selected which share a single target domain (e.g., WAR), but draw on different source domains (e.g., COMPETITION, RACING). The approach allows a direct comparison of stimuli, instantiating each of the mappings, that are identical except for a single word or phrase, the difference determining whether the stimulus instantiates a more or less frequent metaphor (e.g., *he argues like a soldier* vs. *he argues like a racer*).

Based on these three criteria, the ten metaphors selected for use in the study are as follows (examples from the Master Metaphor List are provided below each entry):

Competition

COMPETITION IS WAR

· *The debate team brought out their big guns.*
· *The other team sent in the cavalry against us.*
· *They battled over the chess board every week.*

COMPETITION IS RACING

· *The arms race.*
· *We've been playing chess for years, and he's finally pulling ahead of me.*
· *He's a better writer than I am, and he's widening the distance.*

Hope

HOPE IS A CHILD

· *He fostered hope that the project would continue.*
· *She nourished the hope that he would return.*
· *I'm nursing a hope for a better life.*

HOPE IS LIGHT

· *The clouds were a glimmer of hope that rain might come.*
· *He has bright hopes.*
· *I have a very dim hope that he'll recover.*

Ideas

IDEAS ARE WRITING

· *The mind is a wax tablet.*
· *His words didn't register.*
· *I made a mental note of it.*

IDEAS ARE FOOD

· *His idea was half baked.*
· *Let me chew on that for a while.*
· *It'll take some time to digest that information.*

People

PEOPLE ARE BATTERIES

· *I'm all charged up and full of energy.*
· *He got a charge out of it.*
· *I need a jump start on Monday mornings.*

PEOPLE ARE PLANTS

· *She's a late bloomer.*
· *He's a budding artist.*
· *She's let herself go to seed.*

The Mind

THE MIND IS A MACHINE

· *He has a screw loose.*
· *I could see the wheels turning.*
· *He churns out ideas.*

THE MIND IS A BODY

· *His mind is strong and supple.*
· *In the summer, the mind tends to go flabby.*
· *His mind is decaying.*

4.2 Corpus Methods

For the purposes of the three experiments, it was necessary to determine the overall frequency of a number of metaphorical mappings. The only fully satisfactory method for doing this would be to manually search a corpus for all examples of metaphorical utterances, counting the number of instantiations of each mapping. Any corpus small enough, however, to realistically allow for this approach, is also small enough that the findings will be highly skewed by the content of the corpus. The Santa Barbara Corpus of Spoken American English, for example (a corpus of about 249,000 words, at the upper

limits of what might be realistically hand-coded), comprises 60 discourse segments—mostly informal conversations, but also including other types of events at varying levels of formality. In a corpus of this size, many possible target domains are under- or un-represented, while others are over-represented. One file out of the 60, for example, deals with the work of being a farrier. The ratio of 1:60 presumably far exceeds the actual frequency with which the topic of shoeing horses is discussed in the English language overall. Accordingly, metaphors whereby, for example, horses are referred to as children ('horses … just haven't been disciplined enough. They're just, it's like a kid, they're just… are ornery', SBCSAE 1, lines 854 – 860), or regularly occurring care of a herd is referred to as a 'seasonal dance' (SBCSAE 1, line 930), are over-represented as well. The larger the corpus, the more problems such as these are mitigated.

The increase in corpus size, however, is accompanied by an increasing reliance on automated searches in order to access the corpus. In the case of determining the frequency of a given metaphorical mapping, the problem becomes (given the high productivity of some metaphorical mappings, licensing possible instantiations in any number of linguistic expressions) finding the right search terms. What words ought to be used, for example as search terms in seeking PEOPLE ARE ANIMALS in a corpus? An arbitrary selection from the wide variety of words and expressions that can be used to refer to animals (gorilla, that mutt, Spike) and to people (him, the bartender, Stacey) will provide only an extremely narrow sampling of the domain. The unevenness of metaphorical mappings (see §2.3.5) also means that any given selection is likely to yield a skewed impression of the overall frequency of a mapping.

One possible strategy is to use generally recognized key terms which are often carried over from a given source to a given target. Deignan (2005: 75), for example, adopts the following strategy:

> For my corpus investigation of ARGUMENT IS WAR, I took the following key words from the description in *Metaphors We Live By* [Lakoff & Johnson 1980]: *attack* (noun, verb), *defen(se)*, *defend*, *shoot down*. I also examined citations for the words *(to) fire*, *battle*, and *guns*, which emerged as significant collocates of the key terms in the corpus, as well as inflections of all these words.

Difficulties arise when the goal is to compare results for two (or more) different metaphors, without ending up with an apples-and-oranges comparison. With the search results depending so critically on the choice of search terms, observed differences may reflect, more than any other variable, how well the researcher has plumbed their own intuitions about the basic terms of the source domain.

The goal here, accordingly, is to arrive at objectively achieved search terms to be used in assessing the corpus frequency of metaphorical instantiations of each metaphor. Metaphorical terminology is drawn from the source domain, and so it is source domain terminology which must be used as search terms. A preliminary study, outlined in the following section, is used to attain basic terminology for each of the Lakovian source domains used in the experiments[19]. Two different corpus approaches, while varying in

[19] As Deignan (1995, 2005) points out, terms which are basic to a domain are not necessarily identical to those which are most frequently used metaphorically. Nonetheless, the extent to which those terms most basic to a domain are used metaphorically is hoped to provide a rough measure of the overall frequency with which a given domain is activated as a source.

overall approach, basic terminology selected, and information coded, will share the method of 'fishing' for source domain terms in a corpus. The more instances found, for each domain, of instances of the 'key terminology' for that domain instantiating the metaphor being sought, the more frequent the metaphor is considered to be.

Of the two corpus methods that are presented in the following sections, it is the large-corpus method outlined in §4.2.3 that is used to analyze the results of the experiments reported in Chapter 5. An overview of both methods is presented here, because 1) important insights regarding metaphorical mappings are gleaned from the small-corpus method, especially relating to the overall frequency of terms from a given domain being used metaphorically, and 2) both approaches represent valuable contributions to the field of corpus research on metaphor. While the large-corpus, simile-based method returns results more suited to the study presented here, it is lacking in some regards that are addressed by the small-corpus approach. Methods drawn from both approaches will be valuable in future research on a usage-based approach to metaphor, and to corpus research on metaphor more generally.

Preliminary to both methods is a survey used to ascertain search terms.

4.2.1 Key terms timed survey

A timed survey task was used to arrive at 'basic' concepts and vocabulary for each of the ten metaphorical source domains, to be used as search terms in each of the frequency methods outlined below. The approach is based on both the methods and theoretical underpinnings of classic prototype theory (Rosch & Mervis 1975, Rosch 1978), which maintains that when subjects are asked to list features or examples of a category, those features/examples which are listed most frequently, and which tend to occur higher on

lists, correspond to those examples/features which are closest to the core of a prototypically defined category[20]. The guiding assumption of the approach used here is that the terms which are listed most frequently by participants correspond to concepts that are basic to speakers' understanding of a given domain.

40 students, undergraduates enrolled in introductory coursework in Linguistics at a large research university and offered a small amount of extra credit in exchange for participation, were used as participants in the task. Participants were all adult native speakers of American English.

The survey instrument began with the following instructions, adapted from Rosch & Mervis (1975):

> This is a very simple experiment to find out the things that people think are most important in defining a concept. For example, for *bicycles* you might think of things they have in common, or that you associate with them, like wheels, handlebars, riding, or Schwinn. For *dogs*, you might think of things like barking, fur, or terrier.
>
> There are 10 pages that follow, and each one has a word at the top that describes a concept. For each page, you'll have a minute to write down all the words you can think of. Any word that you think is an important part of the concept is a good answer. You can write down anything that you think of, but try not to just free associate- for example, if bicycles just happen to remind you of your father, *don't* write down *father*.

[20] A similar method has been previously applied to finding words basic to a concept in studies such as Diaz-Guerrero, Rogelio, & Szalay (1991), which adapted the methodology to analyze cross-cultural conceptions of race.

> Okay- I'll give you two minutes for each word. When I say 'go to the next word', read the word printed in bold and write down words that you think are an important part of that concept. Write as many words as you comfortably can, until you're told to go to the next page.

The following ten pages, presented in one of five different random orders to each participant, each bore a heading corresponding to one of the ten metaphorical source domains used in the experiments (COMPETITION, RACING, LIGHT, CHILDREN, FOOD, WRITING, PLANTS, BATTERIES, MACHINES, and THE BODY). The instructions were read aloud, and then participants had two minutes to write a relevant list for each domain.

In analysis, those words occurring on the most participants' lists for each domain were isolated. For LIGHT, for example, the five terms which occurred on the most participants' lists were, in descending order of frequency, *sun*, *bright*, *bulb*, *lamp*, and *dark*. 'Ties', those items occurring on an equal number of participants' lists, were broken on the basis of which term averaged a higher ranking across all subjects' lists. Words differing only in inflection were considered to be the same word, and the most common inflectional form was used in the experiments[21].

4.2.2 Small-corpus Method

The corpus used for the key terms method was a combined corpus, comprising the Michigan Corpus of Academic Spoken English (MICASE) (Simpson, Briggs, Ovens, &

[21] e.g., *leg* and *legs* both came up under 'The Body', and were counted as the same word—the plural form was used in the experiments because it was more frequent.

Swales, 2002), a 1.8 million-word corpus of interactive and monologic speech, occurring with an academic setting, from a variety of academic disciplines at the University of Michigan; the Buckeye Corpus (Pitt, Dilley, Johnson, Kiesling, Raymond, Hume, & Fosler-Lussier, 2007), a 300,000 word-corpus comprising 40 speakers conversing freely with an interviewer; parts I, II, III, and IV of The Santa Barbara Corpus of Spoken American English (Du Bois, Chafe, Meyer, & Thompson 2000, Du Bois, Chafe, Meyer, Thompson, & Martey 2003, Du Bois & Englebretson 2004, Du Bois & Englebretson 2005), a 249,000-word collection of recordings of people from a variety of locations within the United States engaged in natural discourse in a variety of genres; and the Switchboard Corpus (Godfrey & Holliman 1997), a 1.4 million-word collection of about 2,400 two-sided telephone conversations on set topics among speakers from across the United States. The combined corpus, then, was 3,749,000 words. This size corpus was optimal for providing a large enough sample to be useful (representing a wide variety of speakers, registers, and discourse genres) while also returning a small enough number of tokens that they could reasonably be hand-coded.

For the purposes of the small-corpus method, the five most frequently occurring words from the key terms experiment, for each metaphorical source domain, were used as search terms (for example, the search terms used for WAR were *death*, *gun*, *bomb*, *soldier*, and *tank*). Coding took place at three levels. Search returns were initially coded as to whether or not they were consistent with the domain of interest. This subset of the returns was then coded as to whether or not the usage was metaphorical. A final round of coding identified, from all metaphorical uses of the key terms, usages that instantiate the target metaphors. Each category is explained in more detail below.

The category ‘Total uses consistent with the source domain’ is a tally, for each search term, of uses that are used to evoke (whether literally or figuratively) the source domain of which they are being used as a diagnostic.

One function of the category is to exclude, for polysemous forms, everything other than the sense being sought. For ‘tank’, one of the search terms for WAR, the category excludes items such as example (55), while allowing for items such as example (56).

(55) I think it was an old, see the oil here? Isn't that an oil **tank**?
SBCSAE Text 29, lines 148 – 150

(56) well they were and in fact they were so well dug in that uh when we sent our **tanks** around behind them they couldn't get turned around in time
Switchboard Text 2130, lines 135 - 136

The category is also applied rather critically in the case of search terms that have a semantic range which is not limited to the domain the frequency of which is being assessed. For ‘green’, one of the key terms for plants, the category excludes uses such as example (57) in which ‘green’ is used in a way that has have nothing to with PLANTS. (58), on the other hand, is included as an instance of the search term which is consistent with the domain.

(57) actually they came out with a European racing **green** uh Miata that's a limited edition
Switchboard Text 2526, lines 272 - 274

(58) …everything's in bloom you know everything's **green** here till about October and then we lose it all…
Switchboard Text 2307, lines 182 - 183

Example (56) is tallied under ‘Total uses consistent with the source domain’ for WAR, example (58) under the same category for PLANTS.

The category ‘Total metaphorical uses of source terms’ is a subset of ‘Total uses consistent with the source domain’. The category tallies, for all instances of a search term

that are consistent to the domain for which information is being sought, those uses which are used metaphorically. Criteria for metaphoricity are taken from Sanford 2008a[22]:

i) the systematicity of one domain is applied to another domain,

ii) an incongruity exists between domains, and

iii) a literal interpretation of the utterance (in a sense compatible with context) is not possible.

This category counts all instances of the search term invoking the relevant source domain, irrespective of the target to which is being applied.

(59) …yeah I'm sure they will I mean with uh the way things are going I mean uh it's going towards a more **green** type of thing…

Switchboard Text 4107, lines 127 – 129

(60) …actually I'm uh quite a beginner I'm very **green** I just got my first set of clubs this year for Christmas but I have uh gone out to the driving range once or twice…

Switchboard Text 4260, lines 28 - 29

In (59), the source term 'green' is being used as a lexical metaphor that evokes PLANTS, with a conventionalized meaning tied to environmental sustainability. (60) is tallied under 'Total metaphorical uses of source terms' (as well as 'Total uses consistent with the source domain') for PLANTS.

(60) is similar to (59), but here the conventionalized meaning ties people who are novices at something to a young, developing, or unripened plant or plant structure. The

[22] This dissertation argues directly against a view of metaphor as a binary category, arguing that metaphoricity is, rather, graded in nature, with no line separating metaphorical from non-metaphorical utterances (see §2.3.6). The criteria here are used entirely as operational criteria for the purpose of the corpus component of the study.

source of 6 is PLANTS and the target is PEOPLE. Accordingly, (60) is tallied under 'Total instantiations of target metaphor' for PEOPLE ARE PLANTS, while (59) is not.

Table 3.1 presents the results of the small-corpus method for each category. 11,910 tokens were coded in the small-corpus study; of these 7,525 were consistent with the source domain therefore coded for metaphorical status.

SOURCE	Total Uses Consistent with Source Domain	Total metaphorical uses of source terms	Total instantiations of target metaphor		
WAR	67	11	**COMPETITION**	**IS WAR**	0
RACING	37	10		**IS RACING**	10
LIGHT	387	66	**HOPE**	**IS LIGHT**	7
CHILDREN	496	3		**IS A CHILD**	0
FOOD	1300	18	**IDEAS**	**ARE FOOD**	0
WRITING	1575	3		**ARE WRITING**	2
PLANTS	1085	16	**PEOPLE**	**ARE PLANTS**	11
BATTERIES	1	0		**ARE BATTERIES**	0
MACHINES	1260	24	**THE MIND**	**IS A MACHINE**	0
THE BODY	1317	247		**IS A BODY**	4

Table 4.1: Small-corpus method results by domain

Results for all search terms are combined into a single figure for each category. The 67 tokens that are tallied under 'Total uses consistent with the source domain' for WAR, for example, is a total of tokens found for each of the five search terms that were used for the domain WAR. The full results, by individual search terms, can be found in Appendix 1.

The right-most columns of Table 3.1 provide frequencies, for each of the ten target metaphors, of key terms from each source domain being used in way that instantiates the metaphor. The search terms are being used to 'sample' the corpus, and the method does not aim to find all tokens of each metaphor in the corpus. This figure is an indicator, rather than a direct measure, of the frequency of each metaphor in the corpus.

