UC Berkeley
Dissertations, Department of Linguistics

Title
The interaction of syntax and metaphor in gesture: A corpus-experimental approach

Permalink
https://escholarship.org/uc/item/2763w05b

Author
Stickles, Elise

Publication Date
2016-10-01

eScholarship.org Powered by the California Digital Library
University of California

The interaction of syntax and metaphor in gesture: A corpus-experimental approach

By

Elise Stickles

A dissertation submitted in partial satisfaction of
the requirements for the Degree of
Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eve E. Sweetser, Chair
Professor Terry Regier
Professor Mahesh Srinivasan

Fall 2016

The interaction of syntax and metaphor in gesture: A corpus-experimental approach

Copyright 2016
by
Elise Stickles

Abstract

The interaction of syntax and metaphor in gesture: A corpus-experimental approach

by

Elise Stickles

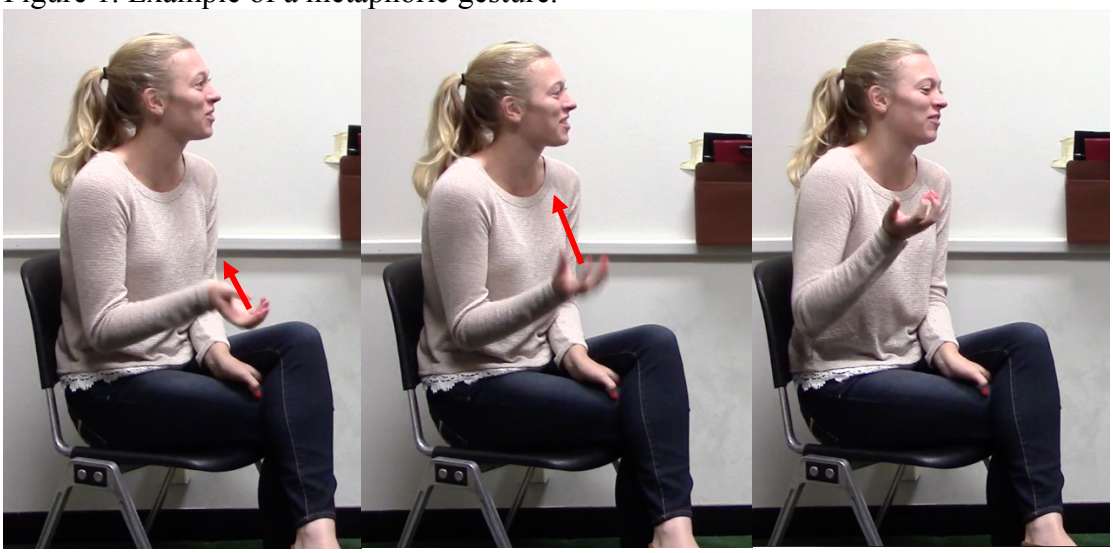
Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Eve. E. Sweetser, Chair

This dissertation is a study of metaphor in usage: metaphor in language, metaphor in gesture, and how they interact. Gesture provides a route to study both the cognition associated with language and the domain-generality of cognitive processes. While English speakers may be *producing* metaphoric manner verbs due to the lexicalization patterns of their language, are they necessarily *thinking* in terms of metaphoric manner? This is difficult to judge when looking at language alone. To answer this question, we turn to metaphoric gesture, as exemplified in Figure 1. In this example, the speaker conveys the metaphor INCREASE IN QUANTITY IS UPWARD MOTION in both her speech *going up* and her upward-moving gesture.

Figure 1. Example of a metaphoric gesture.



so [his confidence was going] up [S4_2X_A_10.20.15_ELV]

Metaphoric gestures, in which the gesture represents the source domain of a conceptual metaphor, are well-known but under-studied (Cienki and Müller, 2008b). Iconic gestures conveying information about a motion event are known to interact with the syntactic and semantic structure of speech; speakers of languages that express manner of motion in the verb gesture differently than speakers of languages that primarily express path of motion in the verb. Metaphoric usages of motion in language – *prices falling*, *hopes rising*, *time flying* – also interact

with the grammatical patterns of language. However, we know little about how metaphoric motion in gesture interacts with grammar.

In Part One of the dissertation, I focus on metaphor in language. In Chapter 2 I propose to represent metaphors as a complex network of frames, mappings, and bindings as implemented in the MetaNet Metaphor Repository (Dodge et al., 2015). This advances the representation of conceptual metaphors to a level that interfaces more accurately with representations of frames and constructions in FrameNet (Ruppenhofer et al., 2006) and Embodied Construction Grammar (Bergen and Chang, 2005). In turn, this enables the detailed analysis of metaphors and metaphor systems, as exemplified by the Location Event Structure Metaphor (Lakoff and Johnson, 1999) case study in Chapter 3. This corpus-based study is one of the first to make use of the MetaNet method for large-scale automatic metaphor identification and annotation. This approach reveals not only how the metaphor system is evoked in language, but further illustrates the conceptual structure of the metaphor. I demonstrate that although English, as a satellite-framed language, privileges manner in its lexicalization of motion events, *metaphoric* English motion backgrounds manner and foregrounds path. The foregrounding of path information in linguistic realization of metaphoric motion runs counter to the privileging of manner in English lexicalization patterns. This finding lays the groundwork for the investigation of the same metaphor system in gesture.

In Part Two, I focus on metaphor in co-speech gesture. I investigate metaphoric motion-evoking metaphoric gestures using two complementary approaches. Chapter 4 uses a corpus approach; I analyze a parallel corpus of video gesture data in which speakers use a motion verb either literally or metaphorically in their speech while producing a co-expressive representational gesture. To analyze the corpora, I develop a set of annotation guidelines and then demonstrate the benefits of taking an image-schematic approach to gesture analysis. I argue that the image schema is the most appropriate level of structure in analyzing the form and meaning of metaphoric gestures. Results of this image schema analysis suggest that, reflecting the English language data in Chapter 3, these metaphoric gestures emphasize path and do not represent the manner of motion.

Chapter 5 is the first study to take an experimental approach to metaphoric gesture that uses non-metaphoric stimuli. Participants were given short stories about state change, such as *prices decreasing* or *grades improving*, to read and re-tell to a friend; half of the stimuli contained metaphoric language and half did not. Results from this study demonstrated the viability of this methodology in eliciting both metaphoric speech and gesture, and supported those of Chapter 4. I find that speakers are more likely produce metaphoric gestures if they are also producing metaphoric language – even if the gesture evokes a *different* metaphor than the speech does.

I unify my analyses of metaphoric motion in speech and gesture in a multi-modal Embodied Construction Grammar analysis of both co-expressive and complementary metaphoric co-speech gestures. I represent both the meaning and form of the gesture and the meaning and form of the speech including frame structure, argument structure, and metaphoric structure. This analysis provides the first formal representation of a multi-modal utterance in a construction grammar and an innovative approach to the unification of the construction of multi-modal meaning.

Dedicated to the memory of my grandmothers

Renee Phillimore

and

Gladys Smith

Contents

Table of Contents ii
List of Figures iv
List of Tables vii
Acknowledgements viii

Table of Contents

Chapter 1: Introduction 1
1. Introduction 1
2.a. Conceptual metaphor theory: An overview 3
2.b. Theoretical foundations: Frames and constructions 6
2.c. A brief history of cognitive linguistics 7
2.d. Why formalize metaphor theory? 8
3. Image schemas 10
4a. Lexicalization of motion events: Manner and path verbs 13
4b. Thinking for speaking 14
5.a. Gesture: An overview 15
5.b. Referential metaphoric gestures 18
5.c. Thinking for speaking and gesturing 20
5.d. Why study metaphoric gesture? 23
5.e. Experimental approaches to metaphoric gesture 26
5.f. Corpus approaches to metaphoric gesture 26
6. Summary 27
7. Organization of the dissertation 28
8. A note on representing the visual-spatial modality 29
9. A note on representing conceptual structures 30

Chapter 2: Formalizing conceptual metaphor theory 31
1.a. Introduction 31
1.b. Roadmap 31
2.a. Development and formalization 31
2.b. Frames 33
2.b.i. Analysis of frame elements and lexical units 33
2.b.ii. Shared frame structures 37
3.a. Frame relations 42
3.b.i. Structure-incorporating relations 42
3.b.ii. Subcase of, special case of, and makes use of frame relations 43
3.b.iii. Incorporates as a role frame relation 44
3.b.iv. Has affordance of and process that makes use of frame relations 44
3.b.v. Profiles part of, is a subscale of, and is a subprocess of frame relations 45
3.c.i. Non-structural relations 46
3.d. Interim summary 49
4.a. Metaphoric construal 50

4.b. Metaphor cascade network 52
5.a. Metaphor relations 53
5.b. Structural metaphor relations 54
5.c. Non-structural metaphor relations 57
6.a. Conclusions 61
6.b. Towards a computational implementation 61

Chapter 3: Location Event Structure Metaphor case study 62
1. Overview 62
2.a. Metaphoric constructions 62
2.b. Metaphoricity evaluation 65
3.a. LESM case study 68
3.b.i. Results and discussion of case study 68
3.b.ii. Lemma frequency 69
3.b.iii. Frames 71
3.b.iv. Constructions 73
3.b.v. The case of 'follow' 77
4. Conclusions 80

Chapter 4: A corpus approach to multimodal metaphoric utterances 83
1.a. Introduction 83
1.b. Current study 87
2.a. Study design and data collection 87
2.b. Annotation 90
2.b.i. Gesture annotation 90
2.b.ii. Speech annotation 98
2.b.iii. ELAN transcription and annotation 99
2.b.iv. Additional annotations 100
3.a. Results 102
3.b. Speech results 103
3.c. Gesture results 106
3.d. Interaction of gesture and speech results 109
4.a. Discussion 113
4.b. Speech results 113
4.c. Gesture results 114
4.e. Image schematic structure 119
5. Conclusions 120

Chapter 5: An experimental approach to multimodal metaphoric utterances 123
1. Introduction 123
2a. Methodology 125
2.b. Stimuli 125
2.c. Norming study 128
2.d. Participants 130
2.e. Design and procedure 130
2.f. Data coding 131

3.a. Results and discussion 134
3.b. Overall results: Metaphoric production 135
3.d.i. Results: Speech 136
3.d.ii. Syntax 136
3.d.iii. Semantics 137
3.e.i. Results: Gesture 143
3.e.ii. Results: Interaction with speech 144
3.e.iv. Results: Gesture metaphors 145
3.f. Results: Summary 152
4. Conclusions 153

Chapter 6: Towards a Multimodal Embodied Construction Grammar 156
1. Summary 156
2.a. Theoretical implications: The Gesture as Simulated Action framework 157
2.b. Theoretical implications: Thinking for Speaking Hypothesis 160
3.a. Multimodal construction grammars 162
3.b. A review of metaphoric speech in ECG 164
3.c. Co-expressive metaphoric gestures in multimodal ECG 166
3.d. Complementary metaphoric gestures in multimodal ECG 176
4. Conclusions 179

References 181
Appendix A: Text of experimental stimuli for Chapter 5. 196

List of Figures

Chapter 1
Figure 1. Gesture accompanying narrative. 1
Figure 2. CMT analysis of IDEAS ARE OBJECTS 4
Figure 3. The Food frame inherits information from its parent, the Object frame. 4
Figure 4. INTELLECTUAL ACCEPTANCE IS SWALLOWING 5
Figure 5. ACCEPTABILITY IS TASTE 5
Figure 6. Early analysis of the Location Event Structure Metaphor 9
Figure 7. Example of an iconic gesture. 16
Figure 8. Example of a metaphoric gesture. 17
Figure 9. Co-expressive metaphoric gesture: Same metaphor in both modalities. 19
Figure 10. Complementary metaphoric gesture: Different metaphors in each modality. 19
Figure 11. Metaphoric gesture with non-metaphoric language. 20
Figure 12. Sylvester the Cat climbing a drainpipe in Episode 2 of *Canary Row*. 21
Figure 13. Backward-moving sagittal gesture. 24
Figure 14. Left-to-right lateral gesture. 25
Figure 15. Metaphoric construals of time relative to speaker's viewpoint. 25

Chapter 2
Figure 1. Roles in the Harm scenario frame. 35
Figure 2. Harm frames subnetwork. Arrows indicate frame relations. 37

Figure 3. Bindings between Purposeful Action, Action, and Desiring frame elements. 43
Figure 4. Relations between Motion Along a Path and causation frames. 48
Figure 5. The lexeme *poverty* evokes the frame Poverty..... 51
Figure 6. Frequency of lemmas r when used in metaphoric expressions related to poverty. 55

Chapter 3
Figure 1. Metaphor analysis of BECOMING IMPOVERISHED IS DOWNWARD MOVEMENT 63
Figure 2. Analysis of the Active Motion Path Construction..... 64
Figure 3. Metaphoric variant of the Active Motion Path Construction. 64
Figure 4. Combination of metaphor and construction analysis of *fell into poverty*..... 65
Figure 5. Relationships between roles in the frames evoked by *fell into a hole.*..... 66
Figure 6. Relationships between the frames evoked by *fell into poverty*. 67
Figure 7. Relative frequency of common source domain lemmas..... 69
Figure 8. Metaphor likelihood measures for frequent lexemes. 70
Figure 9. Most commonly occurring source frames. 71
Figure 10. Relative frequency of frame families. 73
Figure 11. Relative frequency of metaphoric verbal constructions. 74
Figure 12. Constructional frequency in COCA genres. 75
Figure 13. Frame family frequency across constructions. 76
Figure 14. Distribution of *follow* between genres..... 77
Figure 15. Distribution of *follow* tokens in academic and non-academic discourse. 79

Chapter 4
Figure 1. Example of a metaphoric gesture. 83
Figure 2. Iconic gesture evoking the Cycle image schema..... 84
Figure 3. Metaphoric gesture evoking the Cycle image schema. 85
Figure 4. The ANR annotation procedure..... 91
Figure 5. Contrast between (A) static (no path) and (B) translational path gestures..... 92
Figure 6. Basic path direction categories from (A) top view and (B) side view. 93
Figure 7. Manner: straight; Path: down 94
Figure 8. Manner: arc; Path: down-backward..... 94
Figure 9. Manner: circle; Path: forward-down static x3 94
Figure 10. Manner: rotate; Path: left static 95
Figure 11. Manner: chop; Path: down-up forward x2..... 95
Figure 12. Manner: zig-zag; Path: left-right x2 95
Figure 13. Manner: wavy; Path: right-left-up x3 96
Figure 14. Manner: straight; Internal motion of fingers; Path: right 96
Figure 15. (A) Typical adult gesture space; (B) Simplified location coding..... 97
Figure 16. (A) *Up* palm orientation; (B) *Down* palm orientation. 98
Figure 17. (A) *Speaker in* palm orientation; (B) *Speaker out* palm orientation. 98
Figure 18. (A) *Center in* palm orientation; (B) *Center out* palm orientation. 98
Figure 20. Distribution of handshape types between metaphoric and literal data. 100
Figure 21. (A) 5 handshape; (B) Open B handshape. 101
Figure 22. Distribution of co-timing frequencies. 101
Figure 23. Distribution of verb schemas in metaphoric and literal utterances 104

Figure 24. Distribution of verb utterance duration. 105
Figure 25. Handshape type frequency. 106
Figure 26. Gesture manner frequency. 107
Figure 27. Distribution of gesture image schemas. 107
Figure 28. Stroke duration as a function of gesture image schema. 108
Figure 29. Stroke co-timing. 109
Figure 30. Stroke co-timing by type. 110
Figure 31. Interaction of grammatical aspect and stroke duration. 111
Figure 32. Distribution of gesture manner by verb type and metaphoricity. 111
Figure 33. Relationship between metaphoric verb schema and gesture image schema. 112
Figure 34. Distribution of satellites by verb type. 114
Figure 35. American Sign Language classifier predicates. 115
Figure 36. Use of unmarked Open B handshape. 117
Figure 37. Path gesture accompanying metaphoric motion verb. 119
Figure 38. *Arc* gesture evoking a metaphoric boundary crossing. 121

Chapter 5
Figure 1. Dyadic participant configuration 131
Figure 2. Example utterance. 133
Figure 3. Analysis of Figure 2 utterance in ELAN. 133
Figure 4. Lemma frequency by co-expressivity. 139
Figure 5. QUANTITY IS VERTICALITY gesture 147
Figure 6. EMOTION IS VERTICALITY; QUANTITY IS VERTICALITY gesture 148
Figure 7. CONTINUOUS PROGRESS IS CONTINUOUS MOTION gesture 149
Figure 8. Multiple metaphors conveyed in one gesture. 150
Figure 9. STATES ARE OBJECTS gesture. 150
Figure 10. EMOTIONAL STATES ARE PHYSICAL STATES gesture 151

Chapter 6
Figure 1. Lexical construction: Poverty 165
Figure 2. Metaphor construction analysis of *fell into poverty* 165
Figure 3. Metaphoric Active Motion Path Construction 166
Figure 4. Cognitive grammar analysis of a tracing gesture 167
Figure 5. Co-expressive metaphoric gesture 168
Figure 6. Metaphor analysis of *prices drop* 168
Figure 7. ECG analysis of the Metaphoric Intransitive Declarative Construction 169
Figure 8. Metaphor and ECG analysis of *prices drop* 169
Figure 9. Representation of gesture form and image schema. 170
Figure 10. Metaphoric Process Gesture Construction analysis. 171
Figure 11. Complete metaphoric ECG analysis of the multimodal utterance. 171
Figure 12. BentB handshape. 172
Figure 13. Re-analysis of the metaphoric gesture form. 172
Figure 14. Re-analysis of the metaphoric gesture construction. 173
Figure 15. ECG analysis of prices drop to a certain level 174
Figure 16. Complete re-analysis of the multi-modal utterance. 175
Figure 17. Metaphoric gesture with non-metaphoric speech. 176

Figure 18. ECG analysis of the Active Intransitive Construction. 177
Figure 19. ECG and frame analysis of *he feels better*. 178
Figure 20. Complete multimodal analysis. 178

List of Tables

Chapter 2
Table 1. Roles and relations in the Motion along a Path and Corruption frames. 38
Table 2. Relationships between roles in Harm and Corruption frames. 39
Table 3. Causal and aspectual variants of processes. 40
Table 4. Comparison of aspectual and causal variants of Motion Along a Path Frame. 47
Table 5. Summary of frame relations. 49
Table 6. Entity roles and their role types in the Harm to Living Entity frame. 51
Table 7. Mappings between roles and their role types in the ECONOMIC HARDSHIP IS PHYSICAL HARM metaphor. 52
Table 8. Partial structure of the NATION IS A PERSON metaphor. 52

Chapter 3
Table 1. Metaphoric verbal constructions. 73

Chapter 4
Table 1. High frequency motion verbs, sorted in descending frequency 88
Table 2. Gesture manner form categories. 93
Table 3. Lemma frequency measures. 103
Table 4. Distribution of verbal inflection. 104
Table 5. Results of co-timing logistic regression model 110
Table 6. Results of verb type * metaphoricity logistic regression model. 112
Table 7. Results of gesture schema ~ verb schema logistic regression model 113
Table 8. Comparison of effects of utterance metaphoricity 120

Chapter 5
Table 1. Stimuli balance design. 126
Table 2. Stimuli format. 127
Table 3. Stimuli format of an example metaphoric story stimulus. 128
Table 4. Norming study results. 129
Table 5. Distribution of nominal/verbal and aspect by metaphoricity 137
Table 6. Distribution of part of speech by metaphoricity. 137
Table 7. Distribution of *match* and *mismatch* lemma 137
Table 8. Distribution of metaphors between stimuli types. 142
Table 9. Distribution of LESM metaphors. 143
Table 10. Non-metaphoric lemmas co-occurring with metaphoric gestures. 144
Table 11. Gesture co-timing. 145
Table 12. Co-timing categories. 145
Table 13. Metaphoric gesture counts as distributed by stimuli type 146
Table 14. Distribution of LESM gesture metaphors. 147
Table 15. Handshape markedness by metaphor category. 152

Acknowledgements

First and foremost, thank you to my dissertation committee and especially Eve Sweetser, my dissertation committee chair and advisor throughout my time at Berkeley. Eve has had faith in me from the start, when I came into graduate school with no real background in linguistics and was very much in over my head. Every time I walk into her office feeling confused or lost, I walk back out with a new sense of purpose and direction, along with a whole new set of questions to investigate. Eve's uncanny knack for asking the right questions of complicated data, her endless fount of knowledge, and her remarkable generosity have had a profound and deep influence on not only my dissertation, but on my academic development as a whole. Terry Regier taught me how to structure an academic presentation, read an academic paper, and in general think and talk like an academic. From taking his language and cognition seminar in my first year to being a graduate student instructor for his introduction to cognitive science courses, I have seen from him how to be an inspiring and effective instructor. Mahesh Srinivasan has given me a fresh perspective on metaphor, particularly the potential of empirical approaches to metaphor research. Our Friday afternoon meetings were a highlight of the week this past year as I could always look forward to hearing his insights and grappling with his incisive questions. If Eve is my academic "mother", then George Lakoff is my academic godfather. While technically not on my committee, George has taught me metaphor theory from the day I set foot in the department, and his ideas, enthusiasm, and seemingly endless example sentences have driven my interest in metaphor and grammar.

My time at ICSI working at MetaNet has deeply influenced my academic trajectory and reinvigorated my previously dormant interests in applying computational linguistics methods to politics. To all the MetaNet crew, but especially Oana David, Ellen Dodge, Jisup Hong, and Karie Moorman, I think we've managed to create something truly worthwhile. Thank you for all the "deep dives" and the fantastic whiteboard diagrams. I also owe a special thanks to Ellen for teaching me ECG and helping me work out the details of multimodal ECG. Oana, Ellen, and Jisup have kindly agreed to let me use one of our co-authored MetaNet papers as Chapter 2 of my dissertation, for which I am extremely grateful.

I have been so lucky to be in a department with such supportive, intelligent, and fun fellow grad students. From our early days of syntax problem sets in cafes to late hours in the lab, I have leaned on and learned much from so many of them (in no particular order): Emily Cibelli, Oana David, I-Hsuan Chen, Matt Goss, Elisabeth Wehling, Michael Ellsworth, Mischa Park-Doob, Jisup Hong, Shinae Kang, Clare Sandy, Greg Finley, Clara Cohen, Florian Lionnet, Sarah Bakst, Nico Baier, Roslyn Burns, Christine Sheil, Tammy Stark, Geoff Bacon, Auburn Barron-Luztross, Kayla Begay, Kelsey Neeley, Erin Donnelly, Alice Shen – I am so grateful to you all.

I have also benefited enormously from the amazing talent and dedication of the many Berkeley Linguistics undergraduate students who have volunteered their time and energy as my research assistants. They spent many hours watching videos to collect corpus data and annotating in ELAN. They helped design experiments, ran my study trials, and did a grueling amount of data management. At this point, they know this data about as well as I do and their thought-provoking questions and comments have shaped my analysis and understanding of gesture. To Kayla Briones, Nathan Cahn, Shannon Chang, Lauren Cho, Noah Hermalin, Ann Kim, Elizabeth Long,

Natalie Orsi, Rachel Pinkerton, and Emily VanLoo: you rock! I couldn't have done this without you.

A Berkeley Linguistics dissertation acknowledgements section would be remiss not to mention Belén Flores and Paula Floro, and for good reason. They are the heart and soul of our department, tirelessly working their behind-the-scenes magic. Thanks especially to Belén for fixing my mistakes before I even knew I made them.

My work in gesture studies would not have been the same without my broader gesture family, especially the members of the Berkeley Multimodality and Gesture Group, who are always up for watching a 2 second video clip one more time. Kashmiri Stec's insight and dedication to methodological rigor were a large influence on my approach to gesture annotation. Thank you to Jennifer Hinnell and Laura Hirrel, my gesture and construction grammar comrades-in-arms, for sharing your different takes on a brand new field during our many conference chats. Tasha Lewis has been not only a co-author par excellence, but a role model and much-appreciated source of advice on academia and life in general. Thank you especially for agreeing to let me re-use some of our forthcoming paper in Chapter 1.

To my parents, I am grateful that they have been unwaveringly patient and understanding throughout my seemingly interminable graduate school career, and have only rarely asked how my "big paper" was going. (Don't worry, you now officially get to brag to everyone. I know you've been waiting anxiously.) My sister Amanda has supplied me with a steady stream of cat pictures over the years and a sense of perspective that only someone who was nine when I started grad school and is now herself on the verge of college can.

I adopted my cat Indy from Berkeley Animal Care Services a few weeks after finishing my M.A. in 2011; she has kept me grounded throughout my Ph.D. work by always (loudly and insistently) reminding me of the importance of play time, cuddles, and regularly scheduled meals.

Finally, I owe a huge thank you to Kenny Baclawski, who has been by turns my syntax and information structure consultant, copy editor, French translator, research assistant, experiment guinea pig, personal chef, life coach, cheerleader, and above all a constant source of support and positivity in my moments of self-doubt. I look forward to returning the favor.

Chapter 1: Introduction

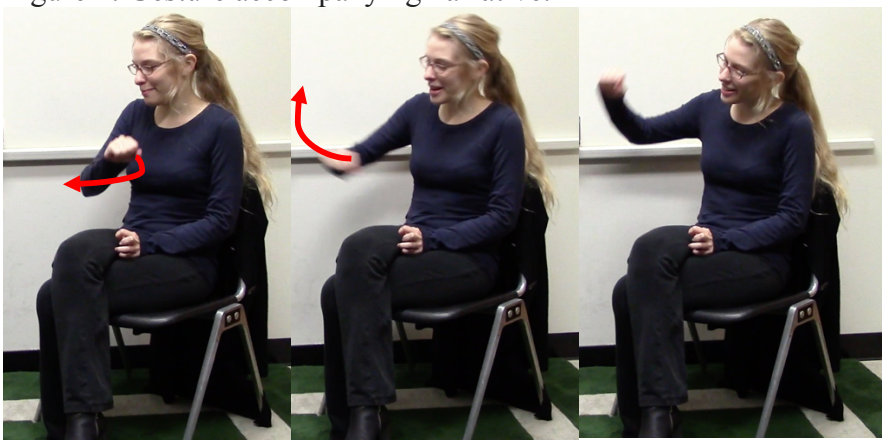
1. *Introduction*

Broadly speaking, this dissertation is about metaphor and gesture in everyday communication, such as this excerpt from a short oral narrative:

> *So, there's a kid named Martin, and he's a high school senior, and he's filling out college applications…He feels pretty good about it. His counselor tells him that he's in the top ten percent of SAT scores.* ***So he's like, he feels better, even better about his college applications.*** *And then he finds out…something causes his hopes to sink.*

Upon initial consideration, the bolded sentence in this story fragment appears to lack any metaphoric language. The narrator simply describes Martin as "feeling better". However, if we consider her gestures, a different picture emerges (Figure 1). First, she makes a sweeping, arcing gesture outward and then upward from her chest to head height as she says "he feels *better*". Then, she makes a second smaller jabbing motion upward as she repeats, "even *better*".

Figure 1. Gesture accompanying narrative.



he feels [better]



even [better]

[S9_2Y_B_10.27.15_ELV]

When we take the narrator's gestures into account, we see that she makes use of a common metaphor, GOOD IS UP: improvement in Martin's emotional state is represented in her gesture as upward movement along a vertical path. The incremented further improvement – "even better" – is matched with a gesture representing a smaller movement upward, suggesting that INCREMENTAL IMPROVEMENT IS INCREMENTAL UPWARD MOTION. So, while the narrator's *speech* appeared to be metaphor-free, her *gesture* is richly metaphoric and reflects how she conceptualizes Martin's changing emotional state. Investigation of this single sentence shows us that studying linguistic data alone doesn't give us the full picture. An accurate investigation of the conceptual structures behind the narrator's understanding of the story, and her subsequent telling of that story, must take her gestural as well as linguistic production into account (e.g., Cienki, 1998; Müller, 1998; Sweetser, 1998; Parrill and Sweetser, 2004).

Why should we care about metaphor at all? Let's return to the narrator's speech, and this time identify all the metaphoric language in her narrative.

> *So, there's a kid named Martin, and he's a high school senior, and he's filling* ***out*** *college applications…He feels pretty good about it. His counselor tells him that he's* ***in*** *the* ***top*** *ten percent of SAT scores. So he's like, he feels better, even better about his college applications. And then he finds* ***out****…something causes his hopes to* ***sink****.*

Metaphoric language actually occurs throughout her story. Martin is located "in" the "top" ten percent of SAT scores. But he is not physically located in the scores, nor are the scores vertically-arranged entities. And yet, our narrator construes Martin's current state as a place he is located in. She portrays the range of SAT scores as a vertical structure, with better scores higher up and worse scores lower down. Throughout her speech, the narrator makes use of spatial language to describe states such as emotions, quantities, and positivity – just as she makes use of space with the motion of her gesture to represent a change of emotional state. In other words, metaphor is pervasive throughout her language *and* her gesture. When such gesture occurs in the absence of metaphoric language, it suggests that these metaphors aren't "just" literary devices, rhetorical flourishes coloring our language. Instead they are something deeper, more fundamental to our cognition, such that they are consistently expressed in both linguistic and non-verbal communication.

This dissertation is an exploration of the nature of these metaphors. Much work has already been done documenting them in both language and gesture, identifying different metaphors, and theorizing the cognitive processes behind them (see Cienki and Müller, 2008b for an overview). However, little work has been done on metaphoric event structure gestures, especially as they relate to grammar. These gestures, which are exemplified by the example in Figure 1, represent a motion event such as 'upward movement', 'cyclic movement', or 'iterative movement', which forms the basis of a metaphoric understanding of a change of state such as 'increase in quantity' or 'ongoing progress'. Language-specific lexicalization patterns influence the words used in the linguistic realization of metaphor, but metaphoric gestures represent only the elements that map from one domain to the other in the metaphor. Here I expand on prior work by addressing *when* metaphor occurs in language and gesture, and which elements of metaphors are realized linguistically versus gesturally. I focus on the *formalization* of metaphor theory, in both language

and gesture; my goal is to develop a more rigorous approach to metaphor analysis in both modalities by using both corpus and experimental methodologies. Throughout the dissertation, I focus on the Location Event Structure Metaphor family, which is a primary, universal metaphor system. For example, nearly every metaphoric utterance in our narrator's story, including all her gestures, evokes variations on the Location Event Structure Metaphor. In the Location Event Structure Metaphor, states of being – such as emotions and relative SAT score rankings – are conceptualized as metaphoric locations. The narrator says that the student is *in* the top ten percent of scores: the SAT score rankings are construed as locations, and the student is "in" a particular (good) location.

In this chapter, I provide an overview of the theoretical and literature background to this dissertation. Given the interdisciplinary nature of this work, an exhaustive literature review would be the length of a dissertation itself. Therefore, in lieu of attempting to completely review the metaphor theory, event structure, and gesture studies literature, I provide a broad overview and history of each field as is necessary for the following chapters. First, I discuss conceptual metaphor theory and the state of contemporary theory; I argue a formalization of the theory is necessary to bring it in line with other aspects of cognitive linguistics. Next, I briefly discuss image schemas as they relate to metaphor theory. I then describe motion event structure lexicalization (Talmy, 1985) and introduce the "Thinking for Speaking" Hypothesis (Slobin, 1987). The remainder of the introduction details aspects of gesture studies. The basics of gesture studies analysis are introduced with a focus on metaphoric referential gestures. I also discuss motivations and methodologies for studying metaphoric gestures. After this background, I lay out the organization of the dissertation and discuss how I represent gestural data throughout this work.

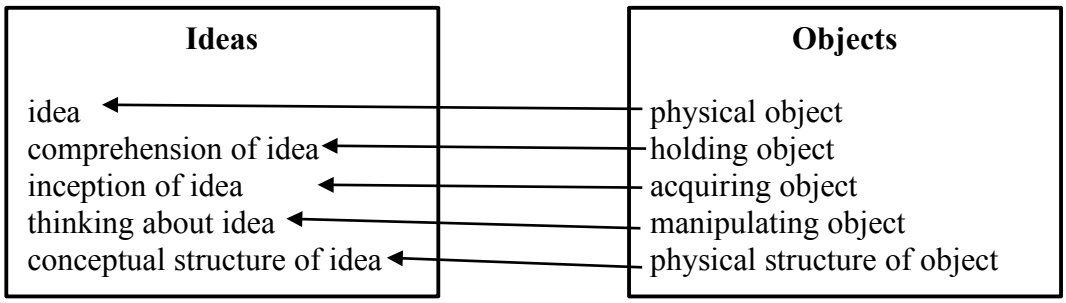
2.a. *Conceptual metaphor theory: An overview*

Conceptual metaphor theory (CMT) is used to model the *structure* of metaphors (Lakoff and Johnson, 1980; 1999). In the CMT framework, metaphors are *mappings* between two semantic domains (experientially-based conceptual gestalts), called *frames* (Fillmore, 1982). One frame, the *source domain*, is frequently (although not always) concrete or physically experienced; for example, spatial or sensory representations are often source domains. They are directly perceived through our interactions with the world. The *target domain* is the semantic domain that is the topic of the metaphor; it is frequently abstract, without a direct physical correlate. States such as emotions, quantities, and notions of "good" and "bad" are often target domains. In primary metaphors, the asymmetry between source and target domains is a result of different levels of *intersubjectivity*: the degree to which a property or state is assumed to be perceptually accessible (Sweetser, 1990; Dancygier and Sweetser, 2014). For example, emotions are highly physically experienced states, but they are highly subjective. In contrast, we expect that assessment of a spatial property like height or size to be fairly objective and that others will agree on the relative height or size of an object as they are readily accessible, perceptual properties.

For example, the metaphor IDEAS ARE OBJECTS maps the source domain of physical Objects onto the target domain of Ideas (Lakoff and Johnson, 1999). Thus, when we talk about *grasping a concept*, *holding a thought*, or *toying with an idea,* we are understanding *concepts, thoughts,* and *ideas* as physical objects that can be *grasped, held,* and *toyed with*. Furthermore, we transfer our

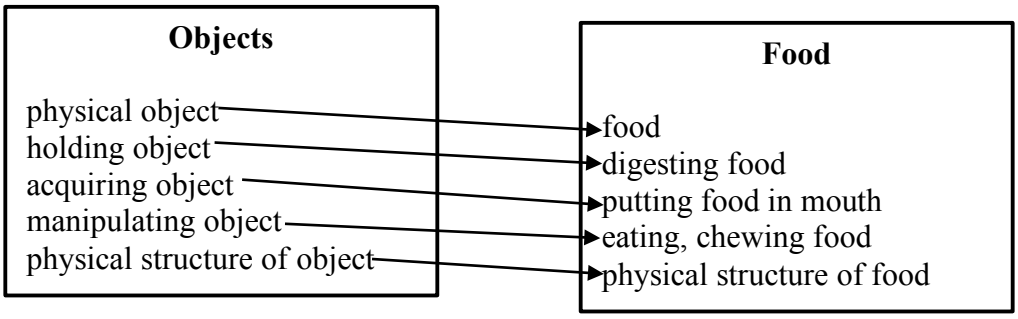
understanding of how objects work onto our understanding of how ideas work. They can be held onto or transferred to someone else: *throw some ideas around*; we can lose them: *the thought escaped me*; and we can acquire them: *a thought just came to me*. In other words, we transfer our knowledge about the Object frame to the Idea frame. In CMT analysis, the structure of the source domain must align with that of the target domain; however, not all elements of the source necessarily map onto the target (Lakoff and Johnson, 1980; Grady, 1997). Figure 2 illustrates such an analysis for the IDEAS ARE OBJECTS metaphor.

Figure 2. CMT analysis of IDEAS ARE OBJECTS (after Lakoff and Johnson, 1999)



As we see from Figure 2, the metaphor IDEAS ARE OBJECTS has several mappings between the source and target domain. These mappings form *entailments*, or conclusions we make based on the relationship between the two domains. For example, if IDEAS ARE OBJECTS, then THINKING IS OBJECT MANIPULATION and UNDERSTANDING IS GRASPING. If we combine this metaphor with the related metaphor the MIND IS A CONTAINER, then thoughts are objects contained in the mind: *keep that thought in mind*. Alternatively, we can make use of a more specific version of IDEAS ARE OBJECTS, IDEAS ARE FOOD. This maintains the structure of IDEAS ARE OBJECTS (since Food is a kind of Object), but adds inferential structure on the basis of specifying the type of object and our interaction with it. When we talk about *food for thought*, *chewing something over*, or finding something *hard to swallow*, we apply our knowledge of Food and Eating to our conceptualization of Ideas (Lakoff and Johnson, 1999). Importantly, because the Food frame is a specific type of Object frame, there is a direct structural relationship between the Object and Food frames (Figure 3).

Figure 3. The Food frame inherits information from its parent, the Object frame.



Furthermore, we can perspectivize the notion of eating food in terms of our physical interactions with it (tasting chewing, swallowing, digesting food) and our social interactions with it

(preparing, sharing, serving food). These different aspects of eating result in different source domains with different entailments (Croft, 2009; Dancygier and Sweetser, 2014). For example, *finding that thought hard to swallow* and *eating up an idea* evoke INTELLECTUAL ACCEPTANCE IS SWALLOWING (Figure 4), which makes use of our understanding of the swallowing process in Ingestion. *Bitter thoughts* and *sweet ideas* make use of Taste, evoking ACCEPTABILITY IS TASTE (Figure 5). As metaphors are compositional and Ingestion and Taste are both associated with Food, they can be combined: *a bitter thought that's hard to swallow*.

Figure 4. INTELLECTUAL ACCEPTANCE IS SWALLOWING (after Dancygier and Sweetser, 2014; Lakoff and Johnson, 1999).