For all five source domains, there are 0 tokens of the source being applied to one target domain, and between 2 and 11 tokens of the source being applied to another (i.e.,

the number of tokens for each of the ten metaphors is between 0 and 11). This in an interesting pattern, and points to two things: one is the possibility that the approach is not sufficiently fine-grained to detect metaphors at relatively low frequencies, and that the use of more search terms and/or a larger corpus would be desirable in achieving more precise results. Two, a corpus approach reveals that many of the metaphors *thought*, by metaphor researchers, to be common, are in fact difficult to detect in spontaneous use. This finding reflects those found in Sanford 2008a, which hand-coded ten conversations from the Santa Barbara Corpus of Spoken American English and found Lakovian conceptual metaphors for emotions to be almost completely unattested, and underscores the necessity of grounding metaphor research in empirical methods.

4.2.3 Large-corpus method

In the second corpus approach, a much larger corpus was searched for similes instantiating the target mappings. Similes are defined here as *x is like y* expressions that fit the criteria for metaphor offered above. The major advantage of using similes in trawling corpora for figurative language is their eminent searchability: the addition of *like* to figurative search terms creates a much smaller set of returns, as well as a set of returns which contains a far higher percentage of target mappings. As an example, a search for *shark* in the 385+ million-word Corpus of Contemporary American English (Davies 2008) yields 3383 hits, of which only a small portion can be expected to be figurative— a sampling of the first 100 hits yielded four figurative uses. *like a shark*, on the other hand, yields 62 hits, 60 of which are figurative. The approach therefore makes a far larger corpus accessible for study.

Several lines of research support the idea that metaphors and similes are functionally equivalent: the career of metaphor hypothesis (Bowdle & Gentner 2005) described above, for example, suggests that both similes and novel metaphors are processed analogically. Chiappe & Kennedy (2000) argue that the two forms are functionally equivalent when they are not purposefully used to contrast one another. In experimental research that does argue for a difference between the two as to processing (Aisenman 1999, Haught 2005), subjects' preference for one form over another seems to be more dependent on the specific source domain terminology used, rather than on the cross-domain mapping itself. The view is taken here, following Conceptual Metaphor Theory, that both similes and metaphors prompt cross-domain mappings (Lakoff & Johnson 1980).

The corpus used for this method was the Corpus of Contemporary American English (COCA) (Davies 2008), a corpus of 385 million words. With respect to content, the corpus is equally divided across the years spanning 1990 to 2008, and also between spoken language, fiction, popular magazines, newspapers, and academic texts.

The simile form favors the use of nouns as source terms. In the similes method, the six most frequently occurring nouns from the key terms experiment, for each metaphorical source domain, were used as search terms (for example, for WAR the search terms used were 'soldier', 'gun', 'bomb', 'army', 'general', and 'tank'). The search was for 'like *x*', (where *x* is the source term), with between 0 and 5 words between 'like' and *x.* This method was used to 'fish' for similes which instantiate the given source-domain mapping, with a given token coded as such provided that it met all of the criteria listed in §4.2.2 for metaphor, and could be plausibly read as an instance of the target mapping.

(61) is an example of a simile tallied as a token for COMPETITION IS WAR, based on a search for 'like… general' (general being one of the six key terms for WAR in the large-corpus method). (62), on the other hand, is not included as a token of COMPETITION IS WAR, because it does not refer to COMPETITION. The 'he' of example (62), in addition, is an army captain, meaning that there is no semantic incongruity between domains in the example (an army captain is being compared to an army general, both the target and source occurring with the domain of WAR, and more specifically, within a military command hierarchy).

(61) Huxley approached the controversy **like** a military **general**, out to upstage an older enemy.

COCA 1995, 'A sea horse for all races'

(62) He loved to repeat his orders as much as he loved to pose **like** a **general** standing in his jeep.

COCA 1990, 'Interior Landscapes'

Table 3.2 presents the results of the large-corpus method, totaling results for individual search terms within the mapping overall. Full results, by individual search terms, can be found in Appendix 1.

Target	Source	Total similes instantiating target mapping
COMPETITION IS…	WAR	111
	RACING	19
HOPE IS…	LIGHT	16
	A CHILD	0
IDEAS ARE…	FOOD	42
	WRITING	26
PEOPLE ARE…	PLANTS	244
	BATTERIES	27
THE MIND IS…	A MACHINE	38
	A BODY	6

Table 4.2: Large-corpus method results by domain

In the large-corpus method, PEOPLE ARE PLANTS vs. PEOPLE ARE BATTERIES provided the largest difference in frequency between mappings sharing a target (the former occurring nearly ten times as often), owing mainly to the high frequency of PEOPLE ARE PLANTS. All of the search terms for PEOPLE ARE PLANTS returned some results; people are compared to trees and flowers, specifically, with a very high degree of frequency (140 tokens and 68 tokens, respectively). The only mapping unattested in the large-corpus method is HOPE IS A CHILD, with HOPE IS LIGHT also coming in at a relatively low number of tokens.

4.3 Synthesis

The validity of both approaches was assessed based on the extent to which they agree in their findings. A high degree of agreement would indicate that both approaches are accurately measuring the frequency with which given metaphorical mappings are

instantiated in natural discourse, while a low degree of agreement would indicate that one of the two approaches is serving as a better measure than the other[23]. The degree of agreement between the two approaches was determined using Pearson's correlation (r) on the total number of instantiations found for each target mapping (the values in the rightmost column of both tables above). This test yielded $r = .43$, indicating a moderate degree of agreement between the two approaches (where 1 would reflect a perfect correlation, and zero a complete lack of correlation)[24].

Despite the moderate degree of agreement between the two approaches, the two approaches yielded different findings with respect to which of the two items were more frequent: for three of the five pairs of metaphorical mappings (those referring to COMPETITION, IDEAS, and THE MIND), the small-corpus method yielded a different ranking than the similes method. Given the need, for the purposes of the experiments reported in the following chapter, to categorize metaphorical stimuli into experimental and control groups based on frequency, a decision is necessary as to which method to make use of.

---

[23] Alternately, of course, this might be interpreted to mean that neither of the two methods is accurately reflecting the overall frequency of metaphorical mappings. For the reasons outlined below, the large-corpus method is taken as a better operationalization of metaphorical frequency.

[24] The test reflects a comparison of the results for mappings overall (e.g., COMPETITION IS WAR vs. COMPETITION IS RACING, rather than results for individual search terms. A comparison for individual search terms is not possible, as the same set of search terms were not used in the two approaches.

The large-corpus method is used to analyze the results of the three experiments reported in Chapter 5 for two reasons. First, the results of the large-corpus method are less dependent on the particular search terms chosen. In the small-corpus method, there is a great deal of variation in the results for each of the search terms used for a particular mapping (See Appendix 1). For IDEAS ARE FOOD, for example, 22 of the 26 instantiations found are for the search term 'bread'. That the omission of this single term could so drastically change the results suggests that the generalizability of the results for each search term is low. The results from the similes study tend to be more consistent across the six search terms used for each mapping[25]. Second, because of the larger size of the corpus and because of the higher rate of returns for figurative usages returned by the similes method, the results of the large-corpus method are far more fine-grained. The most telling case of this is that there are far fewer instances of null returns (in which the method yields no tokens of a particular mapping) for a given search term or mapping.

The small-corpus method does, however, demonstrate a utility that the large-corpus method does not. The smaller size and different approach enables coding at more levels, providing information on 1) the overall frequency of a particular domain being raised in discourse and 2) the overall frequency of a particular domain being invoked as a metaphorical source domain. By sampling a larger portion of the language, the size of the

[25] A related issue is that the findings of the small-corpus method are highly influenced by the frequency of high-frequency idiomatic expressions (reflecting entrenchment, as opposed to conventionalization). In the results for COMPETITION IS RACING, for example, six of the ten token are instances of 'track record'.

corpus used in the simile-based approach provides more powerful and dependable results than the small-corpus method. The use of a smaller corpus, on the other hand, enables coding at a level of detail which is generally not possible when a large corpus is being used.

Chapter 5: Experiments

5.0 Introduction

A series of experiments tests three key predictions of the view of metaphorical cross-domain mappings as cognitive entities upon which frequency effects operate in language. According to the view of metaphor outlined here, every time that an individual is exposed to a metaphorical utterance, it represents a token of use. 'Clouds' of tokens form around cross-domain mappings that are common in a language, as speakers are exposed to metaphorical systems in use. Schemata emerge around common metaphors as speakers generalize across tokens of use, leading to conventional cross-domain mappings and stock metaphorical interpretations. Additional tokens of use further entrench the schema. Metaphors with particular target domains in common (for example, LOVE IS MADNESS, LOVE IS A JOURNEY) are, in some sense, in competition with one another. Speakers make on-line choices, in language use, regarding what source domain to use to refer to a particular target domain. These decisions are based on a host of factors, including the concreteness of the source domain (Stefanowitsch 2005, Sanford 2008a), and are dependent on the speaker's previous exposures to each metaphor. These choices, across time and across many speakers, contribute to the overall frequency, in the language, of each metaphorical mapping.

The dependence of schema formation and entrenchment on frequency means that schematic strength is directly tied to numerical probability: if x is the set of all metaphorical utterances to which a speaker has been exposed, and y the subset of metaphors that instantiate a particular cross-domain mapping, then y over x is the

probability (in a historical, rather than predictive, sense), for that speaker, of a given metaphor occurring.

The literature on frequency effects in language (Bybee 1985, Moder 1992, Pierrehumbert 1994, Dabrowska & Szczerbinski 2006, Wang & Derwing 1994) speaks to three main effects from frequency (see §1.1.4). Accompanying an increase in token frequency (i.e., an increase in the tokens of use to which to the typical speaker of a language is exposed), there is an increase in:

1) Accessibility. The more frequent a schema, the more rapidly it is accessed. Following from this, the speed at which stimuli instantiating a schema are comprehended increases as well.
2) Acceptability. Speakers make decisions as to the acceptability of utterances based on the frequency of the utterance and/or its similarity to frequent utterances.
3) Productivity. High type frequency (in combination with the openness of the schema) determine the likelihood that the schema will be applied to new items.

The experiments reported below test these effects for metaphor, evaluating metaphorical mappings as cognitive representations that are acted upon by linguistic frequency. These experiments demonstrate an effect from frequency not at the level of lexical entrenchment for figurative meanings for individual words and expression, but for schemata that have emerged over multiple stored tokens of use from the sum of a speaker's history of language exposure. As language users encounter novel utterances, they are processed through a process of analogy to existing schemata, with proximity to and the strength of such schemata having a direct effect on how novel utterances are processed.

These effects have not previously been tested for metaphor; doing so requires some rethinking of experimental methods by which they can be addressed. The methods outlined below build on experimental approaches to confirming these effects at other levels of linguistic structure, as well as experimental work addressing other aspects of metaphor.

5.1 Experiment 1: Acceptability

The first experiment is an acceptability-judgment task in which participants are presented with sentences the meaning of which is predicated on an underlying metaphor, and asked to rate each sentence on a scale from 1-5 with respect to its acceptability as a sentence. It is hypothesized that sentences instantiating more frequent cross-domain mappings are deemed more acceptable than sentences instantiating less frequent ones. The effect assumes an approximately similar level of frequency for the instantiations themselves, accomplished here by using novel comparisons (presumed to have a token frequency of '0' in participants' history of exposure to metaphor).This is predicted to occur as a result of acceptability judgments being made on the basis of previous exposure: the more that speakers have been exposed to the metaphorical schema (a particular source-domain mapping) sanctioning a particular utterance, the more likely they are to determine a sentence instantiating the schema to be acceptable. The results of the experiment will be interpreted as supporting this hypothesis if, as a group, the stimuli instantiating more frequent metaphors are judged to be more acceptable than the group of stimuli instantiating less frequent metaphors.

5.1.1 Methods

18 students from the University of New Mexico (11 males and 7 females) participated in the experiment. The participants were native speakers of English between the ages of 18 and 30 and were offered a small amount of extra credit for their participation by the instructors of their introductory linguistics classes.

Stimuli (see Table 1) comprise 40 sentences: for each of five frequent/infrequent metaphor pairs, there are eight stimuli. Four instantiate the more frequent metaphor, applying terminology from a given source domain to a given target domain (e.g., 'When it comes to people, kids are like small plants', instantiating PEOPLE ARE PLANTS), while the other four sentences use source terminology from the less frequent metaphor ('When it comes to people, kids are like AAA batteries', instantiating PEOPLE ARE BATTERIES). Stimuli occur in matched pairs, embedding words triggering either infrequent or frequent mapping in otherwise identical sentences.

All stimuli are of the form 'When it comes to *z*, *x* is like *y*'. The simile form (i.e., the use of *like*) was used to prompt a figurative interpretation, and specifically to avoid a sentence being read as a literal categorization statement, so that numerical ratings for each stimulus reflect the acceptability of the figurative mapping rather than the felicity of the statement as a literal assertion. The functional equivalence of novel similes and novel metaphors is supported by Gentner & Bowdle's Career of Metaphor Hypothesis (2001, 2005), as well as by Chiappe & Kennedy (2000), who note that metaphors and similes have the same strength for speakers when used in isolation.

The introductory clause of each stimulus sets up a clear target domain; the main clause identifies something from the source domain (e.g., *AAA batteries*), likening it to

something analogous in the target domain (*kids*). In selecting source domain terminology, words were selected from the key terms survey which represented domain-specific terminology (i.e., words such as *fun* or *boring*, which do not have a meaning linked to any specific domain of thought, were excluded) and of which there were at least five occurrences across subjects in the key terms survey. From these, four words were selected that represented a variety of concepts from the domain- for example, in writing, the words selected represent an object on which one writes (*paper*), the act of interpreting something written (*reading*), a person who habitually writes (*author*), and a unit used in written language (*sentence*). All the terms used in the stimuli for the experiment, in that they are drawn from the key terms survey, are words representing semantic associations that the population sampled have relating to the source domain; the stimuli embed these terms, which label core concepts from the domain, into semantically coherent English sentences.

To prevent biasing the stimuli in either experimental group towards either a more straightforward or difficult-to-interpret form, all stimuli were written previous to the corpus component of the study. The researcher therefore did not know, at the time that the stimuli were composed, which metaphors were more frequent and which were less frequent.

Subjects were given the following instructions for the experiment:

> The following is a list of items. Each one is a sentence that compares something to something else. What I'm interested in is your thoughts, as a speaker of English, on how acceptable it is.
>
> You'll be ranking these items on a scale that runs between 'not acceptable' at one end, and 'very acceptable' at the other. You'd rate an item as highly acceptable if you think that it's something you might expect a speaker of English to say, and it's clear what the sentence is saying. You'd rate it as 'not acceptable' if you think that it's not a sentence that a speaker of English would ever say, and it's unclear what the sentence means.

Directions For each item below, rank it according to how acceptable you think it is. A ranking of '1' will correspond to items that you think are not at all acceptable. A ranking of '5' will correspond to items that you think are very acceptable. You can rank items anywhere along the scale, from 1 to 5.

Participation was untimed. Stimuli were presented on a two-page instrument, printed to present the stimuli in one of five different random orders to each participant.

5.1.2 Results

Figure 5.1 provides the results of Experiment 1. Averaging across both participants and stimuli, acceptability ratings were .25 higher (5% higher on the 1-5 scale) for the frequent group than for the infrequent group of stimuli. In the subject analysis, a paired t-test of mean acceptability ratings demonstrated that participants judged frequent items to be significantly more acceptable than infrequent items, $t(17) = 3.72$, $p = <.005$.

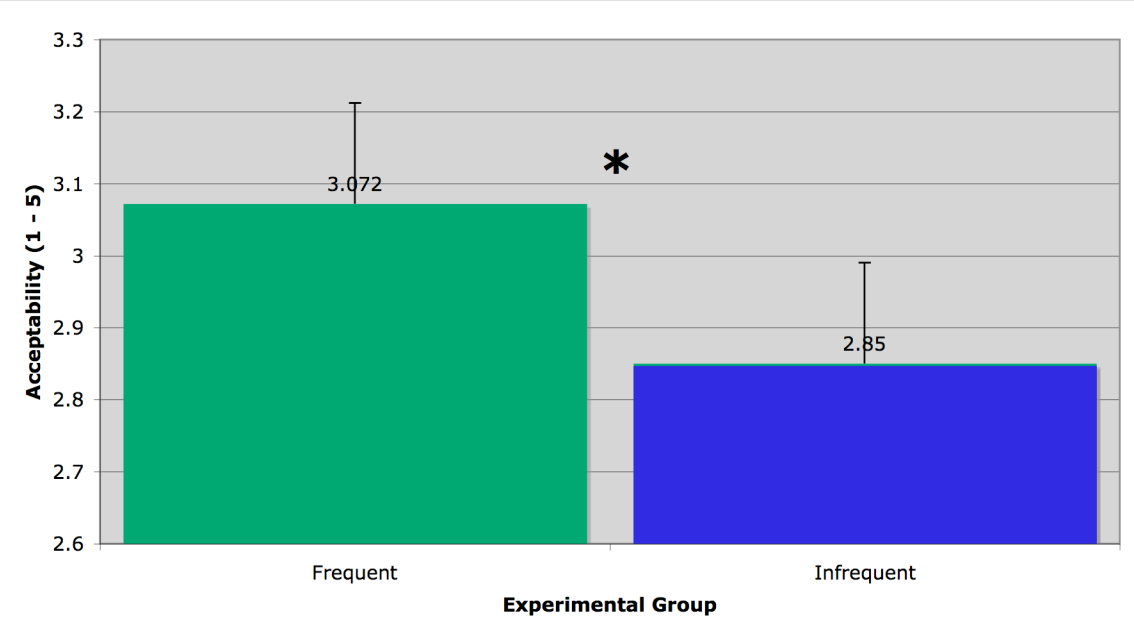

Figure 5.1
Experiment 1 Summary

Table 1 presents a more detailed view of the results of Experiment1. The table is divided into two main data columns, with more frequent mappings on the left and less frequent mappings on the right. Pairs of mappings, reflecting a more and less frequent

metaphorical cross-domain mapping, are presented side by-side, so that the more frequent mapping is on the left, the less frequent mapping in the column adjacent.

The 'stimulus' columns provide, for brevity's sake, only the main clause- the first part of the sentence is 'When it comes to *z*,' where *z* is the relevant target domain. For all ten of the stimuli in the first set, for example (relating to PEOPLE ARE PLANTS and PEOPLE ARE BATTERIES), the sentences begin with 'When it comes to people, …'. In the second set (relating to COMPETITION IS RACING and COMPETITION IS WAR) the sentences begin with 'When is comes to competition, …'. Thus, for example, the complete form of the first stimulus below is 'When it comes to people, she's like a bush.'

'Frequency' provides the corpus frequency for each source-target mapping, based on the results of the corpus component of the study. The 'acceptability' (Acc.) column provides, for each stimulus, the mean acceptability rating across all 18 participants. A higher number corresponds to a higher degree of acceptability. Stimuli are grouped into the source-target mappings that they instantiate, with the final row in each grouping reflecting an average for the mapping overall. The members of each frequent-infrequent pair appear side by side; the right-most column subtracts the 'Infrequent' values from values for the corresponding 'Frequent' values. Positive values are therefore in line with the predicted experimental effect; negative values are not. Standard deviations are reported in parentheses to the right of the average rating, across participants, for each stimulus. The frequency (F) column provides (both here, and also in the data tables for each of the other two experiments) the corpus frequency of each mapping, from the large-corpus study, in order to allow the reader to compare the mean results for each group of stimuli with the frequency of the sanctioning mapping.

Frequent			Infrequent			Frequent – Infrequent
Stimulus: When it comes to [TARGET], _____.	**Freq.**	**Acc.**	**Stimulus** When it comes to [TARGET], _____.	**Freq.**	**Acc.**	
PEOPLE ARE PLANTS		/	**BATTERIES**			
she's like a bush.		1.94(*1.1*)	she's like a Duracell.		2.17(*1.3*)	-0.23
houses are like pots are for plants.		3.94(*1.1*)	houses are like flashlights are for batteries.		2.22(*1.2*)	1.72
cities are like gardens are for plants.		3.5(*1.2*)	cities are like electronics are for batteries.		2.5(*1*)	1
kids are like small plants.		3(*1.6*)	kids are like AAA batteries.		2.78(*1.5*)	0.22
PEOPLE ARE PLANTS (summary)	**244**	**3.1**	**PEOPLE ARE BATTERIES (summary)**	**27**	**2.41**	**0.6775**
COMPETITION IS WAR		/	**RACING**			
they're like armies.		3.5(*1.4*)	they're like race cars.		3.67(*1.3*)	-0.17
winning is like killing the enemy.		3.22(*1.2*)	winning is like coming in first place.		3.83(*1.5*)	-0.61
this company is like World War 2.		3(*1.5*)	this company is like NASCAR.		2.78(*1.4*)	0.22
this place is like a battle.		3.72(*1.5*)	this place is like a race track.		3.44(*1.4*)	0.28
COMPETITION IS WAR (summary)	**111**	**3.36**	**COMPETITION IS RACING (summary)**	**19**	**3.43**	**-0.07**
THE MIND IS A MACHINE		/	**BODY**			
reading is like electricity to a machine.		3.78(*1.1*)	reading is like blood to a body.		3.61(*1*)	0.17
study strategies are like gears.		3.61(*1.1*)	study strategies are like the organs.		2.61(*.8*)	1
creativity is like the engine.		3.56(*1.3*)	creativity is like the heart.		3.5(*1.4*)	0.06
intelligent is like mechanical.		2.06(*1.1*)	intelligent is like muscular.		2.11(*1.2*)	-0.05
THE MIND IS A MACHINE (summary)	**38**	**3.25**	**THE MIND IS A BODY (summary)**	**6**	**2.96**	**0.295**
IDEAS ARE FOOD		/	**WRITING**			
he's like a cook.		2.61(*1.1*)	he's like an author.		2.78(*1.5*)	-0.17
my thoughts are like lunch.		2(*.9*)	my thoughts are like sentences.		2.67(*1.4*)	-0.67
his mind is like a plate.		2.44(*1.1*)	his mind is like a piece of paper.		3.39(*1.1*)	-0.95
going to school is like eating.		2.39(*1.3*)	going to school is like reading.		2.22(*1.3*)	0.17
IDEAS ARE FOOD (summary)	**42**	**2.36**	**IDEAS ARE WRITING (summary)**	**26**	**2.77**	**-0.405**
HOPE IS A LIGHT		/	**A CHILD**			
mine is like an illumination.		2.94(*1.1*)	mine is like a baby.		2.28(*1.1*)	0.66
my pastor is like a lamp.		3(*1.6*)	my pastor is like a parent.		3.11(*1.3*)	-0.11
his encouragement is like a light switch.		3.61(*1.2*)	his encouragement is like a child's toy.		2.28(*1*)	1.33
hers is like a bright light.		3.61(*1.2*)	hers is like a young child.		3(*1.3*)	0.61
HOPE IS A LIGHT (summary)	**16**	**3.29**	**HOPE IS A CHILD (summary)**	**0**	**2.67**	**0.6225**
TOTAL		**3.07**			**2.85**	**.25**

Table 5.1
Experiment 1 Data

Turning to judgments of individual item pairs, there were exceptions to the general pattern for two of the groups of stimuli (relating to COMPETITION and to IDEAS). In explaining these results, it seems likely that participants found some of the sentences for COMPETITION IS RACING and IDEAS ARE WRITING to be more or less literal: sentences predicated on these mappings rated high in acceptability because the terminology of competition applies literally to racing (i.e., winning isn't just *like* coming in first place, it *is* coming in first place) and because ideas are literally expressed in writing.

Large differences, for a given pair of stimuli, between the stimulus instantiating the more vs. the less frequent mappings tend to be accounted for by the higher-frequency item ranking extremely high in acceptability, as is the case for *houses are like pots for plants*, *his encouragement is like a light switch*, *and her [hope] is like a bright light* (which rank 3.94, 3.61, and 3.61, respectively, against an average rating of 2.96 across all stimuli). Those pairs of stimuli which run most contrary to the hypothesized effect are *winning is like killing the enemy*/*winning is like coming in first place*, *my thoughts are like lunch*/*my thoughts are like sentences*, and *his mind is like a plate*/*his mind is like a piece of paper*. In all three cases, item instantiating the less frequent metaphor has an average acceptability at least .5 (10% of the 1 -5 scale) higher than that for the item instantiating the more frequent metaphor. The extremely high ranking for *winning is like coming in first place* accounts in large part for the high average ranking for COMPETITION IS RACING overall, being the exemplar for the problem (outlined above) that sentences in this category tend to read as literal assertions. The second two pairs of stimuli are both members of the IDEAS ARE FOOD/IDEAS ARE WRITING group. *my thoughts are like lunch* received one of the lowest ratings in the survey rating (a 2- only *she's like a bush* ranked

lower, at 1.94). Based on informal follow-up questions, this is apparently because people felt that the interpretation was too open-ended: *lunch* doesn't assign any attributes to food, and it was therefore difficult to transfer any meaning to thought. *his mind is like a plate* was outranked by *his mind is like a piece of paper* due to the high ranking (3.39) of the latter, most likely due its similarity to the highly conventional *X's mind is (like) a blank slate.*

Figure 5.2 shows the results of the study by participant. For 15 of the 18 participants, mean ratings for the frequent stimuli were greater than or equal to mean rating for the infrequent group. The large portion of participants for which frequent stimuli were more acceptable than infrequent stimuli, as well as the closely matched curves for frequent and infrequent stimuli, attest to a clear difference across the two levels of the independent variable 'frequency'.

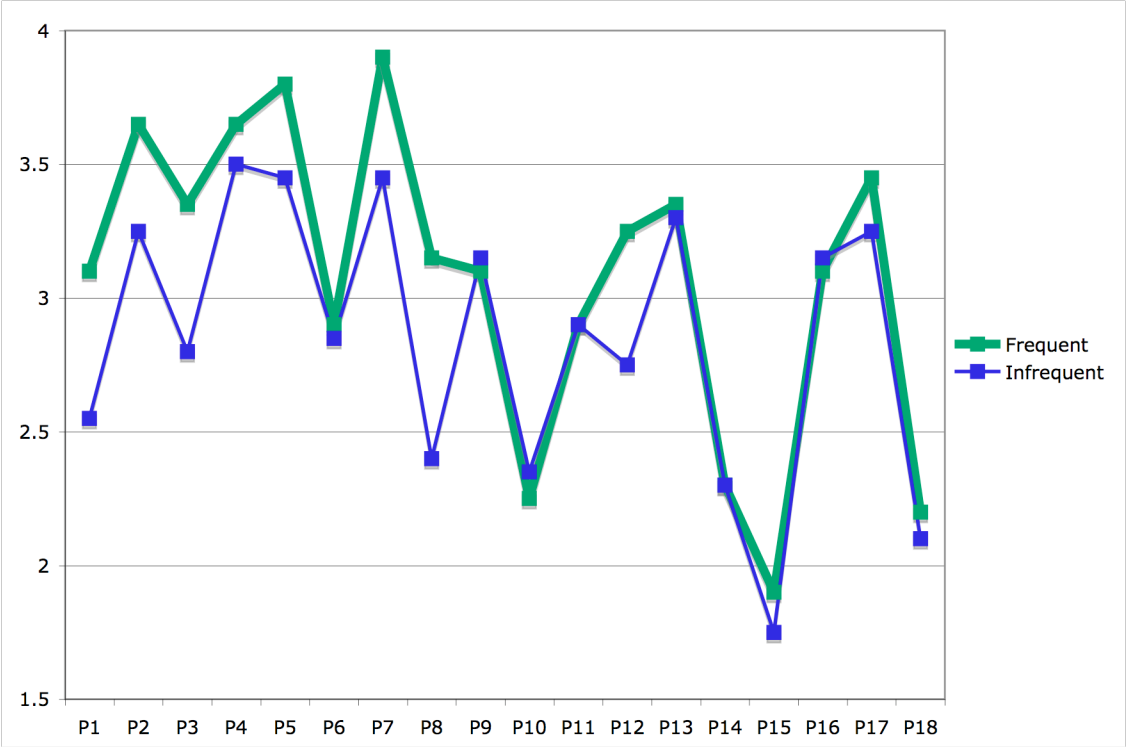

Figure 5.2
Mean acceptability rating by participant

Figure 5.3 shows the results of experiment 1 by stimulus. Within each group of metaphorical stimuli (as defined by target domain), each pair of stimuli is numbered one through four, in the order in which they appear in Table 1.

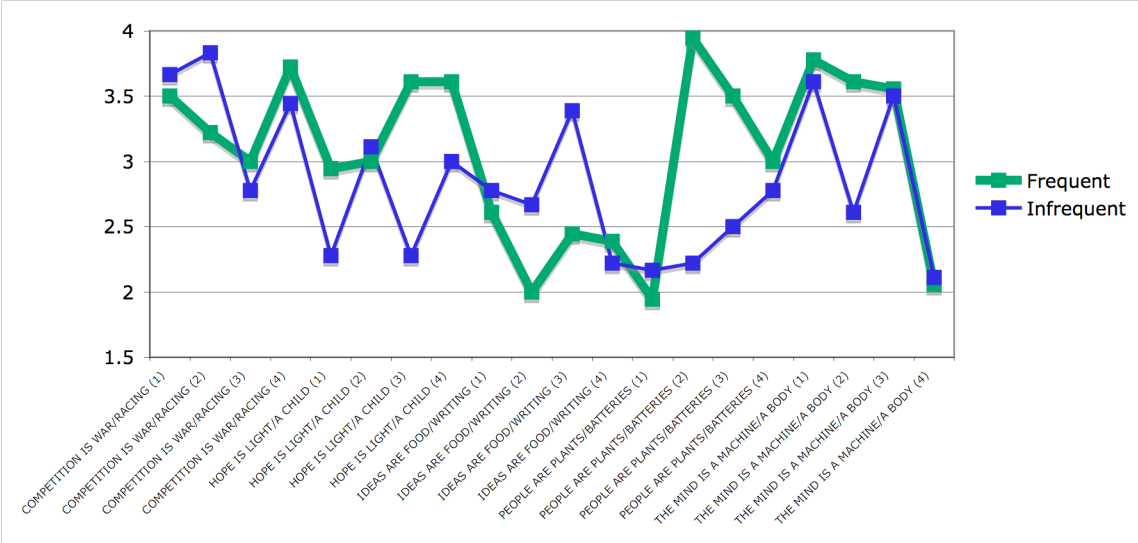

Figure 5.3
Mean acceptability rating by stimulus

For 12 of 20 stimuli, and for three of the five groups of stimuli, frequent stimuli were deemed more acceptable than corresponding infrequent stimuli. A by-stimulus analysis is not significant, $t(19) = 1.49$, $p = .08$. In these results (as in Experiments 2 and 3), a by-subject paired t-test provides a more meaningful interpretation of the results than either a comparison of averages or a by-stimulus t-test. This is the case because no threshold separates the two levels of the independent variable frequency. In each matched pair of stimuli, the frequent item is higher in frequency than the infrequent item (and as a result, the 'frequent' group of stimuli is of higher average frequency than the infrequent group). It is not the case, however, that all of the frequent items are higher in frequency, nor the infrequent items lower, than any particular figure (indeed, many of the 'infrequent' stimuli instantiate metaphors that are more frequent than other metaphors in the

'frequent' category, and vice-versa). A by-stimulus t-test is primarily a test of whether or not there is clear difference, with respect to the independent variable, between the two groups- in this case, there is not. For this reason, a by-subjects analysis is of much greater utility than a by-stimulus analysis in interpreting the data, for all three experiments.