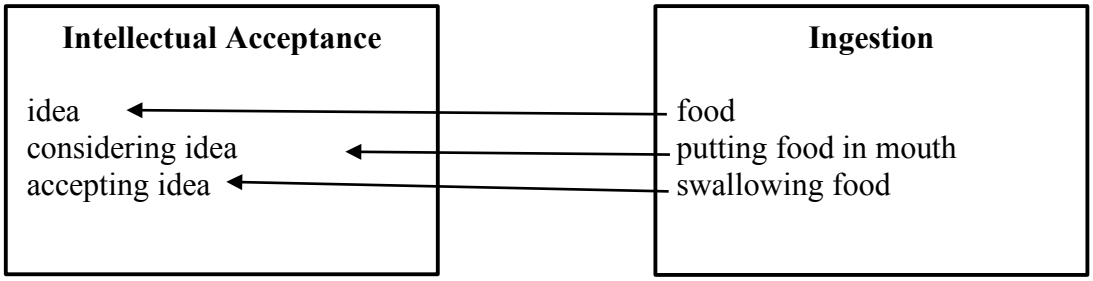
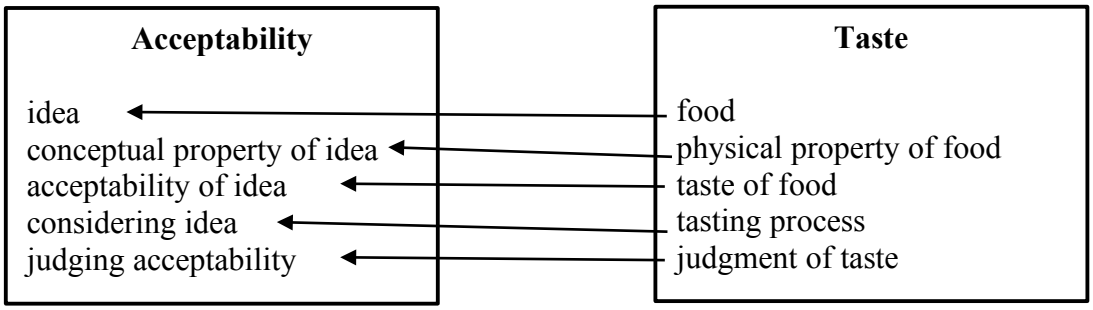


Figure 5. ACCEPTABILITY IS TASTE (after Dancygier and Sweetser, 2014; Lakoff and Johnson, 1999).



Our experiences of things like objects and food are said to be *embodied.* Our mental representations of actions like object manipulation, eating, or running involve activation of parts of the brain associated with physical activity. So when we think about objects, or food, or jogging, the experiences of our actual physical bodies are activated in our minds. However, it is also the case that our experiences of things like ideas and quantities are *also* embodied (Lakoff and Johnson, 1999; Bergen, 2012; Lakoff, 2012). This is perhaps more surprising, since these abstract concepts don't have direct physical correlates the way objects and actions do. The embodiment of abstract concepts is posited to occur due to systematic activation of both the abstract experience and a physical sensorimotor experience at the same time (Lakoff and Johnson, 1999; Lakoff, 2012). For example, when we stack objects on top of each other, we both experience an increase in quantity (more objects), but also an increase in height (the objects stacking). Thus we experience both the concept of *more* and the concept of *up*. Verticality is a

physical experience, and Quantity an abstract construct, but we experience them so frequently together that they become linked in the conceptual metaphor QUANTITY IS VERTICALITY (and thus MORE IS UP; LESS IS DOWN). The expression of conceptual metaphors in speech and in gesture, as in section 1, is evidence of this pervasive system of cognitive structures.

Thus far I have described CMT analyses in prose and diagrammatic form, as is traditionally done. In the next sections, I argue that CMT will benefit from formalization beyond simple diagrams like those in Figures 2, 4, and 5.

2.b. *Theoretical foundations: Frames and constructions*

Before continuing in my discussion of conceptual metaphor theory, I will pause briefly to provide an overview of some basic concepts in cognitive linguistics that will be discussed throughout the dissertation. The previous section has already touched on the notion of *frames*, which were first introduced by Fillmore in his theory of Frame Semantics (1976, 1982). A frame is a conceptual structure representing some coherent collection of knowledge. Frames have gestalt structure, such that evoking a particular frame element evokes the whole, and evoking the whole frame evokes the part. For example, the Commercial Transaction frame can be evoked by such words as *buyer, seller,* and *price*. Using the word *price* evokes the whole Commercial Transaction frame, such that *price* also brings up knowledge of buying, selling, and exchanging money for goods or services. We can only understand the notion of *buying* if we understand the notions of *selling* and *monetary exchange* as well. A lexical frame is a frame evoked by a lexical element. In English, the constituent processes within the Commercial Transaction frame, Buying and Selling, are themselves lexical frames. In Eastern Cham (Austronesian: Vietnam) and Hawaiian (Austronesian: Hawaii), a single word[1] means both *buy* and *sell* – in other words, the Commercial Transaction frame is a lexical frame in Eastern Cham and Hawaiian (Kenneth Baclawski, Jr., p.c., 2016). In the context of conceptual metaphor theory, the source and target domain of the metaphor are frames; metaphors can be understood as mappings between frames and frame elements, and will be represented as such throughout this work.

The second concept critical to this work is the notion of a construction in the context of Construction Grammar (Croft, 2001; Fillmore et al., 1988; Goldberg, 1995; 2006; Kay and Fillmore, 1999). Constructions are mappings between *form* and *meaning* that occur at every level of linguistic structure. Words are considered lexical constructions; bound morphemes like the English past tense suffix *-ed* are also constructions, since they pair meaning (such as 'past tense') to form (such as *-ed*). Some constructions combine syntactic patterns with specific lexical instantiations, such as the *What's X doing Y?* construction, where X and Y are embedded clauses as in *what is it doing raining? or what am I doing writing this dissertation?* (Kay and Fillmore, 1999). As Israel (1996) shows in his analysis of the *way* construction (*dug his way out*; *limped their way across the field*), this usage-based approach to language recognizes that speakers are aware of the general patterns that these constructions follow and the specific common instances

[1] In written Cham, 'buy' is spelled *blei* and 'sell' *pablei*, but in spoken Eastern Cham they are both pronounced as [plèj]. In Hawaiian, the word for 'buy' and 'sell' is *kū'ai*.

of the construction. To be considered a construction, these form-meaning pairs must be conventionalized with frequent and productive use over time, leading to entrenchment of the construction (Croft, 2001).

Argument structure constructions, which we will discuss further in Chapters 3 and 6 in the context of Embodied Construction Grammar, pair syntactic patterns with semantics (Goldberg, 1995; 2006). For example, Goldberg (1995) showed that the English syntactic pattern Subject Verb Object Directional conveys the semantics of Caused Motion in what is now known as the Caused Motion Construction. A typical example is *she pushed the cup off the table*: *she* is the Subject, *pushed* the verb, *the cup* the Object, and *off the table* the Directional. The Subject *she* causes the motion of the Object *the cup* by *pushing* it in the Direction *off the table*. Another less-prototypical example is *she sneezed the foam off the latte*. In typical usage, *sneeze* is not a transitive verb, nor does it have the semantics of Caused Motion. However, in the context of this construction, it changes valence and gains causal semantics due to the syntax and semantics of the construction it is in.

Throughout this work I will be making use of the Embodied Construction Grammar (ECG) framework (Bergan and Chang, 2005; Feldman et al., 2009). ECG is designed specifically to represent an embodied simulation semantics model of language comprehension (Bergen, 2012). Driving this approach is the concept that language comprehension entails activation of the relevant embodied schemas (in particular, image schemas), and the simulation of these embodied experiences. Given the embodied nature of gesture, a framework that considers embodied semantics as the core of language comprehension is a natural candidate for representation of gesture. I discuss image schemas below in section 3 and will discuss the formalisms of ECG and illustrate their application to metaphoric constructions; I apply them to metaphoric gestures in Chapters 3 and 6.

2.c. *A brief history of cognitive linguistics*[2]

In order to understand the importance of formalizing conceptual metaphor theory and its role in gesture analysis, we must first understand the history of its role in cognitive linguistics as a whole. Since the development of Frame Semantics by Fillmore (1976, 1982, among others; see also Lakoff, 1987) and the publication of Lakoff and Johnson's (1980) seminal *Metaphors We Live By*, the field of cognitive linguistics has grown into a mature discipline. Significant advances in Frame Semantics, particularly its instantiation in English FrameNet and versions of FrameNet in languages other than English[3] (Ruppenhofer et al., 2010), and in Construction Grammar frameworks (Fillmore, 1988; Kay and Fillmore, 1999; Goldberg, 1995; Bergen and Chang, 2005; Feldman et al., 2009; Croft, 2001; Fried and Östman, 2004; Boas, 2013) have led

[2] Sections 2.c.-2.d., along with much of Chapter 2, are adapted from Stickles et al. (2016b); furthermore, these ideas are the result of years of collaboration with many members of the MetaNet team, including in particular George Lakoff, Eve Sweetser, and Karie Moorman. I am immensely grateful to all for their contributions, and to my co-authors for their permission to include their work here.

[3] Other FrameNets include: Korean (Nam et al., 2014); Brazilian Portuguese (Salomão et al., 2013); Japanese (Ohara et al., 2004); and Swedish (Borin et al., 2009).

to robust models of frames and constructions. These include established formalized representations and relations between frame elements, frames, and constructions at multiple levels of analysis, from lexeme, to argument structure, to grammatical construct. FrameNet's development of structured representations of semantic information (Petruck, 2013) alongside the computational implementation of Embodied Construction Grammar (ECG) allow not only for codified representation but for analysis that has the potential for verification and replication, which constitute vital elements of any maturing theory that proposes to be scientifically rigorous and cognitively valid. In contrast, conceptual metaphor theory (CMT) has heretofore largely not been subject to the same kind of rigorous formalization necessary to bring it in line with other major aspects of cognitive linguistics. As we will see, a formalization of metaphor analysis is critical not only to enable advancement of the field – such as implementation of corpus approaches to metaphor – but also in order to model the intersection of metaphor and gesture in image schematic structure.

2.d. *Why formalize metaphor theory?*

Major theoretical developments in the history of CMT include the identification of primary and complex metaphors (Grady, 1997; Johnson, 1999), hierarchical levels of metaphor specificity (Lakoff and Johnson, 1999) and metaphor systems (Kövecses, 2010), and the systematic relationship between constructions and metaphors (Croft, 1993; Sullivan 2007, 2013). Additionally, significant advances in metaphor analysis, including development of corpus methodologies (Stefanowitsch, 2005; Stefanowitsch and Gries, 2006; Deignan, 2005) and systematization of metaphor identification (Pragglejaz Group, 2007; Steen, 2007; Steen et al., 2010) have allowed for both validation of metaphors across larger bodies of data and deeper analysis within small data sets. Early ventures in automated extraction (e.g., Mason, 2004) have demonstrated some success in identifying metaphor in corpora, but in limited domains and without a substantial basis in theory. Within the cognitive neuropsychology domain, the embodied underpinnings of conceptual metaphor continue to be validated with both behavioral and neuroimaging studies (see Gibbs et al., 2004; Bergen, 2012; Lakoff, 2012 for overviews).

Despite these advances, several substantive criticisms of contemporary CMT remain (e.g., Gibbs, 2009, 2011; Kövecses, 2008, 2011; Pragglejaz Group, 2007; Steen, 2007; Ruiz de Mendoza Ibáñez and Perez Hernandez, 2011). One main issue is that CMT relies too heavily on the intuitions of the individual linguist at work, and is insufficiently data-driven.[4] Because metaphor identification typically involves a top-down analysis model relying on the analyst to intuitively recognize metaphoric language, it can be perceived as a circular reasoning process by which analysts only identify metaphors they were already looking for, or only those metaphors of which they are already aware. Furthermore, because most metaphor analysis is performed by individuals or small working groups, data analysis must be relatively small-scale and limited to the amount of text a person is capable of parsing. In turn, this leads to criticisms of a lack of scientific rigor and objectivity, as well as the inability to replicate results; using external sources

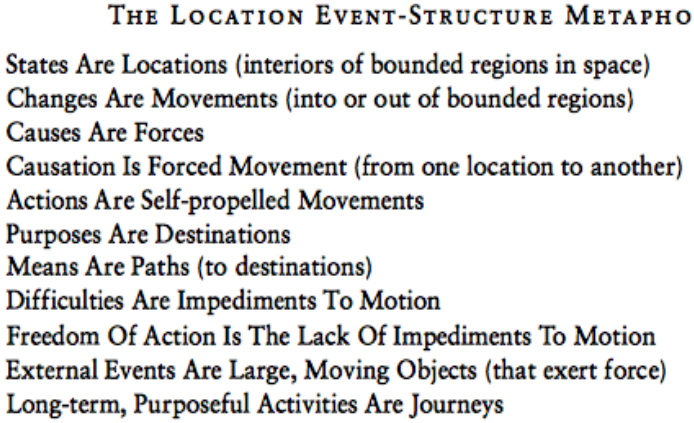
[4] Discussion of broader criticisms regarding the psychological validity of CMT (e.g., McGlone, 2007) is outside the scope of this work, but this has been addressed substantially elsewhere (e.g., Gibbs, 2009, 2011).

rather than purely relying on analyst intuition can increase agreement between analysts and improve consistency (Pragglejaz Group, 2007).

One approach that has somewhat avoided the pitfalls of individual analyst-driven metaphor analysis is the corpus method approach, which relies on pre-determined search terms (i.e., developing lists of source or target domain language to search for within corpora). This, however, can recapitulate the issue of analyses which only discover what the analyst already anticipates. Corpus methods serve to overcome the criticism that top-down metaphor analysis draws general conclusions from few (sometimes constructed) data-points because a bottom-up, corpus-driven approach develops generalizations based on patterns corroborated by naturally-occurring data sets (Kövecses, 2011).

Another criticism of CMT is that metaphor analyses are generally seen as being too prose-based and descriptive; metaphors identified within a text are named in the "TARGET-DOMAIN IS SOURCE-DOMAIN" format first established in Lakoff and Johnson (1980) and then further elaborated upon in prose discussion. Taxonomic, hierarchical relationships between metaphors are similarly under-developed; some classification of types of metaphors has been proposed, based on such criteria as source domain structure, levels of genericity, and types of mappings (Barcelona, 2000; Ruiz de Mendoza Ibáñez and Perez Hernandez, 2011). These taxonomic relations require further development and codification because, while there is apparent systematicity in sets of related linguistic expressions, the *nature* of that systematicity is still under-developed. While the compositional nature of metaphor has been observed (Grady, 1997; Lakoff and Johnson, 1999; Yu, 2011), the details of this compositionality are generally limited to contrasts between primary and complex metaphors. To illustrate, consider an excerpt of Lakoff and Johnson's (1999) analysis of the Location Event Structure Metaphor (Figure 6): it constitutes a list of related metaphors, accompanied by a chapter-length discussion of the metaphor family. While the structure of this list is discernable to the attentive reader, any structural relationships between the metaphors are not made explicit, and thus left to intuition. This work seeks to address the absence of a *formal* representation of the structure of metaphors and relationships between metaphors.

Figure 6. Early analysis of the Location Event Structure Metaphor (Lakoff and Johnson 1999: 179).



THE LOCATION EVENT-STRUCTURE METAPHOR

States Are Locations (interiors of bounded regions in space)
Changes Are Movements (into or out of bounded regions)
Causes Are Forces
Causation Is Forced Movement (from one location to another)
Actions Are Self-propelled Movements
Purposes Are Destinations
Means Are Paths (to destinations)
Difficulties Are Impediments To Motion
Freedom Of Action Is The Lack Of Impediments To Motion
External Events Are Large, Moving Objects (that exert force)
Long-term, Purposeful Activities Are Journeys

Ellen Dodge, Oana David, Jisup Hong, Eve Sweetser, George Lakoff, and I have proposed and described a system for frame and metaphor representations that has been implemented in a metaphor repository. The public repository is available at the MetaNet website at https://metaphor.icsi.berkeley.edu/pub/en/. Our analytical framework provides a formalization of CMT not previously found in the literature, which requires substantial representation of elements within individual frames and relations between frames, as well as relations both within and across metaphors. These relations are not just prose labels but contentful definitions, constituting theory-driven analysis of the structure of frames and metaphors. These definitions are subsequently employed in a computational metaphor extraction and metaphor modeling system, and are subject to consistency checking (Dodge et al., 2015; Hong, under revision).

One advantage of such a formalized system is the generalization of representation at multiple levels of analysis, thus enabling CMT to move beyond localized observations not extensible beyond particular data sets. Such formalization results in a large-scale network beyond what any individual analyst could produce, that is accessible to multiple analysts, such that the network can be searched and visualized at multiple levels of granularity. Its computational implementation allows for corpus-based metaphor analysis and identification of novel metaphors (see Dodge et al. (2015) and Hong (under revision) for automated metaphor extraction; Dodge (under revision) for corpus analysis). By requiring the analyst to adhere to a formal representation, the system is both descriptive and predictive in nature. Thus, "gaps" in the network indicate areas of analysis requiring further investigation, which can then be tested and cross-validated across multiple data sets, including cross-linguistic comparisons. Furthermore, the network's scalability allows for easy addition of metaphors and frames as analysts refine and expand the data in the repository, without affecting the robust nature of the system. Implementation of a formalization scheme also facilitates system-internal accuracy checking as well as external extension of the system into other applications. Formalization of metaphor theory aligns Conceptual Metaphor Theory with the current and future goals of Frame Semantics and Embodied Construction Grammar in terms of representation and analytical power, and allows for the possibility of addressing many of the current criticisms of CMT.

3. *Image schemas*

Central to conceptual metaphor theory is the notion of *image schemas*. I will be focusing on this type of conceptual structure throughout the work, as I will argue in Chapters 4 and 5 that the meaning of metaphoric gestures is best understood as collections of image schemas. In his Preface to his book on image schemas, Johnson broadly describes an image schema as "a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience" (1987: xiv). Langacker focuses on schematization in terms of abstraction: "a schema is abstract relative to its…elaborations in the sense of providing less information and being compatible with a broader ranger of options" (1987: 132). Image schemas, therefore, are abstracted representations of our bodily experiences. These conceptual representations of embodied experiences in turn structure our cognition; Oakley (2007) describes them as "distillers" of spatial and temporal experiences (2007: 215). Johnson (1987:126) identifies a partial list of primary image schemas; I include here those that will be discussed further in this work.

Blockage
Compulsion
Enablement
Path
Cycle
Scale
Iteration
Object

Cienki (1997) further argues that Johnson's (1987) image schemas have subcases, such as the Straight subcase of Path. From these mental representations, we can draw inferences and apply them to metaphoric mappings of the schema. For example, the primary Balance schema includes force vectors arranged equally around an axis. This notion can then be mapped onto morality; justice is conceived of as a countervailing force that reinstates balance after it is upset by a criminal activity. Thus, justice and criminal behavior are each metaphoric forces, with legal justice counteracting criminality to restore balance in the country's moral system (Johnson, 1987).

Beginning with Lakoff's (1990) Invariance Hypothesis (first formulated in Lakoff and Turner, 1989), inferences from image schematic structure have been recognized as the basis of the systematic mapping relationship between the source and target domain. The Invariance Hypothesis states that metaphorical mappings between source and target domain are *partial*, and that the portion of the source domain which *is* mapped preserves the image schematic structure of the source domain, which is topologically consistent with the structure of the target domain. Metaphors only map structure from the source domain that is compatible with the target domain. For example, the topology of the Path schema maps onto that of the Scale schema, which enables the metaphor SCALES ARE LINEAR PATHS. QUANTITY IS VERTICALITY constitutes a mapping between the vertical path of Verticality onto the scalar structure of quantity: points on the vertical path correspond to quantities, higher points to greater quantities, and lower points to lesser quantities. Thus, an increase in quantity can be understood as moving upward on a vertical path.

Sullivan (2007; 2013) extends the Invariance Hypothesis from image schematic structure to frame structure. She argues that constraints on lexical choice in metaphor can be explained by the coherence of mappings between frame elements in the source and target domain and the specific subframes evoked by the lexeme. For example, in American English the metaphoric usage of *brilliant* means 'intelligent' as in *a brilliant mind* and *sunny* means 'cheerful', as in *a sunny disposition*. Although at a high level they both evoke the Light frame, each evokes different subframes of the Light frame. Specifically, *sunny* evokes the Degree of Illumination of a Space by the light, while *brilliant* evokes the Emission of of Light by a Source. Therefore, the frame evoked by *sunny* is coherent with a STATES ARE LOCATIONS metaphor, HAPPY STATES ARE WELL-LIT LOCATIONS; and the frame evoked by *brilliant* is coherent with a KNOWING IS SEEING metaphor, INTELLIGENCE IS LIGHT-EMISSION. Conversely, because *brilliant* does not evoke the Illumination of a Space subframe, and *sunny* does not evoke the Emission subframe, they are

respectively incompatible with HAPPY STATES ARE LIT LOCATIONS and INTELLIGENCE IS LIGHT-EMISSION.

The role of frame element and image schematic coherence is brought to the forefront when studying metaphor in signed languages. In her work on iconicity in American Sign Language, Taub (2001) demonstrates that many signs have an underlying iconic and metaphoric structure. For example, many signs use an iconic representation of a concrete source domain to refer to the Communication frame, evoking the metaphor COMMUNICATION IS OBJECT EXCHANGE. The sign for INFORM consists of a hand grasping at the forehead and then moving forward, away from the signer and towards the addressee. Simultaneously the hand opens, as it "takes" the information from the signer's mind and "sends" it to addressee. The sign iconically represents the domain of Object Transfer, but its *meaning* is Communication. Taub shows that in these signs, the metaphoric mapping is coherent at a semantic level between source and target, *and* the mapping between activated source domain elements and those evoked by the iconic structure of the sign are coherent. She demonstrates that structure must be preserved in both sets of mappings: the metaphoric mapping from source to target, and the iconic mapping from source to the physical form of the sign.

Extending Taub's work, Meir (2010) develops the *double-mapping constraint*. This applies Taub's double-mapping analysis to metaphoric *usage*, as opposed to signs that are inherently metaphoric like INFORM. For example, Israeli Sign Language FLY is signed by "flapping" the arms like wings, iconically mapping the *manner* of flight – flapping wings – onto the physical action of flapping arms. This sign cannot be used in metaphoric contexts like *time flies*. The meaning of *time flies* focuses on the *speed* of the experience of time passing. When time "flies", our experience of time passing is rapid. Conversely, when time "crawls", we experience it as very slow. The metaphor here maps the speed of motion in the Motion frame onto the rate of time in the Time frame: RATE OF TIME IS SPEED OF MOTION. Although the metaphoric mapping is coherent, the mapping between the relevant frame elements of the source domain and the iconic elements of the sign FLY are not. In the context of the metaphor's meaning, *flies* evokes not the *manner* of flight, but rather only the *speed* of flight; the manner of motion does *not* map from source to target. The sign, however, does not iconically represent *speed*; it represents *manner*; the flapping arms of the sign represent the flapping wings of the bird. They do *not* represent the speed of the bird. Thus, the iconicity of the sign – flapping wings – is incompatible with the necessary mapping structure – speed of motion – of the metaphor.

In section 5, I discuss the role of iconicity and metaphor in gesture. As a visual-spatial system, gesture shares many features with signed languages. Like signs, metaphoric gestures are bound by similar constraints of iconicity imposed by the visual accessibility of the modality. In Chapters 4 and 5, I will discuss how frame and image schematic structure interacts with the realization of metaphor in gesture; I argue that metaphoric usage selectively activates certain elements of the source domain, which are then iconically depicted in the gesture.

4a. *Lexicalization of motion events: Manner and path verbs*

Before discussing motion event structure in gesture, I will first discuss the analysis of motion event structure in cognitive linguistic terms. Throughout the dissertation, I focus on the Location Event Structure Metaphor as evoked by two broad categories of English motion verbs: *manner* and *path* verbs. Literal motion events have been widely studied both in language and in gesture, as I describe further in section 5. However, metaphoric motion events in gesture are less-studied despite their prominence (as the source domain of the Location Event Structure metaphor) in metaphoric reasoning. The manner/path distinction draws on a now well-known typological variation in the lexicalization of motion events, as first discussed in Talmy (1985). His sketch of a motion event includes the motion, figure, ground, path, and co-event (including manner)/cause elements. The figure is the entity in motion; the ground, the entity that acts as a reference point for the figure's motion; the path is the motion of the figure's motion; and the manner is the way by which the figure moves. Consider the following:

(1) He ran into the house.

In (1), *he* is the figure and *the house* is the ground. *Ran* provides the manner of motion and *into* is the path of motion.

Talmy showed that languages divide into two categories as based on how they encode the path and manner elements of the event in the verb. (1) exemplifies a typical English sentence in that the manner is provided in the verb *ran*, and the path in a "satellite" (the preposition *into*). He provides three possible combinations of information within the verb: motion+manner (or co-event); motion+path; and motion+figure. These conflation types lead to the basic distinction: either **path** is conflated with motion in the verb (**verb-framed** languages as in (2)) or *manner* is in the verb and the path is in a satellite (**satellite-framed** languages as in (1 and 3)); the conflation of motion+figure falls into the satellite-framed category. In verb-framed languages, the satellite (if present) refers to manner, or a patient; in satellite-framed languages, the satellite conveys path; in figure-conflation languages, the path and ground are conflated.

(2) el hombre **entró** *corriendo* a la casa (Spanish; Slobin and Hoiting, 1994)
the man entered running to the house
'The man entered running to the house' (i.e., 'The man ran into the house')

(3) I *ran / limped / jumped / stumbled / rushed / groped* my way **down** the stairs
(English; Talmy, 1985)

Although Talmy (1985) argues that the 'core schema' of the motion event is the path, one notable fact that falls out from this typology is the variability of representation of manner cross-linguistically. Slobin (2006) recognizes that this results in variation in the salience of manner; languages that robustly encode manner in the verb will place a greater emphasis on manner in a variety of ways than languages that allow it to be optional. He argues that the degree of salience is due both to semantic constraints, such as the boundary-crossing constraint and cognitive processing constraints.

The boundary crossing constraint is a restriction found in verb-framed languages first proposed by Slobin and Hoiting (1994) on the basis of Dutch Signed Language (SLN, Netherlands) but relevant to spoken languages. Single events are broken up into multiple path components when they involve boundary crossings (as in *entering* or *exiting*); in spoken verb-framed languages, the manner verb can be used with path-focused, but not boundary-focused path verbs:

(4) a. El hombre **entró** *corriendo* a la casa (Spanish; Slobin and Hoiting, 1994)
'The man entered running to the house'

b. El hombre *corrió* **hasta** la casa
'The man ran up.to the house'

In (4a), the boundary crossing – entering the house – correlates with the path in the verb 'entró' and the manner in the satellite 'corriendo'. Conversely, (4b) focuses on the path, not the boundary, and therefore it is acceptable to have manner in the verb "corrió". Since Talmy's (1985) original analysis and subsequent revisions (Talmy, 1991, 2000), there has been considerable interest in applying his generalizations cross-linguistically, as he claims the distinction is fairly robust and supports it with data from a considerable number of languages.

4b. *Thinking for speaking*

In particular, Slobin (1987, 1996) hypothesized that our online thought processes during speaking were shaped by language-specific patterns and categories, basing his "Thinking for Speaking Hypothesis" on evidence from cross-linguistic variation in motion event conceptualization. He hypothesizes that language "directs us to attend – while speaking – to the dimensions of experience that are enshrined in grammatical categories" (Slobin, 1996: 71). While speaking (or signing), we activate our language-specific grammatical categories, which lead us to privilege certain types of event structure information over others. For example, speakers of a language without grammaticalized aspect such as Hebrew (5) will be much less likely to include aspectual information in describing a punctual event than speakers of a language such as Spanish (6), which distinguishes perfect and progressive, and in the past imperfective and perfective (Slobin, 1987). Hebrew speakers *can* include aspectual information, but it is not grammaticalized in their language, and so they are less likely to do so.

(5) הוא נפל ו הכלב ברח (Hebrew; Slobin, 1987: 438)
hu nafal ve ha kelev barax
he fall.PAST and the dog run.away.PAST
'he fell and the dog ran away'

(6) se cayó el niño y el perro salió corriendo
NON-AGENTIVE fell-PFV the boy and the dog exit-PFV running
'he fell (perfective) and the dog came out running'
(Spanish; Slobin, 1987: 437)

In (5), the Hebrew speaker does not specify that the boy falling is a punctual event. In (6), the Spanish speaker not only uses the past perfective to indicate the punctual nature of the boy falling, but also contrasts it with the durative nature of the dog running away, as indicated by the gerundive *corriendo* 'running'.

Subsequent work by Slobin and others (e.g., Strömqvist and Verhoeven, 2004) sought to demonstrate how such patterns – in particular, Talmy's motion lexicalization typology – influence both the information speakers encode in discourse and their mental conceptualization of events. Slobin (2006) observes there are several factors that facilitate frequent encoding of manner; particularly, there is the notion that "the rich get richer": a language with a rich manner lexicon causes its speakers to attend to manner more, and thus they attend to finer distinctions in manner, in turn adding more manner verbs to an already rich lexicon. Differences in attention due to manner salience have implications beyond the linguistic, as speakers of satellite-framed languages imagine events differently than speakers of verb-framed languages, who generally don't imagine manner when thinking about motion events. Furthermore, satellite-framed language speakers are better at learning new distinctions in motor patterns than verb-framed language speakers, suggesting a deeper cognitive effect. Thus, not only do linguistic categories influence the information we encode in speech, but they also affect our non-linguistic cognition.

Signed languages have particularly received attention in this regard, given their spatial nature. Slobin and Hoiting's (1994) original work on the boundary crossing constraint focused on SLN as well as American Sign Language (ASL); later, Taub and Galvan (2001) found that ASL signers include more information in speech than English speakers when retelling *Frog, Where Are You?*; they conclude that the highly spatial nature of signed languages allows for greater density of spatial information. Taub and Galvan (2004) further showed that ASL signers produce more manner and path as well as overall motion information than English speakers. However, Taub and Piñar (2001) found that including gestural information in their analysis leveled the playing field. In their comparison of English, Spanish, and ASL narratives, they found path information frequently occurred in gesture in all three languages, and Spanish and ASL speakers included more manner in their gestures. Unsurprisingly, English speakers preferentially lexicalized manner. They conclude that "gesture may be said to play a role in equalizing the amount of information expressed" (471).

Now that I have briefly discussed the necessary theoretical background from conceptual metaphor theory, frame semantics, and construction grammar, as well as the specifics of the motion event typology, I can turn to the primary focus of the dissertation: metaphoric motion event gestures. In section 5 I first provide background to the study of gesture and provide illustrations of the types of gestures I will be investigating in Chapters 4 and 5, before turning to discussing my motivations for studying them.

5.a. *Gesture: An overview*

This dissertation focuses on a specific type of what are typically called *representational* or *referential gestures*, which illustrate speech content using visual and spatial imagery (McNeill and Levy, 1982; McNeill, 1992; Müller, 1998). Although *gesture* is usually taken to mean actions made by the hands, it can include movements made by other parts of the body, such as

the head or shoulders. Representational gestures stand in contrast to nonrepresentational gestures in McNeill's typology, which include conversation-regulating speech act gestures (Bavelas et al., 1992; Wehling, 2017) and rhythmic beat gestures. Emblems are a separate class of gestures, as they have conventional meanings such as the "thumbs up" and "ok" signs. In contrast, the meaning of non-emblems is couched in the gesture's relationship to the overall utterance context. Rather than having a particular fixed meaning, the same gestural *form* could have different *meanings* depending on the context. I focus on a subset of representational gestures, *metaphoric gestures* (also known as *abstract gestures*). One class of representational gestures are *iconic* gestures, in which there is there a relationship between the gesture's form (such as its handshape, trajectory, and manner of motion) and entity and/or action it is depicting. Kita describes this as "a certain degree of isomorphism between the shape of the gesture and the entity that is expressed by the gesture" (2000: 162).

Metaphoric gestures are concrete, physical shapes and motions in space. However, the image schemas represented by the gestural forms refer to elements of the source domain of a conceptual metaphor. Hence, while the *form* of the gesture is iconic, its *meaning* is metaphoric. For example, Figure 7 shows a speaker producing an *iconic* gesture in his description of the motion of snow 'drifting'. He sweeps his right hand from left to right across his body, in a representation of the sideways movement of the snow. In Figure 8, we see another speaker producing a *metaphoric* gesture in his description of someone's attention metaphorically 'drifting'. They use the same verb *drift* and produce a very similar left-to-right gesture, which iconically depicts the movement of an entity away from the deictic center located in front of the speaker. However, Figure 7 depicts *literal* 'drifting', whereas Figure 8 depicts *metaphoric* 'drifting'. Brackets indicate the speech during which the gesture is produced; arrows indicate the trajectory of motion.

Figure 7. Example of an iconic gesture.



the snow is kind of starting to [drift] sideways [m.drift.6.26.2016c]

Figure 8. Example of a metaphoric gesture.



his [attention span] can sometimes drift a bit [m.drift.5.31.2015c]

This comparison illustrates two important points. First, the same (or very similar) representational gesture form can have different meanings depending on context. Second, metaphoric gestures are still *iconic* in the sense that they iconically represent the *source domain* of the metaphor. The speaker in Figure 8 is not producing a gesture that directly evokes the notion of 'paying attention' or 'failing to pay attention', although that is the topic of his speech. Instead, he is metaphorically construing 'failing to pay attention' as a person's attention 'drifting' away from the point of focus. Thus, his gesture evokes the 'drifting' source domain.

McNeill (1992) first identified a type of "abstract gestures" which reference locations in space. This type of abstract gestures makes use of space to establish the relative locations of imagined entities (which may be metaphoric). For example, a speaker comparing and contrasting two theories may "place" one theory to his left and the other to his right. Throughout his speech, as he continues to refer to one theory or the other, he gestures to his left or right accordingly. This use of contrastive spaces – left and right – sets up a visual contrast between the two theories. In his 1992 work, McNeill primarily focused on the use of metaphoric gestures in manipulating concepts in a conversational space. For example, in (7), the speaker "holds" their question in a cupped hand, evoking the COMMUNICATION IS OBJECT TRANSFER conduit metaphor (Reddy, 1980; Lakoff and Johnson, 1980). The hand holds the question (IDEAS ARE OBJECTS; THINKING IS OBJECT MANIPULATION), which can then be transferred into the communicative space (DISCOURSE SPACE IS PHYSICAL SPACE).

(7) I have [a question] (McNeill 1992: 149)
metaphoric: hand forms a cup for the image of a question

Kita (2000: 162) observes that the role of these gestures in communication is aided by their "relatively transparent form-function relationship"; for example, conversational participants readily understand that the "palm up, open hand" gesture conveys the notion of presenting or offering information (Kendon, 2004; Müller, 2004; c.f. Parrill, 2008). In Müller's (2008) gesture typology, she analyzes these as *performative* gestures, which enact speech acts (Searle, 1969). Much subsequent work on these types of metaphoric gestures has further elaborated on the pragmatics and metaphors of such conversation-regulation or "discourse management" conduit

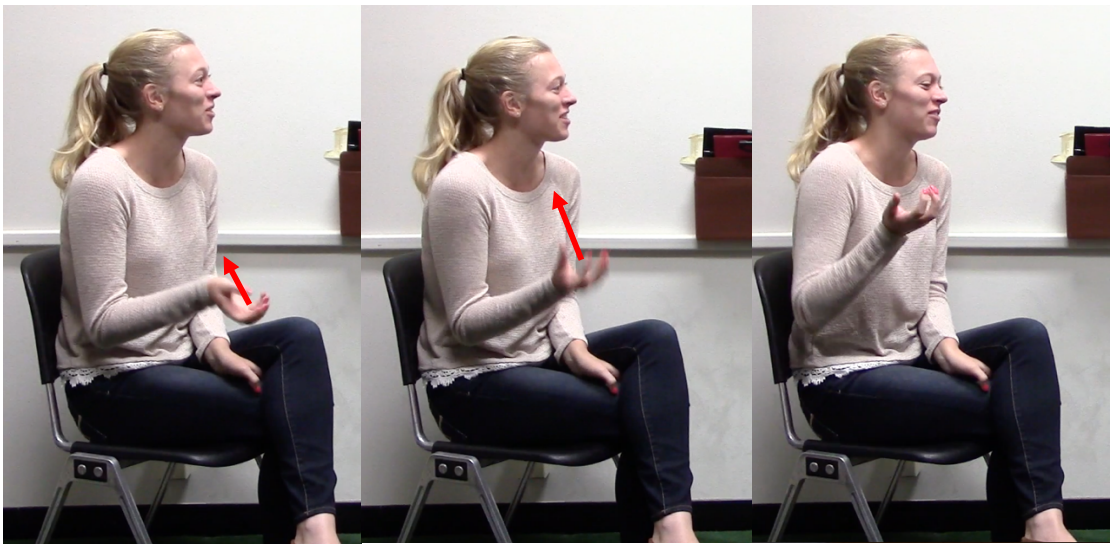
gestures (e.g., Bavelas et al., 1992; Calbris, 1990; Sweetser, 1998; Kendon, 2004; Müller, 2004; Wehling, 2017).

5.b. *Referential metaphoric gestures*

However, in this study I am specifically interested in *speech content* ("narrative-referential") metaphoric gestures, rather than *speech act* ("discourse management") metaphoric gestures. *Speech act* gestures, like those I describe above, govern the flow of information in the conversation (Sweetser and Sizemore, 2008; Wehling, 2017). In this section, I describe *speech content* metaphoric gestures. Müller (2008) considers these and iconic gestures to be *referential* gestures, since they refer to particular entities or actions. While iconic representational gestures have been extensively studied, most of the work on metaphoric representational gestures has focused on discourse-regulating rather than speech content gestures. Part of the reason for this is simply the issue of scale; in conversation, people tend to produce more discourse-management metaphorics than speech content metaphorics. They also produce *families* of discourse-management metaphorics: groups of gestures which are related in both form and meaning (e.g., Müller, 2004; Wehling, 2017). The "palm up open hand" mentioned above is an example of such a family, which occurs frequently in conversation with a reliable repertoire of meaning (Müller, 2004; Parrill, 2008). In contrast, speech *content* metaphorics are by definition governed by the content of the narrative and therefore expected to be more variable and idiosyncratic. A related class of gestures, *metaphoric beat gestures*, serve both the discursive, prosodic functions of beat gestures and convey metaphoric information (Casasanto, 2008). Like linguistic expressions of metaphor, metaphoric gestures exist along a cline of conventionality (Cienki, 2008; Müller, 2008) Thus, drawing conclusions about them requires a much larger body of data simply because the rate of re-occurrence is lower.