The clear difference in acceptability between the two levels of the independent variable (frequency) observed here provides support for the entrenchment of metaphorical mappings as schema which operate on the same principles as schema posited at other levels of linguistic structure.

5.1.3 Discussion

These results are interpreted as providing strong support for the proposed hypothesis, that metaphorical utterances will be judged to be more acceptable by speakers when they are predicated on underlying metaphors to which they have had frequent exposure. The data reported here correspond to experimental findings which have been reported for other levels of linguistic structure, that grammatical judgments aren't absolute, but probabilistic, based on speakers' previous experience with language (Pierrehumbert 1994, Vitevich et al. 1997, Bybee & Eddington 2006). In the case of speakers' acceptability judgments regarding metaphorical stimuli, participants appear to be making determinations as to semantic felicity—essentially, how easily an utterance can be associated with a literal meaning—rather than grammaticality. In both cases, however, language users are making determinations as to how acceptable an utterance is based on

its probability[26]. The less frequent the pattern underlying an utterance (whether said pattern be a particular sequence of segments, a string of morphemes, or a particular source-target metaphorical mapping), and therefore the less its probability, the less such a pattern will have likelihood and strength of entrenchment as a schema. This lack of participation in a highly entrenched schema corresponds to 'oddness' in an utterance- that is to say, a lack of acceptability, whether such a lack of acceptability be labeled grammatical or semantic.

5.2 Experiment 2: Comprehension Time

The second experiment is a computer-based comprehension-time (CT) task in which participants were presented with metaphorical stimuli (e.g., 'When it comes to love, he's a magician') and asked to press a button when they understand the sentence. It is hypothesized that sentences instantiating more frequent metaphors are processed more quickly than sentences instantiating less frequent ones, as a result of frequent mappings having been entrenched in speakers' minds due to frequent activation. This hypothesis is tested in two ways, with the experimental results interpreted as supporting the hypothesis stated above if 1) there is an overall correlation between frequency and reaction time, and

[26] Semantic factors are themselves important in syntactic schemata. Bybee & Eddington (2006), in a study of Spanish verbs meaning 'become' and accompanying adjectives, demonstrate that the token frequency of constructions has a direct effect on how speakers rate the acceptability of sentences. Moreover, they show that expressions semantically similar to highly frequent ones are also judged to be more acceptable, demonstrating that semantic relations play a primary role in formation of exemplar-based representations.

2) as a group, the stimuli instantiating frequent metaphors are processed more quickly than the stimuli instantiating less frequent metaphors.

5.2.1 Methods

26 students from the University of New Mexico (14 males and 12 females) participated in the experiment, a group different from those who participated in Experiment 1. The participants were native speakers of English between the ages of 18 and 30 and were offered a small amount of extra credit for their participation by the instructors of their introductory linguistics classes. Stimuli (see Table 5.2) comprise 30 sentences, occurring in matched pairs: for each frequent/infrequent metaphor pair, there are six stimuli. Three instantiate the more frequent metaphor, applying terminology from a given source domain to a given target domain (e.g., 'Her mind has gears', instantiating THE MIND IS A MACHINE), while the other three sentences are identical except that they use source terminology from the less frequent metaphor ('Her mind has muscles', instantiating THE MIND IS A BODY). Stimuli were constructed using terminology which emerged as basic to the domain in the key terms survey, with the first pair of stimuli constructed around the metaphorical use of nouns from the source domain, the second around the metaphorical use of verbs from the source domain (although in some cases applying them in present participle form), and the third around the metaphorical use of adjectives from the source domain. The words selected for use in each matched pair were analogous with respect to their role in the domain[27]. Stimuli were controlled for mean

[27] For example, the words selected for PLANTS and BATTERIES are *flower* and *Duracell* (a type of plant, and type of battery), *growing* and *charging* (things that plants

lexical frequency, as assessed using the Corpus of Contemporary American English (Davies 2008). While the mean lexical frequencies for the members of each stimulus pair are in many cases unequal, there was no significant difference in mean lexical frequency between the two groups of experimental items overall: $t(14) = .9$, $p = .38$. All stimuli were written previous to the corpus component of the study, so that the researcher did not know which metaphors were more frequent and which were less frequent at the time of their being composed.

The experiment was computer-based, with all participants completing the experiment on the same machine (a Macintosh G4 MacBook). The experiment was designed and run using PsyScope X, and the machine's track button was used as an input device. This setup guaranteed timing accuracy to within 17ms (more than sufficiently accurate for the relatively long reaction times recorded in the experiment).

Participants first read an introductory screen on which they were advised that they would be presented with a series of sentences, and that they were going to be asked to assess each sentences' meaning. On beginning the experiment, participants viewed, for each stimulus, the following series of screens: an initial screen advises them that for the following screen, they are being asked to press a button as soon as they feel that they understand the sentence. Once they press a button to advance to the screen containing the stimulus, a screen appears containing the stimulus. Once the button is pressed a second time, indicating the sentence has been read and understood, a screen appears which prompts them to write in a brief description of the meaning of the sentence, and to press a

---

and batteries do, respectively), and *small* and *rechargeable* (traits that can apply to members of the category).

specific key (cueing the introductory screen for the next stimulus) when their description is complete.

All trials began with a block of 5 'warm-up' items, from which data were not recorded, following which the stimuli were presented to each participant in a different random order. For each stimulus, the time recorded was the time between the stimulus appearing on the screen, and the button being pressed.

5.2.2 Results

Figure 5.4 provides the results of Experiment 2. Comparing results across the two levels of the independent variable, corresponding to the average of responses for the more frequent and less frequent items for each participant, the experiment bears out the prediction that utterances instantiating more frequent metaphors are processed more quickly than utterances instantiating less frequent ones. Mean reaction time for the infrequent stimuli is 5865.2 ms, while mean reaction time for the frequent stimuli was 5371.2 ms—almost 500 ms less. These results are statistically significant for a by-subjects analysis, $t(25) = 2.11$, $p = <.05$.

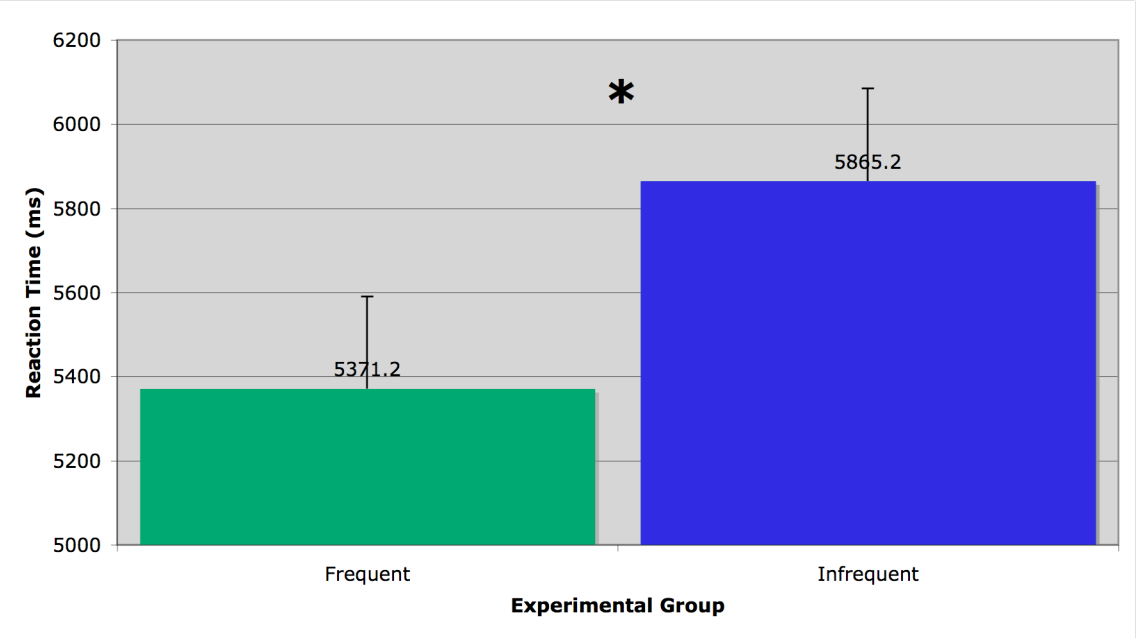

Figure 5.4
Experiment 2 Summary

Table 5.2 presents a more detailed view of the results of Experiment 2, averaged by stimulus. Both Table 5.2 and Figure 5.4 reflect a preliminary cleaning up of the data in which all responses under 100 ms (assumed to represent a 'double-click' from the previous response) and over 100,000 ms (assumed to represent a distraction from the experimental task) were excluded. Data were also excluded from items for which there was a null response in the paraphrase (i.e., items for which participants entered nothing in the field where they were prompted to explain the meaning of the sentence). A total of 35 responses (4% of the data) were discarded based on these criteria. Values for each item are averaged across participants. The values in the right-most column are differences yielded by subtracting values from the first column from values in the second. Positive numbers in the right-most column are therefore in accord with the predicted phenomenon that utterances instantiating frequent metaphors are processed more quickly than those

instantiating less frequent metaphors. Negative numbers reflect instances where the predicted effect is not borne out.

Frequent			Infrequent			Infrequent - Frequent
Stimulus	**Frequency**	**CT**	**Stimulus**	**Frequency**	**CT**	
PEOPLE ARE PLANTS/BATTERIES						
He is a flower.		5361.1	He is a Duracell.		5059.9	-301.2
She is a plant growing.		5073.2	She is a battery charging.		6203.7	1130.6
He is a small plant.		4573.7	He is a rechargeable battery.		6979.8	2406
PEOPLE ARE PLANTS summary	**244**	**5002.7**	**PEOPLE ARE BATTERIES summary**	**27**	**6081.1**	**1078.5**
COMPETITION IS WAR/RACING						
This is the Iraq of debates.		5714.6	This is the NASCAR of debates.		6107.3	392.7
The applicants are fighting in a war for the award.		5535.6	The applicants are running in a race for the award.		6859.8	1324.3
This chess match is a bad war.		5299.8	This chess match is a fast race.		4171.9	-1127.9
COMPETITION IS WAR summary	**111**	**5516.6**	**COMPETITION IS RACING summary**	**19**	**5713**	**196.4**
THE MIND IS A MACHINE/BODY						
Her mind has gears.		4521.2	Her mind has muscles.		5193.6	672.4
His mind is a machine malfunctioning.		4730.5	His mind is a body exercising.		5578	847.5
My mind is an industrial machine.		4552.8	My mind is a strong body.		4907	354.2
THE MIND IS A MACHINE summary	**38**	**4601.5**	**THE MIND IS A BODY summary**	**6**	**5226.2**	**624.7**
IDEAS ARE FOOD/WRITING						
Her ideas are well-cooked pizzas.		4360	Her ideas are poorly written essays.		6288.2	1928.2
They ate my ideas straight from my mouth.		5528.5	They read my ideas straight from my mouth.		6290.7	762.3
His idea is a tasty piece of food.		5742.9	His idea is a fun piece of writing.		7030.1	1287.2
IDEAS ARE FOOD summary	**42**	**5210.4**	**IDEAS ARE WRITING summary**	**26**	**6536.3**	**1325.9**
HOPE IS A LIGHT/A CHILD						
My hope is the sun.		7738.7	My hope is a baby.		6082.3	-1656.4
Her hope is a lamp illuminating the room.		6087.7	Her hope is a child playing a game.		5730.5	-357.2
His hope is a bright light.		5747.1	His hope is a young child.		5495.4	-251.7
HOPE IS A LIGHT summary	**16**	**6524.5**	**HOPE IS A CHILD summary**	**0**	**5769.4**	**-755.1**
TOTAL		**5371.2**			**5865.2**	**494.1**

Table 5.2
Experiment 2 Data

The results corroborate, in several ways, the results gleaned from Experiment 1. Three of the five groups of stimuli agree directly with the results of the acceptability experiment. Mean scores COMPETITION IS WAR/COMPETITION IS RACING, separated by a narrow margin in Experiment 1, are again close in Experiment 2. The final group of stimuli, those instantiating HOPE IS A LIGHT/HOPE IS A CHILD, represent a notable exception to the trend observed elsewhere: for this one group, the items instantiating the less frequent metaphor has, across subjects, a lower mean comprehension time. This may indicate that the frequency method used may not be fine-tuned enough to determine the difference in frequency between relatively low-frequency metaphors, or it may indicate that below a certain threshold of frequency, no meaningful differences in processing take place. When HOPE is excluded from the analysis, the difference between the two groups increases to 645 ms, significance to $t(25) = 2.96$, $p = <.05$. In general, for this experiment, little can be drawn from a comparison at the level of groups of stimuli that share a target domain, or certainly of individual pairs of stimuli: while there is no significant difference between the two groups of stimuli overall, with respect to mean lexical frequency or phrase length, there are differences in these factors for individual stimulus pairs (deviations that are mitigated as larger groups of stimuli are compared).

Figure 5.5 represents the data from Experiment 2 in a chart that plots, for each of the 26 subjects that participated in the experiment, mean comprehension for the frequent against the infrequent sets of stimuli. For 14 of the 26 participants, reaction time for the infrequent items exceeded reaction time for the frequent items (a slim majority, following

from the small margin separating the two groups). The difference between the two groups was significant, as reported above.

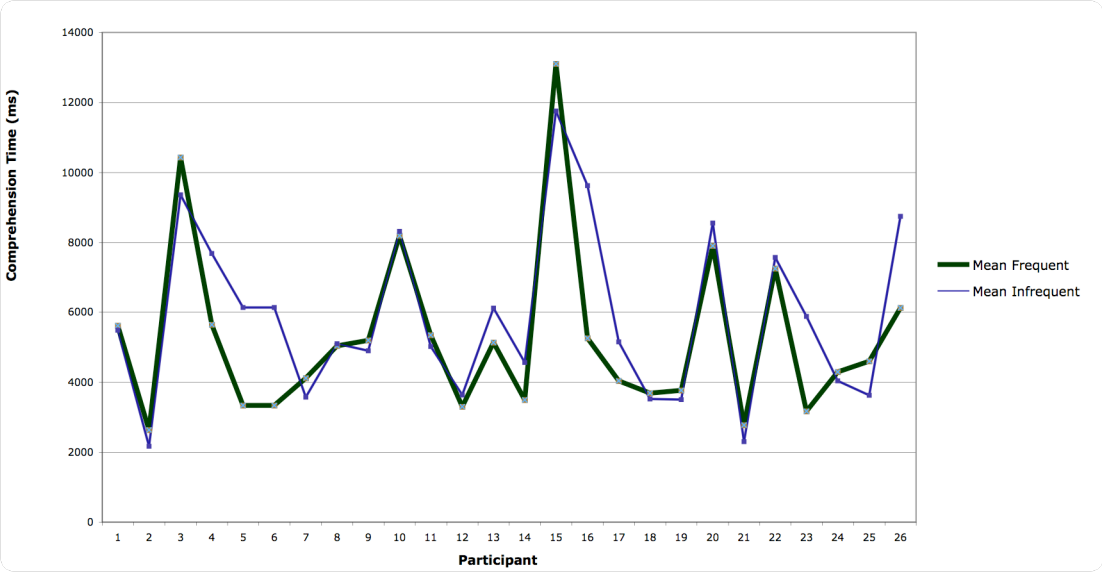

Figure 5.5
Experiment 2; Comprehension Time by Subject

Figure 5.6 plots, for each stimulus pair, the mean comprehension time across participants for the frequent vs. the infrequent stimulus.