Alan Cienki and Cornelia Müller have been two major figures in the study of metaphoric referential gestures. Cienki has written extensively about the role of image schematic structure (1998; 2005; 2008; 2013), whereas Müller has focused on dynamic and semiotic elements (1998; 2004; 2008). Cienki (1998) as well as Cienki and Müeller (2008b) observe that the relationship between metaphor in speech and gesture can take several different forms. Gesture and speech can be *co-expressive* in that both modalities convey the same metaphor. In Figure 9, the speaker conveys the metaphor INCREASE IN QUANTITY IS UPWARD MOTION in both her speech *going up* and her iteratively upward-moving gesture. As Cienki (1998) notes, the gesture may convey additional force-dynamic information not present in the speech. In Figure 10, the progressive aspect of the speech conveys the ongoing nature of the improvement in confidence, but the iterative nature of the gesture suggests that the aspectual structure of the progress may have been punctuated rather than smooth.

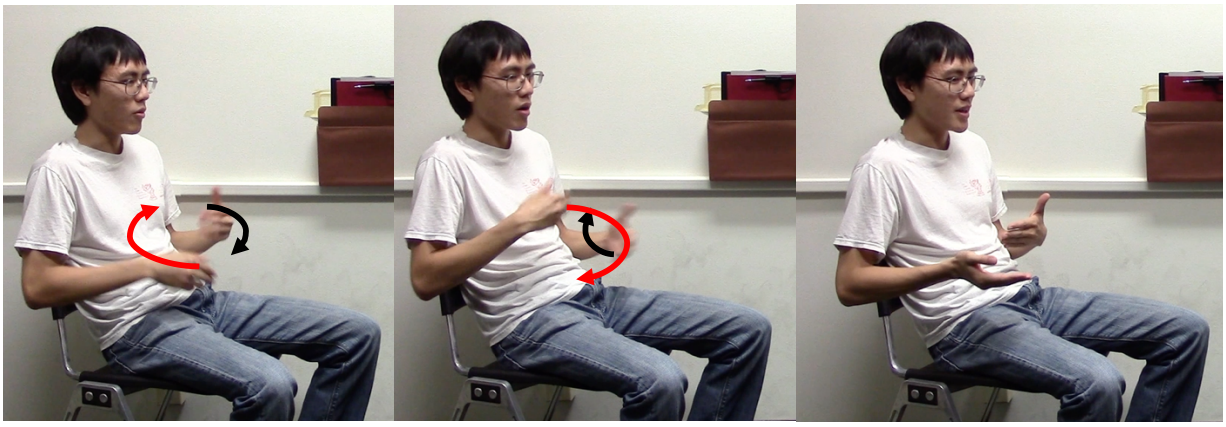
Figure 9. Co-expressive metaphoric gesture: Same metaphor in both modalities.



so [his confidence was going] up [S4_2X_A_10.20.15_ELV]

They also show that *different* metaphors may be realized in speech and gesture simultaneously, such that the gesture is *complementary* to the speech. The metaphors may have the same target domain, but different source domains, as in Figure 10.

Figure 10. Complementary metaphoric gesture: Different metaphors in each modality. (Right hand: red arrows; Left hand: black arrows).



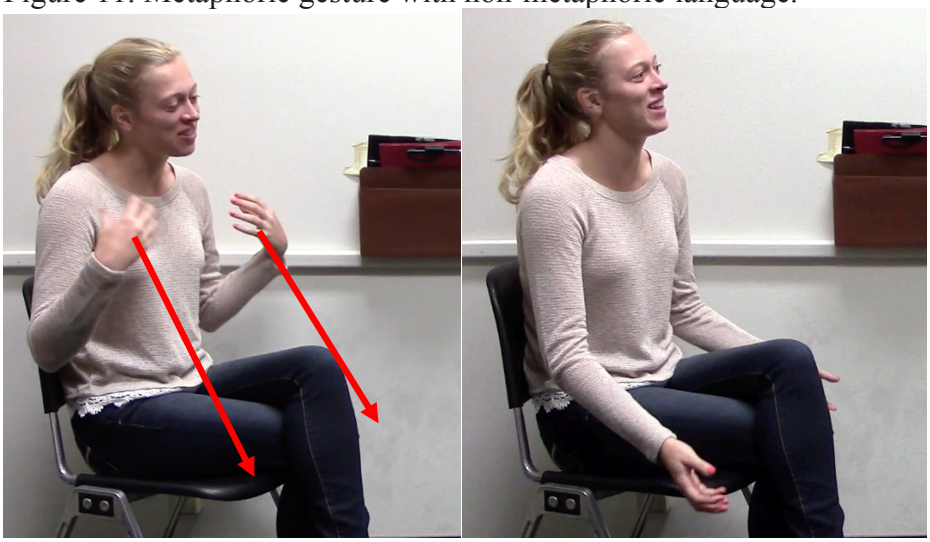
she was getting [more and more confused] [S1_16X_A_10.9.15_EMS]

In Figure 10, the speaker's speech uses the Object Event Structure Metaphor, EXPERIENCING AN EMOTION IS ACQUIRING AN OBJECT: *getting more and more confused*. His cyclic gesture, however, represents the aspectual structure of the event: the continuous nature of the increasing confusion. Thus, it represents an entailment of the Location Event Structure Metaphor, CONTINUOUS CHANGE OF STATE IS CONTINUOUS CHANGE OF LOCATION. This is particularly interesting as the Location Event Structure Metaphor and Object Event Structure Metaphor are metaphoric *duals*; they are normally understood as mutually exclusive in that they constitute opposite viewpoints

on the relationship between Experiencer and State (Lakoff and Johnson, 1999). However, this gesture illustrates that speakers can actually make use of multiple metaphors – here, metaphors about state change – simultaneously.

Third, Cienki and Müller (2008b) show that speakers may produce a metaphor gesturally with no metaphor present in the speech at all. We can conclude that the gesture is metaphoric because it coincides with speech content that is abstract, as in Figure 11.

Figure 11. Metaphoric gesture with non-metaphoric language.



so then his confidence totally [diminished] [S4_2X_A_10.20.15_ELV]

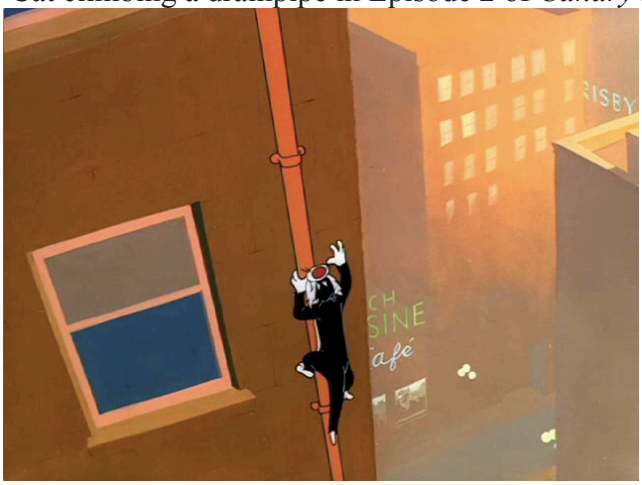
The speaker in Figure 11 continues her story from Figure 9. Previously she used a metaphor with the source domain of Verticality in both her speech and gesture when describing the character's confidence improving. Here as she describes his confidence dropping she uses non-metaphoric language *diminishing*, but still produces a vertical drop in her gesture. The abstract domain of emotional state change in her narrative, along with her prior use of CHANGE IN STATE IS MOTION ALONG A VERTICAL SCALE, can lead us to conclude that here her gesture is metaphoric as well despite the lack of a metaphor in her speech.

5.c. *Thinking for speaking and gesturing*

As the examples in the previous section illustrate, motion event structure plays a key role in metaphor in both language and gesture. My interest in motion event lexicalization comes not only from the primacy of motion in metaphor theory, but also this centrality in both literal and metaphoric event structure. The distribution of manner and path information has been studied in gesture since early seminal work by McNeill (1992), which established it as a main point of inquiry. McNeill (1992; 2005) and McNeill and Duncan (2000) investigated the gestural expression of motion events across a variety of typologically contrastive languages. Following the paradigm established by Slobin and his colleagues (Berman and Slobin, 1994; Slobin, 1996), eliciting motion events cross-linguistically is typically done using visual stimuli such as story books or cartoons that have little to no dialogue (which is beneficial for cross-linguistic studies) and an emphasis on physical actions, such as characters climbing drainpipes, falling into ponds,

and flying out of windows. Linguistic study of motion events often makes use of *Frog, Where Are You?*, a children's picture book (Strömqvist and Verhoeven, 2004). To elicit gestural data in study participants, McNeill (1992) and many others following make use of video stimuli such as *Canary Row*, which is a series of short Sylvester and Tweety cartoons (Figure 12).

Figure 12. Sylvester the Cat climbing a drainpipe in Episode 2 of *Canary Row*.



McNeill (1992) established the dominant paradigm in experimental approaches to gesture studies, typically used by psychologists and linguists. In this design, participants are paired in conversational dyads (or a participant and confederate). One participant is shown a stimulus, such as a video, picture book, or static image, and told to commit it to memory. They then retell or describe what the stimulus depicts to their conversational partner. This approach can also be used to elicit more natural storytelling or dialogue, such as by prompting participants to tell a story from their own lives (e.g., Sweetser and Stec, 2016), or provide their opinions on a topic (Cienki, 1998). This "quasi-experimental" approach, as Coopperrider and Núñez explain, is useful in that it provides a "methodological middle way between, on the one hand, rigorously controlled experimental studies that tend to denature conversation and, on the other, ethnographic studies that offer beautiful specimens but do not always disclose general trends" (2009: 198). They suggest that this approach might be particularly suitable for further study of metaphoric gesture, because the task is controlled; without knowing what imagery prompted a gesture, it is difficult to conclude the meaning of the gesture. This is perhaps particularly true when the topic of discourse is an abstract concept, because the iconicity of the gesture reflects the source domain of the metaphor rather than the actual topic, the target domain of the metaphor.

The realization of manner and path in gesture has been studied in a variety of contexts, particularly in studies of second language acquisition and cross-linguistic comparisons. As Sweetser (2007) observes, the intimate link between gesture, language, and thought makes gesture an excellent route for studying the cognitive processes accompanying language. Thus, a great body of work in gesture studies looks at not only "thinking for speaking", but "thinking for speaking and gesturing". Given the field's historic focus on motion events and gesture's capacity for conveying spatial information, it is unsurprising that motion event structure has played a large role in gesture studies.

Early studies by McNeill (1992, 1997) sought to compare narratives cross-linguistically with a focus on motion events. He predicted that motion information would be packaged differently gesturally, depending on how semantic categories are encoded by the language. In their descriptions of Tweety Bird dropping a bowling ball down a drainpipe, speakers of Georgian (8), Swahili (9), and English (10) each made a downwards gesture (indicated on the speech translation with brackets). However, the syntactic *timing* of the gesture varied: each speaker timed their gesture with the part of speech which corresponded to the semantic category most salient to the event: the downward motion of the ball. The Swahili speaker even isolates the gesture to the verb root in a larger polymorphemic verbal complex.

(8) da uzarmazar rk'inis burts … cha[agdebs] (Georgian; McNeill, 1992: 222)
and enormous iron ball … [throw-down]
'and [threw down] an enormous iron ball'

(9) i-ka-chuku-a li-mpria fulani i-ka … [tum]buk-iz-a (Swahili; McNeill, 1992: 223)
take tire certain … [push down]
'and found a certain tire … and pushed it down'

(10) and Tweety Bird runs and gets a bowling ball and drops [it down] the drainpipe
(English; McNeill, 1992: 225)

McNeill and Duncan (2000) later showed that distributions in co-timing of representational gestures and variation in the manner encoded in the gesture reflected differences in verb framing. For example, when the manner is a core element of the event, English speakers co-time manner gestures – which depict the manner of motion – on the motion verb, which also encodes manner. Otherwise, they can omit manner of motion in the gesture and instead gesture path elsewhere in the sentence. Because English is a satellite-framed language, the manner will still be included in the verb. In contrast, Spanish speakers may produce gestural manner in the absence of lexically encoded manner. Because Spanish is a verb-framed language, including manner requires a bit more work syntactically (i.e., adding an adjunct). Thus, an easy "workaround" is to express manner gesturally instead.

Following McNeill's example, many researchers have since made use of these lexicalization patterns in studying language and gesture, particularly in the context of understanding the greater underlying conceptual structure at work. For example, in her study of path gestures made by Spanish speakers learning English, Stam (2006) found Spanish thinking-for-speaking patterns in their gestures despite making use of correct English patterns in their speech. This showed that the English learners still made use of their native Spanish *conceptualization* patterns – as belied by their gestures – in spite of acquiring English *linguistic* patterns. Their first language's grammatical structures still affect their cognition as demonstrated by their gestures, even when speaking in their second language. More broadly, the gestural realization of manner and path has been studied to understand the interaction between gesture and language in the context of "thinking for speaking". Kita and Özyürek's (2003) study of differences in gestural expression of motion events in Turkish, Japanese, and English demonstrates that gesture production aids the packaging of spatial information. Özyürek et al. further demonstrated how language-specific patterns in Turkish and English motion event construal predict differences in gestural

representations, arguing that "online speaking gestural and linguistic representations interact in such a way that gestures reflect the underlying online conceptualization that fits appropriate semantic and syntactic encoding of events" (2005: 237).

Despite the central role of motion in both metaphor and gesture studies, remarkably little work has been done on metaphoric event structure gestures, especially as they relate to grammar. For example, the interaction of aspect and gesture is only beginning to be explored, particularly cyclic gestures (Parrill et al., 2013; Hinnell, 2013; Mittelberg et al., 2015). At the most recent International Society of Gesture Studies conference as of this writing (2016), Cienki organized a symposium on the interaction of aspect and gesture across French, German, and Russian (Cienki, 2016). Additionally, Tong and Cienki (2015, 2016) have begun to compare image schematic structure in literal and metaphoric object manipulation events. It is worth noting that nearly all of these citations are conference presentations from the last three years; clearly, the field is only getting started in this regard.

5.d. *Why study metaphoric gesture?*

The publication of Cienki and Müller's (2008a) edited volume on metaphor and gesture coincided with the development of new tools in gesture research (in particular, video databases and motion capture), which together ushered in an era of increased research in the field. Metaphoric gestures have come to play an important role in cognitive linguistics because they provide non-linguistic evidence for the existence of conceptual structures posited to underlie our language and cognition (Núñez and Sweetser, 2006; Sweetser, 2007; Cienki, 2008; Cienki and Müller, 2008b; Langacker, 2008b; Cooperrider and Núñez, 2009; Gibbs, 2011). One common criticism of conceptual metaphor theory is that much of its support is linguistic in nature; given that conceptual metaphor is claimed to be *conceptual*, it should be observed in non-linguistic expression as well (Kövecses, 2008; Gibbs, 2009). As gesture and language production are tightly linked at the conceptualization stage (e.g., Alibali et al., 2000; McNeill, 2005), it holds that the cognitive processes behind linguistic metaphors are those driving gestural expression of metaphor as well. Thus, by studying metaphoric gesture, we observe evidence for the conceptual nature of metaphor itself. This extends beyond conceptual metaphor specifically to other points of interest in cognitive linguistics; for example, it provides a useful route to study mental spaces via Mental Spaces Theory (Parrill and Sweetser, 2004; Sweetser, 2007). More generally, the embodied nature of gesture provides evidence for the embodied nature of cognition. As Langacker aptly puts it: "The form and application of metaphoric gestures give palpable indication of a basic claim of cognitive semantics, namely the embodied nature of meaning and the grounding of abstract conceptions in perceptual and motor experience" (2008b: 249).

Further evidence for the importance of gesture in metaphor research comes from gesture perception studies. Metaphoric gestures have been shown to influence the metaphors *addressees* use to reason with: not only are these gestures evidence of the metaphors the *speaker* is using, but they also have an impact on the metaphors the *addressee* uses too. For example, Jamalian and Tversky (2012) showed that addressees' answers to the ambiguous question, *Wednesday's meeting has been moved forward two days. What day is the meeting on now?* can be primed by pairing that question with a sagittal forward or backward gesture. This question has been extensively used in studying metaphor priming (e.g., Boroditsky, 2000; Boroditsky and Ramscar,

2002; Matlock et al., 2005); it has two possible answers, Monday or Friday, each of which reflects a different metaphoric construal of time. A Monday answer indicates that the addressee is using the metaphor TIME IS A MOVING ENTITY metaphor, with time conceptualized as an entity in motion towards the individual, who is at the present. Thus, moving the meeting "forward" moves the meeting closer to the individual and hence the present time, i.e., earlier. A Friday answer indicates the use of the EXPERIENCING TIME IS MOVING ALONG A PATH metaphor, wherein the individual sees herself as moving through time, with the future ahead of her. In that case, moving the meeting forward moves it farther ahead of the individual – farther into the future. Jamalian and Tversky's (2012) work showed that addressees' use of the Monday or Friday-evoking metaphors could be primed by gesturing either forward – evoking the Friday metaphor, as the motion is moving farther ahead of the speaker – or backward – evoking the Monday metaphor, as the motion is moving closer to the speaker (Figure 13).

Figure 13. Backward-moving sagittal gesture.[5]



My work with Tasha Lewis (Lewis and Stickles, in press) extends this work by testing the same question with metaphoric gestures in different contexts. While Jamalian and Tversky (2012) only tested addressees side-by-side with speakers, to ensure that they shared the same perspective, we also tested them face-to-face, such that a gesture away from the speaker was a gesture *toward* the addressee, and vice-versa. In such face-to-face discourse, the addressee can either perceive a forward sagittal gesture as movement away from the speaker (priming the Friday-evoking metaphor), or movement toward the addressee (priming the Monday-evoking metaphor). We found that addressees maintained their own perspectives: the forward gesture primed the TIME IS A MOVING ENTITY metaphor, and the backward gesture the EXPERIENCING TIME IS MOVING ALONG A PATH metaphor. These results are the opposite of those in the shared perspective condition in both Jamalian and Tversky's (2012) and our results. We even found that *lateral* gestures (Figure 14), which are not iconic for the source domain of forward motion from the speaker's perspective because they move to the speaker's sides, still prime use of spatiotemporal metaphors in addressees.

[5] Figures 13 and 14 are from Lewis and Stickles (in press). Tasha Lewis has given me permission to make use of this work here.

Figure 14. Left-to-right lateral gesture.



While the ambiguous test question evokes metaphoric construals of time in relation to the position to the speaker (Figure 15), the lateral gesture evokes a metaphor that does not relate time to the speaker. Instead it evokes a notion of time passing, independent of the speaker's viewpoint (Figure 16). Thus the forward motion of time can occur on an individual's lateral axis as it "passes by" her (Dancyiger and Sweetser, 2014). This is realized in English gesturally as a lateral motion across the body; this metaphoric gesture for time actually occurs more often than the sagittal variant (Cooperrider and Núñez, 2009; Casasanto and Jasmin, 2012). In general, both sagittal and lateral spatiotemporal gestures are common cross-linguistically (e.g., Calbris, 1990; Núñez and Sweetser, 2006).

Figure 15. Metaphoric construals of time relative to speaker's viewpoint. (A) TIME IS A MOVING ENTITY; (B) EXPERIENCING TIME IS MOVING ALONG A PATH.

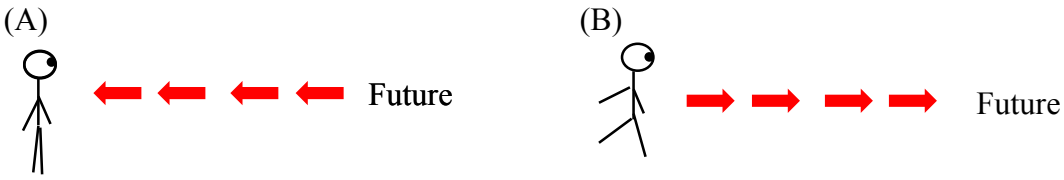
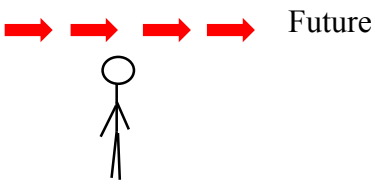


Figure 16. Metaphoric construal of time independent of speaker's viewpoint.



Tasha Lewis and I found that a lateral gesture, when produced in conjunction with the ambiguous "Wednesday's meeting" question, primes a Monday or Friday response (depending on the gesture's direction: leftward gestures move earlier in time and rightward, later). Despite representing a different conceptualization of time than the metaphor in speech, these gestures were still taken as informative by addressees and incorporated into their understanding of the

question. From these findings we can conclude that not only do we regularly produce multiple metaphors across speech and gesture, but we also perceive and comprehend multiple streams of metaphoric information across modality as well.

5.e. *Experimental approaches to metaphoric gesture*

In the prior section, I discuss some of the current experimental research on metaphoric gesture *comprehension.* Relatively little work has been done on elicitation of metaphoric gesture *production* in experimental or quasi-experimental settings. Cienki (1998) videotaped students discussing morality and ethical issues as part of a larger project on morality; this resulted in the production of metaphoric gestures, due to the frequency with which issues of "goodness" are described in metaphoric contexts. Casasanto (2008) similarly elicited open-ended narratives by asking participants to tell autobiographical stories with either positive ("tell me about a time you felt good about yourself") or negative prompts. Chui (2011) collected data from a corpus of videotaped free-form conversations between Mandarin speakers; she then focused on two conversations for detailed analysis. None of these studies were experimental approaches per se in that they did not attempt to control the linguistic or gestural output with a particular stimulus (beyond open-ended prompts). However, they did entail elicitation of naturalistic speech and gesture by study participants. Because Cienki (1998) and Casasanto (2008) prompted participants to focus on topics related to abstract states – morality, positive and negative emotions – they were likely to elicit metaphoric language and, hence, metaphoric gesture.

A more targeted study of metaphoric gesture by Cooperrider and Núñez (2009) was designed to elicit spatiotemporal gestures by having participants study and then describe a large poster representing the time-course of the history of the universe. They concluded that this quasi-experimental design holds promise in studying metaphoric gestures in well-defined semantic domains. Casasanto and Jasmin (2012) focused on eliciting both transverse and sagittal spatiotemporal metaphors, both by explicitly instructing participants to produce gestures and by using the familiar "storytelling" format.

5.f. *Corpus approaches to metaphoric gesture*

Corpus approaches to referential metaphoric gesture are a relatively recent phenomenon, due to previous technological limitations. An early exception is Calbris (2003), who analyzed a corpus of gestures produced by a politician over six television interviews. With the advent of television video databases which allow searching over closed captions, gesture researchers can now collect much larger amounts of data over multiple speakers and demonstrate broader trends that generalize beyond the idiosyncrasies of individuals in more naturalistic contexts than experimental settings. One commonly used resource is the Little Red Hen UCLA NewsScape Archive co-directed by Francis Steen and Mark Turner, which provides closed-caption network television from around the world from 2005 to the present. Currently it has over 250,000 hours of searchable videos. Hinnell (2013, 2014) and Tong and Cienki (2015, 2016) both make use of the Little Red Hen database to develop their corpora. Hinnell (2013, 2014) focuses on gestural correlates of aspect by searching for inflections of semi-auxiliaries (*continue, keep, start,* and *stop*). Tong and Cienki (2015, 2016) collected data by collecting instances of object

manipulation verbs (such as *put, pull, lift, hold*) and searching for co-speech gestures in both metaphoric and literal verb usage.

Another resource is the TV News Archive, which also provides over 1,106,000 searchable closed-caption television starting from 2012. Winter et al. (2013) make use of the TV News Archive to investigate metaphoric construal of numeric quantity by searching for terms combining spatial adjectives or gerunds such as *rising* and *tiny* with *number*, as in *tiny numbers* or *rising number*. All of these studies demonstrate the sheer labor costs and low return in developing metaphoric corpora from databases. For example, Winter et al. (2013) searched for 20 expressions over a two-month period (over 350,000 news programs), resulting in 552 videos. Of those, many had no visible gestures; as a result, they chose 27 to analyze. In my corpus study in Chapter 4, I have 130 metaphoric and 49 literal gestures resulting from a search of nearly two years' worth of Little Red Hen data.

6. *Summary*

This dissertation is a wide-ranging exploration of the interaction of metaphor, gesture, and grammar in usage; consequently, its diverse findings have implications for conceptual metaphor theory, theories of language and cognition, and gesture studies. In Chapter 2 I provide a formalization of conceptual metaphor theory that advances the representation of conceptual metaphors to a level that interfaces more accurately with representations of frames and constructions in FrameNet (Ruppenhofer et al., 2010) and Embodied Construction Grammar (Bergen and Chang, 2005). In turn, this enables the detailed analysis of metaphors and metaphor systems, as exemplified by the Location Event Structure Metaphor (Lakoff and Johnson, 1999) case study in Chapter 3. This corpus-based study is one of the first to make use of the MetaNet method for large-scale automatic metaphor identification and annotation.

In Chapter 4, I demonstrate the benefits of taking an image-schematic approach to metaphoric gesture analysis. I argue that the image schema is the most appropriate level of structure in analyzing the form and meaning of metaphoric gestures. Results of this image schema analysis suggest that these metaphoric gestures emphasize path and do not represent the manner of motion. Chapter 5 is the first study to take an experimental approach to metaphoric gesture that uses non-metaphoric stimuli. I find that speakers are more likely to produce metaphoric gestures if they are also producing metaphoric language – even if the gesture evokes a *different* metaphor than the speech does. These results are discussed in terms of the Gesture as Simulated Action framework and provide evidence for the Thinking for Speaking Hypothesis. Finally, I unify my analyses of metaphoric motion in speech and gesture in a multi-modal Embodied Construction Grammar analysis of both co-expressive and complementary metaphoric co-speech gestures. This analysis provides the first formal representation of a multi-modal utterance in a construction grammar and an innovative approach to the unification of the construction of multi-modal meaning.

7. *Organization of the dissertation*

Broadly speaking, following this introductory chapter this dissertation is divided into two parts. Part One – Chapters 2 and 3 – focuses on the formalization of conceptual metaphor theory. Chapter 2 is a based on a paper forthcoming in *Constructions and Frames* co-written by myself, Ellen Dodge, Oana David, and Jisup Hong (Stickles et al, 2016b). It describes our proposal for a formalization of metaphor theory as instantiated in the MetaNet metaphor repository system. We describe a series of semantic relations between metaphors and frames, similar to but expanding upon those in the current revision of FrameNet (Ruppenhofer et al., 2010) and incorporating details of frame and construction representations from Embodied Construction Grammar (Bergen and Chang, 2005; Feldman et al., 2009). We argue that this approach allows for not only computational approaches to metaphor computation and extraction, (Dodge et al., 2015; Stickles et al., 2016a; Hong, 2016) but also a more robust approach to metaphor analysis itself. As I discuss earlier in this chapter, conceptual metaphor theory has lagged behind other major elements of cognitive linguistics (frame semantics, construction grammar) in its lack of formalization. We argue that this approach enables all three of these major elements of cognitive linguistics to unify in a coherent representation. This chapter then concludes with an analysis of the Location Event Structure Metaphor network, which leads to Chapter 3.

Chapter 3 is a detailed case study of the Location Event Structure Metaphor. In order to properly analyze the metaphoric gestures in Chapters 4 and 5, it is necessary to first fully understand the motion-based metaphors those gestures evoke. This chapter builds on Chapter 2 by expanding its analysis of the Location Event Structure Metaphor into an Embodied Construction Grammar account that incorporates metaphoric representations. I then discuss results of a corpus search using the MetaNet metaphor extraction engine to identify instances of this metaphor network in the Corpus of Contemporary American English (Davies, 2008). Analysis of these results reveals a privileging of the path component of the Translational Motion frame (and more-specific subcases of that frame). I conclude that while English, as a satellite-framed language, privileges manner in its lexicalization of motion events, *metaphoric* English motion backgrounds manner and foregrounds path. Nonetheless, manner is still nearly obligatory in English speech. For this reason, in Part Two I turn to analyzing metaphoric gestures that evoke the Location Event Structure Metaphor.

As many cognitive linguists have argued, gesture provides a route to study the cognition associated with language. While English speakers may be producing metaphoric manner verbs due to the lexicalization patterns of their language, are they necessarily *thinking* in terms of metaphoric manner? It's hard to judge this by looking at language alone. While gestures are influenced by a language's structures, they are not bound by them. Part Two (Chapters 4 and 5) investigates Location Event Structure-evoking metaphoric gestures using two complementary approaches. Chapter 4 uses a corpus methodology; over a span of about two years I collected short video clips from searchable video databases of English speakers using a motion verb metaphorically in their speech while producing a co-expressive representational metaphoric gesture. I also developed a parallel corpus of the same verbs and gestures in literal usage contexts. To analyze the corpora, I develop a set of annotation guidelines and then demonstrate the benefits of taking an image-schematic approach to gesture analysis (when specifically interested in analyzing the frame elements evoked by the gesture, as is the case here). Results of

this image schema analysis suggest that, reflecting the English language data in Chapter 3, these metaphoric gestures emphasize path over manner. Furthermore, we also see a backgrounding of the metaphoric moving entity in the Translational Motion frame.

However, there are natural limitations in this approach to gesture analysis, especially since we are focused on metaphoric gestures. The corpus approach necessitates collection of gestures which are co-expressive and speech which in which both the source and target domain of the metaphor are lexically instantiated. It is well-established that speakers frequently do not instantiate the source domain of metaphors and leave it to discourse context or frame instantiation to fill in the gaps. Furthermore, gesturers often produce complementary gestures that express different information than their co-speech. Thus, the corpus approach in Chapter 4 is a skewed representation of how English speakers actually use the Location Event Structure Metaphor in their gestures. This issue is addressed by Chapter 5, which takes an experimental approach. In this study, participants were given short stories to read and re-tell to a friend. The stories all contained state change events culturally familiar to American English speakers. For example, in one story a student worked to improve his test scores; in another story, an investor in the stock market lost his life savings. In both cases, a state (grades; stock market prices) changes either positively (grades improving) or negatively (prices decreasing) and can be described using metaphoric language (grades *climbing*; prices *dropping*). Results from this study demonstrated the viability of this methodology in eliciting both metaphoric speech and gesture, and supported those of Chapter 4. Intriguingly, I find that speakers are more likely produce metaphoric gestures if they are also producing metaphoric language – even if the gesture evokes a *different* metaphor than the speech does. I discuss implications of these findings in Chapter 6, the conclusion.

The concluding chapter of the dissertation (Chapter 6) seeks to draw together the results of Parts One and Two in a unified approach to metaphor analysis that combines linguistic metaphor and gestural metaphor. First I summarize the results of Chapters 2-5 and discuss the implications of results from Part Two for conceptual metaphor theory and the Gesture as Simulated Action hypothesis (Hostetter and Alibali, 2008). Then I conclude with a sketch of an Embodied Construction Grammar analysis that incorporates meaning from metaphoric gesture into the representation of the utterance.

8. *A note on representing the visual-spatial modality*

Throughout the past seven years, I have struggled to adequately represent signed and gestured data in print. In conference presentations and classrooms, we have the luxury of playing and replaying videos at fractions of the actual speed, or can at least resort to recreating the gestural data ourselves. Unfortunately, we don't yet have the option of embedding playable video in PDFs; this leads to the challenge of translating dynamic three-dimensional information to a static two-dimensional page. Given my focus throughout the dissertation on manner and path, it is critical that those elements of gesturing are made explicit in data presentation. To do so, I make use of multiple stills, screenshotted from the video data. The path transcribed by the hand's movement through space is indicated by red or black arrows overlaid on the photos, along with in-text description of the gesture. The speech utterance accompanying the gesture is provided as a caption beneath the images; text [in brackets] indicates when in the gesture stroke was produced. Each video has a unique identifier associated with it, as cited in the text. The data will

be made publically available to the extent possible on my website at elisestickles.com/dissertation. Some of the data from the experiment in Chapter 5 will not be publically released at the request of study participants. This link is permanent and should work even if my university-associated website address changes.

Following signed language linguistics convention, signs will be represented as the closest English gloss and written in all capitals, such as the sign INFORM 'to inform'.

9. *A note on representing conceptual structures*

Throughout the dissertation I follow several conventions in cognitive linguistic representation. In order to distinguish lexical and morphological structures from conceptual structures, conceptual-level entities have different capitalization conventions. Frames, image schemas, and constructions are capitalized, as in the Commercial Transaction frame, the Container image schema, or the Caused Motion Construction. Names of individual metaphors appear in small capitals, as in KNOWING IS SEEING. However, names of metaphor families are capitalized as proper nouns, such as the Location Event Structure Metaphor.

Chapter 2: Formalizing conceptual metaphor theory[6]

1.a. *Introduction*

In this chapter and the next, I briefly discuss the theoretical background of conceptual metaphor theory (CMT) and then describe the MetaNet approach to the formalization of CMT. We propose to refine the notion of hierarchical organization in frames and metaphors to include a variety of relations between frames and metaphors. By not only positing these relationships but implementing them in a computational framework, we can model the conceptual structures in a level of complexity beyond the ability of the average analyst working with a whiteboard or pen and paper. Furthermore, this implementation – the MetaNet Metaphor Repository – can be leveraged in an automated metaphor identification and annotation process, which given a text corpus and queries (such as particular source or target domains), can return a corpus annotated for metaphoric language including frame and metaphor identification and metaphoricity rating. In this chapter I focus on the Location Event Structure Metaphor (LESM), as translational motion constitutes the source domain structure of that metaphor family. The LESM is a likely cognitive universal. This chapter first describes the theoretical background and formalization of CMT with the LESM as a case study. The next chapter briefly discusses the crucial role of constructions in metaphor analysis, and then demonstrates an application of the MetaNet system to identifying the LESM in the spoken sub-corpus of the Corpus of Contemporary American English (Davies, 2008). In doing so we can observe patterns in use frequency at the lexical and syntactic levels, and in turn better understand how the LESM is realized in English usage.

1.b. *Roadmap*

We first describe in brief the theoretical underpinnings of CMT, including the nature of semantic frames and their relation to metaphors. From this we detail how frame analyses are developed in service of metaphor analysis; these frames and their ensuing conceptual network form the basic conceptual structure on which metaphor theory depends. After discussing the internal structure and logic of frames, we define and illustrate a set of relations between frames, which provides the structure of the network itself. Given that metaphors are bundles of mappings between frames, we then turn to conceptual metaphors themselves. We show that internal frame structure, including analysis of individual frame elements, is crucial to formalizing metaphor structure. Finally, we present an improved understanding of the relationships between metaphors at varying levels of specificity, as driven by the underlying relationships between the frames that comprise the source and target domains of the metaphors.

2.a. *Development and formalization*

The framework developed here reflects several fundamental premises shared across cognitive semantics (Clausner and Croft, 1999). The main level of analysis, variously referred to as a frame, base, or domain in the literature (Fillmore, 1976; Langacker, 1987, 2008a; Lakoff, 1987),

[6] This chapter appears as Stickles et al. (2016b); all co-authors have kindly provided their permission for me to reprint it here. I use "we" throughout to indicate that this portion of the dissertation is co-authored.

comprises the cognitive structures which support and define basic semantic concepts. Central to frame semantics is the notion that frames are structured, not just associated lists of semantic concepts or conditions; they are schematized experiential representations of world knowledge (Petruck, 1996). Moreover, frames are taxonomically related to one another; there are varying degrees of generality and specificity, such that some frames instantiate more general ones, and certain basic-level frames constitute primary bodily experiences (Clausner and Croft, 1999). As metaphors are cross-domain mappings between frames, they are similarly structured, both within an individual metaphor and in relation to other metaphors. The goal of our project is to formalize and implement these frame and metaphor structures.

Current CMT does not typically incorporate Frame Semantics in any codified manner; individual analyses identify particular roles within source and target domains, but the issue of which frame elements to include relies on the discretion of the analyst without any agreed-upon state of the art, including what kind of and how many frame elements to include in the metaphoric mappings. At present, representation of relations between frames, such as inheritance of frame elements, is typically not incorporated into CMT analysis. Instead, there is an accepted general understanding in CMT that the source and target domains of metaphors are composed of frames or image schemas. In recent years, increasing attention has been paid to the need to incorporate more detailed frame analysis into CMT (e.g., Bouveret and Sweetser, 2009; Moore, 2011; Sullivan 2006; Sullivan 2013).

The development of MetaNet frames owes much to the instantiation of Frame Semantics in FrameNet. Since its establishment in 1997, FrameNet has come to define the structure of semantic frames, including core and peripheral frame elements and relations between frames. As will be discussed in Section 3.a, FrameNet's frame-to-frame relations informs our formalization of frame relations. The notion of the frame as a cohesive experiential gestalt underlies both ours and FrameNet's frames. While they share these central conceptual tenets, instantiation of a particular concept may vary between the two databases. Consider, for example, the English verbs *rise* and *fall*. Together both lexemes evoke the concept of *directional motion*. Both FrameNet and our repository have frame entries reflecting such a concept: Motion_directional (FrameNet) and Motion Along a Path (MetaNet). FrameNet defines it as "a Theme moves in a certain Direction which is often determined by gravity or other natural, physical forces"; the analysis in our repository informally describes it as "A moving entity (the mover) starts out in one place (Source), and ends up in another place (Goal), having covered some space between these two (Path)." In both cases there is a notion of translational motion by an entity in a direction. However, while *rise.v*[7] and *fall.v* are both included in the Motion_directional lexical entries, they are not in the list of Motion Along a Path lexical units. Instead *rise.v* is in Upward Motion, and *fall.v* is in Downward Motion, both of which are more specific frames that inherit their structure from Motion Along a Path. In our system, splitting frames into these finer-grained concepts is driven by their frequency of use in metaphoric constructs; this process is described in detail in Section 2.b.i. and illustrated specifically for vertical motion in Section 5.b.