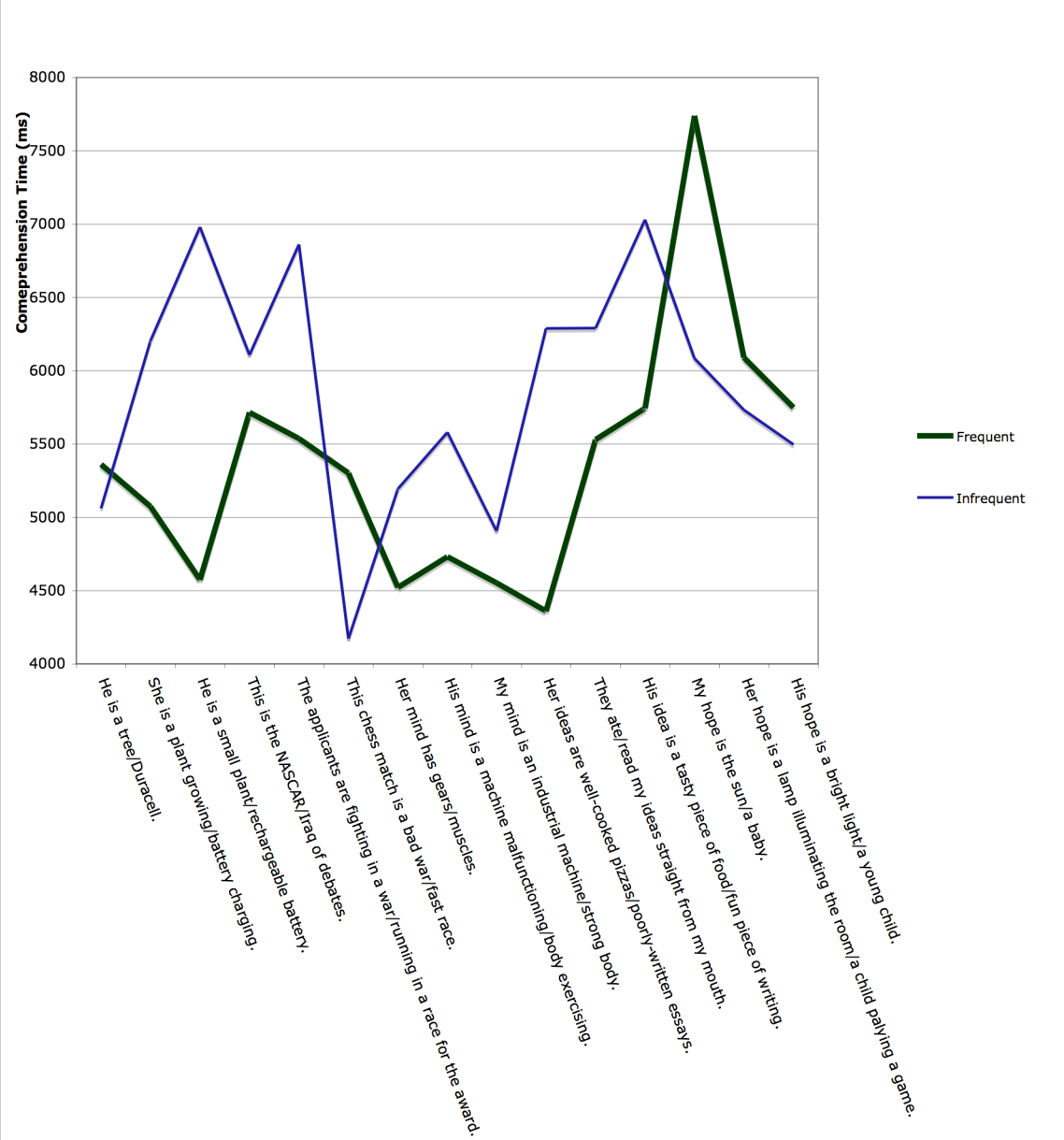

Figure 5.6
Experiment 2: Comprehension Time by Stimulus

In a by-stimulus analysis, the difference between the two groups of stimuli approached, but failed to reach, significance, $t(14) = 1.74$, $p = .051$.

5.2.3 Discussion

These results are interpreted as supporting the stated hypothesis that sentences instantiating more frequent metaphors are processed more quickly than sentences instantiating less frequent ones as a result of their having been entrenched by previous usage. These findings mirror those such as Vitevich et al. (1997) and Hare et al. (2001), which indicate for phonotactic and morphological sequences that processing speed is more rapid when such sequences instantiate patterns which are frequent. This suggests a common explanation for both sets of findings: that patterns to which subjects have had frequent previous exposure are entrenched as schemata, and that increasing frequency is accompanied by increasing speed of access. With respect to accessibility, frequency has the same effect on metaphorical schemata as it has on schemata at other levels of linguistic structure.

5.3: Experiment 3: Productivity

The third experiment is a timed sentence completion task in which participants are presented with the first portion of a sentence which specifies an entity from a given target domain, and they are prompted to write down as many metaphorical completions as they can think up in a specified period of time. It is hypothesized that completions instantiating more frequent metaphors will appear with more frequency, across subjects, than completions instantiating less frequent ones, as a result of frequent mappings having been entrenched in speakers' minds due to frequent activation. This experiment will be taken as supporting the hypothesis if there is a significant difference in the number of completions instantiating frequent vs. infrequent metaphors, such that there are significantly more completions instantiating the group of frequent metaphors identified in

the corpus component of the study relative to the corresponding related infrequent metaphors.

5.3.1 Methods

18 students from the University of New Mexico (8 males and 10 females) participated in the experiment, a group distinct from those who participated in Experiments 1 or 2. The participants were native speakers of English between the ages of 18 and 30 and were offered a small amount of extra credit for their participation by the instructors of their introductory linguistics classes. The five stimuli used in the experiment (see Table 3) correspond to the five target domains (PEOPLE, COMEPTITION, THE MIND, IDEAS, HOPE) that are attested in the ten metaphors (five matched pairs) the frequency of which were assessed in the corpus component of the study. The initial portion of the stimulus sets up a situation, describing literally an attribute of a member of a category defined by a particular source domain (e.g, PEOPLE). The second portion of the stimulus, an uncompleted simile of the form "[Target] is like ____," prompts a figurative description of the assertion made literally in the first portion of the stimuli. The first stimulus, then, reads as follows:

> Mary was tired, but she took a nap and is now feeling more awake.
> Mary is like _____.

The first portion of the stimulus makes a literal assertion, attributing a quality (awake) to a member (Mary) of a category aligning with a metaphorical source domain (PEOPLE). The second portion prompts the figurative use of a term to provide an alternate description. The information being coded for analysis is the source domain of the terms used to complete the sentence: *a budding flower* would be a use of terminology from the

source domain PLANTS (and therefore of the mapping PEOPLE ARE PLANTS, while *a recharged battery* would be an instance of BATTERIES (and therefore of PEOPLE ARE BATTERIES)[28].

Participants were provided with the following instructions:

> When people use metaphors and similes, they compare one thing to another thing. For example, if I say "My car is like an elephant", I'm comparing my car to an elephant. I might to do this to indicate that my car is large, or that it's slow moving. I could also say the same thing, or something similar, by using a different metaphor- for example, I could say "My car is like a boat."
>
> In this experiment, I'm interested in what things make good metaphors. On each of the following pages, there is a brief description of a situation. At the end of the description, there is a sentence that ends with a blank. I'd like you to try to come up with things that finish the sentence metaphorically.
>
> For example, a description of a situation might be something like:
>
> > Michael and Lisa are very much in love, and they think about each other all the time.
> > They are like _____.
>
> To make the last sentence a metaphor, you could write something like 'teammates', or 'magnets', or 'lovebirds', or 'people who are on a journey together', or anything else you can think of.
>
> When I say "Go to the next page", read the description, and then write down all the metaphorical completions that you can think of for the sentence that ends in a blank.
> Start at the top of the page, and work your way down.

[28] In an initial version of the experiment, the prompt portion of the stimulus took the form of a metaphor rather than a simile, lacking the word 'like' (i.e., "Mary is ____"). Despite instructions which directed participants towards figurative completions, metaphorical completions were almost unattested in trials of the experiment in its original form. The current version is much more successful in prompting figurative responses, which follows from and provides support for Gentner & Bowdle's (2001, 2005) assertion that the simile sentence form prompts figurative processing.

Subjects had two minutes to complete each page of the instrument. The five stimuli were presented in a different random order to each participant. Data coded are the number of items, totaling across all participants' lists, items instantiating each of the two metaphors (the frequent and the infrequent one) for each target domain.

5.3.2 Results

Figure 5.7 provides a comparison of the mean production tokens (a single instantiation, within a list of completions for a given target domain, of a particular mapping) across all participants, and all stimuli. The average number of production tokens is higher, averaging across subjects, for the frequent than the infrequent metaphorical mappings (7.8:2.8), which bears out the predicted experimental effect. These results are significant, $t(17) = 3.08$, $p = <.005$.

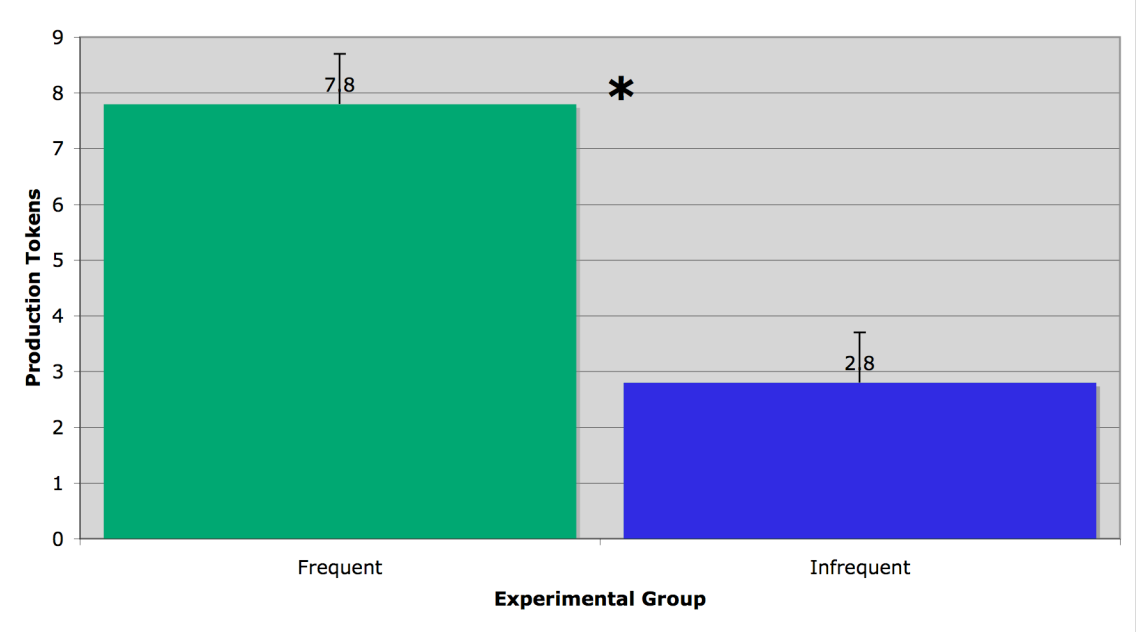

Figure 5.7
Experiment 3 Summary

Table 3 presents a more detailed version of the results of Experiment 3. There is a single prompt corresponding to each of the five target domains; the ‘frequent’ and ‘infrequent’ columns correspond to the source domains which are used more and less frequently in relation to the given target domain. The column labeled PT provides a sum, across the 18 participants, of production tokens for each mapping.

Prompt	Frequent			Infrequent		
	Metaphor	F	PT	Metaphor	F	PT
Mary was tired, but she took a nap and is now feeling more awake. Mary is like _____.	PEOPLE ARE PLANTS	244	**2**	PEOPLE ARE BATTERIES	27	**5**
John and Tyler are students, and they are extremely competitive over who is better at math. They have a test coming up, and both of them want to do better than the other one on the test. John and Tyler are like _____.	COMPETITION IS WAR	111	**2**	COMPETITION IS RACING	19	**3**
Jessica is extremely intelligent. Her mind is like _____.	THE MIND IS A MACHINE	38	**23**	THE MIND IS A BODY	6	**2**
I just had a really good idea. My idea is like _____.	IDEAS ARE FOOD	42	**3**	IDEAS ARE WRITING	26	**3**
My friends and I are in a very bad situation, but we're starting to hope that things will get better. Our hope is like _____.	HOPE IS A LIGHT	16	**9**	HOPE IS A CHILD	0	**1**
Average		90.2	**7.8**		15.6	**2.8**

Table 5.3
Experiment 3 Data

Looking at the results across each of the five target mappings, it's the case for three of the five target domains that PTs for the more frequent mapping are lower than, or equal to, PTs than for the infrequent mapping. In the two categories for which there are, as predicted, more PTs for the frequent than for the infrequent mapping, the difference across the two categories is extremely high relative to the other categories (23:2, 9:1). In the three categories for which PTs for the infrequent mappings exceed PTs for the frequent ones, the ratio is at most 5:2. This pattern accounts for the higher average PTs for the category of frequent mappings.

As an example of the types of entries that were included as tokens of the target mapping:

The prompt corresponding to PEOPLE ARE PLANTS and PEOPLE ARE BATTERIES is:

Mary was tired, but she took a nap and is now feeling more awake.

Mary is like _____.

Across all subject's lists, the two completions that were coded as instantiations of PEOPLE ARE PLANTS were 'a fresh lettuce' and 'a daisy'. The five items that were coded as instantiations of PEOPLE ARE BATTERIES were 'a recharged battery', 'a charged battery', 'a device that needs to be recharged', and two instances of 'battery'. Metaphorical completions that don't instantiate either target metaphor abound (as they do across all ten target metaphors), and responses such as the following were common: 'a new penny', 'a million bucks', 'a deer after being shot at', and 'a lightbulb'. So, too, were non-metaphorical responses, such as 'a person who had just slept for 24 hours', or 'a workaholic freak'. Across all 18 subjects and all five target domains, the average number of responses for each list was seven.

Figure 5.8 shows the results of experiment 3 by subject.

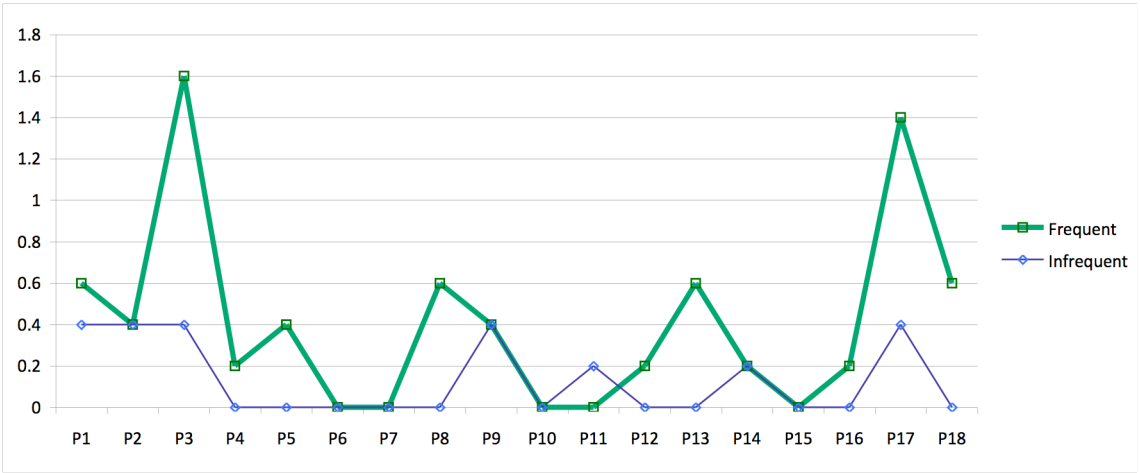

Figure 5.8
Results of Experiment 3 by subject

For 3 subjects, there were no PTs for either metaphorical category. For 7 subjects there were no responses for the infrequent mapping and at least 1 for the frequent mapping, while the reverse is true for only 1 subject. 14 subjects gave a pattern of

responses in line with the predicted effect, productivity for the frequent category exceeding productivity for the infrequent category.

A by-stimulus analysis fails to reach statistical significance, $t(4) = 1.13$, $p = .16$.

5.3.3 Discussion

The results of Experiment 3 are in line with the predicted experimental effect, providing direct support for the hypothesis that frequent metaphors are more productive than infrequent metaphors. The difference in averages between the two groups, as well as the by-subjects analysis, are highly consistent with the predicted effect that more frequent metaphors will, relative to less frequent metaphors, more commonly serve as templates for new constructions.

Overall support notwithstanding, in some cases results for specific frequent-infrequent metaphor pairs fails to corroborate the general pattern of frequent metaphors being more productive than their less-frequent ones, pointing to two weaknesses in the experimental design that should be addressed in follow-up experiments. First, the low numbers of responses for each metaphor (eight of the ten having five responses or fewer), even totaling across all participants, makes the results difficult to interpret. This issue is inherent to the nature of the study- the prompts were very open-ended, and metaphorical completions of all sorts, and in very high numbers, were attested. Only a small subset of these (those completions, for each target domain, that instantiated one of the metaphors the frequency of which was assessed in the corpus component of the study) were coded, which necessarily means that only a small number of responses were recorded. A follow-up study, or any experiment with a methodology similar to the one reported here, would do well to seek a much higher number of participants. Second, there was only a single

prompt corresponding to each target domain, which made the results for each domain highly dependent on the particular stimulus used. The results would be more powerful were they averaged, for each domain, across a number of prompts for the domain. This change will also be instituted in any follow-up studies.

5.4 Conclusion

The experiments reported here indicate that metaphorical entrenchment can and does take place at the level of underlying metaphorical schemata. The corpus frequency of a given mapping provides a diagnostic of speakers' previous exposure to an utterance, which is demonstrated here to be a factor in processing, with respect to acceptability accessibility, and production. The finding that manipulation of utterances with respect to the frequency of the metaphors underpinning them has a direct effect on processing indicates that speakers' previous exposure to particular metaphors influences their on-line processing of such metaphors. Experiments 1 and 2 directly support the predictions that instantiations of higher-frequency metaphors are deemed more acceptable and processed more quickly (respectively) than instantiations of lower-frequency mappings. Experiment 3 confirms that higher-frequency mappings are more productive than lower-frequency mappings, suggesting a strong role for the token frequency of a metaphorical mapping in determining such productivity.

Both experiments 1 and 3 make use of stimuli in the grammatical form of similes, rather than of metaphor. The functional equivalence of metaphors and similes, with respect to processing, does not mean that metaphors and similes are identical, nor should any statement made above be interpreted as claiming as such. As outlined in §2.3.4, different grammatical forms have associations with different types of metaphorical affect

and pragmatic function- the piloting of stimuli for all three experiments indicated that the simile form is more effective than the metaphor form in biasing language users towards a figurative interpretation. Croft & Cruse (2004) note that the categories of metaphor and simile overlap to a great extent, but have distinct prototypes. Similes tend towards restricted mappings, making a single, narrow assertion about the target, and prototypically treat the source and target being treated as discrete. Metaphors tend towards open mappings, inviting a limitless number of inferences about the target based on the source, and prototypically 'mix' the source and target into a single conceptual space. Nonetheless, several lines of research (Bowdle & Gentner 2005, Chiappe & Kennedy 2000) support the view that novel similes and novel metaphors are processed in essentially the same way, supporting the view from Conceptual Metaphor Theory that both forms represent linguistic instantiations of cross-domain mappings. The most relevant evidence, however, comes from the experiments reported here: Experiments 1 and 3 make use of grammatical similes as stimuli, while Experiment 2 uses grammatical metaphors. All three experiments, however, point towards the same relationship between processing for the stimuli, and the metaphorical schemata upon which they are based.

In several cases, results for specific pairs of domains point to cases in which the metaphors used in the study were not stated at a level of schematicity that accurately captures the productive range of the metaphor (a criticism elaborated in Clausner & Croft 1997), leading to stimuli that are questionable in their relation to an overarching schema. In Experiment 2, for example, results relating to the source domain HOPE run contrary to both the corpus study, and the other two experiments- apparently, because utterances such as 'my hope is the sun', predicated directly on HOPE IS LIGHT, run contrary to a

metaphor that might be more accurately described as profiling light sources as things that nourish hope. PEOPLE ARE BATTERIES might be more accurately phrased PEOPLE RUN ON BATTERIES, leading to issues with the aptness of a stimulus such 'kids are AAA batteries'. Metaphors were drawn from the Master Metaphor List (Lakoff, Espenson, & Schwartz 1991) for reasons elaborated in §4.1 (in short, if the metaphors on which the experiments in based are in error, they at least aren't flawed in a way that reflects researcher bias). The issue underscores the need for all posited metaphors to be evaluated closely against corpus usage; that the experiments found an effect despite such issues points towards a significant effect outweighing issues relating to specific groups of stimuli.

These results are interpreted as providing strong support for the view of metaphor outlined here. Speakers' repeated exposure to utterances predicated on a particular cross-domain mapping license the formation of a metaphorical representation—a conventionalized link between two domains, corresponding to 'X IS Y' Lakovian conceptual metaphors. For language users in the act of engaging in figurative speech, previous exposure to metaphorical systems has a direct, measurable effect on the way in which they use and process metaphorically predicated utterances.

Chapter 6: Conclusion

In traditional and generative models of language, the facts of language production are dependent on underlying mental representations (Chomsky 1965, 1980, Peters & Ritchie 1973, Jackendoff 1974, May 1977, Lappin, Levine & Johnson 2000). Stored units and rules (competence) together account for the facts of language as it is produced (performance), with the goal of linguistic analysis being to posit the correct rules and stored units to accurately and completely account for language as it exists in the mouths of speakers. A speaker's knowledge of language, in this view, wholly precedes any actual use of language: the abstract structure of language exists in speakers' minds, and it is this abstract structure which is the essence of language. Language, whether realized in production or not, exists as a complete system in the mind. When utterances are produced and perceived, the underlying system—capital 'L' Language—is activated in order to do so.

The perceived benefit, and indeed the elegance, of a system such as this one is its perceived cognitive efficiency. Speakers don't need to 'worry', so to speak, about remembering all of the possible combinations of morphemes and words in a language. They need not waste the storage space, because complex linguistic units such as sentences and morphologically derived words can be formed on the basis of stored units and rules. No matter how many times a given word or sentence is formed, the mind goes through the same process of putting it together based on underlying structures. In this processing-heavy approach to language, people are like calculators: instead of storing the entire multiplication table, calculators compute the necessary values as needed. Instead of memorizing all of the possible sentences in a language, humans put them together as

necessary. Just as with early computers, processing is exploited to make up for a lack of storage, and the beauty of language conceptualized thus is that a relatively small amount need be stored.

A usage-based model of language (Bybee 1985, Givón 1984, Hawkins 1994, Lindblom, MacNeilage, & Studdert-Kennedy 1983) turns such a conception of language directly on its head, by ascribing primacy to the facts of language production itself. When the regular in language is conceptualized as patterns generalized across actual tokens of use, the utterance—an actual speech event—becomes, rather than just an output of an underlying system, the unit upon which linguistic structure is predicated. Language, as it exists in a system in speakers' minds, emerges as an organic system as speakers are exposed to language in use.

Storage is highly redundant. Not only are individual tokens of use stored, but stored schemata can overlap and encompass one another, with representations for units at lower levels in no way precluding representations at a higher level. A frequently occurring string of words, for example, need not be built anew from its constituent elements with each repetition. A sentence such as 'How was your weekend,' while analyzable in terms of each of the four words comprising the sentence, can also be stored and processed as a whole. If the generative model of language is analogous to early computers, then the way that media is stored on many more recent personal computers provides a template for thinking about a usage-based model: memory is cheap. With hundreds of gigabytes of storage well within the means of a typical computer user, the existence of a particular item of information at multiple levels (for example, a single song appearing multiple times in a music library, as compilations containing the same song are uploaded) is in no

way unacceptable. Storage is similarly cheap in the mind, with humans capable of storing incredible amounts of information, and similar (and even far more extreme, in line with the extent to which human memory exceeds that of a typical computer) redundancies occurring. Such a model may lack the sparse elegance of a theory of language that draws a sharp line between rules and the entities upon which they operate, positing utterances as the output of a process whereby a regular (if highly elaborate) set of rules operates on a limited set of stored entities. It has the great advantage however, of not treating language as something set wholly apart from other aspects of cognition—and boasts its own aesthetic in a vision of language as a self-organizing system. At every level, repetition shapes how language is stored and processed. The recurring use of one domain of thought to describe another creates conventionalized connections between cognitive domains.

For metaphor, the usage-based model handles well precisely those aspects of metaphor which have raised the most eyebrows among critics of Conceptual Metaphor Theory, and of cognitive theories of metaphor in general: that idiomatic and formulaic metaphors are treated as essentially the same, that certain aspects of mappings are fleshed out while others remain unexplored, and above all else, that metaphor researchers see metaphor everywhere. Real metaphors do indeed underlie formulaic and idiomatic utterances, but the extent to which they are activated in on-line processing is dependent on the autonomy of the expression, and on the priming effect of the preceding context. That some aspects of a mapping are attested in common utterances while others aren't follows from the nature of schemata as abstractions over tokens of use, rather than as *a priori* constructs that sanction utterances. And while metaphor is indeed pervasive and prevalent, not all metaphors are equally metaphorical. That many utterances profile a

particular aspect of a mapping, with a conventionalized meaning often not predictable based on the mapping overall, in no way detracts from the fundamental way in which metaphor is a fundamental feature of human language and cognition.

Conceptual Metaphor Theory made the key insight, reflected in both further iterations of CMT and in subsequent cognitive theories of metaphor, that metaphor is a conceptual, rather than a strictly linguistic, system. In individuals and on a cultural level, the systematicity of one domain of thought is used to structure another. This is a domain-general cognitive phenomenon, instantiated in, but not limited to, language. The usage-based approach to metaphor put forward here preserves this view of metaphor, but takes a more dynamic view of the conceptual system that is the essence of metaphor: in that metaphorical schemata are created and strengthened as metaphorical utterances are processed, linguistic metaphor—the facts, as they pertain to figuration, of language in use—has direct input into the conceptual system that motivates language. If it can be said that linguistic metaphor is predicated on a more general conceptual system, it can as easily be said that it is the conceptual system that is motivated and shaped by language. Neither statement is entirely true. The cognitive structures that are metaphor emerge over instances of linguistic metaphor in use, which are in turn produced, judged, and processed on the basis of the emergent cognitive structure of metaphor.

The methods and tools developed within the frequency-based approach to language are ideally suited to handling a property of metaphor which has been almost universally observed, and almost universally underexplained: the conventionalization of cross-domain mappings, which has been described here as the strengthening of conceptual schema that operate across cognitive domains. Along the way, other key features of

metaphor are explained as well. The automatic and idiosyncratic interpretations associated with lexical, formulaic, and idiomatic metaphors are a function of autonomy for specific instances of a mapping, effected by a high degree of entrenchment for some aspects of a mapping relative to the cross-domain mapping overall. Metaphorical strength, the ability of a term to evoke a given source domain, decreases alongside increasing autonomy. 'Families' of related metaphors are tokens of use over which lexical connections have formed, based on shared properties with respect to source and target domain. For metaphorical idioms, tokens of use instantiate syntactic schemata as well, creating complex relationships among utterances with related meanings.

Both corpus and experimental research methods are essential to continued work on a frequency-oriented approach to figuration. From a corpus perspective, the usage-based approach to metaphor will be corroborated by studies on the model of, for example, Poplack & Tagliamonte 1999, Krug 2000, Zilles 2005, Travis & Silveira 2009, and Aaron 2006: studies that note variation in parallel linguistic forms, track differences in frequency between the variants, and tie such differences to diachronic change as frequency effects the displacement of one form by another (and to other change as well). For metaphor, diachronic corpus-based studies demonstrating the usage-based model would do so most effectively by tracking changes over time in the source domains used to refer to a given target, tying such change to the variable frequency of the respective mappings. Diachronic corpus methods are well-suited, as well, to inquiry into the emergence of metaphorical autonomy, as forms begin as novel instantiations of frequent (and, accordingly, productive) metaphorical mapping, and by virtue of high token

frequency lose connections to other forms instantiating the same mapping and take on idiosyncratic meanings.

Conceptual Metaphor Theory (Lakoff & Johnson 1980) asserts that any given metaphorical mapping is instantiated in a wide range of utterances. While this may be true, corpus research on metaphor points towards a hugely disproportionate distribution of forms across a mapping, a large number of tokens for a given mapping accounted for by a small number of expressions (Sanford 2008a). A small-corpus study could demonstrate, across a variety of mappings, how even for highly productive mappings a large proportion of instantiating tokens are accounted for by a handful of highly entrenched forms. Such a study would provide important support for the emergent nature of metaphorical structures.

The experiments reported in Chapter 5 address 'competition', between overall metaphorical schemata, based on differences in their degree of entrenchment (as assessed by their type frequency in a corpus), such that more frequent mappings are more accessible, more acceptable, and more productive than less frequent mappings. Another explicit prediction of the model is a similar competition for forms within a schema. For two forms sanctioned by the same metaphorical schema (for example, *men are dogs* and *she's a dog*), the form higher in token frequency should be processed more quickly, and possibly be found more acceptable, by speakers of a language.

Moreover, since the frequency of the sanctioning schema is the same for all such forms, their token frequencies can be used to make meaningful comparisons as to each form's degree of autonomy. Autonomy effects a weakening of the connection between a form and its sanctioning schema. A sufficiently sensitive priming experiment should find

a difference in the extent to which priming using forms of different frequencies speeds processing for other forms sanctioned by the same schema (for example, the difference between *men are dogs* and *she's a dog* in priming *my banker is a dog*).

The results of experiment 3 (as reported in Chapter 5) broadly confirm the effect of frequency upon metaphorical productivity. More powerful support, however, might be gathered from a study taking the design reported here as a starting point, improving on it (as recommended in §5.3.3) by exploring different approaches to how stimuli are crafted, and seeking a much higher number of participants. For idiom, important support for the dual nature of idioms as instantiations of metaphorical and syntactic schema would come from a series of experiments seeking priming effects across utterances related by each type of schema.