[7] The .v suffix on the lexical entry indicates it is a verb; .n indicates a noun.

2.b. *Frames*

As Grady (1997) and Johnson (1999) first demonstrated, primary metaphors such as KNOWING IS SEEING and STATES ARE LOCATIONS consist of experientially-based embodied universals learned early in life, including primitives variously referred to as *image schemas* or *cogs* (Talmy, 1983; Lakoff, 1987; Lakoff and Turner, 1989; Johnson, 1987; Gibbs and Colston, 1995; Lakoff and Johnson, 1999; Dodge and Lakoff, 2005; Gallese and Lakoff, 2005). For example, MORE IS UP (or, more generally, QUANTITY IS VERTICALITY), comprises the notion of Quantity in the target domain and the image schema of Verticality in the source domain, with the experientially-based inference that upwards motion constitutes an additive property. In contrast, many conceptual metaphors are not composed of relations between experiential universals, but between culturally-constrained semantic frames. LOVE IS A JOURNEY (or, A ROMANTIC RELATIONSHIP IS A JOURNEY) (e.g., Lakoff and Johnson, 1980) requires culturally specific notions of what constitutes both a romantic relationship and a journey with an intended destination. Hence, frames are a complex gestalt comprising culturally-specific elements with experiential grounding; in contrast cogs are taken to be conceptual primitives grounded in universally shared human bodily experiences.

From its initial design stages, our project has been intended for cross-linguistic comparison, in part because CMT is theorized to derive from the cross-cultural universal experiences driving embodied cognition. Given that metaphoric source and target domains can be made up of universal primitives, of culturally-specific frames, or of some combination of the two, MetaNet treats both universal and culturally-specific of structures as *frames*; however, throughout the dissertation I differentiate between frames and image schemas in order to focus specifically on image schematic structure. MetaNet views frames as coherent semantic and cognitive structures, formed from bodily interaction with the world. In the case of culturally-specific frames, this interaction includes one's sociocultural experiences. Frames are then proposed to be analyzed as either culturally-bound frames or image schemas in order to enable the validation of these universals and cross-cultural comparison of frames. When a particular conceptual metaphor is validated cross-linguistically, it provides evidence not only for the universal nature of the metaphor but also for the image schemas that make up its source and target domains.

2.b.i. *Analysis of frame elements and lexical units*

In the currently-described system, frames are developed specifically in the service of metaphors, rather than independently; hence, a frame is created or derived from another more general frame when a metaphor is identified that has that frame as its source or target domain. For example, the Motion Along a Path frame and the Action frame are established in order to model the metaphor ACTION IS MOTION ALONG A PATH. In other words, frames are developed as a result of a continuous process of annotation and analysis of metaphors. Furthermore, frames can serve as the source or target of multiple metaphors, whereby different elements of that frame may be activated in each of the metaphors. For example[8], compare (1a) and (1b) (source domain language italicized, target domain language underlined):

[8] Examples in are from the English GigaWord corpus (Graff and Cieri (2003) or the British National Corpus (2007) except where otherwise noted.

(1) a. crime and corruption have *infected* virtually every aspect of the Russian economy[9]
b. High taxes and wage freeze (*sic*) *rob* people of a meaningful living[10]

In (1a), crime is metaphorically conceptualized as an infectious process; it is the target domain of the metaphor CRIME IS A DISEASE. In contrast, crime (i.e., robbery) is the source domain of the metaphor evoked in (1b), TAXATION IS THEFT, where high taxes are conceptualized as the criminal activity of theft. Hence, these analyses do not distinguish between "core" and "non-core" roles for a particular frame, because different elements are involved in different metaphor mappings, instead of a core set of elements involved in all mappings that the frame contributes to semantically. The focus on metaphor analysis in the creation of frames has thus led to the reinforcement of the notion of a frame as a bundle of roles, which interact with other roles in metaphoric mappings, as emphasized in the original works in both frame semantics and CMT.

Additionally, the Invariance Principle – which states that a metaphor only maps elements of the source domain onto the target domain such that they remain coherent in the context of the target -- dictates that metaphors always involve *partial* mapping between source and target domain (Lakoff, 1990). Given this partial schematicity inherent to metaphoric mapping, not *all* elements of a frame are expected to be engaged in the internal structure of a metaphor. Our project's frame roles are added and elaborated upon as additional metaphors making use of particular frame roles are analyzed. For instance, the vehicle role may be added to the Journey frame only once it is evident that RELATIONSHIPS ARE VEHICLES is a mapping in the LOVE IS A JOURNEY metaphor. Only as many frame elements are initially added to a frame as are needed to account for the mappings that frame and its roles participate in, with room for expansion once additional roles are discovered via the analysis process to be needed in other mappings.

The coherence of mappings between source and target domain is maintained in part by constraining which roles can map onto one another. These constraints are accomplished via *role typing*, in which the roles of a frame can only be filled by role fillers of a particular type; types are themselves high-level (highly schematized) frames. For example, in the Harm scenario frame shown in Figure 1, the thematic harmed_entity role, which undergoes the experience of harm, must be something that *can* be harmed, i.e. an entity of some kind. In contrast, a process cannot be harmed because it is not a "thing" of some kind; rather, it is a sequence of events with temporal duration. Because the Harm scenario is a very generalized frame, it does not specify *what* the harmed_entity is; it could be any type of physical entity, regardless of animacy or personhood. Hence, the harmed_entity role is just typed as an Entity.

[9] ENGW_apw_eng_199611:44236
[10] ENGW_xin_eng_200402:35272

Figure 1. Roles in the Harm scenario frame.

Roles:

Role Name:	harmed_entity
Role Type:	Entity
Definition/Comments:	
Role Name:	harmful_process
Role Type:	Process
Definition/Comments:	
Role Name:	effect_of_harm
Role Type:	
Definition/Comments:	damage caused by process
Role Name:	consequence_of_harm
Role Type:	
Definition/Comments:	results of damage, e.g. loss of function
Role Name:	cause
Role Type:	
Definition/Comments:	

We can link the harmed_entity role within the Harm frame to the harmed_entity role within the Physical Harm frame and the victim role within the Harm to Living Entity frame. As will be further illustrated, Physical Harm and Harm to Living Entity are essentially progressively more specific versions of the more general Harm frame. Each adds further specification to the roles of the frame such that the harmed_entity is a physical entity and subsequently an animate, living entity of some kind. The harmed entity of the Harm to Living Entity frame is thus termed a "victim", which appropriately invokes a notion of harm to some entity that undergoes a detrimental, negative experience as a result of the harm. We may contrast this to the harmed_entity that, while damaged, may not have a "negative" experience if it is some inanimate object that cannot conceptualized as a true semantic patient experiencing something detrimental to itself. The victim role is further constrained to be of type Animate Entity, which provides this specification. Additionally, the Harm scenario does not specify that the harm be *physical* harm to a *physical* entity; this is determined in the Physical Harm frame and the roles within are typed accordingly. (Other frames that specify different types of harm include Environmental Harm and Psychological Harm, each with their own role typing-based constraints.) Notably, just as the victim is a particular kind of harmed_entity, so too is the Animate Entity *type* a particular kind of Entity *type.* This is a direct result of the fact that the Animate Entity and Entity types are themselves frames, and hence are subject to the same frame relations.

Just as FrameNet frames are evoked by lexical units, our frames similarly are associated with lexical units. Association of lexical units with frames is developed via linguistic metaphor analysis: when a linguistic metaphor (i.e., an instance of a conceptual metaphor) is identified, the metaphor-evoking language is noted. This may consist of words evoking the source domain or both source and target. These lexical units are then associated with the appropriate frame constituting the source or target of the conceptual metaphor (CM). Lexical units are only assigned to a particular frame with data-based verification that a specific lexical unit evokes that

particular frame in a linguistic expression of metaphor. These linguistic expressions that instantiate conceptual metaphors lexically, to be distinguished from visual metaphors or, more generally, conceptual metaphors. As metaphors are further analyzed and added to the network, so too are additional frames; these additions result in the re-assignment of lexical units to different frames as a reflection of refined analyses. For instance, *push.v* may initially be assigned to the Caused Motion frame, but may subsequently be reassigned to a Forceful Caused Motion frame if such a split is justified in the frame structure (whereas Caused Motion is left neutral as to the forcefulness of the causal action).

Consider the conceptual metaphor ONGOING NEGATIVE STATE IS EXPERIENCING PHYSICAL HARM, as illustrated by the following LMs:

(2) a. my debt is *killing* me[11]
b. alleviate the debt *burden* on poor countries[12]
c. poverty *robs* people of dignity and health[13]
d. a region wracked by a *crushing* poverty[14]
e. debt *endangers* Iraq's long-term prospects for political health and economic prosperity[15]

An initial analysis of the linguistic metaphors in (2) might reveal that the continuing experience of a negative state (being in debt or being poor) is conceptualized metaphorically as the experience of some kind of harm. Hence, we develop an Ongoing Negative State frame to denote the experience of a negative state like indebtedness or poverty, and a Harm frame to denote the experience of Harm. We assign the corresponding lexemes to the frame Harm:

Harm
kill.v
burden.v
crush.v
rob.v
harm.v
suffer.v

Further analysis, bolstered by discovery of additional metaphoric instances, leads the analyst to reconsider ONGOING NEGATIVE STATE IS EXPERIENCING PHYSICAL HARM, proposing additional CMs[16]:

[11] http://www.popsugar.com/smart-living/Maxed-Out-My-Debt-Killing-Me-2425114
[12] ENGW_afp_eng_200410:6142
[13] http://borgenproject.org/impact-poverty-life-expectancy/
[14] ENGW_apw_eng_200207:1175
[15] ENGW_xin_eng_200312:12106
[16] Most likely, the analyst would also propose that Ongoing Negative State be further divided into into Debt and Poverty frames as well; this is reflected in the current network.

ONGOING NEGATIVE STATE IS BEING THE VICTIM OF A CRIME
ONGOING NEGATIVE STATE IS EXPERIENCING PHYSICAL HARM

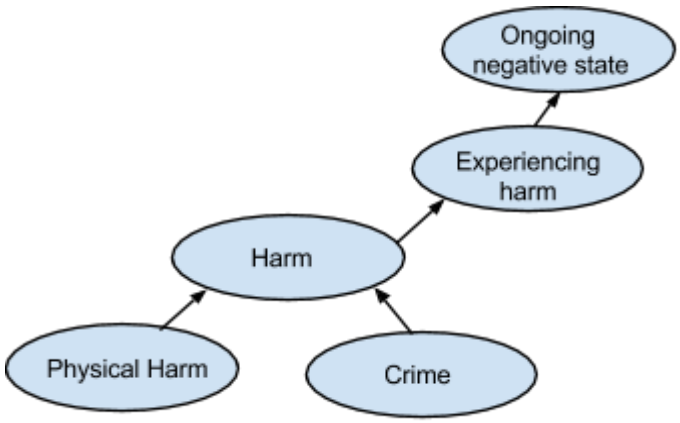
These new CMs, in turn, lead to the development of additional frames, where the lexemes evoking these frames are reassigned:

Physical Harm
kill.v
burden.v
crush.v

Crime
rob.v

Notably, because the lexemes *suffer.v* and *harm.v* don't particularly evoke the frames of either Physical Harm or Crime, they remain in the Harm frame. The new frames, together with the original Harm frame, instantiate a Harm subnetwork as shown in Figure 2.

Figure 2. Harm frames subnetwork. Arrows indicate frame relations.



As additional LMs are discovered, either by the analyst or via the automatic extraction system, this aspect of the frame network would be elaborated further with additional frames and lexemes.

2.b.ii. *Shared frame structures*

A feature of the representation of frames is the systematization of the structural and temporal nature of frames themselves, similar to the FrameNet implementation of relations between frames. We observe that for any particular scenario, a frame's structure must include particular structural components: entity role(s); non-entity role(s); relations between those roles if there are more than one; and an *executing schema*, or *x-schema* (Bailey, 1997; Narayanan, 1997; Bergen and Chang, 2005)[17], which is a process frame that specifies the temporal structure of the state or

[17] X-schemas are formally based on Petri nets (Reisig, 1985; Murata, 1989).

event the frame encodes.[18] We can compare two frames, the universal image schema Motion Along a Path and the cultural frame Corruption, to further illustrate. For example, consider the sentence *The dog is running from his bed to his food bowl*. This sentence instantiates the Motion Along a Path frame with the verb *running*. The *dog* is the entity in motion – the mover – and he follows a path from his bed (the source of the path) to his food bowl (the goal of the path). The progressive aspect of the verb conveyed by the *-ing* suffix indicates that the motion process is ongoing. In comparison, consider the sentence *The congressional bribery scandal is damaging public trust in government.* Here the sentence evokes the Corruption frame. This frame has both a corrupt actor (congressional politicians) and an element affected by that corruption (public trust). The ongoing process of corruption has an effect (damage to public trust). Whereas semantically the Motion Along a Path and Corruption frames are quite different, we observe structural similarities, as summarized in Table 1. They both have entity roles (the dog, the corrupt officials); image schemas (the path of the dog running, the cause and effect of the corruption); x-schema processes (the motion process, the corruption activity), and defined relationships between roles and x-schema processes.

Table 1. Roles and relations in the Motion along a Path and Corruption frames.

	Motion Along a Path	**Corruption**
Entity role(s)	mover	corrupt_actor; harmed_entity; corruption_affectee
Non-entity role(s)	Source-Path-Goal schema	corruption_effect
X-schema	motion process	corruption_activity
Relation between roles and processes	mover begins at source; mover moves along path; mover progresses towards goal; mover performs motion process	corruption_effect effects corruption_affectee; corrupt_actor performs corruption_activity

Both Motion Along a Path and Corruption are scenarios, despite their differences, i.e. the former is a universal image schema based on embodied experience, and the latter is a highly culturally defined frame. Still, both frames comprise the same basic structural elements; and these components define any given frame. Returning to the Harm frame presented earlier, note that many of the elements of Corruption correspond to elements of Harm (Table 2):

[18] While not conceptualized as events per se, following Comrie (1976) we take the approach that statives are eventualities that incorporate temporal structure; ongoing or durative states must necessarily be experienced over some expanse of time.

Table 2. Relationships between roles in Harm and Corruption frames.

	Harm	**Corruption**
Entity role(s)	cause_of_harm; harmed_entity	corrupt_actor; harmed_entity; corruption_affectee
Non-entity role(s)	effect_of_harm	corruption_effect
X-schema	harmful_process	corruption_activity
Relation between roles and processes	effect_of_harm harms harmed_entity; cause_of_harm causes harmful_process	corruption_effect effects corruption_affectee; corruption_effect harms harmed_entity; corrupt_actor performs corruption_activity

Essentially, we see that Corruption is an instance of a Harm frame with semantically elaborated roles and additional structure. Note in particular that these frames thus far are underspecified with regards to the dynamic or causal state of the given scenario. For example, while the Motion frame has a mover agent to *perform* the motion x-schema, a scenario, such as Harm, does not specify whether the mover is currently in motion, if it is about to be motion, or if the motion is caused to occur, and so on. Scenario frames constitute structural, rather than dynamic, information. A set of related frames constituting aspectual and causal alternatives provide these x-schematic options for scenarios. Additionally, scenario frames do not take a particular perspective on a scene, but can have multiple perspectives. This is similar to the structure in FrameNet, where, for instance, Buying and Selling are perspectives on a Commercial Transaction Scenario.[19]

While scenarios provide structural and semantic information, alternative frames specify the values of the x-schema role of the related scenarios, which constitute the possible temporal and causal values of the state dynamics of an eventuality, as listed in Table 3 and illustrated in part for Motion Along a Path. Aspectual variants are exemplified by the phrase *run to the food bowl* in (3). The verb *run* specifies the manner of motion and the directional *to the food bowl* the path and goal of the motion.

(3) a. Begin Motion Along a Path — *began to run to the food bowl*
b. Motion Along a Path — *run to the food bowl*
c. Interruption of Motion Along a Path — *paused running to the food bowl*
d. Resumption of Motion Along a Path — *resumed running to the food bowl*
e. Stop Motion Along a Path — *stopped running to the food bowl*

[19] Our frame relation "profiles part of" is similar to the FrameNet frame-frame *perspective on* relation in this regard.

Alternatives include x-schema stages (Feldman, 2006; Narayanan, 1997; Feldman and Narayanan, 2004) as well as causal variants of those stages. Aspectual and causal alternatives are integral to the conceptual network as a whole; given that more specific Processes, such as Motion Along a Path, share the same basic conceptual information as the generic Process schema, it follows that they also have all the same alternatives as well.

Table 3. Causal and aspectual variants of processes.

	Alternatives of a Process	**Motion Along a Path Alternatives**
X-schema stages	Begin process Initiate a state	Begin moving along a path
	Ongoing process Ongoing state	Motion along a path
	Interrupt process	Interruption of motion along a path
	Resume process	Resumption of motion along a path
	End process Cease experiencing a state	Stop moving along a path
Causal variants	Cause to begin process Cause to enter a state Enable to begin process Enable entering a state Prevent from beginning a process Prevent from entering a state	Cause beginning motion along a path Enable beginning process Prevent beginning motion along a path
	Cause interruption of process Enable interruption of a process Prevent interruption of process	Cause interruption of motion along a path Enable interruption of motion along a path Prevent interruption of motion along a path
	Cause resumption of process Enable resumption of process Prevent resumption of process	Cause resumption of motion along a path Enable resumption of motion along a path Prevent resumption of motion along a path

	Cause end of process Cause to cease experiencing a state Enable end of process Enable cessation of experiencing a state Prevent end of process Prevent cessation of experiencing a state	Cause end of motion along a path Enable end of motion along a path Prevent end of motion along a path

This kind of detailed event structure breakdown of frames is resonant with the detailed force-dynamic and causal image schemas introduced and elaborated by Talmy (1983, 1985, 2003). He notes that complex causal structures and finer-grained force-dynamic distinctions such as these are necessary to account for not only the vast range of conceptualization and profiling patterns within a language, but cross-linguistically comparable patterns as well. The typology of frames and roles in our ontology is intended to capture these valuable generalizations, much as subsequent developments of force-dynamic and causal schemas do (such as Narayanan's (1997) x-schemas, and the grammatical structures in Embodied Construction Grammar).

Specifying additional alternative stages to fulfill the x-schema parameter of a particular scenario provides additional aspectual information for that scenario. While this temporal and causal structure underlies all eventualities, and hence all scenarios, it is the specifications of the metaphors under analysis that drive the analysis of particular alternatives within our frame network. For example, the metaphor matching system identifies the conceptual metaphor CORRUPTION IS A DISEASE, as illustrated by the following expressions in (4):

(4) a. Chinese corruption will *infect* Hong Kong's orderly administration[20]
b. his fight to stop the *cancer* of corruption in Papua New Guinea[21]
c. The Congress is *plagued* by corruption[22]

This conceptual metaphor requires the structure of Corruption and Disease frames to align such that the inference of slow, ongoing action is available. Disease is a slow degenerative natural process that occurs over a period of time whose effects are at first undetectable, but eventually result in destruction of the entity, if not stopped. Likewise, after the invariance principle maps the proper inferences to the target domain, corruption is a slow degenerative social process that occurs over a period of time whose effects are at first undetectable, but eventually result in destruction of social structure, if not stopped. In the metaphor analysis, the source domain Disease x-schema maps onto the target domain Corruption x-schema. To provide for the inference of these processes as slow and continual, the system elaborates alternative stages of Corruption and Disease: Ongoing Corruption and Development of Disease. Each alternative stage provides the x-schematic information of ongoing_process and slow_rate.

[20] ENGW_apw_eng_199603:11938
[21] ENGW_afp_eng_200407:4591
[22] ENGW_apw_eng_199604:35926

While these stages may only be realized in our network as the result of a particular metaphor analysis, we emphasize that conceptually, the causal and dynamic alternatives represented by alternative stages are generalized across all scenarios, because all frames filling the source domains of metaphors are grounded in embodied primary experiences that are structured by very predictable force-dynamic patterns we encounter in our interactions with the world. Just as all scenarios ultimately inherit from shared highly schematized frames, all scenarios too have the same causal and dynamic alternatives. These conceptual building blocks originate with the most general, high-level frames, and are inherited by all subsequent frames of increasing specificity. This generalizability reflects the notion that conceptual structures are formed by composition of more basic, schematized structures, which is a basic tenet of embodied cognition. For example, Motion Along a Path is in part composed of the more general Source Path Goal frame, and Source Path Goal itself, in part contains the more general still Trajector Landmark frame. The next section explores the nature of these frame relations in further detail.

3.a. *Frame relations*

Much like FrameNet's hierarchically-structured frame network, our system defines relations between the nodes (i.e., the frames, frame elements, and metaphors) of its network. As MetaNet constitutes a proposed model of cognitive conceptual structure, connections between frames (and metaphors, as discussed later) elaborate a hypothetical model of the brain's activation complex as described in Cascade Theory (David et al. (2016); Lakoff, 2008). Frame relations fall into two broad categories: structural relations, and non-structural relations.

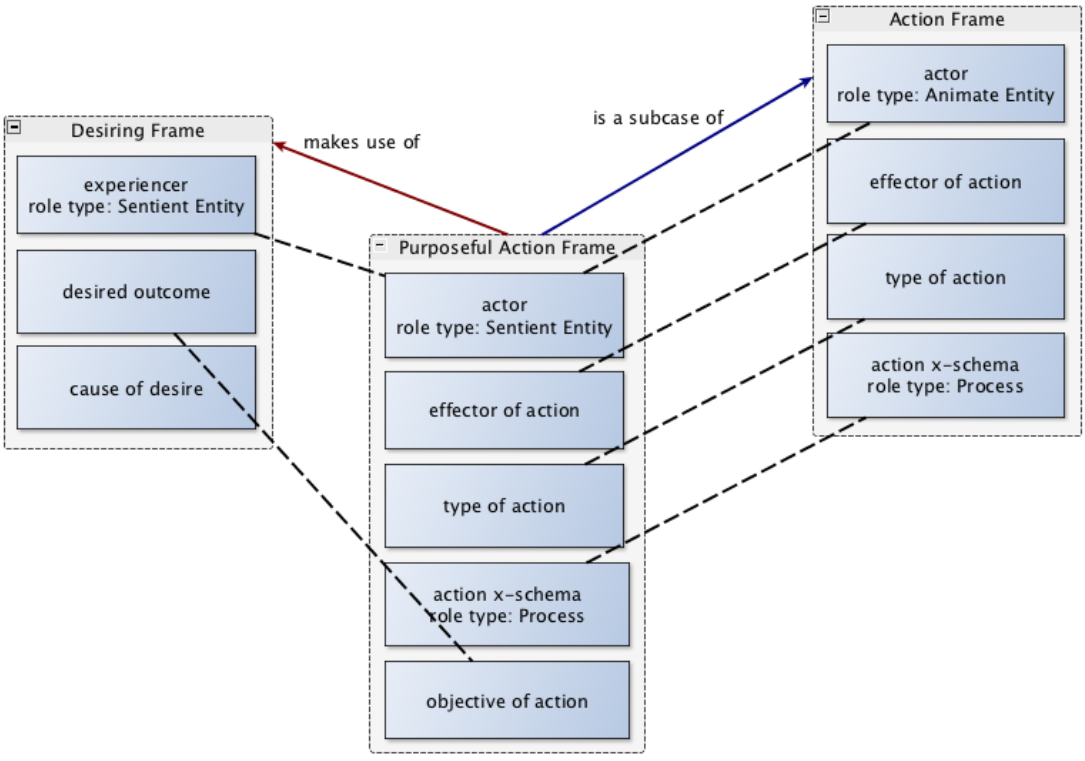
3.b.i. *Structure-incorporating relations*

Structural frame relations broadly resemble common ontological relations such as inheritance ("is-a") and composition ("has-a"). Like other inheritance relations, such as FrameNet's "inheritance" relation, the MetaNet *is a subcase of* frame relation denotes a hierarchical structure that connects elements of a more general parent frame onto corresponding elements of a more specific child frame. However, we do not consider this "inheritance" in the typical ontological sense, as the "inherited" elements are not "copied" or "duplicated" in the structure of the child frame. Rather, the child frame directly incorporates the semantics of its parent frame, and that of its parent's parent's frame, and so on; each corresponding element is not copied, but rather it is underlyingly a single conceptual structure with each child frame adding more semantic information to it. Thus, corresponding elements within this hierarchical structure are connected via *bindings*, which indicate that they are essentially activating the *same* concept. This is intended to accord with the hypothesis in neurocognitive linguistics that more complex conceptual structures are the result of neural networks that simultaneously activate multiple bound primitive structures located in various areas of the brain, such as motor-action neural bundles, sensory neural bundles, and visual neural bundles (Feldman, 2006;, Feldman and Narayanan, 2004; Lakoff, 2008). Those neural structures do not "move" or "copy"; rather, they are activated as part of the spreading of neural activation as a whole.

3.b.ii. ***Subcase of***, ***special case of***, *and* ***makes use of*** *frame relations*

Is a subcase of indicates a frame relation in which the elements of the more specific frame, which *is a subcase of* the more general frame, bind to all (*subcase of*) in the general frame. This relation can be compared with FrameNet's *inheritance* relation. We illustrate these relations with the high-level Purposeful Action frame (Figure 3), which is an image schema representing a scenario of some action that an actor performs. It is a *subcase* of the even more general Action scenario; thus, each element of the Purposeful Action frame binds to each element of the Action frame. Importantly, because Purposeful Action is a subcase of Action, all of the inferential structure associated with Action binds to Purposeful Action as well. This includes all of the type constraints specified for frame elements. The Action scenario specifies that the actor role must be an Animate Entity; this constraint dictates that the actor of Purposeful Action must also be an Animate Entity. Because subcase frames fully bind the elements of their parent frames, they contain all the elements of their parent frames. However, because they are more specific instances of the parent, subcases must add some semantic information, such as elaborating on the parameters of roles or adding additional roles. Purposeful Action adds the notion of intention or purpose such that the actor becomes an agentive actor and the action is performed for a purpose. To add this information, Purposeful Action *makes use* of the Desiring scenario, which defines the semantics of purpose. Only some of the elements of Purposeful Action bind to those of Desiring, which distinguishes the *is a subcase of* relation from *makes use of*.

Figure 3. Bindings between Purposeful Action, Action, and Desiring frame elements. Dotted lines indicate bindings and arrows indicate relations.



Furthermore, when an element binds to multiple elements in other frames, the accompanying inferences must remain semantically coherent. Hence, when the actor of Purposeful Action binds

both to the actor of Action and the experiencer of Desiring, the role constraints of Action's actor and Desiring's experiencer must be compatible. In this case, Action specifies an Animate Entity and Desiring specifies a Sentient Entity; because Sentient Entity and Animate Entity are both subcases of Entity, this double binding aligns properly, given the absence of incongruence for the semantics of each binding. Thus the actor role of Purposeful Action is of type Sentient Entity, which provides the inference of the ability to have the state of intention and purpose. In contrast, an element could not bind to both an Animate Entity role and a Machine role, as those roles are mutually incompatible frames. For an entity to be both living (Animate Entity) and an inanimate object (Machine) is self-contradictory in a semantic ontology of frame structures. (However, it is exactly this kind of construal that occurs in the context of metaphoric language, as will be seen in Section 4).

The *special case of* frame relation is a type of *subcase of* relation in which the child frame only binds to one parent frame, with no additional frame relations; it only adds information by providing role fillers or otherwise filling the value of a parameter that is underspecified in the parent frame. Thus, the *Boat* frame is a *special case of* the Seafaring Vehicle frame.

3.b.iii. ***Incorporates as a role*** *frame relation*

While child frames typically incorporate some or all of the structure of the parent frames by binding some or all of their frame elements to those of the parent frame, it may also be the case that a frame incorporates the entire frame itself as an element. While the *subcase of* relation creates a frame that is a more-specific version of the general frame, the *incorporates as a role* relation indicates that one frame fully includes another frame as an *element*. To put this another way, sometimes frame elements are entire frames in their own right, and exist independently of the frame to which they are currently bound as a role. This situation occurs when two frames capture different perspectives on a scenario, usually because one is focused on an Entity within the scenario and the other on a Process occurring in the same scenario. This approach allows the modeling of perspectives on scenes and information packaging. For example, Eating, which is a type of Process, is a scenario which has a structure including the eater, the eaten_object, and the x-schema process of eating. Eating, as a Process, focuses on the interactive process between the eater and the eaten_object. The eaten_object (prototypically) is an edible object or substance, here termed Food. Food is modeled as a frame in its own right, which has much of the same structure as the Eating scenario, but focuses on the properties of the Entity being eaten and includes roles for type of food, amount of food, and properties such as taste and texture. Thus, the network models the relation between Eating and Food by conceptualizing Eating as *incorporating* Food *as a role* within it; each frame portrays the same overall scenario differently. In summary, *subcase* relations are between two frames of the same type – for example, Food and Beverage are both Entities, *subcases of* Consumable_Entity, whereas Eating is a Process that *incorporates as a role* Food, which is an Entity.

3.b.iv. ***Has affordance of*** *and* ***process that makes use of*** *frame relations*

Conversely, whereas Eating *incorporates* the Food frame, we must also observe the relationship between Food and Eating. Food is inherently understood as something that can be eaten. While Processes *incorporate* Entities as participant roles, Entities similarly *incorporate* other frames as

properties. Given this observation that certain Entity frames are defined in part by intrinsic properties, the *has affordance of* relation encapsulates the idea that not only is the second frame fully incorporated into the first, it is central to the Entity's conceptual structure, but it is not a singular role within the frame. For example, Liquid *has affordance of* Fluid Motion; the motion of fluid defines something as a Liquid. As the name suggests, the *has affordance of* relation is intended to formally capture the physical affordances of various real-world entities, and hence capture systematic generalizations about possible interaction scenarios individuals may have with them.

A third type of frame incorporation occurs when a Process makes use of a scenario or Entity, providing a processual perspective on it. As described in section 2.b.i., the Harm to Living Entity frame includes both the harmful_entity and the victim. However, the metaphor ONGOING NEGATIVE STATE IS EXPERIENCING PHYSICAL HARM captures the notion of harm from the perspective of the victim (or "harm experiencer"). As metaphors often foreground the experience or perspective of particular entity roles within a non-perspectivized scenario, the *process that makes use of* frame relation indicates that a frame such as Experience Harm is a process variant of a scenario such as Harm to Living Entity, which incorporates the structure of the base scenario and adds information from the dynamic perspective of a particular role, in this case the victim experiencing harm.

3.b.v. ***Profiles part of***, ***is a subscale of***, *and* ***is a subprocess of*** *frame relations*

As metaphors often foreground the experience or perspective of particular entity roles within a non-perspectivized scenario, the *profiles part of* frame relation indicates that a frame such as Curriculum foregrounds an element of a complex scenario like Education. This process of *profiling*, wherein some element or backgrounded component of a frame is brought into focus, is denoted with the *profiles part of* relation. Hence, Curriculum includes elements of Education related to the goals of lessons, information to be taught, and so forth; but it may not incorporate elements such as the students or administration.

Given that metaphor analysis drives the development of frame structure, one aspect of analysis entails categorizing metaphors into common types as well. The well-known orientation metaphors such as QUANTITY IS VERTICALITY, GOODNESS IS VERTICALITY, and CONTROL IS VERTICALITY[23] that Lakoff and Johnson (1980) described all rely on mappings between scalar structures in the source domain of Verticality and various target domains. Our approach emphasizes developing generic or generalized frames that can be elaborated into specific cases, thus recognizing that the entailments of the orientation metaphors (i.e., MORE IS UP and LESS IS DOWN are entailments of QUANTITY IS VERTICALITY) profile the end regions of the source and target domains. GOODNESS IS LIGHT/EVIL IS DARK, while not orientation metaphors, follow the same pattern of target and source frames with scalar structure – Goodness and Luminosity – with entailed metaphors that profile either end of each scale. Hence, GOODNESS IS LIGHT profiles the good, or light, regions of the Goodness and Luminosity scales, respectively; and EVIL IS DARK profiles the evil, or dark, regions of the Goodness and Luminosity scales. Observing this trend

[23] These are non-perspectivized variants of the more traditional MORE IS UP, GOOD IS UP, and CONTROL IS UP.

throughout a variety of metaphors relying on mapping between regions of scales dictates the need for the frame relation *is a subscale of*, to relate frames like Good or Light to the scalar scenarios Goodness and Luminosity. This frame relation indicates that the child frame profiles one end region of the parent frame's scale. *Subscale of* constitutes a type of profiling, in that an element (one scalar perspective) of a larger frame (the entire scale) is brought into focus.

The *is a subprocess of* relation also constitutes a form of profiling because it indicates that a frame is a particular stage of a multi-step complex process. Subprocess frames are developed as metaphor analysis identifies entailed metaphors that make use of these subprocesses within a larger metaphor network. For example, the Physical Affliction scenario might include the processes of affliction diagnosis, treatment, and cure. Identification of the metaphor SOCIAL PROBLEMS ARE PHYSICAL AFFLICTIONS makes use of Physical Affliction in the source domain, as seen in the following:

(5) a. poverty is *epidemic*[24]
b. Illegal drugs are a *cancer* on this and every other community in this country[25]
c. How do we *diagnose* poverty?[26]
d. the *misdiagnosis* of unemployment[27]
e. the best way to *cure* unemployment is to punish the unemployed[28]

Further analysis of these metaphors produces entailments of the main metaphor: ANALYSIS OF SOCIAL PROBLEMS IS DIAGNOSIS OF AFFLICTION (5c, 5d) and ADDRESSING SOCIAL PROBLEMS IS TREATMENT OF AFFLICTION (5e). Thus, the source and target domains of these entailed metaphors constitute processes within the source and target domain of the metaphor, i.e. analyzing and addressing social problems in the target domain, and diagnosing and treating disease in the source domain. Identifying these entailments leads to the analysis of the subprocesses of the Social Problems and Physical Affliction frames, such that Diagnosis of Physical Affliction and Treatment of Physical Affliction are each *subprocesses of* the Physical Affliction scenario frame.

3.c.i. *Non-structural relations*

Non-structural frame relations stand in contrast to structural frame relations in that they are not structure-incorporating or hierarchically related to each other. Instead these relations exist between sister frames with a common parent; their shared structure is via bindings to the shared parent, rather than as bindings between the frames themselves. These types of relations indicate that either the frames interact in a causal or temporal manner, or that they are variants of another frame. Hence, they are similar to FrameNet's "precedes" and "causative_of" and "inchoative_of" relations, which capture temporal ordering or stative/causative and state change relations.

[24] ENGW_nyt_eng_199412:23202
[25] http://www.news-gazette.com/news/local/2009-10-21/brothers-sentenced-20-years-heroin-case.html
[26] http://www.shapingdestiny.org/leadership/sdmleadership/defining-the-problem/
[27] http://en.wikipedia.org/wiki/Keynesian_economics
[28] BNC:CAJ:1062

3.c.ii. ***Is in causal relation with*** *frame relation*

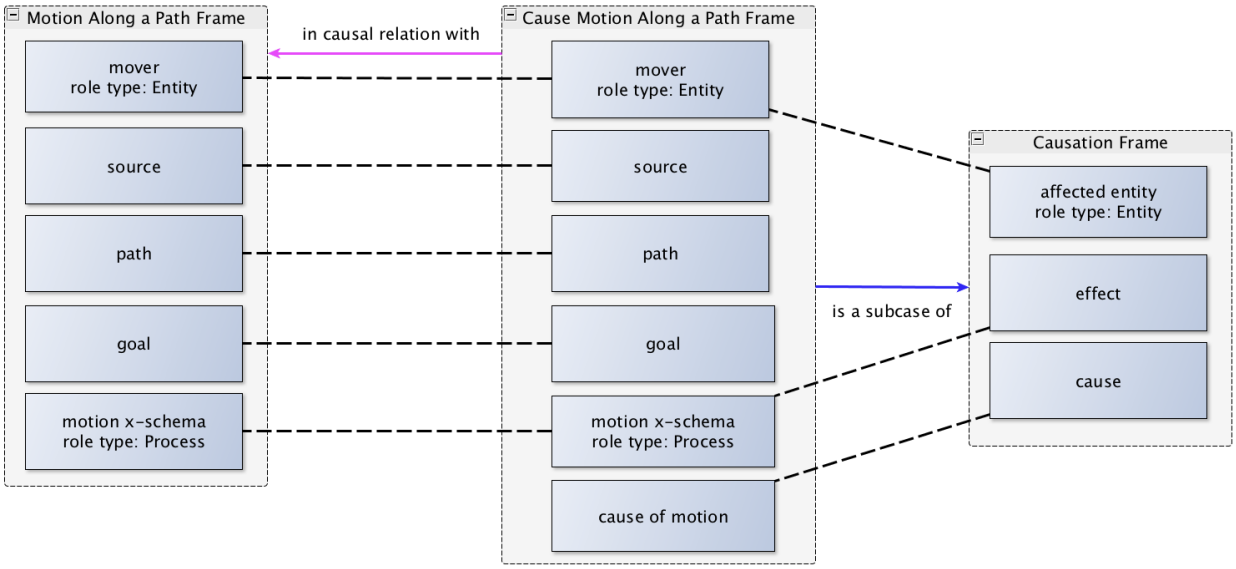
Processes have causal variants, as previously illustrated in Table 3; any frame that inherits from Process will have alternative scenarios that add causal information. For example, Table 4 recapitulates the causal variants of Motion Along a Path in comparison to the non-causal aspectual forms.

Table 4. Comparison of aspectual and causal variants of Motion Along a Path Frame.

Motion Along a Path Frames	**Caused Motion Along a Path Frames**
Motion along a path	Cause motion along a path Enable motion along a path Prevent motion along a path
Begin motion along a path	Cause beginning motion along a path Enable beginning motion along a path Prevent beginning motion along a path
Interruption of motion along a path	Cause interruption of motion along a path Enable interruption of motion along a path Prevent interruption of motion along a path
Resumption of motion along a path	Cause resumption of motion along a path Enable resumption of motion along a path Prevent resumption of motion along a path
End motion along a path	Cause end of motion along a path Enable end of motion along a path Prevent end of motion along a path

The main process scenario, Motion Along a Path, has several aspectual forms elaborating various stages of the x-schema; each has causal versions that include information regarding the cause or prevention of activating that x-schematic information. The *is in causal relation with* relation describes the relationship between the non-causal and causal variants: Caused Motion Along a Path *is in causal relation with* Motion Along a Path. Cause Motion Along a Path does not provide any additional semantic information to the Motion Along a Path frame; the frame structures are the same, as are the core semantics of the frames themselves. The only difference is that the causal variant provides causal information, as illustrated in Figure 4.

Figure 4. Relations between Motion Along a Path and causation frames. Dotted lines indicate bindings and arrows indicate relations.



3.c.iii. ***Precedes*** *frame relation*

Whereas a causal relationship between two frames inherently implies a temporal relationship in that cause precedes effect, frames can be also be temporally ordered without causation when they are different stages of a process. For example, the Motion to a Destination frame includes the semantics of the stages of a journey: departing from the source location, travelling along a path towards the destination, and arriving at the destination. Each of these elements constitute an element of the overall process; hence, Departing and Arriving are both *subprocesses* of Motion to a Destination. These sister frames are temporally ordered in relation to one another: Departing has to happen before Arriving. This temporal relation is codified in the *precedes* relation, such that Departing *precedes* Arriving.

3.c.iv. ***Mutually inhibits*** *and* ***is in scalar opposition to*** *frame relations*

Finally, some sister frames are alternatives of a scenario that are inherently in opposition to one another. These frames have a parent with an under-specified valence, which are parameterized in the child frames. The entity frames Aids to Motion and Impediments to Motion both describe Motion-affecting Objects, as evoked by such lexemes as *ladder, trampoline, trap, shackle*. Something that enables movement inherently does not impede movement, and vice versa; thus the notions of enablement and impediment are in mutual opposition. The *mutually inhibits* relation is a symmetric relationship between two such frames: Aids to Motion *mutually inhibits* Impediments to Motion, and Impediments to Motion *mutually inhibits* Aids to Motion.

A special type of this opposition occurs when the two frames are each subscales of the same scale; Good and Bad are both *subscales of* Goodness, and Light and Dark are both *subscales of*

Luminosity. As discussed in Section 3.b.v., the frequency with which these oppositions occur in basic metaphors, such as GOOD IS UP and BAD IS DOWN or GOOD IS LIGHT and BAD IS DARK, drives the analysis of the *subscale* relationship. Similarly, the core nature of these oppositions leads to the development of the *is in scalar opposition to* frame relation, which specifies this opposition between two subscales. As a specialized variant of *mutually inhibits*, it is also a symmetric relation: Good *is in scalar opposition to* Bad, and Bad *is in scalar opposition to* Good.

3.d. *Interim summary*

The preceding sections describe the ontology of frame-frame relations developed by the MetaNet Project. As in other relational semantic systems, relations may be changed or added over time as analyses are refined and additional data prompts further development of the system. Table 5 summarizes the project's current frame relations.

Table 5. Summary of frame relations.

	Frame Relation	**Description**	**Example**
Structural relations	*is a subcase of*	full incorporation of frame structure	Purposeful Action *is a subcase of* Action
	is a special case of	no additional structure added	Boat *is a special case of* Seafaring Vehicle
	makes use of	partial incorporation of frame structure	Purposeful Action *makes use of* Desiring
	incorporates as role	frame fully included as an element	Eating *incorporates as a role* Food
	has affordance of	frame is an intrinsic property	Food *has affordance of* Eating
	is a process that makes use of	scenario is incorporated into a perspectivized process	Experience Harm *is a process that makes use of* Harm to Living Entity
	profiles part of	foregrounds element of the frame	Curriculum *profiles part of* Education
	is a subscale of	profiles element of a scale	Light *is a subscale of* Luminosity
	is a subprocess of	profiles stage in a process	Diagnosis of Physical Affliction *is a subprocess of* Physical Affliction

Non-structural relations		*is in causal relation with*	causal variants	Cause Motion Along a Path *is in causal relation with* Motion Along a Path
		precedes	temporal ordering between frames	Departing *precedes* Arriving
	mutually inhibits		frames are in semantic opposition to one another	Aids to Motion *mutually inhibits* Impediments to Motion
		is in scalar opposition to	opposite elements of a scale	Good *is in scalar opposition to* Bad

4.a. *Metaphoric construal*

Fundamentally, conceptual metaphors are mappings between frames that arise from our embodied experiences. The core of the conceptual network of metaphors are primary metaphors, such as STATES ARE LOCATIONS and CAUSES ARE FORCES, some of which are posited to be (near) universals (Grady, 1997; Lakoff and Johnson, 1999; Kövesces, 2005). These basic metaphors comprise the backbone of more complex or culturally specific metaphors via cascades (David et al., 2016). Drawing from formalisms developed in Embodied Construction Grammar, we propose that conceptual metaphors are evoked by linguistic metaphors via *role type constraint violations*, where a role type mis-match between individual lexical units in a given construct triggers a metaphoric interpretation of the sentence. In the following, we see a return to metaphorical construals of the lexical unit "poverty" as introduced in Section 2.b.i:

(6) a. poverty crushes *people*[29]
b. marginalized people already *injured* by poverty[30]
c. a country already handicapped by *crippling* poverty [31]

These phrases constitute three different grammatical constructions: (6a) the Subject-Verb construction, with the target domain lexeme in the subject and the source *crushes* as the verb; (6b) the Passive construction, with the target in the *by*-phrase adjunct and source *injured* in the verb; and (6c) Adj-Noun construction, with the target in the noun and source *crippling* in the adjective. While the syntax of each construction is different, the semantics of (6a) and (6b) clearly have an agent, either in canonical subject position or demoted in the passive. Further, the semantics of all three source-domain lexemes -- *crush, injure, cripple* -- convey the notion of a semantic patient being physically harmed by some injurious cause. Hence, these lexemes all evoke the Harm to Living Entity frame, which has the entity roles and attendant role types depicted in Table 6, below.

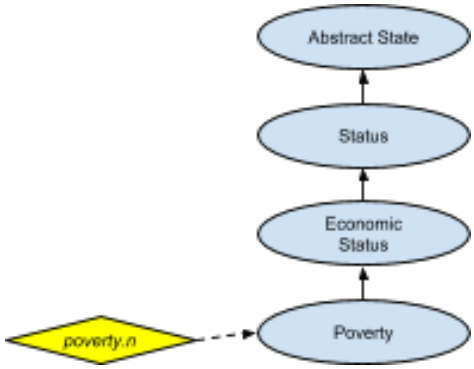
[29] http://www.imva.org/Pages/stories/viewstory.asp?The+Burden+of+Poverty
[30] http://dyinginhaiti.blogspot.com/2013/04/illness-as-moral-experience.html
[31] http://www.concernusa.org/story/rebuilding-haiti-3-years-after-the-earthquake/

Table 6. Entity roles and their role types in the Harm to Living Entity frame.

Entity role	**Role type**
cause_of_harm	Physical Entity
victim	Animate Entity

The frame element causing harm to the victim in the Harm to Victim frame is of type Physical Entity[32]. However, in (6a-c) above, the agent causing harm is the lexeme *poverty*, which evokes the Poverty frame; the corresponding role with this frame is of type Abstract State. Turning to an abbreviated form of the high-level frame network (Figure 5), we observe that *poverty* does not evoke a Physical Entity or one of its more specific descendants. Rather, *poverty* evokes an Abstract State frame.

Figure 5. The lexeme *poverty* evokes the frame Poverty, which descends from an Abstract State frame.



In non-figurative usage, the lexeme filling the slot bound to the cause_of_harm in the Harm to Living Entity frame should be of type Physical Entity, such as *rock* in *the rock crushed the bug*. Hence, a type mismatch exists between Harm to Living Entity as evoked by *crush*, *injured*, and *crippling*, and the role type (Abstract State) of the frame evoked by the lexeme *poverty*. The conflict that arises between the incompatibility of Poverty and Harm to Living Entity triggers the metaphoric interpretation of the above linguistic expressions of metaphor, and hence the conceptual metaphor, ECONOMIC HARDSHIP IS PHYSICAL HARM with its attendant mappings, as shown in Table 7.

[32] This state of affairs does not mean that the cause_of_harm must be of type Physical Entity; rather, cause_of_harm must be of type Physical Entity or a more specific frame that inherits from it, such as a Biological Entity or Animate Entity.

Table 7. Mappings between roles and their role types in the ECONOMIC HARDSHIP IS PHYSICAL HARM metaphor.

Poverty			Harm to Living Entity	
entity role	**role type**		**entity role**	**role type**
impoverishment	Abstract State	<==	cause_of_harm	Physical Entity
impoverished_entity	Person	<==	victim	Animate Entity

Notably, the impoverished_entity of the Poverty frame and the victim of the Harm to Living Entity frame *are* type-matched, because Person inherits from Animate Entity. The specific type mis-match of impoverishment and cause_of_harm lead to the metaphoric evocation.

Further, (6c) shows a secondary type mis-match stemming from the lexeme *country*, which evokes the Social Group frame. Given that the target domain of ECONOMIC HARDSHIP IS PHYSICAL HARM is an impoverished_entity of type Person, the entity *crippled* by *poverty* should be of type Person. However, Social Group as evoked by *country* does not match this role type. Thus, another metaphoric reading is required for (6c) to produce the CM above correctly. The NATION IS A PERSON metaphor as illustrated partially in Table 8 conceptualizes the nation as an entity of type Person. Hence, by construing the country as an individual, (6c) evokes this metaphor in order to conceptualize the nation as an entity that can undergo harm.

Table 8. Partial structure of the NATION IS A PERSON metaphor.

Nation			Person	
entity role	**role type**		**entity role**	**role type**
nation_whole	Social Group	<==	person_self	Person

4.b. *Metaphor cascade network*

Given that Poverty is a negative state, ECONOMIC HARDSHIP IS PHYSICAL HARM must be closely related to the CM ONGOING NEGATIVE STATE IS EXPERIENCING PHYSICAL HARM, as discussed above. The Poverty frame is a subcase of the Negative State frame; therefore, we posit a more general metaphor: NEGATIVE STATES ARE PHYSICAL HARM. Thus, ONGOING NEGATIVE STATE IS EXPERIENCING PHYSICAL HARM is an experiential variant of NEGATIVE STATES ARE PHYSICAL HARM. These hierarchical relationships between metaphors are formalized in similar fashion to the frame relations described in Section 3. One of the major innovations of formalizing CMT includes the development of the metaphor cascade network, which parallels the frame network.

The core of the metaphor network is the primary/experiential metaphors such as the Event Structure metaphors and orientation metaphors. The following section illustrates how (primary) metaphors are related to one another in a conceptual lattice-like network, rather than an

unstructured list. Metaphor relations are largely driven by the relations between the frames that constitute their source and target domains. By adding inferential structure to either domain, additional entailed metaphors of increasing specificity are produced. Starting with fundamental primary metaphors, a network that also includes culturally specific non-experiential CMs is built up. In the following section, we discuss metaphor relations and further explore the nature of the network in the context of a fragment of the Location Event Structure metaphor (LESM).

5.a. *Metaphor relations*

Lakoff and Johnson's (1999) discussion of the LESM includes the examples in (7), among many others.

(7) a. I'm *in* love. (Lakoff and Johnson, 1999: 180)
b. He's *in* a deep depression.
c. She's *out* of her depression.
d. She's *close* to insanity.

Upon examining the frames evoked by the target domain lexemes *love*, *depression*, and *insanity*, these four sentences all have target domains which are subcases of the highly schematized State frame: Emotion (7a, 7b, 7c) and Mental State (7d). Thus, these data illustrate perspectivized variants of the primary metaphor STATES ARE LOCATIONS; states are conceptualized as locations in which entities can exist in (7a, 7b) or out of (7c, 7d), with scalar physical closeness to the location corresponding to scalar degree of the state (7d). Thus these sentences make use of variants of the frames State and Location in evoking variants of the metaphor STATES ARE LOCATIONS[33]. Given our prior analysis of the frame relationship between the State scenario Negative State and the perspectivized Ongoing State, we observe that there is a consistent frame relation between experiential variants of scenarios, such that Being at a Location *is a process that makes use of* the Location frame just as Ongoing State *is a process that makes use of* the State frame.

As semantic frames and the concepts that define them are the primary units of analysis in cognitive semantics, frame structure and relations between frames provide the foundation of metaphor structure and relations as well. Just as the internal structure of a metaphor maintains components of the structures of the frames that constitute its source and target domains, relations between two metaphors are informed by the relations between their respective frames. Metaphors are composed in a network with hierarchical, structure-incorporating and non-structural relations similar to the network governed by relations between frames. The source domains of metaphors tend to be those based in shared embodied experiences, whereas target domains are typically more conceptually abstract and viewpoint-dependent. Furthermore, because they are embodiment-based, source domain frames are more likely to have coherent and elaborated internal structure *independent* of metaphor. In contrast, highly abstract or intangible

[33] More specifically, STATES ARE LOCATIONS is atemporal; STATE EVENTUALITY IS LOCATION EVENTUALITY adds the temporal semantics of a process that can have some duration. Hence, aspectual and causal variants of STATES ARE LOCATIONS actually descend from STATE EVENTUALITY IS LOCATION EVENTUALITY, as illustrated later on in Figure 7.

concepts like "poverty" may have their frame structure more shaped by their usage in metaphoric construal, due to the more frequent tendency to reason about them metaphorically. For this reason, because the structure and and relations of the source domain frames are less subject to conceptual contestation, it is likely that metaphor relations are driven largely by the more structurally specified source domain. (See David et al. (2016), for a more detailed discussion of the relatively contested nature of target vs. source domains.)

5.b. *Structural metaphor relations*

Returning to the analysis of the metaphor illustrated in (7), the relationship between ONGOING STATE IS STAYING AT A LOCATION and STATES ARE LOCATIONS can be explored in terms of relations between their constituent frames. Just like structural frame relations, structural metaphor relations describe a hierarchical relationship between metaphors such that the structure of the higher-level metaphor -- its frames, their roles, and the mappings between roles -- is bound to those of the more specific inheriting metaphor. Again, because of the nature of the cascade, this information is not "inherited" in the traditional sense, but rather it is directly incorporated via bindings. Hence with each more elaborate case, less elaborate cases are simultaneously active, rather than copied into the more specific metaphor. Here, as Ongoing State and Being at a Location are identified as *processes that make use of* State and Location respectively, the conclusion can be drawn that the metaphors that use them as target and source domains are in a structural relation, given that *is a process that makes use of* is a structure-incorporating frame relation. Specifically, ONGOING STATE IS STAYING AT A LOCATION *structurally incorporates* STATES ARE LOCATIONS; the *structurally incorporates* metaphor relation describes any metaphor relation wherein one's frames are in structural relations with the other's frames. Further examples illustrate a specific variation on ONGOING STATE IS STAYING AT A LOCATION:

(8) a. *living in* poverty[34]
b. *mired in* crime and corruption[35]
c. caught in the debt *trap*[36]

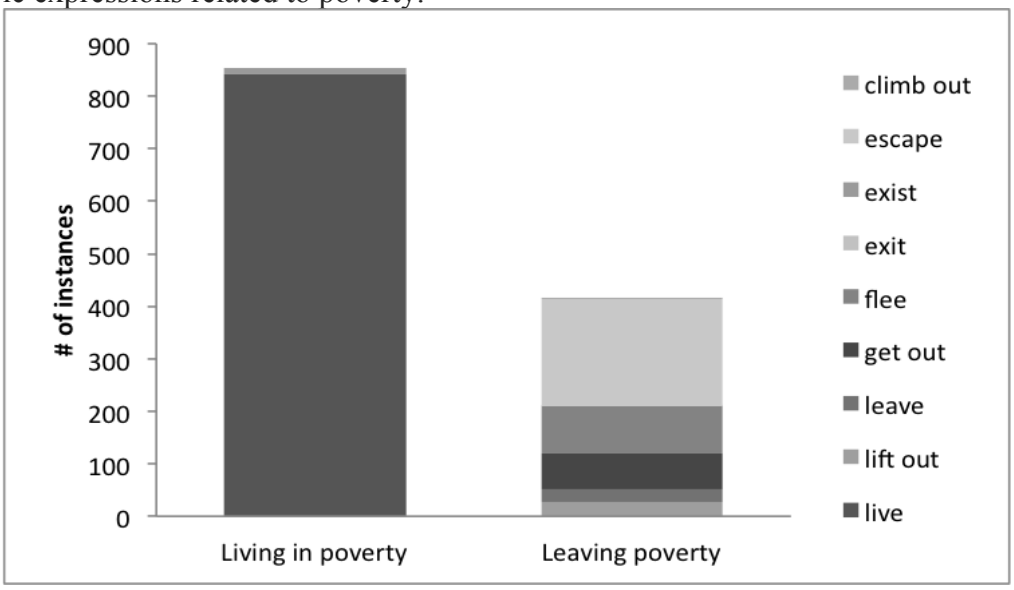
Again we see the experience of negative states (Poverty, Crime, Corruption, Debt), this time conceptualized as locations in which entities exist (and in some cases are unable to get out of).

The role of temporal duration also bears noting here. The examples in (8) vary in their morpho-syntactic realization of temporal information: (8a) is in the progressive, (8b) is a perfective atelic particle, and (8c) is a perfective telic particle. However, all of them convey a notion of the experience of a state for some duration of time. This aspectual construal is enabled in part by the compatible lexical semantics of the source and target domains of each lexeme; *living*, *mired*, and *caught* are all verbs indicating an ongoing event. Similarly, *poverty*, *crime*, and being in *debt* are all eventualities as well. Returning to the frame analysis of x-schema structures in Section 2.b.ii., we see that all frames which incorporate x-schematic temporal structure will have aspectual variations available to them. Hence, the lexical semantics of the individual lexical unit

[34] ENGW_afp_eng_199405:9440
[35] http://www.straighttalktt.com/media/?cat=21
[36] ENGW_xin_eng_199511:26479

participating in the linguistic metaphor must evoke frames with compatible x-schematic structure for the metaphoric mappings to be consistent. Which aspectual alternatives are most commonly occurring, however, may be driven more by the target domain. For example, in our corpus our system has extracted 13,658 metaphoric expressions with the target domain lemma "poverty". There are over twice as many linguistic expressions about living or existing in poverty (854) as there are about leaving poverty (416), as shown in Figure 6.

Figure 6. Frequency of lemmas related to existing in or exiting a location when used in metaphoric expressions related to poverty.



From this distribution of source domain language, we can conclude that poverty, when metaphorically conceptualized as a location, is more commonly thought of as a location that people exist in (i.e., a durative state) rather than a location people leave (i.e., cessation of a state).

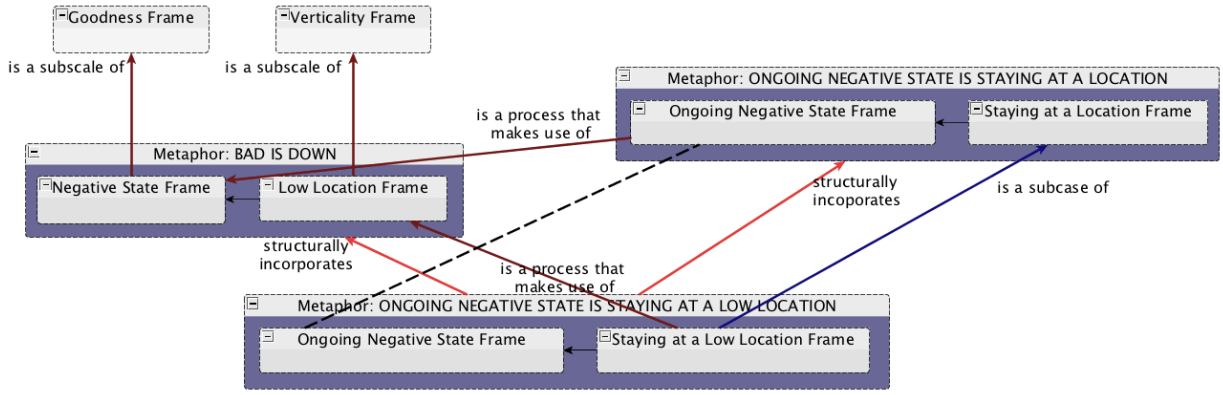
The fact that all the expressions in (8) evoke a negative state suggests a more specific metaphor: ONGOING NEGATIVE STATE IS STAYING AT A LOCATION. Ongoing Negative State is a *subcase* of Ongoing State, the target domains of their respective metaphors. Given this *subcase* relation between the two target domains of ONGOING NEGATIVE STATE IS STAYING AT A LOCATION and ONGOING STATE IS STAYING AT A LOCATION, it follows that the more-specific metaphor *structurally incorporates* the more general metaphor, whose target domain does not specify the type of State. Compare the above with the following:

(9) a. the *pits* of poverty[37]
b. The Philippines has stumbled over reform and is now *sunken in* corruption[38]
c. help Mexico *climb out* of debt[39]

[37] http://www.cihadf.org/The-Edge-Initiative
[38] http://www.freerepublic.com/focus/f-news/1044148/replies?c=1
[39] ENGW_apw_eng_199606:73315

In these examples, Negative States continue to be metaphorically construed as Locations, but they are specifically *Low* Locations via an additional basic metaphor, BAD IS DOWN. In this case, the source domain Staying at a Low Location is a *subcase* of Staying at a Location. Given this *subcase* relation between the two frames, the consequent metaphor ONGOING NEGATIVE STATE IS STAYING AT A LOW LOCATION *structurally incorporates* ONGOING NEGATIVE STATE IS STAYING AT A LOCATION. The frame Staying at a Low Location *is a process that makes use of* Low Location (*a subscale of* Verticality), which is the source domain of BAD IS DOWN. Ongoing Negative State *is a process that makes use of* Negative State (*a subscale* of Goodness), the target domain of BAD IS DOWN. Both frames incorporate the scalar structure of the general scalar scenarios by specifying the low/negative end of each scale. The scalar information in the source and target frames of BAD IS DOWN are accessed via *is a process that makes use of* relations between Ongoing Negative State and Negative State; and Staying at a Low Location and Low Location. Hence, the *is a process that makes use of* relations between the frames involved in each metaphor result in ONGOING NEGATIVE STATE IS STAYING AT A LOW LOCATION incorporating the subscale valences of BAD IS DOWN, meaning that ONGOING NEGATIVE STATE IS STAYING AT A LOW LOCATION must also *structurally incorporate* BAD IS DOWN; the partial network is illustrated in Figure 7.

Figure 7. Metaphor and frame relations between ONGOING NEGATIVE STATE IS STAYING AT A LOCATION, ONGOING NEGATIVE STATE IS STAYING AT A LOW LOCATION, and BAD IS DOWN. Dotted lines indicate bindings and arrows indicate relations.



Further comparison of the above source lexical units reveals additional entailments of the general Being in a Location source domain: *mired, trapped,* and *sunken* all carry the inference of the inability to (easily) escape from the Location, whereas *climb out* brings with it the opposite inference, that of the ability to get out of the Location. Analysis of these variations requires beginning with the primary metaphor ACTION IS MOTION, which provides the basic structure of the metaphor PURPOSEFUL ACTION IS MOTION TO A DESTINATION via the intermediate metaphor ACTION IS MOTION ALONG A PATH. In (9b), the source lexeme evokes the Upward Motion frame, which is a *subcase* of Motion Along a Path. Thus, the Purposeful Action of becoming debt-free is Upward Motion, because Debt is already metaphorically construed as a Low Location. From this analysis we might propose the highly specific metaphor IMPROVEMENT OF STATE IS UPWARD MOTION as an entailment of ONGOING NEGATIVE STATE IS STAYING AT A LOW LOCATION. Entailed

metaphors are highly specific variants that explicate the inferential structure of a given metaphor. In this case, IMPROVEMENT OF STATE IS UPWARD MOTION draws on the inferential structure of GOODNESS IS VERTICALITY incorporated into ONGOING NEGATIVE STATE IS STAYING AT A LOW LOCATION: if a Negative State is construed as a Low Location and negative qualities the low region of a vertical scale, then an improvement of the State corresponds to upward movement from that low point.

Why propose such detailed entailments of the LESM? Our elaboration of the nature of these specific entailed metaphors is directly driven by our corpus data. By investigating differences in usage frequencies between individual lexemes which evoke different components of related source domain frames, we can uncover crucial differences in how people conceptualize and reason about target domains. Dodge (2016) provides such a detailed exploration of the subtle distinctions in metaphoric construals of poverty. In short, consider for example the nuanced difference between *trapped in the pit of poverty* and *mired in poverty*. Both are from the viewpoint of the individual experiencing poverty, and both portray poverty as a location that is difficult to escape, with the inference that successful motion out of this location would mean the individual has improved their economic status. Nonetheless a *pit*, however deep, does not include the semantics of being truly physically stuck as in *mired*, which conveys the notion of some kind of extremely sticky situation. If the individual has a *rope* or a *ladder*, he or she might *climb* out of the pit. In contrast, it would be more difficult to extricate themselves if *mired* in a particular location. As these sorts of subtle variations in lexical choices in the data are observed by the analyst, additional entailed metaphors are in turn developed.

5.c. *Non-structural metaphor relations*

Stopping the metaphor analysis at this point would miss a crucial generalization. The entailed metaphor IMPROVEMENT OF STATE IS UPWARD MOTION also derives from a much more general metaphor CHANGE OF STATE IS CHANGE OF LOCATION, an *alternative of* ONGOING STATE IS STAYING AT A LOCATION. There is an additional relation needed, the *alternative of* metaphor relation, that accounts for non-structural relations. Whereas the *structurally incorporates* metaphor relation indicates that elements of one metaphor's frames are incorporated into the other's, the *alternative of* metaphor relation indicates that they are alternatives, like aspectual or causal variants. For example, CHANGE OF STATE IS CHANGE OF LOCATION is an aspectual variant of ONGOING STATE IS STAYING AT A LOCATION; CAUSED CHANGE OF STATE IS CAUSED CHANGE OF LOCATION is a second variant that further builds in causation.

The alternative nature of these metaphors follows from the simple inference that if an entity changes from one state to another, it no longer experiences the first state. This logic also applies to the primary experiential scene of being in a location: if an entity experiences the properties of a location while *in* that location, then when *out* of the location those properties are no longer experienced. From this observation of alternative stages of both the experiences of states and experiences of locations, these alternatives also apply when state change is construed metaphorically as an alternative of the primary metaphor STATES ARE LOCATIONS. We may further conclude that if CHANGE OF STATE IS CHANGE OF LOCATION, and CAUSED CHANGE OF STATE IS CAUSED CHANGE OF LOCATION, then PREVENTED CHANGE OF STATE IS PREVENTED CHANGE OF LOCATION. This notion of prevention of state change accounts for the entailed

semantics of *mired, trap,* and *sunken* from (8b-c) and (9b-c), restated as (10a-d) with additional examples (10e-f).

(10) a. *mired in* crime and corruption
b. caught in the debt *trap*
c. The Philippines has stumbled over reform and is now *sunken in* corruption
d. help Mexico *climb out* of debt
e. *pulled from* the *abyss* of unemployment, hunger and poverty[40]
f. public education has long been a *ladder out* of poverty[41]

The lexical units in (10a-c) each evoke a *subcase* of the Impediments to Motion frame: a specific kind of Low Location which prevents a change of location by impeding motion.

Finally, we consider some more complex examples, in (10e-f). (10e) mirrors (9c, restated as 10d) in that the former involves Upward Motion, as does the potential of upward movement in (10d). However, in contrast to *climb* (10d), the verb *pull* (10e), evokes Caused Upward Motion rather than a self-propelled motion. Based on prior analyses, the addition of causation suggests the metaphor ENABLED CHANGE OF STATE IS ENABLED CHANGE OF LOCATION, an alternative to PREVENTED CHANGE OF STATE IS PREVENTED CHANGE OF LOCATION. Whereas (10e) involves an external cause of upward movement, the *ladder* in (10f) provides a *means* to upward movement out of a location. This can be compared to (10b), where the *trap* is an instrument *preventing* movement out of location. Thus, as (10b) evokes PREVENTED CHANGE OF STATE IS PREVENTED CHANGE OF LOCATION, in contrast (10e-f) evoke ENABLED CHANGE OF STATE IS ENABLED CHANGE OF LOCATION. The lexeme *ladder* evokes the enabling_source role in a Aids to Upward Motion, a *subcase* of the Aids to Motion frame. *Pulled* does not specify the means, but similarly evokes externally-caused motion. Impediments to Motion have the opposite force dynamics to Aids to Motion and hence are *mutually inhibitory*.

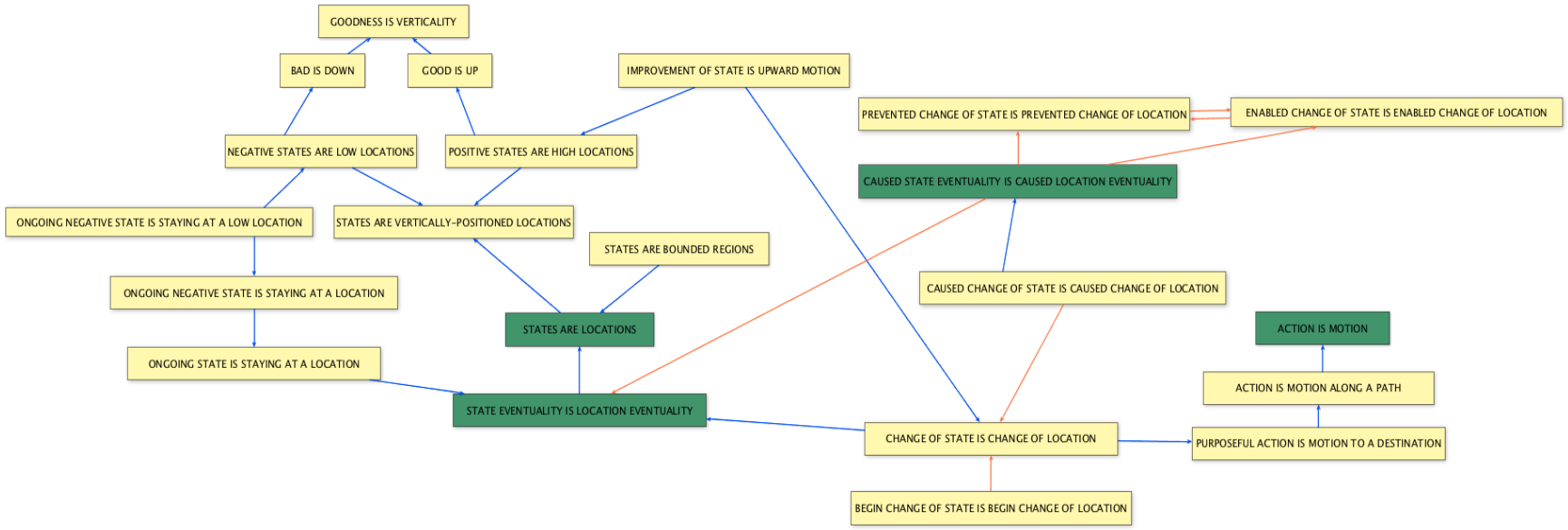
Prevented Change of State and Enabled Change of State similarly *mutually inhibit* each other as incompatible alternative stages of Change of State: by preventing change, enablement of change is also made impossible, and vice-versa. Thus, the two sister entailed metaphors PREVENTED CHANGE OF STATE IS PREVENTED CHANGE OF LOCATION and ENABLED CHANGE OF STATE IS ENABLED CHANGE OF LOCATION are *alternatives of* one another. This metaphor relation reflects the mutually incompatible inferential structure of preventing and enabling state change. Because the semantics of preventing and enabling both imply a cause of the prevention or enablement, the frames of these metaphors constitute alternative variants of causation. In turn, the metaphors are *alternatives of* CAUSED CHANGE OF STATE IS CAUSED CHANGE OF LOCATION.

In addition to causal variants, the *alternative of* metaphor relation also describes the relationship between metaphors whose frames are aspectual variants. As analyzed in Table 3, Begin Process is a *subprocess* (i.e., an alternative) of a Process; hence Begin Change of State and Begin Change of Location are *subprocesses* of Change of State and Change of Location. The entailed metaphor BEGIN CHANGE OF STATE IS BEGIN CHANGE OF LOCATION provides an aspectual variant of CHANGE

[40] http://www.eurozine.com/articles/2014-09-03-majcen-en.html
[41] ENGW_apw_eng_199512:21646

OF STATE IS CHANGE OF LOCATION that profiles the start of the metaphoric process. As described in Sections 2.b.ii. and 3.b.v., subprocesses are stages of an x-schema, or rather, stages of processes (Narayanan 1997). These can be generalized as the beginning, ongoing stage, and ending of processes, with several intermediate stages. These *subprocess* relations between the sources and targets of the two metaphors shows that BEGIN CHANGE OF STATE IS BEGIN CHANGE OF LOCATION *is an alternative of* CHANGE OF STATE IS CHANGE OF LOCATION. This partial analysis of the LESM is summarized in Figure 8.

Figure 8. Relations between metaphors in the Location Event Structure Metaphor Family. Blue connections are *structurally incorporates* relations; orange represents *alternative of* relations.



6.a. *Conclusions*

The above analysis of the Location Event Structure Metaphor Family is inherently a partial one. By treating x-schematic process variation and causal variation as separate alternatives that apply to any process, the conceptual network allows for expansion of any Process scenario into combinations of aspectual and causal variations. For example, Ongoing State has a causal variant Caused Ongoing State, which in turn has further variants Enabled Ongoing State and Prevented Ongoing State, with aspectual variations such as Enabled Beginning of Ongoing State or Enabled Resumption of Ongoing State, and so on. The combinatorial nature of the scenarios, causal variants, and aspectual variants provides for an understanding of a how the semantics of such a conceptual network relate. It is daunting for a human analyst to attempt to list or comprehend the nature of such a complex structure, particularly as additional scenarios are added as data necessitates that more frames are developed.

6.b. *Towards a computational implementation*

These challenges for the human analyst are addressed in part by the computational implementation of CMT in our system repository. As each frame and each metaphor is given its own page in a Semantic MediaWiki-based database (Krötzsch et al., 2007), they are easily viewed and edited by individual analysts. Live links between frames, metaphors, and frame elements instantiate the relations described in this paper. Hence, the network lattice is built into the repository, and as analysts add new frames and metaphors the network grows with it. Additional visualization tools such as automatically-generated graphs and software which allows the analyst to perform advanced search queries further enable the metaphor analyst to gain a larger view of different aspects of the network. Because the wiki is an easily editable, familiar format, it is accessible to other linguists and researchers in a variety of fields; the underlying data can be downloaded and exported into other databases as well, which provides the possibility for greater collaboration among cognitive linguists and in allied disciplines. Furthermore, because the representations of frames and metaphors as instantiated in the wiki require the analyst to complete certain fields (such as specifying the type of relations between related frames, or entering the source and target frames in metaphor entries), performing metaphor analysis in this context is by nature required to be rigorous and thorough. As analyses are refined and additional structure added to frames, these changes can be automatically propagated throughout the network.

In addition to a tool for metaphor analysis, this repository also provides the backbone of our system's automated metaphor extraction pipeline, as described in Dodge et al. (2015) and Hong (2016). The network of frames and metaphors is accessed first via the lexical units associated with frames; a defined set of metaphoric constructions in conjunction with the metaphors and metaphor relations defined in the repository lead to identification of linguistic metaphors in a corpus. For example, the analyst can specify a set of source or target domains, and the system will identify and annotate all the metaphoric linguistic expressions in a corpus that use the specified frames. This metaphor extraction system will thus enable larger-scale, finer-grained corpus approaches to metaphor. A detailed example of this approach follows in Chapter 3.

Chapter 3: Location Event Structure Metaphor case study

1. *Overview*

In this chapter, I provide a case study of the Location Event Structure Metaphor (LESM) as used in a corpus of spoken American English (COCA). First, I discuss a key component of the system: the metaphoric constructions as implemented in Embodied Construction Grammar which are used for metaphor identification. Then I analyze a dataset of metaphoric utterances in which the source domain of translational motion is evoked by manner and path verbs.

2.a. *Metaphoric constructions*

The MetaNet system comprises two main parts: The Repository, which contains the result of frame and metaphor analyses by a team of cognitive linguists, and the Identification System, which is an automated system that uses the Repository along with a set of constructions and frame/metaphor relations to identify metaphoric language in text. The constructions are based on the observation by Croft (1993) and Sullivan (2007; 2013) that metaphoric source and target domain language reliably occurs in the same syntactic positions. Metaphoric constructions are comprised of an argument structure construction combined with metaphor-evoking language (Goldberg, 1995). The argument structure construction specifies the form of a syntactic pattern and its meaning. For example, the Ditransitive Construction has the form Subject Verb Object1 Object2. The Subject, Object1, and Object2 are the sender, recipient, and theme which is transferred from the sender to the recipient; the construction itself conveys the semantics of the transfer of the theme from sender to recipient (Goldberg, 1995). In *He baked her cookies*, "bake" is not a verb of transfer, but in the context of the Distransitive Construction we understand that *baking someone cookies* entails not only making the cookies but also giving them to the intended recipient. In a metaphoric argument structure construction, the construction additionally specifies that the syntactic positions of the construction are filled by lexemes that evoke the source or target domain of a metaphor. For example, in (1) below, the source domain-evoking language occurs in the verbs and the target domain-evoking language in the nouns.