The effects noted in all three experiments might also be sought by comparing two groups of subjects who differ in the frequency to which they have been exposed to a given metaphorical mapping. This might be accomplished artificially (exposing subjects, prior to the experiment, to utterances instantiating one of the mappings), but could also be done by comparing participants from any given in-group to the greater language community. A group of tutors in a university writing center, for example, are likely to have had greater exposure than the general population to the metaphor whereby the transition sentences between paragraphs, and other organizational indicators, are described as 'signs' for the reader. An experiment demonstrating, for example, that the tutors process utterances instantiating the metaphor more quickly than the control (non-tutor) group would clearly indicate an effect from an individual's personal language history on metaphor processing.

The usage-based model hypothesizes that highly fixed, idiosyncratic interpretations for lexical, formulaic, and idiomatic metaphors result from a high degree of entrenchment for particular forms relative to the entrenchment of overall metaphorical mappings. This assertion could be tested using an experimental method prompting literal paraphrases for such forms, analyzing responses for the extent, across responses, of variation from a 'core' interpretation. A corpus study analyzing the token frequency of such forms, and the type frequency of overall mappings, would be used to inform the analysis of the experimental results.

Methodologically, solid approaches to determining the frequency of metaphorical mappings are prerequisite to the approach. Chapter 4 outlines two possible angles on the problem; many more are possible, and at least a few, it is much hoped, would be capable of more fine-grained results.

Over 30 years of cognitive and psycholinguistic inquiry into metaphor have provided invaluable insight into metaphor, metaphorical processing, and metaphorical systems, deeply enriched by a wealth of empirical data from experimental and, more recently, corpus methods. Consequent theories of metaphor are valuable to the extent that they account for this wealth of knowledge, even as they question fundamental assumptions of previous approaches. One such assumption, an immense hindrance to the field as a whole, is the view of metaphor as a 'deep' system of which language in use provides us with glimpses. Moving forward, metaphor is more productively viewed as an organic, dynamic system that emerges over and follows from language in use.

Appendix: Frequency Tables

1.1 Key-terms (small-corpus) method

COMPETITION

	Total Uses Consistent with Source Domain	Total Metaphorical Uses of source term	Total Instances of …
WAR			**COMPETITION IS WAR**
death	6	0	0
gun	2	0	0
bomb	36	6	0
soldier	1	1	0
tank	22	4	0
TOTAL	**67**	**11**	**0**
RACING			**COMPETITION IS RACING**
car	9	0	0
fast	6	0	0
track	16	10	10
NASCAR	6	0	0
competition	0	0	0
TOTAL	**37**	**10**	**10**

HOPE

	Total Uses Consistent with Source Domain	Total Metaphorical Uses of source term	Total Instances of …
LIGHT			**HOPE IS LIGHT**
sun	193	0	0
bright	82	42	6
bulb	26	0	0
lamp	8	0	0
dark	78	24	1
TOTAL	**387**	**66**	**7**
A CHILD			**HOPE IS A CHILD**
small	4	0	0
innocent	0	0	0
play	7	3	0
young	482	0	0
playful	3	0	0
TOTAL	**496**	**3**	**0**

IDEAS

	Total Uses Consistent with Source Domain	Total Metaphorical Uses of source term	Total Instances of …
FOOD			**IDEAS ARE FOOD**
fruit	118	6	0
meat	153	1	0
pizza	107	0	0
restaurant	442	0	0
eat	480	11	0
TOTAL	**1300**	**18**	**0**
WRITING			**IDEAS ARE WRITING**
pen	47	1	1
paper	491	0	0
pencil	37	0	0
book	500	2	1
word	500	0	0
TOTAL	**1575**	**3**	**2**

PEOPLE

	Total Uses Consistent with Source Domain	Total Metaphorical Uses of source term	Total Instances of …
PLANTS			**PEOPLE ARE PLANTS**
green	141	8	5
flower	238	1	1
tree	500	6	4
leaf	194	1	1
water	12	0	0
TOTAL	**1085**	**16**	**11**
BATTERIES			**PEOPLE ARE BATTERIES**
power	0	0	0
AA	1	0	0
AAA	0	0	0
rechargeable	0	0	0
duracell	0	0	0
TOTAL	**1**	**0**	**0**

THE MIND

	Total Uses Consistent with Source Domain	Total Metaphorical Uses of source term	Total Instances of …
A MACHINE			**A MACHINE**
computer	500	0	0
car	500	0	0
metal	18	24	0
robot	20	0	0
technology	222	0	0
TOTAL	**1260**	**24**	**0**
A BODY			**THE MIND IS A BODY**
leg	192	14	0
arm	161	17	0
head	407	54	0
hand	446	154	2
finger	111	8	2
TOTAL	**1317**	**247**	**4**

1.2 Similes (Large-Corpus) Method

COMPETITION IS…

WAR		**RACING**	
source term (x)	**'Like … x' similes that instantiate target metaphor**	**source term (x)**	**'Like … x' similes that instantiate target metaphor**
soldier	4	NASCAR	6
gun	16	race track	3
bomb	13	Indy 500	3
army	43	race car	3
general	24	racer	1
tank	11	checkered flag	1
TOTAL	**111**		**19**

HOPE IS…

LIGHT		**A CHILD**	
source term (x)	**'Like … x' similes that instantiate target metaphor**	**source term (x)**	**'Like … x' similes that instantiate target metaphor**
sun	3	baby	0
bulb	5	kid	0
lamp	2	parent	0
day	2	school	0
illumination	0	toy	0
ray	4	Playground	0
TOTAL	**16**		**0**

IDEAS ARE…

FOOD		**WRITING**	
source term (x)	**'Like … x' similes that instantiate target metaphor**	**source term (x)**	**'Like … x' similes that instantiate target metaphor**
meat	7	pen	0
restaurant	4	paper	3
dinner	11	pencil	0
bread	12	book	22
chef	5	letter	0
breakfast	2	author	1
	42		**26**

PEOPLE ARE…

PLANTS		**BATTERIES**	
source term (x)	**'Like … x' similes that instantiate target metaphor**	**source term (x)**	**'Like … x' similes that instantiate target metaphor**
flower	68	Duracell	3
tree	140	electricity	7
garden	8	charge	0
roots	10	energy	16
leaves	17	power	1
soil	1	Energizer	0
TOTAL	**244**		**27**

THE MIND IS…

A MACHINE		**A BODY**	
source term (x)	**'Like … x' similes that instantiate target metaphor**	**source term (x)**	**'Like … x' similes that instantiate target metaphor**
computer	22	blood	1
engine	2	bones	1
cog	0	legs	0
car	12	arm	0
gear	1	muscles	3
robot	1	stomach	1
TOTAL	**38**		**6**

References Cited

Aaron, Jessi. 2006. *Me voy a tener que ir yendo*: A Corpus-Based Study of the Grammaticization of the *ir a* + INF Construction in Spanish. In Sagarra, Nuria and Toribio, Almeida Jacqueline (Eds.) *Selected Proceedings of the 9th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project.

Aisenman, Ravid. 1999. Structure-Mapping and the Simile—Metaphor Preference. *Metaphor and Symbol* 14(1), 45-51.

Aitchision, J. 1987. *Words in the Mind: An introduction to the mental lexicon*. London: Blackwell.

Allbritton, D.W., McKoon, G., & Gerrig, R.J. 1995. Metaphor-Based schemata and text representations: Making connections through conceptual metaphors. *Journal of experimental Psychology: Learning, Memory, and Cognition* 21, 612-625.

Baayen, R. & Lieber, Michelle. 1991. Productivity and English Derivation: a corpus-based study. *Linguistics* 29, 801-843.

Barlow, Michael and Kemmer, Suzanne. 1994. A schema-based approach to grammatical description. In Luna, S.D., Corrigan, R.L., and Iverson, G.K. (Eds.) *The Reality of Linguistic Rules.* New York & Philadelphia: John Benjamins.

Beckner, Clay & Joan Bybee. 2009. A Usage-Based Account of Constituency and Reanalysis. *Language Learning* 59, 27-46.

Berlin, B. & Kay, P. 1969. *Basic Color Terms: Their Universality and Evolution*, Berkeley: University of California Press.

Billow, Richard M. 1975. A Cognitive Developmental Study of Metaphor Comprehension. *Developmental Psychology* 11, 415-423.

Billow, Richard M. 1981. Observing Spontaneous Metaphor in Children. *Journal of Experimental Child Psychology* 31, 430-445.

Blank, Glenn David. 1988. Metaphors in the lexicon. *Metaphor & Symbolic Activity* 3(1), 21-36.

Bock, J. Kathryn. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18:355–387.

Boers, Frank. 1999. When a bodily source domain becomes prominent: the joy of counting metaphors in the socioeconomic domain. In Raymond W. Gibbs & Gerard J. Steen (eds.) *Metaphor in Cognitive Linguistics: papers from the Fifth International Cognitive Linguistics Conference*. Amsterdam: John Benjamins.

Bosman, N. 1999. Verb Particle Constructions with km and the Motivation Theory. *South African Journal of Linguistics*, 1999, 17, 4, 219-233.

Bowdle, Brian F. and Gentner, Dedre. 2005. The Career of Metaphor. *Psychological Review* 112(1), 193-216.

Branigan, Holly P., Pickering, Martin J., & Cleland, Alexandra A. 2000. Syntactic co-ordination in dialogue. *Cognition* 75: B13–25.

Broderick, Victor. 1991. Young Children's Comprehension of Similarities Underlying Metaphor. *Journal of Psycholinguistic Research* 20(2), 65-81.

Brooks, L. 1978. Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization (pp. 169-211). Hillsdale, NJ: Erlbaum.

Bybee, Joan L. 1985. Morphology: a study of the relation between meaning and form. Philadelphia: John Benjamins.

Bybee, Joan L. 1995. Regular Morphology and the Lexicon. *Language and Cognitive Processes* 10, 425-455.

Bybee, Joan. 2000. The Phonology of the Lexicon: evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.) *Usage-Based Models of Language.* Stanford: CSLI, 65-85.

Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, Joan. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4), 711-733.

Bybee, Joan and David Eddington. 2006. A usage-based approach to Spanish verbs of 'becoming'. *Language* 82, 323-355.

Bybee, Joan & Scheibman, Joanne. 1999. The effect of usage on degrees of constituency: the reduction of *don't* in English**.** *Linguistics* 37(4), 575–596.

Bybee, Joan & Thompson, S. 1997. Three frequency effects in syntax. *Berkeley Linguistics Society* 23: 378-388.

Cacciari, C. & Glucksberg, S. 1995. Understanding idioms: Do visual images reflect figurative meanings? *European Journal of Cognitive Psychology*, 7, 283-305.

Cameron, L. 1999. Operationalizing ‘metaphor’ for applied linguistic research. In L. Cameron & G. Low (eds.) *Researching and applying metaphor*, pp. 59-71. Cambridge, UK: Cambridge University Press.

Cameron, Lynne & Alice Deignan. 2006. The emergence of metaphor in discourse. *Applied Linguistics* 27(4), 671-690.

Carr, Sonya & Bruce Thompson. 1996. The Effects of Prior Knowledge and Schema Activation Strategies on the Inferential Reading Comprehension of Children with and without Learning Disabilities. *Learning Disability Quarterly*, 19, (1), 48-61.

Chafe, W. 1998. Language and the Flow of Thought. In M. Tomasello (Ed.) *The new psychology of language.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

chain(n). 2008. In *Oxford English Dictionary* (Draft addition, 2006). Retrieved January 26, 2010 from Oxford English Dictionary Online database.

Charteris-Black, Jonathan. 2004. *Corpus Approaches to Critical Metaphor Analysis*. New York: Palgrave Macmillan.

Chiappe, Dan L. & John M. Kennedy. 2000. Are metaphors Elliptical Similes? *Journal of Psycholinguistic Research* 29(4), 371-398.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.

Church, Daniel L. and Barbara A. 1992. Auditory priming: implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 18, 915-930.

Clausner, Timothy and Croft, William. 1997. Productivity and Schematicity in Metaphors. *Cognitive Science* 21(3), 247-282.

Clausner, T.C. 1993. Cognitive structures in metaphor comprehension. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Cooper, D. 1986. *Metaphor*. London: Blackwell.

Company Company, Concepción. 2006. Subjectification of verbs into discourse markers: Semantic-pragmatic change only? In Bert Cornillie and Nicole Delbecque (Eds), *Topics in subjectification and modalization*, 97-121. Amsterdam: John Benjamins.

Croft, William. 1990. *Typology and universals.* Cambridge: Cambridge University Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, William. 2007. *Exemplar semantics*. Unpublished manuscript.

Croft, William and Cruse, Alan. 2004. *Cognitive Linguistics.* Cambridge: Cambridge University Press.

Cronk, Brian, Susan Lima, and Wendy Schweigert. 1993. Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research* 22, 59-82.

Cruse, D. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.

Dabrowska, Ewa. 2006. From formula to schema: The acquisition of English questions. *Cognitive Linguistics* 11(1-2), 83–102.

Dabrowska, Ewa & Szczerbinski, Marcin. 2006. Polish children's productivity with case marking: the role of regularity, type frequency, and phonological diversity. *Journal of Child Language* 33, 559-597.

Davies, Mark. (2008-) The Corpus of Contemporary American English (COCA): 385 million words, 1990-present. Available online at http://www.americancorpus.org.

Deignan, Alice. 1995. *Cobuild Guides to English 7: Metaphor*. London: HarperCollins.

Deignan, Alice. 1999a. Corpus-based research into metaphor. In L. Cameron & G. Low (eds.) *Researching and applying metaphor*, pp. 37-48. Cambridge, UK: Cambridge University Press.

Deignan, Alice. 1999b. Linguistic metaphors and collocation in nonliterary corpus data. *Metaphor and Symbol*, 14, 1, 19-36.

Deignan, Alice. 2005. *Metaphor and Corpus Linguistics*. Amsterdam, Philadelphia: John Benjamins.

Diaz-Guerrero, Rogelio and Lorand B. Szalay . 1991. *Understanding Mexicans and Americans: cultural perspectives in conflict.* Plenum Publishing Company: New York.

Dirven, R. & Taylor, J. R. 1988. The conceptualisation of vertical Space in English: The Case of Tall. In Rudzka-Ostyn, B. (Ed): *Topics in Cognitive Linguistics*. Amsterdam: John Benjamins.

Du Bois, John W., Stephan Schuetze-Coburn, Susanna Cumming, and Danae Paolino. 1993. Outline of discourse transcription. In Jane Edwards and Martin Lampert (eds.) *Talking data: transcription and coding in discourse*, pp. 45-89. Hillsdale, NJ: Lawrence Erlbaum Associates.

Du Bois, John W., Chafe, Wallace L., Meyer, Charles, and Thompson, Sandra A. 2000. Santa Barbara corpus of spoken American English, Part 1. Philadelphia: Linguistic Data Consortium. ISBN 1-58563-164-7.

Du Bois, John W., Chafe, Wallace L., Meyer, Charles, Thompson, Sandra A., and Martey, Nii. 2003. Santa Barbara corpus of spoken American English, Part 2. Philadelphia: Linguistic Data Consortium. ISBN 1-58563-272-4.

Du Bois, John W., and Englebretson, Robert. 2004. Santa Barbara corpus of spoken American English, Part 3. Philadelphia: Linguistic Data Consortium. ISBN 1-58563-308-9.

Du Bois, John W., and Englebretson, Robert. 2005. Santa Barbara corpus of spoken American English, Part 4. Philadelphia: Linguistic Data Consortium. ISBN: 158563-348-8.

Erman, Britt and Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text* 20(1), 29–62.

Estes, W. K. 1986. Array models for category learning. *Cognitive Psychology* 18, 500-549.