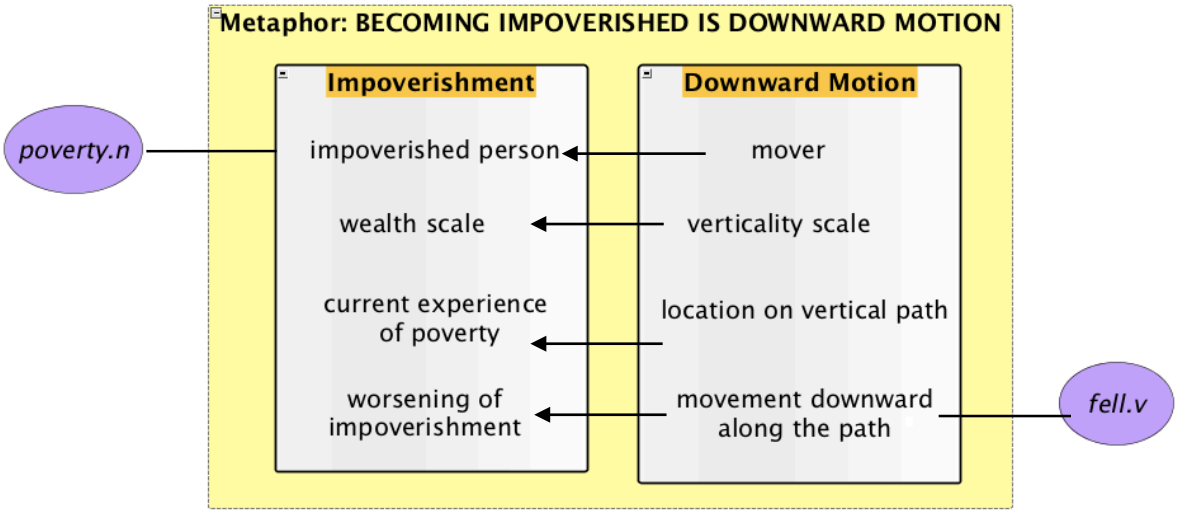
(1) a. prices *plummeted*
b. *enter* an agreement
c. they *fell* into poverty

The phrases in (1) represent three different types of metaphoric verbal constructions. In each case, the verb evokes the source domain of the metaphor and the noun the target domain of the metaphor (see Chapter 1, section 2.a. for an overview of conceptual metaphor theory). In (1a) the subject *prices* evokes the target domain of Quantity and the verb *plummeted* the source domain of Downward Motion; (1a) evokes the metaphor decrease in QUANTITY IS DOWNWARD MOTION. (1a) is an example of the Subject (Target) – Verb (Source) construction as the Subject is filled by a target-domain evoking lexeme and the Verb a source-domain evoking lexeme. (1b) is an example of the Verb (Source) – Object (Target) construction: the verb *enter* evokes the frame Motion Along a Path and *agreement* a Communication frame, and *enter an agreement* as a whole evokes the metaphor BEGINNING TO AGREE IS ENTERING A SHARED LOCATION. In (1c) we see an example of the Verb (Source) – Preposition – Noun (Target) construction; the verb again evokes

a source domain frame Downward Motion and the noun a target domain frame Poverty, this time inside a prepositional phrase.

In each case, the verbs evoke a translational, Motion Along a Path frame; *plummeted* and *fell* evoke Downward Motion, whereas *enter* evokes Motion Along a Path with the additional specification of a Location as the Goal of the Trajector. The verbal arguments, in turn, each evoke a different target domain: Quantity (1a), Communication (1b), and Poverty (1c). In all three cases, the path verbs evoke variations of the CHANGE OF STATE IS CHANGE OF LOCATION entailment of the LESM. In (1a) the Change of State is a Decrease in Quantity; in (1b) it is a change in the state of Agreement; and in (1c) it is Becoming Impoverished. Setting aside for a moment the BAD IS DOWN component of (1c), we can represent the metaphor analysis for BECOMING IMPOVERISHED IS DOWNWARD MOVEMENT as in Figure 1.

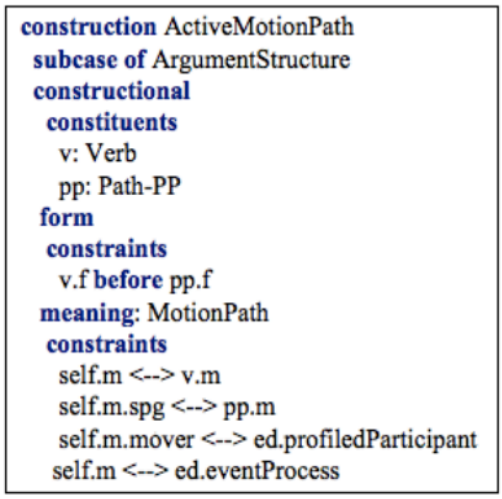
Figure 1. Metaphor analysis of BECOMING IMPOVERISHED IS DOWNWARD MOVEMENT, evoked by *fell into poverty.*



The noun *poverty* evokes the Impoverishment frame, whereas the verb *fell* evokes the process of motion in the Downward Movement frame. The impoverished_entity is conceptualized as a physical entity, the mover; the process of changing degree of impoverishment is the process of motion along a path. Lower locations on the Vertical path map onto increased degrees of impoverishment, due to a combination of QUANTITY IS VERTICALITY (poorer people have less money) and GOODNESS IS VERTICALITY (being poor is bad, so it is a low point on the vertical scale).

To incorporate this metaphor analysis into a constructional analysis, we first make use of the analysis by Dodge and Petruck (2014) of the Active Motion Path Construction (Figure 2), which is neutral as to whether it is literal or metaphoric. The metaphoric construction will inherit this construction's structure.

Figure 2. Analysis of the Active Motion Path Construction by Dodge and Petruck (2014: 42).



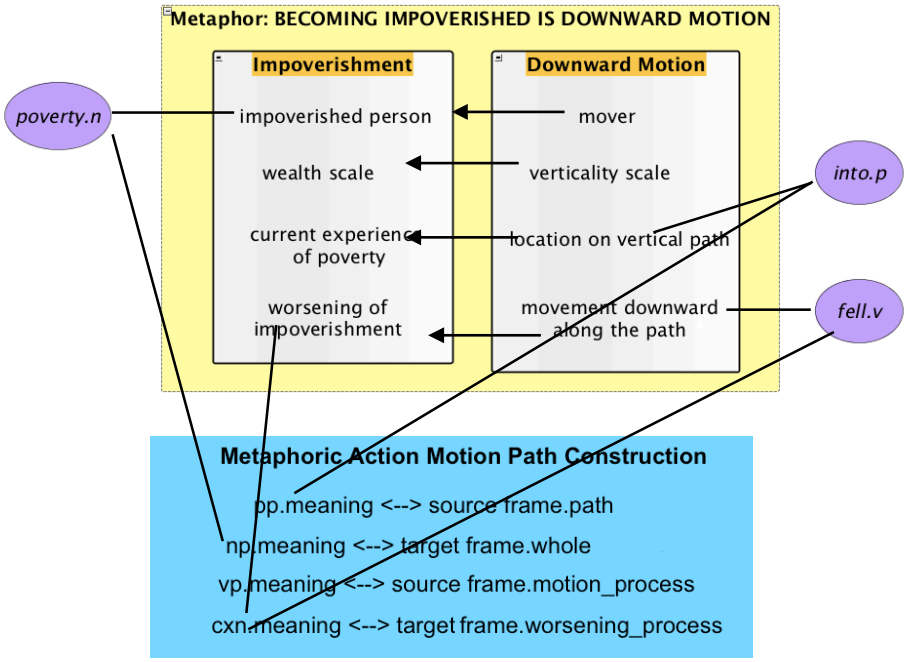
In this analysis, the construction comprises a verb and a prepositional phrase; its meaning is the Motion along a Path frame. The mover of the frame is the Participant in the moving Process, which is linked to the verb. The Source-Path-Goal image schema of the Motion along a Path frame links to the preposition, which provides the path information. In the case of the *literal* Active Motion Path phrase *he fell into a hole*, the analysis is as follows. Note that in construction grammar, lexemes are themselves constructions as well; hence *he, fell,* and *hole* are lexical constructions. *He* is part of a NP construction, and *into a hole* comprises the Path-PP construction which constrains the prepositional phrase to a Source-Path-Goal reading. The NP *he* is the mover of the construction; *into a hole* specifies the path (*into*) and goal (*a hole*) of his movement. To extend this to the metaphoric *they fell into poverty*, we can add constraints to the meaning of the construction that link the verb and Path-PP to the source and target domain of the metaphor (Figure 3).

Figure 3. Metaphoric variant of the Active Motion Path Construction.

```
Construction Metaphoric Active Motion Path
Subcase of Active Motion Path
Meaning: Metaphoric Motion Path
  Constraints
   pp.meaning <--> source frame.path
   np.meaning <--> target frame.whole
   vp.meaning <--> source frame.motion_process
   cxn.meaning <--> target frame.worsening_process
```

The metaphoric construction combines the two lexemes into a configuration that licenses a metaphoric reading, by mapping the frame evoked by the verb onto the source domain of the metaphor and the frame evoked by the PP onto the target domain. This analysis is represented diagrammatically in Figure 4.

Figure 4. Combination of metaphor and construction analysis of *fell into poverty*.



We can observe a similar pattern with manner verbs in (2).

(2) a. his attention span *drifts*
b. *sway* someone's vote
c. *roll* out a new program

The examples in (2) represent the same three types of metaphoric constructions as in (2). While the verbs in (2) are manner verbs rather than path verbs, they still evoke some kind of Motion along a Path; (2b) and (2c) are additionally Caused Motion constructions. Due to the hierarchical nature of constructions, a metaphoric Caused Motion construction incorporates the Caused Motion construction, which itself inherits information from the Action Motion Path construction (Dodge and Petruck, 2014). Again, in all these cases we see that the metaphoric construction links the verb to the source domain and the argument to the target domain.

2.b. *Metaphoricity evaluation*

These systematic relationships between syntactic structure and metaphor as licensed by the metaphoric constructions are a critical component of automatically identifying metaphoric language in text. When tasked when finding metaphors with a Motion source domain, the MetaNet extractor can focus on only those utterances in which a motion-evoking lexeme occurs in the source domain slots. In the case of verbal constructions as shown in (1) and (2), the motion-evoking lexeme will be a verb. Conversely, if tasked with finding metaphors with a Poverty target domain, the extractor will find utterances with poverty-evoking lexemes in target domain slots (in this case, the noun or prepositional phase). Once it has identified a set of utterances with the appropriate combination of frame-evoking lexemes and syntactic positions, the system can then make use of the relations between metaphors and frames in the Repository to evaluate the metaphoricity of each utterance.

For example, *fell into a hole* would be given a low metaphoricity score as *hole* evokes a physical location, and *fell* a physical path. There is a close connection between the two in the frame network, which reflects their closely-related semantics, and the *type match* between the two reflects the fact that this is a literal utterance (Figure 5).

Figure 5. Relationships between roles in the frames evoked by *fell into a hole.*

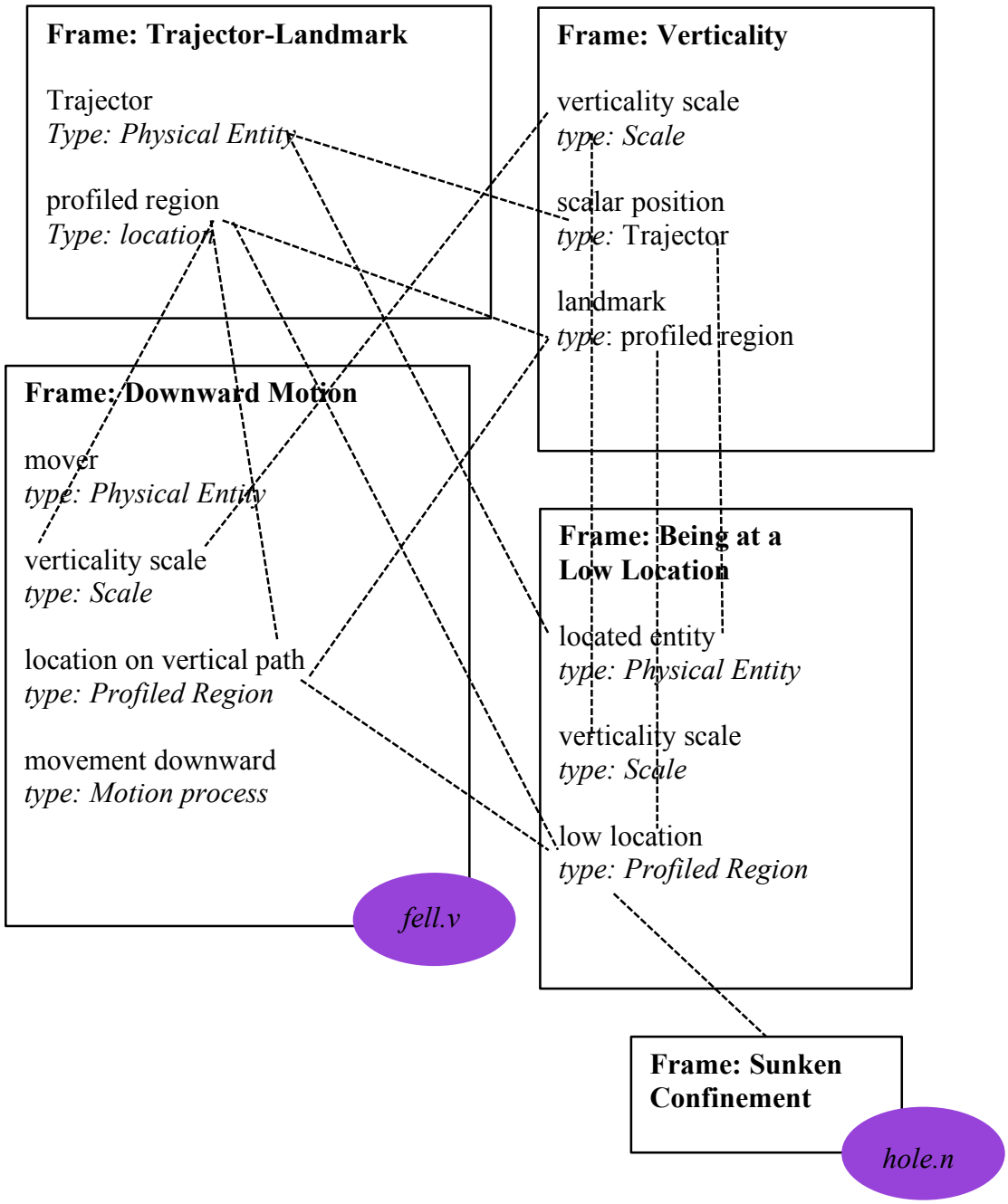
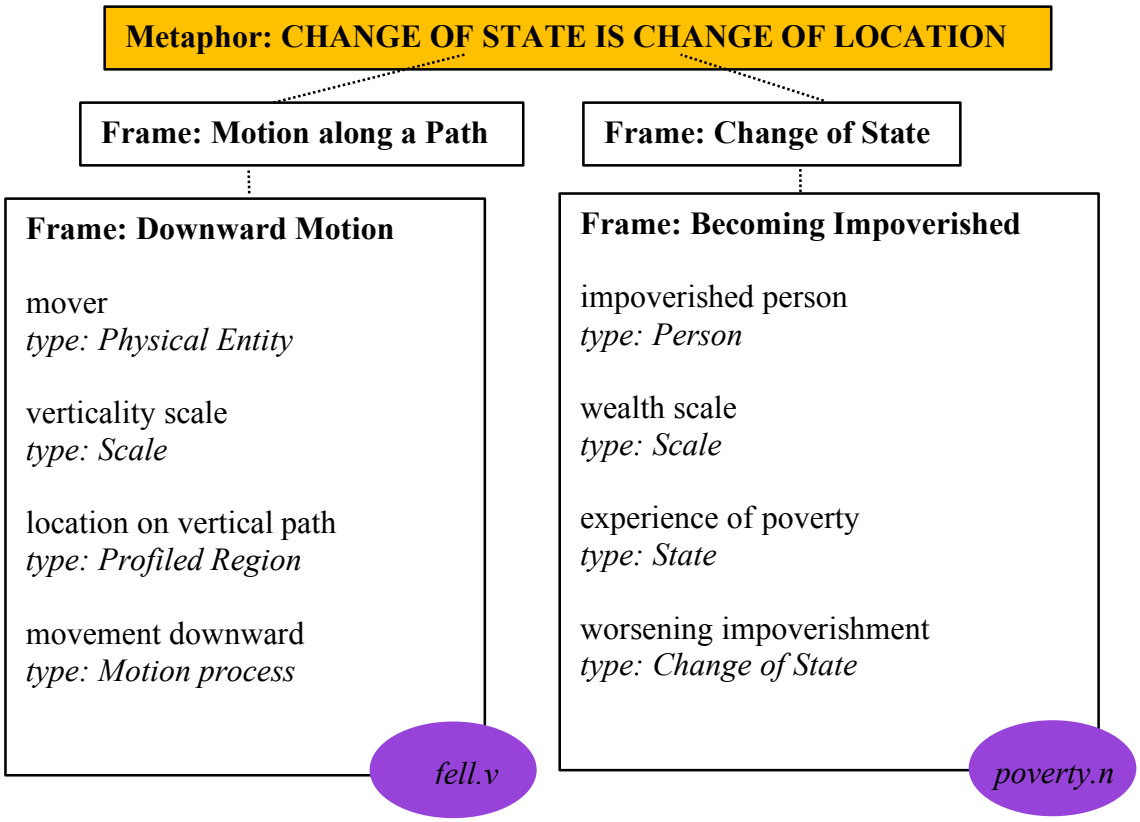


Figure 5 provides a partial analysis of the relationships between the Downward Motion frame evoked by *fell* and the Sunken Confinement frame evoked by *hole*. Sunken Confinement is a

special case of a Low Location, which inherits from both Verticality and Trajector-Landmark. Downward Motion also inherits information from Verticality and Trajector-Landmark; in a complete analysis, there would be several intermediate levels of frames including Source-Path-Goal and Motion along a Path. This analysis illustrates that the frames evoked by *fell into a hole* share a Physical Entity (the mover or located entity) and Vertical locations which incorporate a Trajector-Landmark relationship.

In contrast, *fell into poverty* would be given a high metaphoricity score due to the fact that a metaphor intervenes between the frames evoked by each lexeme in the network (Figure 6).

Figure 6. Relationships between the frames evoked by *fell into poverty*.



In contrast to *fell into a hole*, the utterance *fell into poverty* is highly likely to be metaphoric given the relationship between the Downward Motion frame evoked by *fell* and the Becoming Impoverished frame evoked by *poverty*. While both the mover and impoverished person are Physical Entities (as a Person is a type of Physical Entity) and the image schematic structure of the verticality scale maps onto the Wealth Scale, the *processes* are of different types. Whereas movement downward is a type of physical motion, the process of becoming impoverished is a change of state. Furthermore, a location on a vertical path is a physical region, whereas being impoverished is a state. In other words, there is a mis-match between the states and locations: this an evocation of the CHANGE OF STATE IS CHANGE OF LOCATION metaphor. And so we see that in the metaphor network, that metaphor intervenes between the frames Downward Motion and Becoming Impoverished. This intervening metaphor indicates to the extraction system that this utterance is highly metaphoric.

This is a basic description of the logic behind the metaphoricity system; in practice it makes use of a complex set of relational configurations (Hong, 2016), which weighs relationships between frames depending on the type of relation (for example, a *subcase* relation is scored higher than a *makes use of*). Once the system has evaluated each of the potentially-metaphoric utterances, it produces an annotated corpus as the result of the search process. The annotated information includes the metaphoric construction, the lexemes evoking the source and target domains, the relevant frames and metaphor, and the metaphoricity score.

3.a. *LESM case study*

In this study, I focused on the LESM as realized in metaphoric verbal constructions like those discussed in section 2.a. above. To do so, I compiled a set of 359 translational motion manner and path verbs; the details of this process are discussed in further detail in Chapter 4. Given this set of verbs, I used the MetaNet extractor to identify utterances in which one of those verbs appeared in the source domain slot of one of the three metaphoric verbal constructions: Subject (Target) – Verb (Source); Verb (Source) – Object (Target); and Verb (Source) – Preposition (Target). This ensured that all of the metaphoric utterances identified by the extractor would have a Translational Motion source domain. The extractor excluded all utterances in which a Motion or Location frame-evoking lexeme occurred in the target domain constructional slot, as that would indicate a literal motion event (e.g., *run into a barrier*). It also excluded utterances in which Physical Entities occurred in the target domain constructional slot, as that would also indicate literal motion (e.g., *the cat sauntered*). The search was done on the Corpus of Contemporary American English (Davies, 2008), which includes 109 million words of spoken English and 424 million words of written English across different genres (academic, newspaper, magazine, and fiction writing).

3.b.i. *Results and discussion of case study*

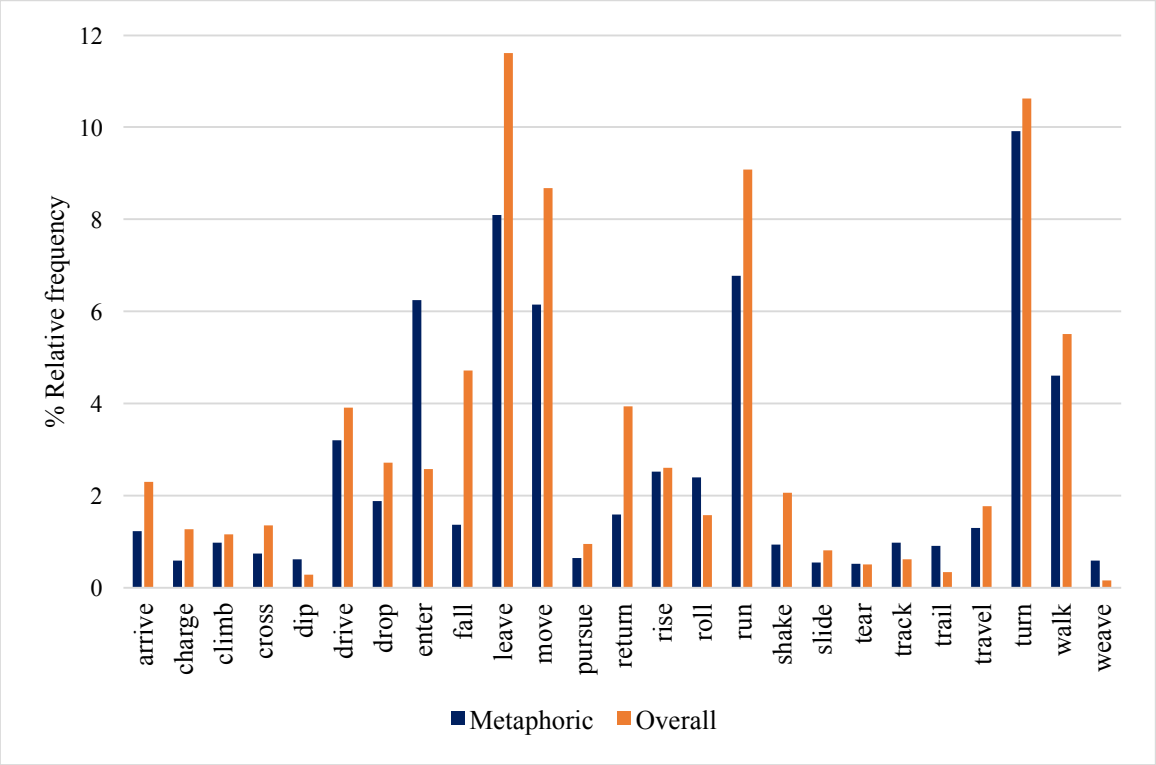
The metaphor extraction process identified 3,080 utterances that fulfilled the following criteria: (a) it contained a motion-evoking verb; (b) the motion-evoking element filled a source domain slot in a metaphoric verbal construction; and (c) the target domain slot was *not* filled by a motion- or location-evoking word in a metaphoric construction. If a motion or location-related word occurred in the target domain position in the construction, that is highly *unlikely* to be metaphoric.

Given the exploratory nature of this case study, I will first describe several characteristics of the resulting metaphoric corpus. By analyzing common source domain lemmas and comparing them to overall use frequency, we can understand how English speakers (and writers) selectively choose to evoke the LESM using specific words. Considering relative use frequency of metaphoric constructions, and how they interact with word choice, can provide further illumination of the ways in which speakers are understanding state change and the biases of speakers towards profiling and backgrounding certain frame elements. We can also consider more broadly how speakers make use of different motion frames to evoke different variants of the LESM. Finally, I make note of some intriguing results and further pursue them in a short case study of the verb *follow*.

3.b.ii. *Lemma frequency*

Beginning with basic word choice, we can observe tendencies in linguistic realizations of the LESM. 65 different motion-evoking lemmas were found; Figure 15 shows the relative frequency of the most common lemmas (at least 0.5% of tokens) in the metaphoric data in comparison to their overall relative use frequency in the COCA data. Overall use frequency in COCA was calculated by combining the COCA frequency counts for the present tense (*walk, walks*), past tense (*walked*), and present participle (*walking*) forms of the verb. Relative use frequency was then calculated by comparing these counts against one another; so, here use frequency is relative to the other verbs in the data, rather than overall corpus frequency in COCA. Note that Figure 7 purposefully excludes the data for the verb *follow*, which was an extreme outlier in the metaphoric data: it constitutes fully 27.05% of the metaphoric data (and 5.82% of literal data). To avoid skewing the graph, it is excluded from Figure 7.

Figure 7. Relative frequency of common source domain lemmas in comparison to overall relative frequency (excluding *follow*).

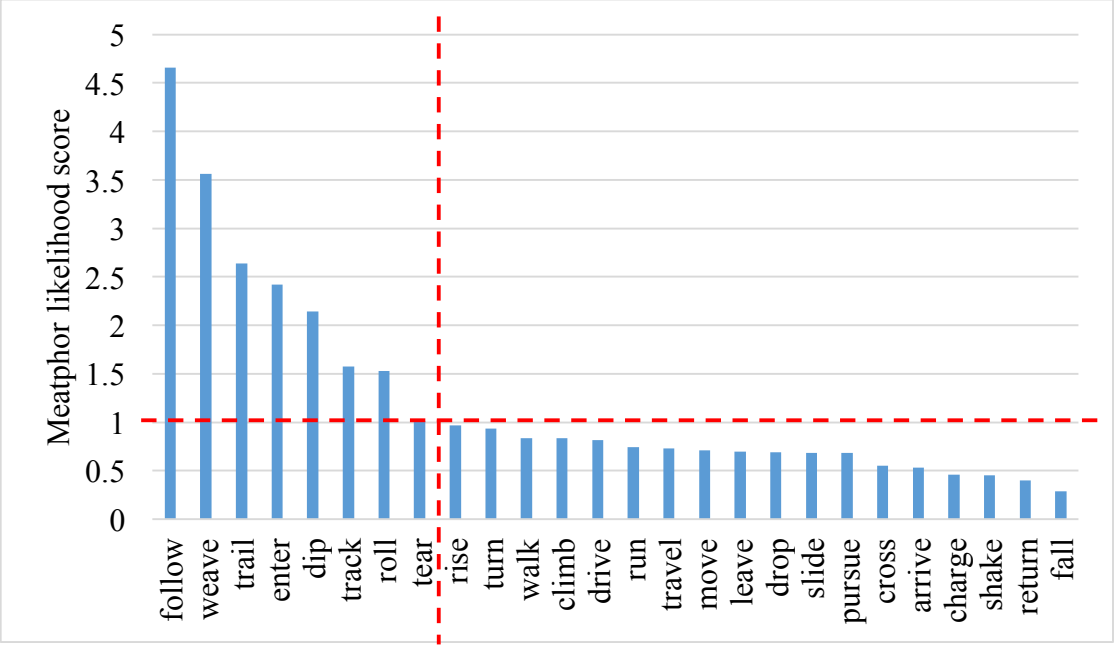


Impressionistically, we can observe from Figure 7 differences between lexemes in terms of use frequency. Some lemmas, such as *rise* and *travel* appear to have relatively similar use frequencies in both the metaphoric and overall data. In addition to *follow*, *enter* occurs more frequently in the metaphoric data than in the overall data. In contrast, although *leave* and *run* are relatively common in both contexts, they are more frequent in the overall data, indicating a lower metaphoric usage rate. Prior work in corpus approaches to metaphor has shown that use frequency for specific collocations varies depending on lexical choice (Deignan, 2006). For example, *heavy price* is more frequent than *weighty price*, although they make use of the same target domain-evoking lexeme *price* and both *heavy* and *weighty* evoke the same source domain.

While Deignan (2006) focused on specific collocations using concordance-based methods combined with hand annotation, the MetaNet approach combines a wider search with automatic annotation, producing a larger amount of data. Here we see similar results for individual LESM-evoking lexemes across an array of target domain lexemes and syntactic patterns. For example, *run* occurs more frequently than *walk* in both the overall and metaphoric data.

To compare relative metaphoricity directly, we can calculate what I term a lexeme's *metaphoricity likelihood score*, which compares a lexeme's metaphoric frequency to its overall frequency (Figure 8). A score > 1 indicates it is more frequent in the metaphoric data than overall, and a score < 1 indicates the opposite. This approach is similar to that in Lederer (2015a) who compares use frequency in a topic-specific corpus to overall use frequency in COCA, following the methodology established in Ahmad (2005). This allows her to identify highly salient lexemes in the topic-specific corpus. While I use the same basic formula as Lederer (2015a) and Ahmad (2005), who respectively call it a "keyness" and "weirdness" score, I prefer the term *metaphoricity likelihood* to reflect that these lexemes are not particularly "weird" or necessarily important to the LESM, but happen to be more frequently employed in metaphoric contexts in COCA.

Figure 8. Metaphor likelihood measures for frequent lexemes. Dotted lines indicate the division between score > 1 and score < 1.

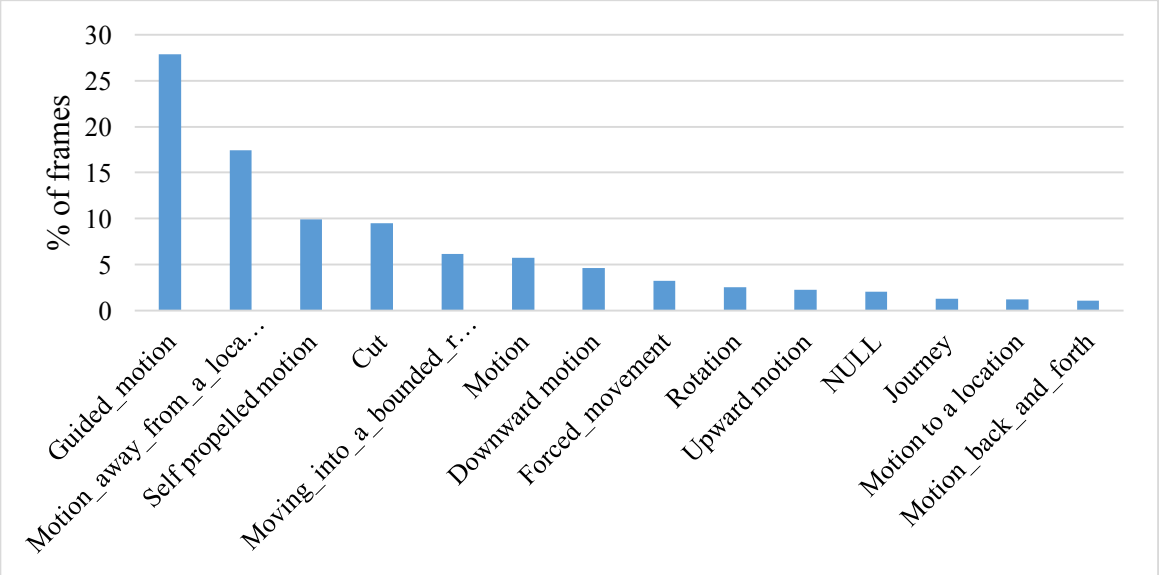


From Figures 7 and 8, we can observe that although *run* is more common overall, *walk* is more likely to be metaphoric, with a likelihood of 0.8359 in comparison to *run*'s 0.7451. However, both are relatively less frequent in the metaphoric data than overall. *Follow* again dominates the data, with a score of 4.6547. *Trail*, which also evokes Guided Motion, is also highly metaphoric with a score of 2.6410. The high frequency of both *follow* and *trail* suggests that the Guided Motion frame is likely to be evoked in metaphoric contexts, leading us to further investigate frequency at the frame-based rather than lexical level.

3.b.iii. *Frames*

In addition to Source domain lemmas, the MetaNet system also identifies the source domains evoked by those lexemes; the most common source domain frames (at least 1% of tokens) are reported in Figure 9. By considering the frames and frame families evoked by the source domain language, we can better understand the *semantics* of the LESM source domain. For example, it may be that a few high-frequency lexemes (such as *follow* and *trail*) here, evoke the same source domain; in that case, the lexical-level information is useful for identifying a common source domain. In contrast, if several middling-frequency lexemes evoke the same source domain, it is not apparent from lexical counts alone how frequent that frame actually is.

Figure 9. Most commonly occurring source frames. NULL indicates no source domain frame was found by the extractor.



As Figure 9 shows, Guided Motion is indeed the most commonly occurring frame in this data set. However, we also see that Motion Away from a Location, which is evoked by *leave* and *turn* is the second most-common with Self Propelled Motion following behind. Self Propelled Motion is evoked by several lemmas: *crawl*, *lurch*, *run*, *stagger*, *swim*, and *walk*. Besides *run* and *walk*, none of those constitute a high-frequency lexeme. Together, however, they evoke a common source frame. From this point we can begin to paint a picture of the ways in which the LESM is linguistically evoked in the COCA data. The entailment GUIDED ACTION IS GUIDED MOTION occurs very often; examples from the search results are provided in (3).

(3) a. *follow* the simple advice of one poet in Poetry by Heart
b. *trailing* by three touchdowns in the third quarter
c. FACE *pursued* all subsequent negotiations

Variations on direction-related metaphors are very frequent as well: collectively, Motion Away from a Location, Cut, Moving into a Bounded Region, Downward Motion, Upward Motion, Journey, and Motion to a Location constitute 42.44% of the data, as shown in (4); one example is provided for each frame.

(4) a. *turning* away from Ruth El Saffar's traditional reading (Motion Away from a Location)
b. *cutting* through those assertions is not easy (Cut)
c. most *entered* the profession during the 1970s (Moving into a Bounded Region)
d. many terraces *fell* into disrepair (Downward Motion)
e. revenue *rose* 8.8 percent (Upward Motion)
f. *traveling* with watercolors (Journey)
g. let us *arrive* faster at solutions (Motion to a Location)

Considering only the subcases of CHANGING STATE IS MOTION ALONG A VERTICAL PATH, Downward Motion and Upward Motion together comprise 6.85% of the data, as further illustrated in (5). We can compare these measures to the non-path ACTION IS MOTION-evoking Self Propelled Motion and Motion, which together are 15.62% of frames (6). In (6a), we can observe the additional path information conveyed by the preposition *through*, which evokes a complex image schematic structure that includes Source-Path-Goal and Containment. While *walk* evokes Self Propelled Motion, which entails translational motion as it incorporates the Trajector Landmark image schema, *through* provides the directional element of Path. Similarly, *move* in (6b) only evokes Motion (which in this case may be self-propelled motion, although it is not clear from context), and the prepositional phrase headed by *beyond* relays the Landmark (*definitions of empathy*) that fills the Source role of the Source Path Goal schema evoked by the preposition.

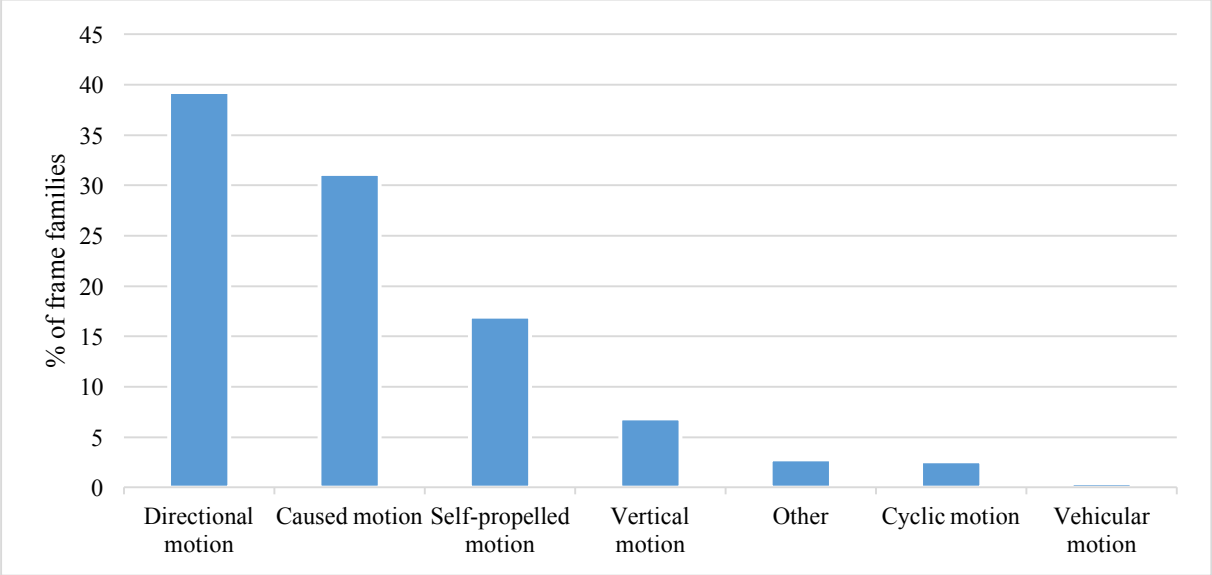
(5) a. *falling* below investment grade costs an issuer 75 basis points
b. a trend that experts believe will continue *to climb*

(6) a. let's just *walk* through the numbers here
b. *moving* beyond the more traditional definitions of empathy

From this we can conclude that, when only considering verb-evoked source domains, instances of the LESM in English are more likely to explicitly evoke directional information via the Trajector-Landmark or Source-Path-Goal image schemas in Motion frames. In fact, the rate of path-related information in LESM usage may be even higher, as this does not take into account path information conveyed by verbal satellites (particles and prepositions) as seen in (6). In contrast, the non-translational motion Motion Back and Forth is very infrequent, at just 1.07%. From this we can infer that verbal instance of the LESM focus on state *change* rather than *stasis*. Translational motion in the LESM evokes CHANGE OF STATE IS CHANGE OF LOCATION via the inference in the source domain that motion along a path changes the location of the moving entity.