Evans, Mary Ann and Gamble, Dianna Lynn. 1988. Attribute saliency and metaphor interpretation in school-age children. *Journal of Child Language* 15(2), 435-449.

Fauconnier, Gilles. 1997. *Mappings in Thought and Language*. UK: Cambridge University Press.

Fauconnier, Gilles and Turner, Mark. 2002. *The Way we Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Fauconnier, Gilles and Turner, Mark. 2008. Rethinking Metaphor. In Gibbs, Ray (Ed.) *Cambridge Handbook of Metaphor and Thought.* Cambridge: Cambridge University Press.

Fillmore, Charles J., Kay, Paul, and O'Connor, Mary K. 1988. Regularity and idiomaticity in grammatical constructions: the case of let 'alone.' *Language*, 64, 501-38.

Fowler, Carol and Jonathan Housum. 1987. Talkers' signalling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 25, 489-504.

The Generator Bog: Euphemism Generator. Retrieved January 24, 2010 from <http://generatorblog.blogspot.com/2006/03/euphemism-generator.html>.

Gentner, Dedre and Wolff, Phillip. 1997. Alignment in the Processing of Metaphor. *Journal of Memory and Language*, 37, 331-355.

Gentner, Dedre and Bowdle, Brian. 2001. Convention, Form, and Figurative Language Processing. *Metaphor & Symbol* 16(3&4), 223-247.

Gibbs, Raymond W. 1994. *The Poetics of Mind.* Cambridge: Cambridge University Press.

Gibbs, R.W. and O'Brien, J.E. 1990. Idioms and mental imagery: the metaphoric motivation for idiomatic meaning. *Cognition*, 36, pp. 35-68

Gibbs, Raymond W., Bogdanovich, Josephine M., Sykes, Jeffrey R., & Barr, Dale J., 1997. Metaphor in idiom comprehension. *Journal of Memory and Language*, 37, 141-154.

Givón, T. 1984. *Syntax: A functional-typological introduction.* Volume I. Amsterdam: John Benjamins.

Glucksberg, Sam, Brown, Mary, & McGlone, Matthew S. 1993. Conceptual metaphors are not automatically accessed during idiom comprehension. Memory & Cognition 21(5), 711-719.

Godfrey, John J. & Holliman, Edward. 1997. Switchboard-1, Release 2. Linguistic Data Consortium, Philadelphia.

Goldberg, Adele. 1995. *Constructions: a construction grammar approach to argument structure.* Chicago: The University of Chicago Press.

Goldberg, Adele. 1998. Patterns of Experience in Patterns of Language. In M. Tomasello (ed.) *The new psychology of language*, pp. 203-219. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Goldberg, Adele E. 2006. Constructions at work: the nature of generalization in language. Oxford: Oxford University Press.

Grady, Joseph. 1997. THEORIES ARE BUILDINGS revisited. *Cognitive Linguistics* 8(4), 267-290.

Greenberg, Joseph. 1966. *Language universals, with special reference to feature hierarchies.* The Hague: Mouton.

Gregory, Michelle, William Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *Papers from the Regional Meetings, Chicago Linguistic Society* 35, 151-166.

Gries, Stefan Th. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34, 365–399.

Haiman, John. 1994. Ritualization and the development of language. In William Pagliuca (Ed.), Perspectives on Grammaticalization, Amsterdam: Benjamins, 3-28.

Hamblin, Jennifer & Ray Gibbs. 1999. Why you can't kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research* 28(1), 25-39.

Hare, Mary L., Michael Ford, and William Marlsen-Wilson. 2001. Ambiguity and frequency effects in regular verb inflection. In Bybee, J. and Hopper, P. (Eds)

Frequency and the Emergence of Linguistic Structure. Amsterdam: John Benjamins. 181-200.

Haught, Catrinel. 2005. A Tale of Two Tropes: On the Relation between Metaphor and Simile. Princeton, NJ: Princeton Dissertation.

Hawkins, J. A. 1994. A Performance Theory of Order and Constituency. Cambridge: Cambridge University Press.

Hay, Jennifer & Harald Baayen. 2002. Parsing and productivity. *Yearbook of morphology, 2001*, 203-235.

Hintzman, D. L. 1986. "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93, 411-428.

Hoffman, Robert. 1984. Recent Psycholinguistic Research on Figurative Language. *Annals of the New York Academy of Sciences* 433 (1), 137–166.

Jackendoff, Ray. 1974. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, Ray.1992. Babe Ruth homered his way into the hearts of America. In Stowell, Tim, and Eric Wehrli (eds.) *Syntax and the Lexicon*. San Diego, CA: Academic Press, Inc., 155-178.

Johnson, Keith. 1997. Speech perception without speaker normalization. In K. Johnson & J. Mullennix (Eds) *Talker Variability in Speech Processing*. San Diego, CA: Academic Press, 145-165.

Keysar, Boaz, Shen, Yeshayahu, Glucksberg, Sam, and Horton, William. 2000. Conventional Language: How Metaphorical is it? *Journal of Memory and Language* 43, 576-593.

Koller, Veronika. 2005. Critical discourse analysis and social cognition: evidence from business media discourse. *Discourse & Society*, 16, 2, 199-224.

Kovecses, Zoltan. 2002. *Metaphor: A Practical Introduction*. New York: Oxford University Press.

Krug, Manfred G. 2000. *Emerging English modals: A corpus-based study of grammaticalization*. Berlin, New YorK: Mouton de Gruyter.

Labov, William. 1975. *What is a Linguistic Fact?* Lisse: de Ridder.

Lakoff, George. 1987a. The death of dead metaphor. *Metaphor and Symbolic Activity* 2(2), pp. 143-147.

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

Lakoff, George. 1993. The contemporary theory of metaphor. *Metaphor and Thought 2'nd ed.)*, ed. Andrew Ortony, 202-251. Cambridge: Cambridge University Press.

Lakoff, George, Espenson, Jane, & Schwartz, Alan. 1991. The master metaphor list. Draft 2nd ed. Technical Report, University of California at Berkeley.

Lakoff, G. & Johnson, M. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, George, & Kovecses, Zoltan. 1987. The cognitive model of anger inherent in American English. In D. Holland, & N. Quinn (Eds.) *Cultural models in language and thought*. Cambridge, England: Cambridge University Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Volume 1: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Lappin, Shalom, Robert Levine and David Johnson. 2000. Topic ... Comment. *Natural Language & Linguistic Theory* 18: 665–671.

leg(n). 2008. In *Oxford English Dictionary* (Second Edition, 1989. Retrieved January 26, 2010 from Oxford English Dictionary Online database.

Lindblom, Bjorn, Peter MacNeilage, & Michael Studdert-Kennedy. 1983. Self-organizing processes and the explanation of phonological universals. *Linguistics* 21(1), 181-204.

May, Robert C. 1977. *The Grammar of Quantification*. Phd Dissertation, MIT.

McClelland, James & Jeffrey Elman. 1986. The TRACE model of Speech recognition. *Cognitive Psychology* 18, 1-86.

McGlone, Matthew, Glucksberg, Sam, and Cacciari, Cristina. 1994. Semantic Productivity and Idiom Comprehension. *Discourse Processes* 17, 167-190.

Medin, D. L., & Edelson, S. M. 1988. Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General* 117, 68-85.

Miller, Joanne L. 1994. On the internal structure of phonetic categories: a progress report. *Cognition* 50, 271-285.

Moder, Carol Lynn. 1992. Productivity and categorization in morphological classes. Buffalo, NY: SUNY Dissertation.

Morton, J. 1969. The interaction of information in word recognition. *Psychological Review* 76, 165-178.

Musolff, Andreas. 2004. The heart of the European body politic: British and German perspectives on Europe's central organ. *Journal of Multilingual and Multicultural Development*, 25, 5&6, 437-452.

Musolff, Andreas. 2006. Metaphor scenarios in public discourse. *Metaphor and Symbolic Activity*, 21, 1, 23-38.

Nippold, Marilyn A. and Sullivan, Michael P. 1987. Verbal and Perceptual Analogical Reasoning and Proportional Metaphor Comprehension in Young Children. *Journal of Speech and Hearing Research* 30(3), 367-376.

Nosofsky, R. M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115, 39-57.

Noveck, Ira, Maryse Bianco, & Alain Castry. 2001. The Costs and Benefits of Metaphor. *Metaphor & Symbol* 16(1&2), 109-121.

Nunberg, G. 1978. *The Pragmatics of Reference.* Bloomington: Indiana Linguistics Club.

Nunberg, Geoffrey, Sag, Ivan and Wasow, Thomas. 1994. Idioms. *Language* 70(3), 491-538.

Ohala, John. 1992. The segment: primitive or derived? In Gerald J. Doherty & D. Robert Ladd (Eds), *Papers in Laboratory Phonology II: gesture, segment, prosody, etc.* Cambridge: Cambridge University Press, 166-189.

Palermo, David S. 1986. Metaphor: A Portal for Viewing the Child's Mind. In Lipsitt, Lewis P. and Cantor, Joan H. (Eds.) *Experimental Child Psychologist: Essays and Experiments in Honor of Charles C. Spiker.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Palmeri, Thomas J., Stephen D. Goldinger, and David B. Pisoni. 1993. Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 19, 309-328.

Peters, Stanley; R. Ritchie (1973). "On the generative power of transformational grammars". *Information Sciences* 6, 49–83.

Pierrehumbert, Janet. 1994. Syllable structure and word structure: a study of triconsonantal clusters in English. In Patricia Keating (Ed.) *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III.* Cambridge: Cambridge University Press. 68-190.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds) *Frequency and the Emergence of Linguistic Structure.* Amsterdam: John Benjamins,137-157.

Pitt, M.A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E. and Fosler-Lussier, E. 2007. Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).

Poplack, Shana. 1980. The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. InW. Labov (ed.), *Locating language in time and space*. New York: Academic Press. 55– 67.

Poplack, Shana & Sali Tagliamonte. 1999. The grammaticization of *going to* in (African American) English. *Language Variation and Change* 11, 315-342.

Renouf, A. 1992. What do you think of that? A pilot study of the phrasology of the core words of English. In Leitner, G. (Ed.), *New Directions in English Language Corpora: Methodology, Results, Software Developments*, 301-317. Berlin: Mouton de Gruyter.

Renouf, A. & Sinclair, J., 1991. Collocational frameworks in English. In Aijmer, K. & Altenberg, B. (Eds.), *English Corpus Linguistics: Studies in the Honour of Jan Svartvik*, 128-143. London: Longman.

Reddy, Michael. 1970. The Conduit Metaphor: a case of frame conflict in our language about language. Ortony, Adrew (Ed.) *Metaphor and Thought.* Cambridge: Cambridge University Press.

Rosch, Eleanor. 1973. Natural categories. *Cognitive Psychology* 4, 328-350.

Rosch, Eleanor and Carolyn B.Mervis. 1975. Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology* 7, 573-605.

Rosch, Eleanor. 1978. Principles of Categorization. In E. Rosch & B Lloyd (Eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.

Rubenstein, H., Garfield, L., Milliken, J. A. 1970. Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior* 9, 487-492.

Sanford, Daniel R. 2008a. Discourse & Metaphor: A Corpus-Driven Inquiry. *Corpus Linguistics & Linguistic Theory* 4(2), 209-234.

Sanford, Daniel R. 2008b. Metaphor and phonological reduction in English idiomatic expressions. Cognitive Linguistics 19(4), 585–603.

Schacter, Daniel L. and Barbara A. Church. 1992. Auditory priming: implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 18, 915-930.

Scherre, Maria Marta Pereira, & Naro, Anthony J. 1991. Marking in discourse: 'Birds of a feather'. *Language Variation and Change* 3, 23–32.

Szmrecsanyi, Benedikt. 2005. Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1, 113–149.

Seitz, Jay A. 1997. Metaphor, Symbolic Play, and Logical Thought in Early Childhood. *Genetic, Social, and General Psychology Monographs* 123(4), 373-391.

Shank, Christopher. 2007. Figurative language, transitivity, and the development of epistemicity: a cognitive linguistic case study of the English verbs of tactile perception**.** Dissertation, University of New Mexico.

Siltanen, Susan A. 1989. Effects of Three Levels of Context on Children's Metaphor Comprehension. *Journal of Genetic Psychology* 150(2), 197-215.

R. C. Simpson, S. L. Briggs, J. Ovens, and J. M. Swales. 2002. The Michigan Corpus of Academic Spoken English. Ann Arbor, MI: The Regents of the University of Michigan.

Stefanowitsch, Anatol. 2005. The function of metaphor: developing a corpus-based perspective. *International Journal of Corpus Linguistics*, 10, 2, 161-198.

Strassler, J. 1982. *Idioms in English*. Tübingen: Gunter nar Verlag.

Svanlund, Jan. 2007. Metaphor and Convention. *Cognitive Linguistics* 18(1), 47-89.

Sweetser, Eve. 1990. *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure.* Cambridge, UK: Cambridge University Press.

Sweetser, Eve. 1992. English Metaphors for Language: Motivations, Conventions, and Creativity. *Poetics Today*, 13, 4, 705-724

Taylor, John. 1998. Syntactic constructions as prototype categories. In M. Tomasello (Ed.) *The new psychology of language.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Tissari, Heli. 2001. Metaphors we love by: on the cognitive metaphors of love from the 15th century to the present. *Studia Anglica Posnaniensia: international review of English Studies* Annual 2001, 217-226.

Travis, Catherine. 2006. Dizque: a Colombian Evidentiality Strategy. *Linguistics* 44(6), 1269-1297.

Travis, Catherine & Silveira, Agripino. 2009. The Role of Frequency in First-Person Plural Variation in Brazilian Portuguese: *Nós* vs. *a gente*. *Studies in Hispanic and Lusophone Linguistics* 2(2).

Turner, Mark & Fauconnier, Gilles. 1995. Conceptual Integration and Formal Expression. *Metaphor & Symbolic Activity* 10(3), 183-204.

virus. (2008. In *Oxford English Dictionary* (Draft Revision, March 2008. Retrieved November 9, 2009 from Oxford English Dictionary Online database.

Vitevich, Michael S., Paul A. Luce, Jan Charles-Luce, and David Kemmerer. 1997. Phonotactics and Syllable Stress: implications for the processing of spoken nonsense words. *Language and Speech* 40, 47-62.

Vosniadou, Stella, Ortony, Andrew, Reynolds, Ralph E. and Paul T. Wilson. 1984. Sources of Difficulty in the Young Child's Understanding of Metaphorical Language. *Child Development* 55, 1588-1606.

Waggoner, John E. and Palermo, David S. 1989. Betty is a Bouncing Bubble: Children's Comprehension of Emotion-Descriptive Metaphors. *Developmental Psychology* 25(1), 152-163.

Wang, H. S. & Derwing, Bruce L. 1994. Some vowel schemata in three English morphological classes: experimental evidence. In M. Y. Chen and O.C. L. Tzeng (Eds.) *In honor of Professor William S.-Y. Wang: interdisciplinary studies on language and language change.* Taipei: Pyramid Press. 561–575.

Wilson, Damián Chase Vergara. 2009. Formulaic language and adjective categories in eight centuries of the Spanish expression of 'becoming' /quedar(se)/ + ADJ. Dissertation, University of New Mexico.

Wray, Allison & Perkins, Michael. 2000. The functions of formulaic language: an Integrated Model. *Language & Communication* 20 (1-28).

Zilles, Ana. 2005. The development of a new pronoun: The linguistic and social embedding of *a gente* in Brazilian Portuguese. *Language Variation and Change* 17**,** 19-53.