This emphasis on directional motion becomes even more apparent when we group the source domain frames into *frame families*, by categorizing frames together via their shared parent frames (Figure 10). Motion Away From A Location, Motion To A Location, and Journey, all inherit from Moving Into A Bounded Region, so they are categorized together as Directional Motion: nearly 40% of the data. Caused Motion (comprising Guided Motion and Forced Movement) is also very common at 31.15%, showing that English speakers view a change of state as having an observable cause (whether internal or external).

Figure 10. Relative frequency of frame families.



3.b.iv. *Constructions*

Frames alone do not tell the whole story, however. As Ellen Dodge and I have demonstrated in our ongoing case study of English poverty metaphors (Stickles, Dodge, and Hong, 2014; Stickles and Dodge, 2016; Dodge, 2016), there is also considerable variation in usage of syntactic patterns. Constructional analysis can reveal more details about the conceptual structure of the metaphor at work by helping to identify the roles that individual frame elements play as well as which are foregrounded or background (Stickles et al., 2016a; Dodge, 2016). This study focused on four metaphoric verbal constructions (Table 1); in all cases, the source domain language is in the verb and the target domain language in the noun.[42] The first two constructions relate the verb and an argument; the latter two relate the verb with a prepositional phrase headed by either *by* or another preposition. We separate out the *by*-headed prepositional phrases to aid in distinguishing agent-demotion passives from active sentences with prepositional complements or adjuncts.

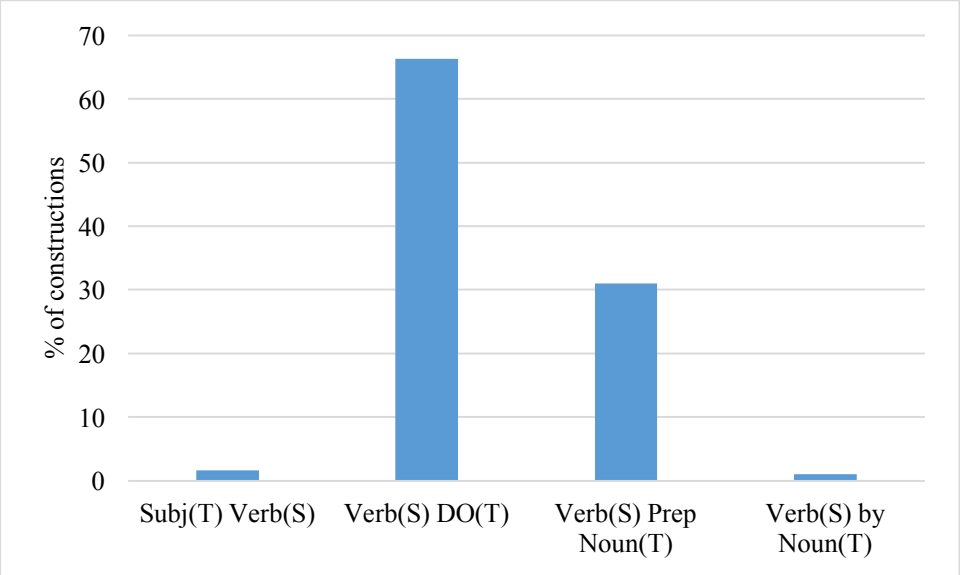
Table 1. Metaphoric verbal constructions.

Construction	**Example**
Subject (Target) Verb (Source)	Netscape *rose* that day from $28 to a close of $58.25
Verb (Source) Direct Object (Target)	*weaving* a quiet, sometimes meandering story of terror
Verb (Source) Preposition Noun (Target)	*jumping* on the "tax relief bandwagon"
Verb (Source) By Noun (Target)	*driven* by the lure of the new technology badge

[42] Sullivan (2007; 2013) argues the tendency for source domain language to occur in verbs and target domain language to occur in arguments is a result of the relationship between conceptually autonomous and dependent elements. In metaphor, the target domain is the autonomous element and the source domain is constrained as it depends on the structure of the target domain. In verbal constructions, the head verb is the dependent element and its argument(s) are autonomous.

Figure 11 illustrates the relative frequency of the four verbal constructions. We can see that two clearly dominate: The Verb (Source) Direct Object (Target) construction at 67.17% of the data and the Verb (Source) Prep Noun (Target) at 31.41%. Further examples of these constructions are provided in (7) and (8).

Figure 11. Relative frequency of metaphoric verbal constructions.



(7) a. *Follow* coverage and blogging of the event on Twitter
b. *Leave* the low life to the low lifers
c. *Following* today's presentation, …
d. *Pursuing* a monastic freedom dream in prison, he gradually learns…
e. *Driving* these projects…is an ever-increasing need for suitable performance space

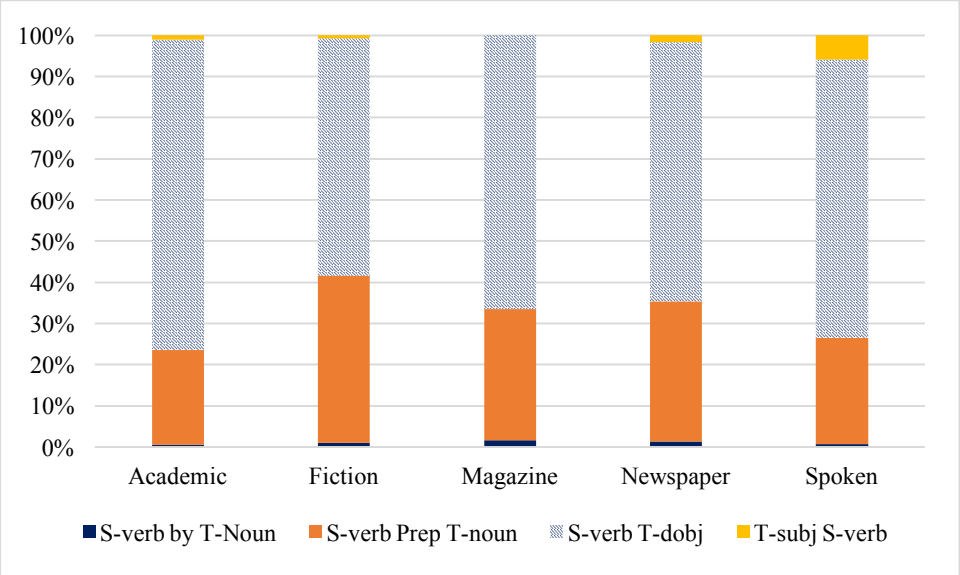
(8) a. *dipping* into that investment account
b. *moving* into film had seemed intimidating
c. *leaping* into the playoffs in a conference so loaded it should be featured in Forbes
d. *walked* away with 9 and a half percent on their money
e. *walking* in election wonderland
f. *returning* from five and a half years of torture

The Verb-Direct Object examples in (7a-c) demonstrate that the syntactic subject is frequently dropped from the sentence; the semantic agent is instead implied by the discourse context. (7a-b) are imperative constructions, which drop the subject; in (7a) the sentence speaks directly to the reader, whereas (7b) the addressee is a third party given in earlier discourse. (7c) is an example of a very common construction *Following X, Y*, in which the subject is null-instantiated in a case of genre-specific argument omission. This construction will be discussed further in 3.b.v. (7d-e) illustrate another common construction similar to that in (7c); instead of a null subject, the cataphoric subject of the head verb is given in the following clause. All these variants of the Verb-Direct Object construction – particular those with argument omission – indicate an emphasis on the *process* image schema of the frame, rather than the entity performing the process. Semantic agents and experiencers have been argued to be highly topical (e.g., Goldberg,

2006), which makes them prime targets of argument omission; furthermore, topics have been argued to constitute a subset of the background (Vallduví and Engdahl, 1996). This leads us to the conclusion that the agent or experiencer in these instances of the LESM is a backgrounded element of the Motion source frame.

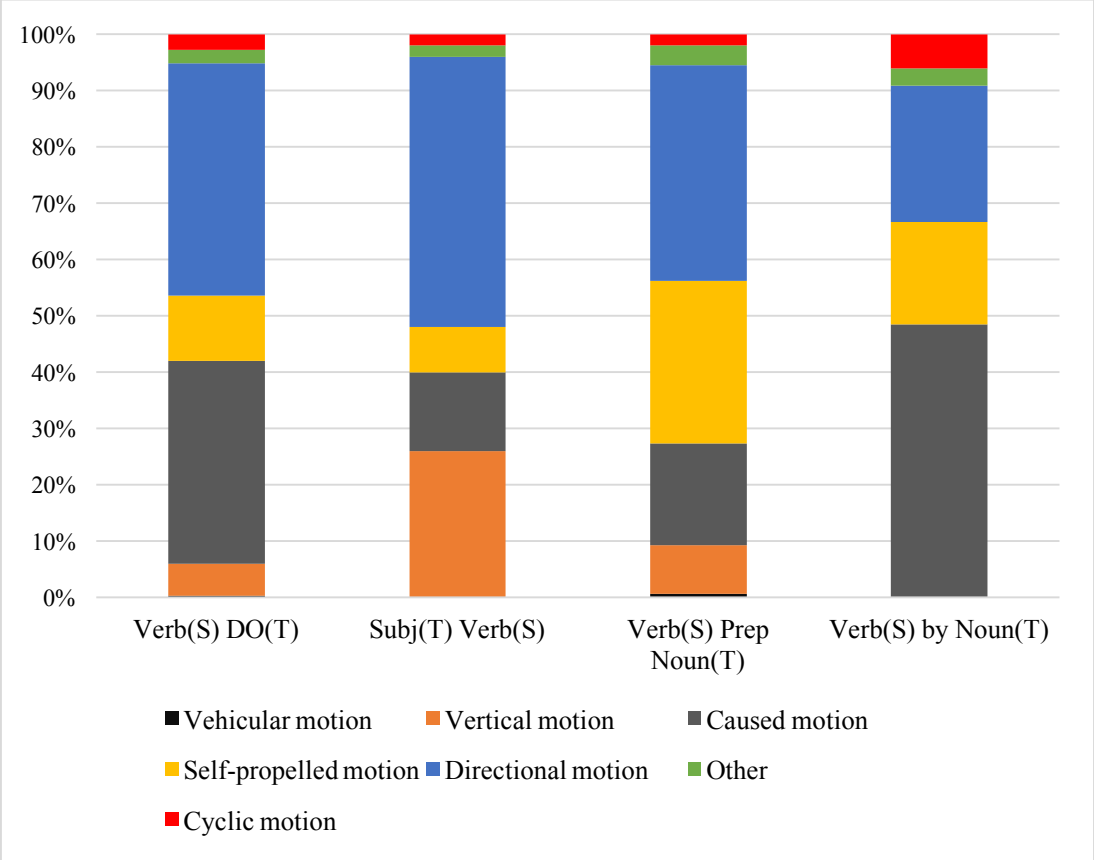
Figure 12 shows the distribution of the four constructions between the genres in COCA. Combining the written genres together, a comparison with the spoken genre reveals there is a significant difference in construction frequency (Fisher's Exact Test, $p < 0.0001$). As we see in Figure 12, the spoken genre has a greater number of Subject Verb construction instances and fewer Verb Prep Noun, whereas the academic genre has the highest rate of the Verb Direct Object construction. The genre-specificity of the argument omission/cataphora of (7c-e) can be seen in this distribution of the constructions between genres; fiction writing is less likely to make use of the construction than more formal writing genres. I will discuss this genre-specificity further in 3.b.v.

Figure 12. Constructional frequency in COCA genres.



Finally, we conclude our investigation of the interaction of syntax and semantics in the LESM COCA data by looking at the types of frames, as organized in frame families, that occur in each construction (Figure 13).

Figure 13. Frame family frequency across constructions.



It is immediately apparent that Caused Motion is much more frequent in the Verb By Noun construction than the Verb Prep Noun construction, showing the importance of separating the two.[43] The English passive, in which the agent/cause is demoted to a by-phrase, is typically considered in functional approaches to reflect a greater prominence of the patient over the agent (e.g., Shibatani, 1985; Givón, 2001; Croft, 2002). However, if we consider the Caused Motion examples of the by-passive in (9), we observe neither agent demotion nor patient promotion due to the cataphoric subjects; in other words, both the agent *and* patient of the sentences in (9) have been reduced in prominence. Instead the verbs in (9) appear to be focused.

(9) a. *driven* by hostility to the Western institutions and ideals of their former European colonizers, they are equally contemptuous of the autocratic regimes
b. *moved* by his memories, Roger reached out
c. *swayed* by these numbers, Congress in 1988 passed an amendment

[43] Future development of the MetaNet extractor will introduce more fine-grained prepositional analysis for this reason.

We also see Caused Motion is frequent in the Verb Direct Object construction, as in (10):

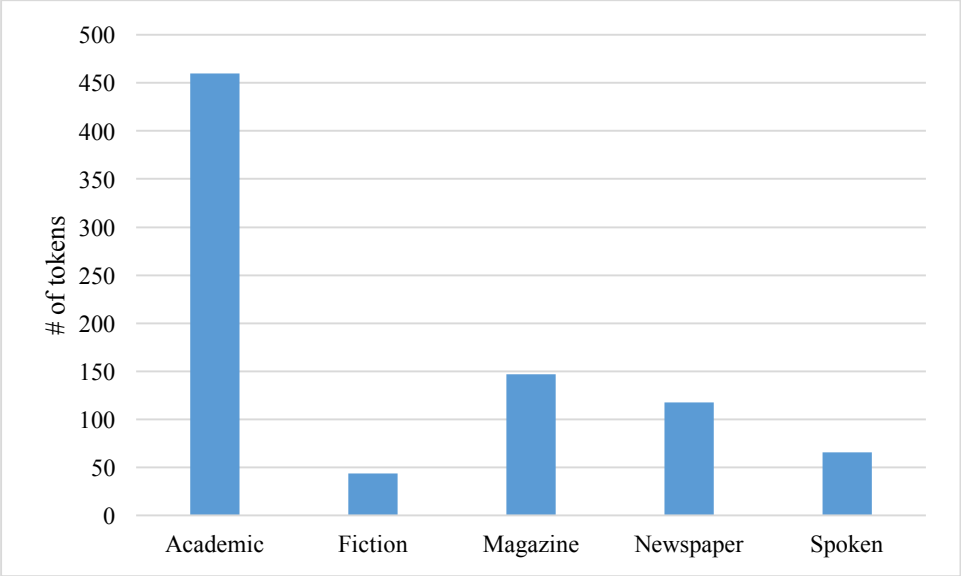
(10) a. *following* a short break for lunch, the committee heard from its two chief investigators
b. *follow* the same protein- and fiber-rich eating habits for lunch and dinner
c. *driving* each other crazy, as sisters will do
d. *driving* this process, I argue, is the struggle for authentic translation

These all follow a similar pattern: the direct object is either the temporal precedent, which comes before the subsequent event in the next clause (10a-b); or the direct object is the cause of the metaphoric change of location (10c-d). The majority, however, pattern after (10a-b); a full 76.56% have the lemma *follow*. Hence, the most common use case of Caused Motion in the LESM Verb Direct Object construction is GUIDED ACTION IS GUIDED MOTION rather than CAUSED CHANGED OF STATE IS CAUSED CHANGE OF LOCATION.

3.b.v. *The case of 'follow'*

It is interesting that the lexeme *follow* heavily dominates the data, at 27.05% of lexical tokens; as a result, the Guided Motion frame is the most commonly occurring frame, and Caused Motion makes up a large portion of the most common construction, Verb Direct Object. As we will see, both the frequency of *follow* and its influence on the data as a whole is largely an effect of genre. Breaking down *follow* by the genres in COCA, we see the following distribution (Figure 14):

Figure 14. Distribution of *follow* between genres.



Clearly the academic genre, which comprises articles from peer-reviewed journals covering the range of the Library of Congress classification system (Davies, 2008), dominates the *follow* data at 55.10% of all *follow* tokens. (In contrast, the academic genre overall is around 20% of the total

COCA data.) Diving deeper into the metaphoric academic *follow* data, we see that all but 17 of the *follow* wordforms in that subcorpus are actually "following", as in the examples[44] below:

(11) a. *Following* the interview, member checking occurred by asking each participant to read his or her transcribed interview.

b. *Following* Oppenheim's (1966) recommendation, several statistical tests were performed…

c. *Following* Dewey's notion that content and teaching method are inseparable, Martorella notes that instruction can not simply put knowledge and skill into peoples' heads.

d. *Following* the examples provided by her cooperating teachers, Lisa planned to take the approach that "students should never have time to misbehave"

The examples in (11) each illustrate a different usage of the Guided Motion frame evoked by *follow*. In (11a), it is used in a spatiotemporal metaphor: sequential events follow one after another. The methodology section of an academic article typically discusses the step-by-step actions the authors took in performing their research; in (11a), they describe doing "member-checking" after performing an interview. Given that the "checking" cannot occur until after the interview is completed, it must *follow* the interview. (11b-c) is highly constructionalized in academic discourse; it is commonly used when making use of another author's scholarship. In (11b) the author is performing statistical tests due to Oppenheim's recommendation; in (11c), Martorella's work is described as being in the same intellectual tradition as Dewey's. Both cases explicitly evoke the guided component of Guided Motion by construing the actions of the author in (11b) and Martorella in (11c) as being guided by the prior work of Oppenheim and Dewey. Their actions are both intellectually and temporally posterior to those of Oppenheim and Dewey. (11d) similarly evokes Guided Motion as it describes Lisa performing an action modeled after that of her teachers, thus evoking the metaphor GUIDED ACTION IS GUIDED MOTION, but does not fall in the same discourse-constructional category as (11b-c) as Lisa's actions are not themselves a scholastic pursuit.

100 of the academic *follow* sentences and 100 written non-academic *follow* sentences were randomly selected and hand-annotated for constructional information. They were annotated as constructional if they made use of the academic *Following X, Y* construction where *X* is a previous academic work and *Y* is an action or theoretical analysis implemented on the basis of that prior work, as exemplified by (11b-c) above. They were annotated as *temporal* if they indicated one event occurring after another in time, as in (11a); and as *other* if they fell into neither category. Results are shown in Figure 15.

[44] I initially wrote here "the **following** examples" – an example of both the pervasiveness of *following* in academic discourse, as well as the priming effects of discourse context.

Figure 15. Distribution of *follow* tokens in academic and non-academic discourse.

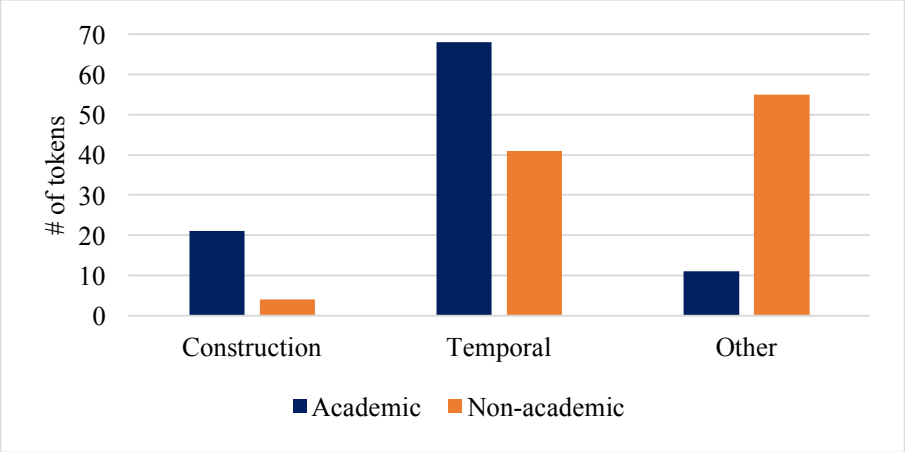


Figure 15 illustrates that the non-academic written corpus had a wider variety of *follow* usage, with over half analyzed in the Other category. Only 4% of the non-academic discourse made use of the *Following X, Y* construction, in comparison to 21% of the academic data. The academic data was predominantly temporal (68%), most likely reflecting the emphasis in academic writing on describing one's research methodology in detail. Many of the non-academic Other usages were in the same vein as the academic construction, as illustrated by examples in (12).

(12) a. *Following* the manufacturer's directions, fuse the iron-on adhesive to the fabric rectangle.

b. *Following* orders, I made a long cast to the rear with one rod and a medium-length cast with the second.

c. *Following* their example, Romans kept regular records, but added personal comments.

d. *Following* a now well-established pattern, Republicans in Congress demanded an independent counsel investigation.

e. *Following* this line of reasoning, Phaethon may be a comet masquerading in its old age as an asteroid.

In the examples above, we see the *Following X, Y* construction as realized in non-academic contexts. Although they lack the specifics of academic research – *X* is not an academic, and *Y* is not an implementation of research methods or theory – they follow the same overall pattern. *X* is an external example to base further action on, such as instructions (12a-b), prior actions by external agents (12c), prior actions by the same agent (12d), or more abstractly, logic (12e) leading to a conclusion. *Y* is the resulting action that is implemented on the basis of *X*. Hence, the *Following X, Y* pattern we observed in the academic genre, and subsequently uncovered in the non-academic genre, fits the profile of a construction as described by Kay and Fillmore (1999).

The construction follows a consistent morpho-syntactic pattern, consisting of the antecedent verb *following* with an object NP, combined with a consequent clause. It has consistent causal

semantics between the antecedent and consequent; the consequent action follows from the antecedent model. The *Following X* is often a non-restrictive relative clause, whereby its subject NP is in the consequent clause as in (11c) and (12b-e). However, in academic writing, the subject is often elided to avoid use of the first person; instead the passive voice is used to eliminate the author from the work, as in (11a-b). In those cases, the syntax of the construction has become further grammaticalized, such that the consequent clause does not realize the subject of the antecedent.

The subject of the antecedent is instead frame instantiated by the academic discourse context, in which null subjects are understood to be the article's author(s); this constitutes a case of constructional null instantiation (Fillmore, 1986; Goldberg, 2006; Ruppenhofer and Michaelis, 2010). Notably, this construction follows the pattern described by Ruppenhofer and Michaelis (2010) for other types of genre-based constructional null instantiation: the omitted argument is a globally prominent referent (the author); argument omission is canonical and lexeme-specific (i.e. it is licensed by *following*); and has grammatical restrictions (the omitted argument must be the subject). In sum, we have observed the process of genre-based constructionalization[45] in the academic data, whereby the *Following X, Y* construction has developed a genre-specific special case that specifies the role-fillers of the frames evoked by the noun phrases in *X* and *Y*, and furthermore allows for zero anaphora in the subject due to genre-based frame instantiation.

4. *Conclusions*

This chapter has presented an exploratory case study of the Location Event Structure Metaphor. Whereas Chapter 2 lays out much of the theoretical foundation and formalization of the LESM, this chapter explores how the LESM is actually used in a balanced corpus of contemporary American English, and demonstrates via a case study of *follow* how the MetaNet system can be used to discover new patterns and generalizations. We have seen that certain lexemes are highly likely to be used in metaphoric contexts, whereas others are balanced or less likely to be used metaphorically. These variations in usage patterns at the lexical level show how use frequency is influenced by both metaphoricity and genre. Bouveret and Sweetser (2009) have argued that this type of lexical choice, in which certain lexemes with similar semantics are used in different metaphoric contexts, reflects which frames those lexemes foreground or background; David (2016) extends this analysis in a FrameNet and MetaNet corpus approach. Here, I claim that the high frequency of Directional Motion-evoking lexemes indicates that the LESM foregrounds path and backgrounds manner.

We have also seen variation in use frequency at the semantic level, both in considering individual frames as well as sub-networks of frame families. By looking at frame-level semantics, rather than lexical-level variation, we gain a deeper understanding of which specific entailments of the LESM are more frequently evoked; here, we see that GUIDED ACTION IS GUIDED MOTION occurs quite often, especially in academic writing. This type of inquiry can also be performed on the target domain side; for example, Stickles et al. (2014) show that in the British National Corpus, Disease metaphors for Poverty are more likely to emphasize Disease

[45] Constructionalization, as described by Traugott and Trausdale (2013), is the process by which a new form-meaning pairing develops.

Treatment than the Disease itself, suggesting that written British English is more focused on addressing poverty (ADDRESSING POVERTY IS TREATING A DISEASE) than simply describing it (POVERTY IS A DISEASE).

Through constructional use frequency, we begin to understand how these variations in lexical and frame frequency interact with argument structure. Both the overall high frequency of the Verb Direct Object and highly co-occurring Caused Motion Verb *by* passives indicates that the agentive entities involved in the metaphor are topics or otherwise backgrounded elements. Further evidence for the prominence of processes over entities in the LESM will be seen in Chapters 4 and 5.

I have also argued that when manner verbs are used to evoke translational motion in the LESM, the manner itself is not a foregrounded element of the metaphoric mapping. Sullivan (2006) has argued that, similar to Lakoff's (1990) Invariance Principle for image schematic structure in metaphoric mapping, mappings between frame elements are similarly is constrained to systematic relationships in which the semantics of the source and target elements align. My analysis of the relationship between source and target domains in the LESM follows Lakoff (1990) and Sullivan (2006) in that I argue the manner of motion is either backgrounded or unmapped information. For example, *drive up prices* does not evoke the manner of the motion frame (Vehicular Motion); only the frame element of external causation, inherited from Caused Motion, is relevant to the metaphoric mapping. We can see evidence of the centrality of Path in the LESM – even when manner verbs are used – in the high co-occurrence of prepositional constructions with manner-oriented frames (i.e., the Self Propelled Motion family). Simply categorizing all the source domain verbs as manner or path reveals a high rate of manner verbs occurring in the Verb Prep Noun construction (47.01% of manner verbs) in comparison to path verbs (23.47%); hence, even when using a manner verb, English speakers have a high rate of including path in a verbal satellite.

There are some limitations to the MetaNet extraction system. Chief among them is the reliance on the collocation of both source and target domain language. Speakers will frequently drop target-domain language, especially in ongoing discourse where the target domain is an established topic (Lederer, 2015a; 2015b). This search method will miss those instances of metaphoric language. The system is also not as precise as hand-annotation; errors in the dependency parser will lead to errors in constructional analysis. Mistakes or gaps in the MetaNet repository network will also lead to flawed analyses. For example, one issue with the current analysis is that it returns the Wilderness frame for the lexeme *trail*, even though here it is used as a verb in the Guided Motion motion sense rather than a noun in the Wilderness sense of a *hiking trail*. This can be dealt with in a post-hoc check of the automatic extraction data; given that we know we are focusing on verbs, we can safely re-categorize the *trail* lemmas as evoking Guided Motion rather than Wilderness. This is confirmed by the actual data, such as in (13).

(13) a. *trailing* him in Georgia is Sen. John McCain of Arizona
b. *trailing* by 26-17 with just under nine minutes left in the final quarter
c. *trailing* it: PricewaterhouseCoopers and the Federal Bureau of Investigation

In (13), we see two types of metaphoric races: political campaigns (13a) and sports competitions (13b). In (13a-b), a specific subcase of the LESM, COMPETITIONS ARE FOOT RACES, makes use of the relative distance between the runners in the Foot Race frame to conceptualize the "distance" between competitors in elections or American football games. The Foot Race frame makes use of Guided Motion in that the second runner follows the first. (13c) evokes Guided Motion more directly, in that an FBI investigation of PricewaterhouseCoopers is construed as physically following a guilty party. Therefore we can conclude that the original analysis was incorrect to associate the lexeme *trail* with Wilderness, as the appropriate frame in this context is Guided Motion.

Setting aside such errors in the automatic metaphor identification system and limitations of the collocation search, flaws remain in this approach to metaphor research. Generally, a major issue in conceptual metaphor theory is in developing empirical evidence for the psychological reality of metaphors as conceptual structures. One benefit to studying metaphoric *gesture*, as opposed to metaphoric *language*, is that it allows us to side-step many of the issues inherent in studying linguistically-realized conceptual metaphors, as I have done in Chapters 2 and 3. Linguistic evidence only demonstrates that metaphors are pervasive in language; conceptual metaphors are claimed to be domain-general mental representations. Therefore, we should see evidence for them in non-linguistic cognition as well. Language data alone cannot demonstrate domain-generality; to do so, it is necessary to study non-linguistic – i.e., gestural – expression of metaphor.

As discussed at length here and elsewhere (e.g., Lakoff, 1990; Sullivan, 2006) the source domain of a metaphor only partially maps onto the target domain. However, one may argue a given lexeme still evokes the semantics of a particular frame; how, then, can we tell that only the "relevant" structure of that frame is in fact activated in the speaker's mental representation when using it in a metaphoric context? In this chapter, I have argued from usage-based evidence that the LESM tends to profile directional motion. One might simply counter that argument by pointing to the high frequency of manner verbs in American English; why would English speakers use them so often if they actually intended to focus on the path of motion? By studying gestures evoking the source domain of the LESM, we can actually see the mental conceptualization of motion at work in the manual modality, and avoid the complexities of lexical semantics entirely. One benefit of gesture is that it provides a "back door to cognition" (McNeill, 1992); it can reveal the underlying conceptual structure of the metaphor at work. Thus in Part Two of the dissertation we turn to studying the LESM in metaphoric gesture.

Part Two: The Realization of the LESM in Metaphoric Gesture

Chapter 4: A corpus approach to multimodal metaphoric utterances

1.a. *Introduction*

In this chapter and the next, I turn my focus from the realization of metaphor in linguistic utterances to metaphor in multimodal utterances. In particular, these chapters investigate a class of gestures called *referential metaphoric gestures* (see Cienki and Müller, 2008 for a review); I describe them in detail in section 5.b. of Chapter 1. These gestures, exemplified by Figure 1, do not represent concrete, physical entities or actions. Instead, they represent the source domain of a conceptual metaphor. Hence, while the *form* of the gesture is physical, is *meaning* is abstract. In Figure 1, we see a speaker producing a *metaphoric* gesture in his description of someone's attention metaphorically *drifting*. His gesture iconically depicts the movement of an entity away from a location in front of the speaker. However, the content of his speech makes it clear that he is not talking about the literal motion of a physical entity; rather the topic of his speech is 'failure to pay attention'. Yet the speaker in Figure 1 is not producing a gesture that directly evokes the notion of 'paying attention' or 'failing to pay attention', although that is the topic of his speech. Instead, he is metaphorically construing 'failing to pay attention' as a person's attention "drifting" away from the point of focus. Thus, his gesture evokes the notion of metaphoric 'drifting'.

Figure 1. Example of a metaphoric gesture.

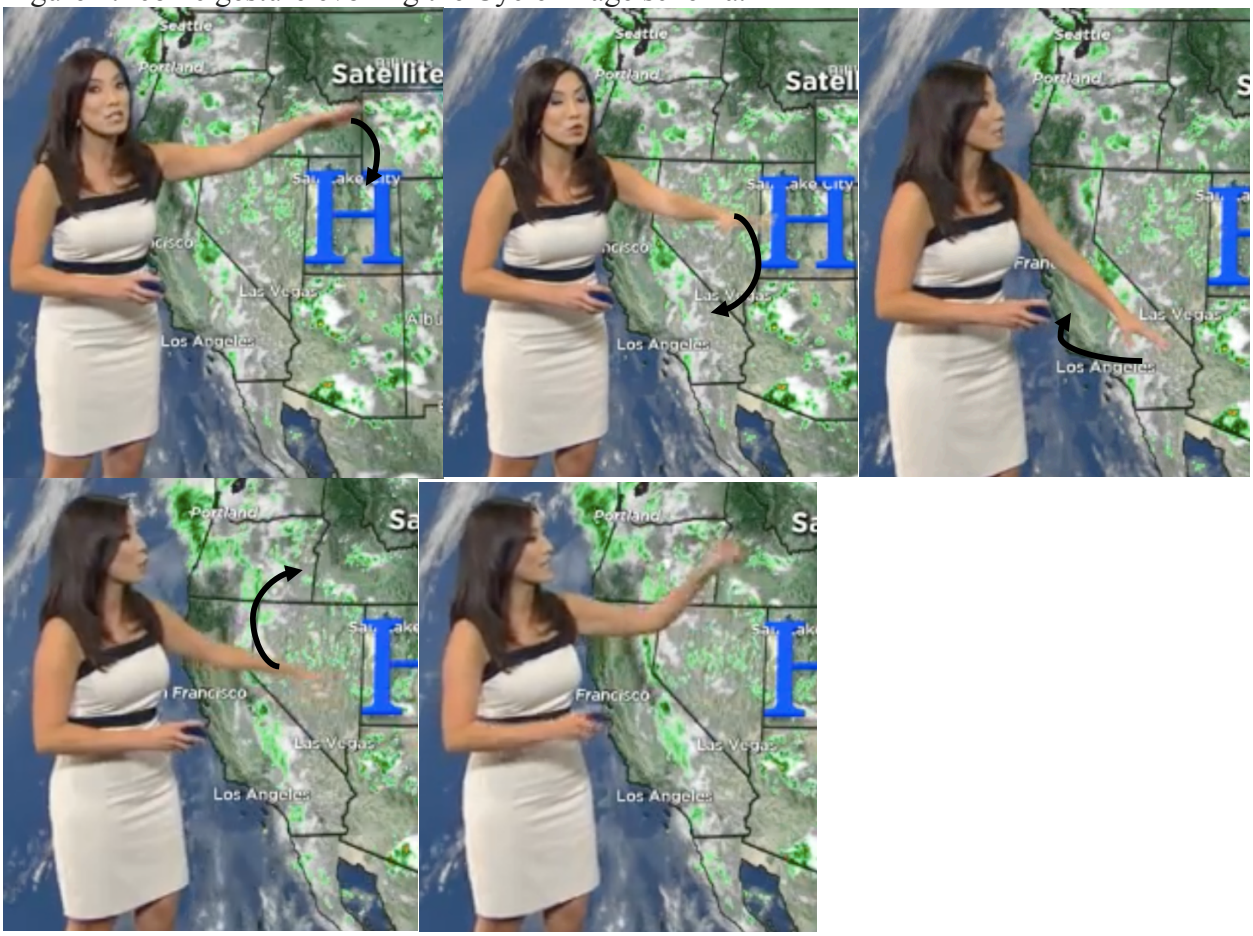


his [attention span] can sometimes drift a bit [m.drift.5.31.2015c]

In the following studies, I will demonstrate that *image schemas* are the most appropriate level of structure in analyzing the form and meaning of metaphoric gestures. An image schema (discussed in detail in section 3 of Chapter 1) is an experientially-grounded conceptual representations of bodily is experience; Johnson describes it as "a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience" (1987: xiv). Cienki (2005) argues that there is a likely connection between the imagery in gesture and the structure of image schemas. First, gesture is analog and holistic (McNeill, 1992), as are image schemas, which exist "in a continuous, analog fashion in our understanding" (Johnson, 1987: 23 as cited in Cienki, 2005: 422). He further observes that both

image schemas and gestures can be viewpointed as dynamic or static; for example, the Container image schema can be experienced as a Container entity or the process of Containing. The motions and holds of referential gestures can similarly represent either actions or entities. For example, the literal iconic gesture in Figure 2 traces a circular shape. In her speech, the meteorologist describes the motion of a wind pattern with the verb *rotate*, which evokes the Cycle image schema. Because her focus is on the *process* of motion, this evokes the dynamic variant of the Cycle image schema. Her gesture also represents the Cycle image schema as it represents the motion of the wind.

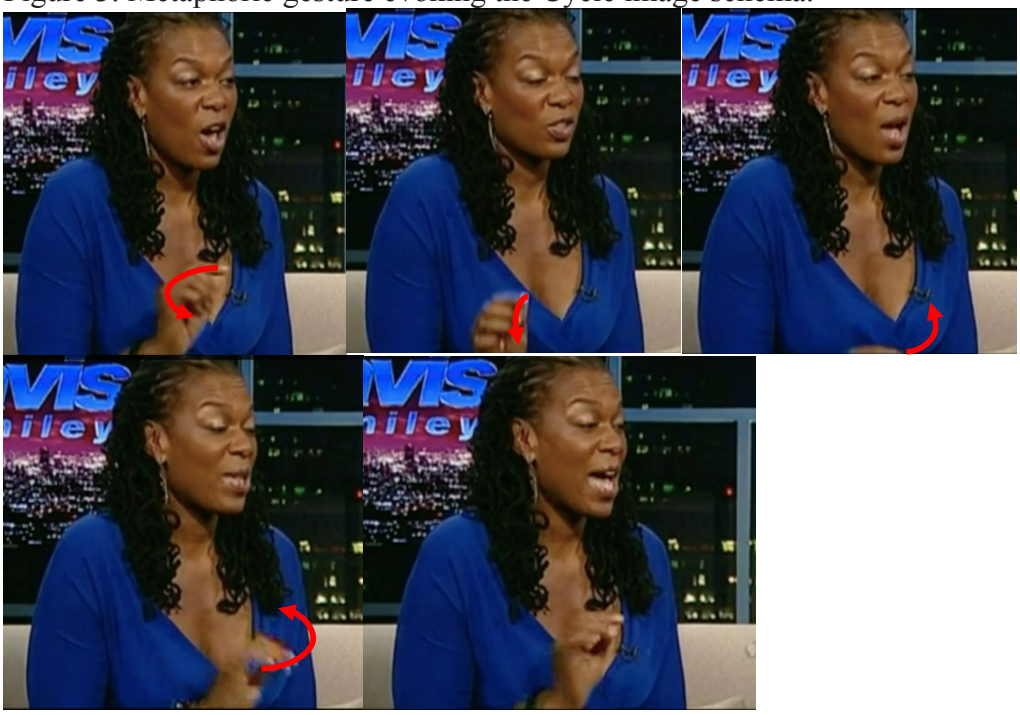
Figure 2. Iconic gesture evoking the Cycle image schema.



it [continues to rotate clockwise] [m.rotate.5.28.2016b]

The speaker in Figure 3 also makes use of the Cycle image schema in both speech and gesture.
Figure 3. Metaphoric gesture evoking the Cycle image schema. The verb *revolve* evokes the Cycle schema, as it refers to circular motion. However, she uses the verb metaphorically; the research findings are not literal entities in physical rotation, but metaphorically 'revolve' around the content *topic* of Black men and pain medication. Like the speaker in Figure 2, her gesture traces a circular shape and evokes a dynamic perspective on the Cycle image schema. However, while the *form* of her gesture is a physically rotating entity – her hand tracing the circle – the *meaning* of her gesture is not literal rotation. By pairing this gesture with metaphoric speech, the speaker is using this gesture to evoke *metaphoric* cyclic motion.

Figure 3. Metaphoric gesture evoking the Cycle image schema.



one of the really interesting findings [revolves around] Black men and pain medication
[m.revolve.5.5.2015a]

As demonstrated in Cienki (2005), image schemas in metaphoric gestures (or in his terms "abstract" gestures, following McNeill [1992]), like literal ("concrete") gestures, can be readily recognized. His experimental work shows that when provided a list of potential image schemas, study participants will agree on which image schema most characterizes a referential gesture. This suggests that, while gesturing is an idiosyncratic process (McNeill 1992), referential gestures reflect an underlying structure: an image schematic representation in the mind of the gesturer which is then accessible and interpretable by the addressee. Similarly, Calbris' (2003) semiotic analysis of cutting gestures discusses the variety of image schemas, such as Separation and Stoppage, evoked by referential gestures in both iconic and metaphoric contexts. She further breaks down four cutting gesture variants in terms of hand orientation, palm orientation, hand movement, finger movement, and movement repetition; different variations comprise combinations of these features, the composite of which evokes a particular mimesis (e.g., the prototypical axe blow) or schema(s). These in turn evoke particular semantic elements for the addressee. While Cienki (2005) demonstrates the accessibility, and Calbris (2003) the relation between form and meaning of these image schemas, their systematicity and structure require further investigation.

Cienki (2005) focused on the image schematic nature of abstract gesture in the absence of speech content, as his focus was on the shared visual accessibility of the gesture. However, this leaves open the question of the shared systematicity of conceptual structures between speech and gesture. Given the claim that image schemas are domain-general non-linguistic mental representations (Johnson, 1987; Lakoff, 1987), and furthermore the claim that gesture and speech are related at a conceptual level (e.g., McNeill, 1992; Kita and Özyürek, 2003) co-expressive

metaphoric gestures produced in conjunction with metaphoric speech should reflect the same image schemas which are at work in the metaphor-evoking speech. Co-expressive gestures represent the same information as the speech they co-occur with. Figures 1 and 3 are examples of co-expressive metaphoric gestures as they both represent the same metaphors as expressed in their accompanying speech. In Figure 1, the speaker uses the word *drifting* metaphorically; his gesture also represents a 'drifting' motion. In Figure 3, the speaker uses the phrase *revolves around* metaphorically; her gesture also represents a circular 'rotating' motion. Therefore, given similar metaphoric language, speakers should produce similar structures in gesture; the realization of the same image schema in metaphoric speech should co-occur with similar realizations in metaphoric gesture. To be clear, I am not claiming that all metaphoric gestures will always reflect the image schemas evoked in metaphoric co-speech; rather, that there should be observable systematicity in a subset of metaphoric gestures produced with co-expressive co-speech.

To address these questions, we should first consider the relationship between metaphoric language and literal language on the one hand and metaphoric and literal gesture on the other. When language is used metaphorically, it evokes the image schematic structure of a frame relevant to the metaphoric mapping (Lakoff's (1990) Invariance Principle). As discussed in Chapter 3, when a verb such as *drive* is used in the context of the Location Event Structure Metaphor, as in *drive up prices*, it evokes the relevant image schemas of verticality and forced movement, but not the manner itself of the vehicular motion frame. In contrast, when the same word is used literally, as in *drive up the street*, it may evoke the manner of vehicular motion, as well as those structures evoked by the metaphoric usage. The "up" in *drive up prices* evokes a prototypical representation of Verticality, as it conceptualizes an increase in price as upward movement. The "up" in *drive up the street* also evokes Verticality, but transforms it onto a horizontal axis that extends outward from the viewpoint of the speaker. Extending this to gesture, we would expect that the image schemas in a metaphoric gesture accompanying a metaphoric utterance should be similar to those accompanying the same language used literally, but only reflect those elements of the source frame relevant to the metaphoric mapping. In other words, we should expect that metaphoric gestures and literal gestures evoking the same frames should share image schematic structure.

In contrast, co-expressive literal gestures may represent image schemas present in the source frame which are *not* mapped in metaphoric contexts. For example, when the verb *fly* is used literally, the speaker may be referring one of several elements in the semantics of Flight: the *manner* of motion, flapping wings; the *path* of the entity in motion; the *speed* of the motion; and so on. Therefore, a gesture accompanying literal 'flying' could represent any or all of the above: flapping the arms to represent bird wings; tracing a line to represent the path of motion; varying the speed of the arm movement to represent the speed of flight. However, in metaphoric contexts the *manner* of flight – wings flapping – is not relevant and does not map from the source to the target domain. Rather, using the semantics of Flight in a metaphoric context refers to rate and does not evoke manner; compare metaphoric *the day flew by* and literal *the bird flew by*. In *the day flew by*, the rate of the passage of time is construed as the rate of a motion in entity. It is *not* construed as an entity with flapping wings; the only relevant aspect of Flight that applies to Time is the rate of motion. A "flapping arms" gesture would be infelicitous with the metaphoric usage, as it represents the manner of flight. In contrast, the literal notion of Flight in *the bird flew by* can

evoke not only rate of motion, but also the manner of flight. Therefore, a flapping arms gesture would be licit in conjunction with the literal usage. Conversely, a straight *path* gesture representing an entity moving along a path would be acceptable in either context, as it could either be time (conceptualized as an entity) moving, or a bird moving. The speed of the gesture would then represent either the speed of time passing in the metaphoric usage, or the speed of the bird flying in the literal one. For more on the semantics of manner and path and metaphoric motion in the Location Event Structure Metaphor, see section 4 of Chapter 1 and Chapters 2 and 3.

1.b. *Current study*

This chapter approaches the above issues by developing and analyzing a corpus of co-expressive metaphoric gestures like those in Figures 1 and 3 and a parallel corpus of co-expressive literal gestures as in Figure 2. Both corpora are comprised of short utterances in which the speaker produces an American English manner or path verb and simultaneously produces a referential translational motion gesture, co-timed with the speech utterance. Manner verbs, such as *fly, rotate,* and *drive,* primarily express the manner of motion; they do not specify the path of motion. Conversely, path verbs, such as *fall, rise,* and *enter*, express the direction of the motion but not the manner. For a more detailed discussion of the semantics and realization of manner and path, see section 4 in Chapter 1. By using linguistically parallel corpora, this enables us to directly compare literal and metaphoric utterances, controlling for variations in speech. Furthermore, by using the same set of verbs for both the literal and metaphoric data, we ensure that the same core frame elements are at work in both corpora. This should lead to metaphoric gestures representing a subset of image schemas used in literal contexts.

2.a. *Study design and data collection*

Development of the parallel corpus followed these basic steps:

1. Identify search terms
2. Search closed-captioned databases for videos containing utterances using the search terms
3. Identify data containing the co-occurrence of a search term and co-expressive gesture
4. Annotation

I now discuss steps 1-3; step 4 (annotation) will be described in further detail in section 2b.

1. *Identify search terms.* First, in order to ensure Location Event Structure Metaphor-evoking data, data collection focused on the verbal portions of utterances. A list of potential verbs to search for was compiled by combining the verbs listed in manner- and path- related sections of Levin's (1993) work on English verb classes[46] and FrameNet frames[47] (Ruppenhofer et al., 2010). This produced a total of 359 potential search terms. High-frequency verbs were initially selected based on English SUBTLEX (Brysbaert and New, 2009) frequency scores;

[46] Sections 47.3, 47.7, 51.1, 51.2, 51.3.1, 51.3.2, 51.4, 51.4.2, 51.5, and 51.6 of Levin (1993)
[47] Motion, Moving_in_place, Fluidic_motion, Mass_motion, Self_motion, Fleeing, Travel, Intentional_traversing, and Motion_directional FrameNet frames (Ruppenhofer et al., 2010).

SUBTLEX was chosen as it reflects spoken English use frequency. This resulted in an initial list of 17 manner and 17 path verbs (Table 1, column A). Later, this search list was revised by extracting metaphoric motion language from the English SUBTLEX corpus using the MetaNet metaphor extraction engine (Table 1, column B) and retaining high-frequency verbs. Results in Table 1 are sorted in descending frequency; generic motion verbs ("go" and "move") are excluded. Whereas the initial results retained the same number of manner and path verbs, the revised results retained more manner (21) than path verbs (14), more accurately reflecting English use frequencies.

Table 1. High frequency motion verbs, sorted in descending frequency.

A. Initial SUBTLEX results	**B. Revised SUBTLEX results**
Manner Verbs	
run	run
cut	shake
walk	drive
drive	walk
follow	roll
charge	jump
dance	slide
fly	swing
jump	rock
roll	speed
rock	leap
wind	float
shake	sail
speed	drift
rush	rotate
tumble	creep
crawl	sway
	stumble
	revolve
	track
	crawl
Path Verbs	
come	come
leave	fall
drop	rise
fall	enter
return	drop
cross	cross
escape	arrive
enter	flee
rise	escape
desert	climb

arrive	plunge
climb	dip
advance	topple
exit	tumble
angle	
plunge	
drop	

2. *Search databases.* This set of high-frequency lexemes was used to search through the closed captions of two video databases. Data was collected from October 13, 2014, to June 26, 2016 by the author and undergraduate research assistants trained in metaphor and gesture analysis[48]. Initial searches used the Television News Archive until September 2015 and then focused on the Little Red Hen TV News Archive from May 2015 on. Data collection focused on metaphoric data until April 2016, when the focus switched to literal data. As the databases allow searches to be limited by date range, researchers maintained a list of search terms and date ranges searched for each term, as well as the number of pages of results reviewed. For each search, researchers first read the closed captions for entry in a page of results. Each result was initially evaluated on the following criteria: (a) Is the search lexeme a verb? (b) Is the usage metaphoric? (c) Is the utterance accompanied by a visible gesture? If the result met all three criteria, it was retained for review in step 3. A similar process was followed to collect literal data, except that criteria (b) asked if the usage was literal, not metaphoric. Literal data searches were over a subset of the pages initially searched for metaphoric data, to ensure they were taken from the same pool of potential data. In total, 4,970 video clips in the Red Hen database were evaluated for metaphoric data and 3,330 of those were evaluated for literal data.

3. *Identify co-expressive data*. The next step was to identify gestures co-expressive with the the verbal lexical affiliate. The lexical affiliate is the word(s) whose meaning matches that of the gesture (Schegloff, 1984); hence, this search focused on gestures with the same meaning as that of the verb. Gestures were considered "co-expressive" if the form of the gesture: (a) was not emblematic (e.g., shrugging, thumbs-up, the "ok" sign); (b) included motion by a body part (typically the hands, but could also be the head or torso; legs were very rarely visible in videos); (c) was not an interactive/control conversation-regulating gesture (see Bavelas et al. 1992; Wehling, 2017, for a discussion of such gestures). As McNeill (1992) and Wehling (2017) explain, conversation-regulating gestures are metaphoric in that they either 'present' speech content into the conversation by use of the metaphor COMMUNICATION IS OBJECT TRANSFER, or metaphorically move conversational participants themselves in and out of the conversational 'space'. However, they are not speech content gestures in that they are about the conversation itself, not the content of the conversation. The focus of this study is those gestures which represent the literal or metaphoric motion referred to in the co-speech. Ambiguously-categorized gestures were discussed by the research team and resolved.

Results of this lengthy search process were 130 metaphoric gesture data points and 49 literal data points. While these results appear to be quite small, they reflect two factors: (a) the relative

[48] Special thanks to my undergraduate research assistants who performed data collection and annotation: Kayla Briones, Nathan Cahn, Shannon Chang, Lauren Cho, Noah Hermalin, Ann Kim, Elizabeth Long, Natalie Orsi, and Rachel Pinkerton.

frequency of metaphoric use in television; (b) the visibility of gestures in televised sources. For comparison, Winter et al. (2013) searched the 2009-2013 TV News database for expressions containing the word "number" and spatial adjectives (e.g. "high", "tiny"); they found 552 videos with their search phrases, but retained only 27 with numerical-related discourse and clearly visible gestures. Their low rate of return reflects the challenges in identifying metaphoric gestures in television databases. Tong and Cienki (2015, 2016) used another approach to collecting a metaphoric gesture corpus, by searching for object manipulation motion verbs in the Little Red Hen database and collecting a predetermined number of data points. Results were then evaluated for metaphoricity using the MIPVU metaphor identification procedure (Steen et al., 2010). Less than 10% of results both contained a referential gesture and a metaphoric utterance.

2.b. *Annotation*

First, I will discuss my approach to gesture annotation and how it was developed; then I describe the process of speech transcription. Next I discuss the formal annotation process in ELAN. Finally, I describe a set of post-hoc additional annotations to consolidate the fine-grained speech and gesture annotations into broader categories.

2.b.i. *Gesture annotation*

A major challenge in gesture annotation is developing a codified process for systematically describing the form of the gesture due to the holistic nature of gesture (McNeill 1992); as Duncan (2008) puts it, "a speech transcription can approach a state of completion. Gesture annotations to it (likely) never do." In developing this annotation scheme I consulted several different gesture coding procedures[49], in particular the ANR Multimodality Research Project annotation and ELAN[50] (Brugman and Russel, 2004) transcription procedures (Colletta, Jean-Marc et al., 2009) and Duncan's (2008) discussion of the McNeill lab annotation procedure. An important caveat to gesture annotation is that transcription of the gesture form will use coding conventions "…that captures dimensions relevant to the target of the particular analysis" (Duncan 2008). In other words, the amount of detail and formalization included in the annotation will be determined by the goals of the project, much as the degree of phonetic information in a speech transcription can vary from a basic text with no attention paid to the actual articulated speech, to a broad phonemic treatment, to a detailed narrow phonetic transcription with varying amount of detail as to turn-taking, non-linguistic sounds, etc. (Ladefoged and Johnson, 2010). Annotation was done in two stages: a first pass in which the researcher provided a brief description recorded at the time of data collection, and a second pass in ELAN performed separately on a later date. The first pass was used to develop a formal coding scheme to be used in the second pass ELAN annotations. Given the overall goal of eventually developing a framework that can be incorporated into an ECG representation, it was important to codify the

[49] My gratitude to Kashmiri Stec for compiling a bibliography of multimodal annotation methodologies.

[50] ELAN is an annotation tool for video and audio resources, developed by the Max Planck Institute for Psycholinguistics, The Language Archive, The Netherlands and widely used in multimodal research. It is downloadable at http://tla.mpi.nl/tools/tla-tools/elan.

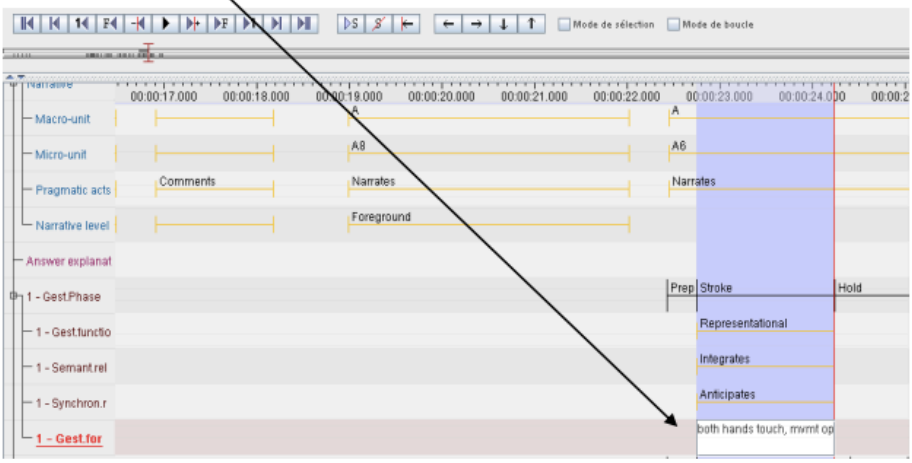
gesture form in as much detail as could be represented in discretized image schematically-informed categories rather than prose descriptions. This stands in contrast to many coding schemes, which are either detailed but prosaic in their approach to describing form (Figure 4); or require discrete but broad categories, such as whether the gesture's trajectory was straight or arced (Kita and Özyürek, 2003), or movement was iterated or not (Parrill et al., 2013).

Figure 4. The ANR annotation procedure (Colleta et al., 2009:30).

ANR-05-BLANC-0178-01 et -02

Stage 5 : < Gest.form >

Double click on the place where you wish to annotate, then type the information you wish to input in the block:



We give a brief description of the annotated movement sticking to its most salient points:
- body part of movement: head, chest, shoulders, 2 hands, left hand, right hand, index, eyebrows, mouth….
- if there is a trajectory: direction of the movement (towards the top, bottom, left, right, frontal, behind…..)
- if there is a hand shape: the form of the hand (flat, cutting, closed in a punch-like form, curved, palm up, palm down, fingers pinched, fingers in a circle….)
- the movement itself: head nod, beat, circular gesture, rapid or not, repeated or not…

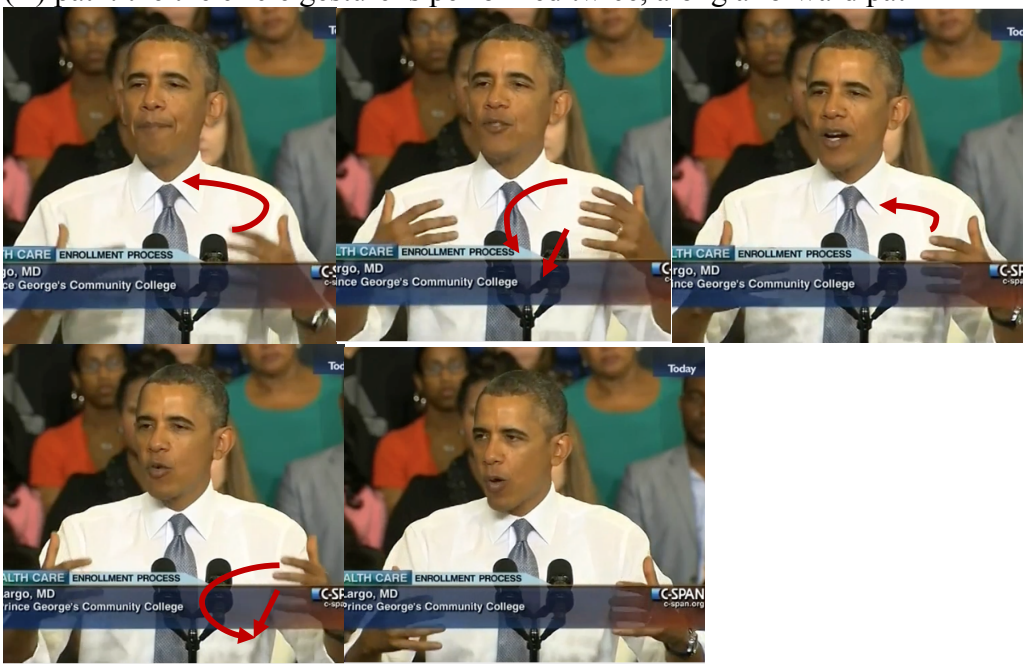
Thus, the first pass annotation, while informed by the goals of the study, was also open-ended in order to allow the researcher to note what they felt were the most salient elements of the gesture, as described in Figure 4. Results of these annotations were then reviewed by the research team and discussed; this allowed us to determine what elements of the gesture form needed to be annotated to convey enough information as to the gesture's content, and how they could be codified into categories. Effectively, annotation took a "bottom-up" approach, in which the coding scheme emerged from the collective results of the early descriptive analyses. Early annotations focused on handshape, palm orientation, and movement. For example, early movement annotations included such descriptions as "arc outwards", "straight away", or "circular". This demonstrates the saliency of both the "shape" of the movement (arc, straight, circular, etc.) and its trajectory (outwards, away, not present, etc.). From the movement descriptions we identified several types of movement shapes, which I will henceforth refer to as "manner". These can be contrasted with the path of the gesture, which if present was almost always described separately by the annotator along six directions within two basic planes, horizontal and vertical. Furthermore, we noted that some gestures could be produced "in place" (Figure 5A), which I term *static*, or along a path (Figure 5B).

Figure 5. Contrast between (A) static (no path) and (B) translational path gestures.
(A) static: the circle gesture is performed three times "in place"



she [didn't come right out] and say it [p.come.12.11.2014c]

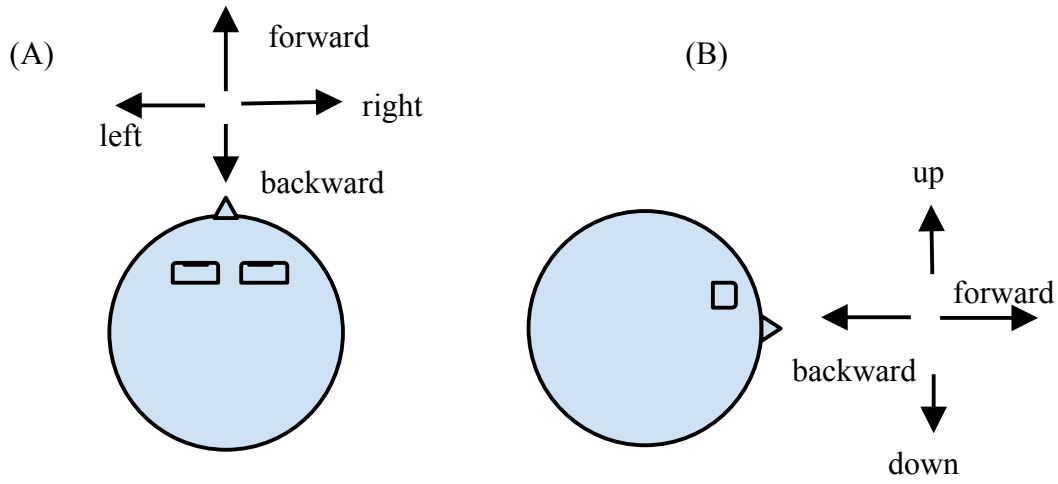
(B) path: the the circle gesture is performed twice, along a forward path



[*that happens whenever you roll] out a new program* [m.roll.10.20.2014b]

After determining these categories, I developed an annotation practices handbook with illustrative examples. Figure 6 provides an illustration of the path categories.

Figure 6. Basic path direction categories from (A) top view and (B) side view.



While speakers often produced simple paths, with a trajectory on a single path category axis, they also frequently produced complex paths, along multiple axes or repeating a motion. They also sometimes produced gestures without a true translational path, as in a circle traced in space; in that case, the hand concludes where it began. To differentiate these types of gestures, paths are *combinatorial.* Multiple axes can be *concatenated* ("forward-up", "down-right-backward", and so on), *repeated* ("forward x2"), and produced *in place* ("static"). Table 2 provides a summary of the manner descriptions available in the handbook.[51] Figures 7-14 provide examples of each manner category, with path descriptions illustrating the various combinatorial possibilities.

Table 2. Gesture manner form categories.

Manner category	*Description*
Straight	a movement with a straight path; no curves
Arc	a movement tracing part of a curve
Circle	a complete curve all the way around
Rotate	the hand rotates at the wrist
Chop	a series of small, punctuated gestures
Zig-zag	a curvilinear path that goes back and forth with sharp turns
Wavy	a curvilinear path that goes back and forth with curved turns
Internal motion	fingers move or wiggle

[51] The full handbook may be accessed at elisestickles.com/dissertation.

Figure 7. Manner: straight; Path: down



if oil prices [drop] to a certain level [p.drop.11.29.2014b]

Figure 8. Manner: arc; Path: down-backward



temperatures have [really begun] to drop [p.drop.11.29.2014a]

Figure 9. Manner: circle; Path: forward-down static x3



[once we creep up to fifty percent] [m.creep.12.5.2015a]

Figure 10. Manner: rotate; Path: left static



the ability to rotate [three-dimensional objects] [m.rotate.5.28.2016d]

Figure 11. Manner: chop; Path: down-up forward x2



we meet and [stumble towards] love [m.stumble.2.12.2016a]

Figure 12. Manner: zig-zag; Path: left-right x2



[you shaking your] head [m.shake.5.24.2016b]

Figure 13. Manner: wavy; Path: right-left-up x3



it [doesn't wind through] [m.wind.12.10.2014]

Figure 14. Manner: straight; Internal motion of fingers; Path: right



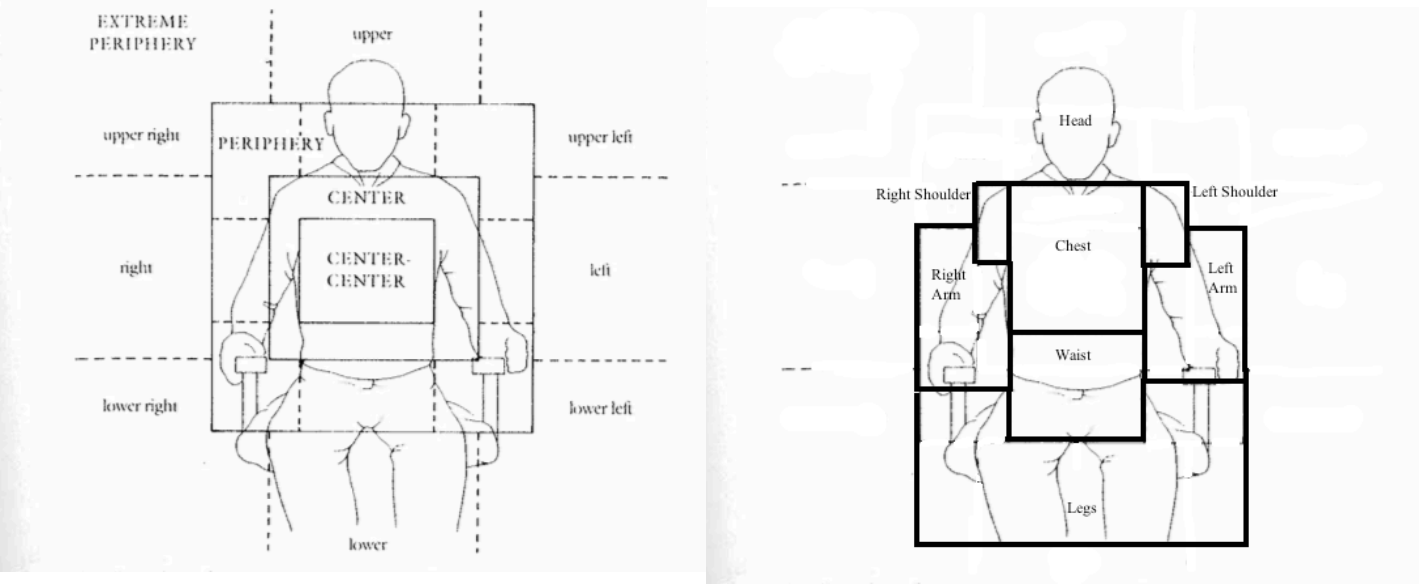
you don't [dance around] it [m.dance.12.2.2014]

These manner categories are more fine-grained than a basic set of image schemas such as that used in Cienki (2005); however, they can be grouped together into sets of foundational schemata as laid out in Johnson (1987). *Straight* and *Arc* constitute basic Paths, and *Zig-zag*, and *Wavy* Complex Paths. Complex Paths are compound schemas, wherein the gesture traces the shape of a trajectory (the Path schema) but it changes directions multiple times such that it comprises multiple internal segments (the Multiplex schema). While *Circle* and *Rotate* evoke the Cycle schema, *Chop* is an Iteration schema. *Internal motion* is a special category as it combines with the others, but it is also an Iteration schema. Furthermore, any of these manner categories may themselves be produced multiple times, which constitutes the combination of any schema with the Iteration schema.

In addition to movement, annotators also noted handshape, palm orientation, and location. To codify handshape, we made use of basic American Sign Language handshape terminology[52]. Location was based on a simplified version of the location coding scheme in McNeill (1992), as shown in Figure 15.

Figure 15. (A) Typical adult gesture space (McNeill 1992:89); (B) Simplified location coding.

(A) (B)



Palm orientation was coded in orientation to the speaker; in addition to *up* and *down* (Figure 16), it can be *speaker in* or *speaker out* (Figure 17) and *center in* or *center out* (Figure 18).

[52] A list of American Sign Language handshapes with examples is available at http://en.wiktionary.org/wiki/Appendix:Sign_language_handshapes.

Figure 16. (A) *Up* palm orientation; (B) *Down* palm orientation.

(A) (B)

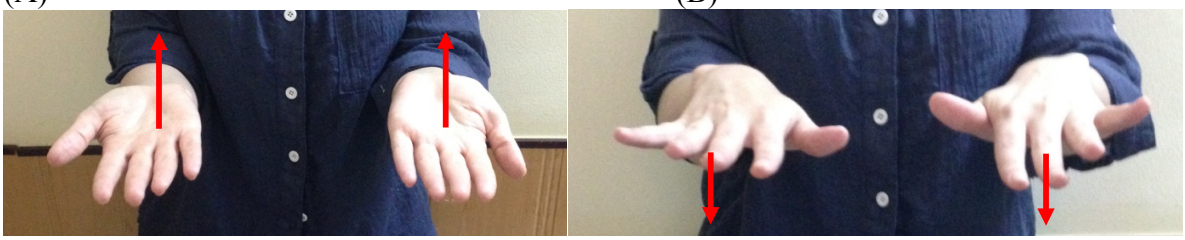


Figure 17. (A) *Speaker in* palm orientation; (B) *Speaker out* palm orientation.

(A) (B)

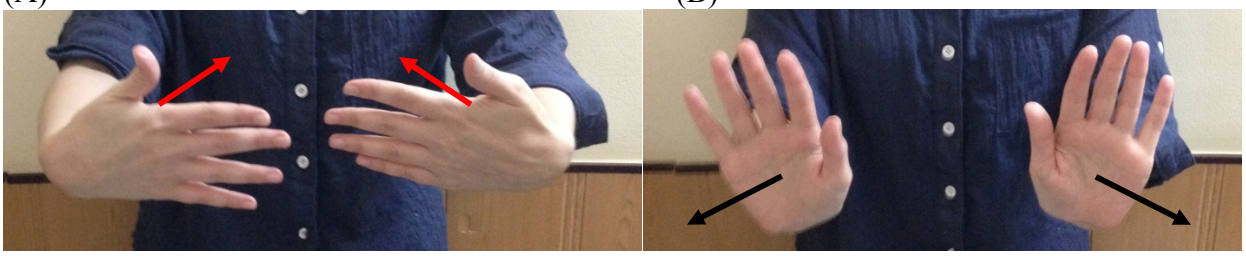
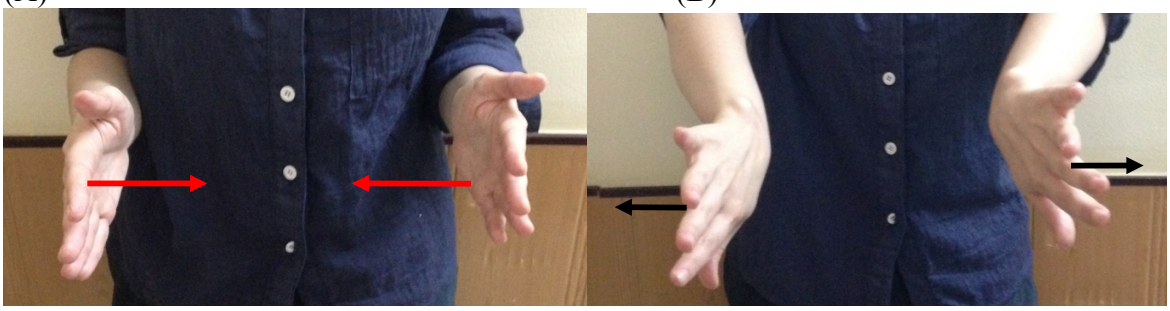


Figure 18. (A) *Center in* palm orientation; (B) *Center out* palm orientation.

(A) (B)



2.b.ii. *Speech annotation*

Speech utterances were transcribed by hand and tokenized into individual words using the ELAN tokenizer. Annotators focused on the clause co-occurring with the gesture, resulting in utterances 1-2 seconds in length. Word part of speech was annotated using a set of tags adapted from the Penn Treebank Project[53] (Marcus et al., 1993), with additional tags to distinguish *gerunds* from *present participles* and for *infinitive verbs*. The *verb type* – manner or path – of the search term was also recorded, along with its *lemma*. *Valency* of the verb's constituents as well as nominal adjuncts was also coded. Subjects were recorded as either A (subject of transitive) or S (subject of intransitive, including those with obliques). Nouns in the VP were either O (direct object, including those in particle verbs) or OBL (obliques in prepositional phrases). Later, subjects of passive sentences were re-coded as such, although there were only three of these, all in the metaphoric corpus.

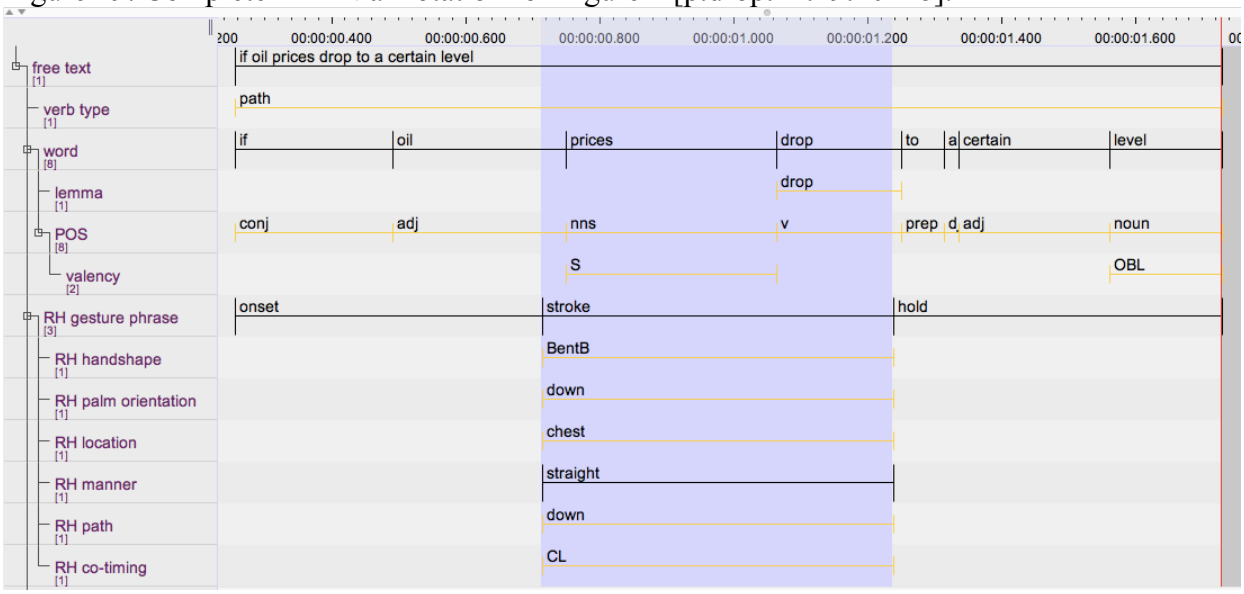
[53] A list of the Penn Treebank tags is available here: https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html.

2.b.iii. *ELAN transcription and annotation*

After the coding scheme was developed, data was coded in ELAN. Each video clip was coded separately. Linguistic tiers encoded the following: *verb type*; *free text*; *word*; *part of speech*; *verb lemma*; *valency*. The *word* tier is time-aligned, such that each word in the tier is aligned temporally with the speech utterance. Following McNeill (1992) and Kendon (2004), the base level of the gesture was the *gesture phrase*; this gesture tier was tied to the *free text* tier, and subsequent gesture-related tiers were related to the *gesture phrase*.

The gesture phrase, which roughly corresponds to a speech utterance, is typically comprised of the *onset,* or preparation of the gesture; the *stroke* of the gesture – the main movement – a *hold* in which the hand is held in space; and a *recovery* as the hand returns to rest. In practice a gesture may or may not have an onset and pre- or post-stroke hold, and instead of recovery the speaker may launch into another gesture phrase instead of putting her hands back in her lap or at her sides. The components of the gesture phrase are also time-aligned in ELAN. Gesture form information tiers – *handshape*, *palm orientation*, *location*, *manner*, and *path* – are tied to the *gesture* phrase tier. A separate tier allows for coding of internal motion of the *fingers.* Finally, stroke *co-timing* was coded: this annotated the syntactic phrase(s) of the co-speech co-timed with the gesture stroke. Each hand received its own set of annotations; if a non-hand body part produced the gesture, the right hand gesture tiers were used for the body part and renamed accordingly (e.g., *Head gesture phrase*). An example of a complete annotation is shown in Figure 19.

Figure 19. Complete ELAN annotation for Figure 4 [p.drop.11.29.2014b].



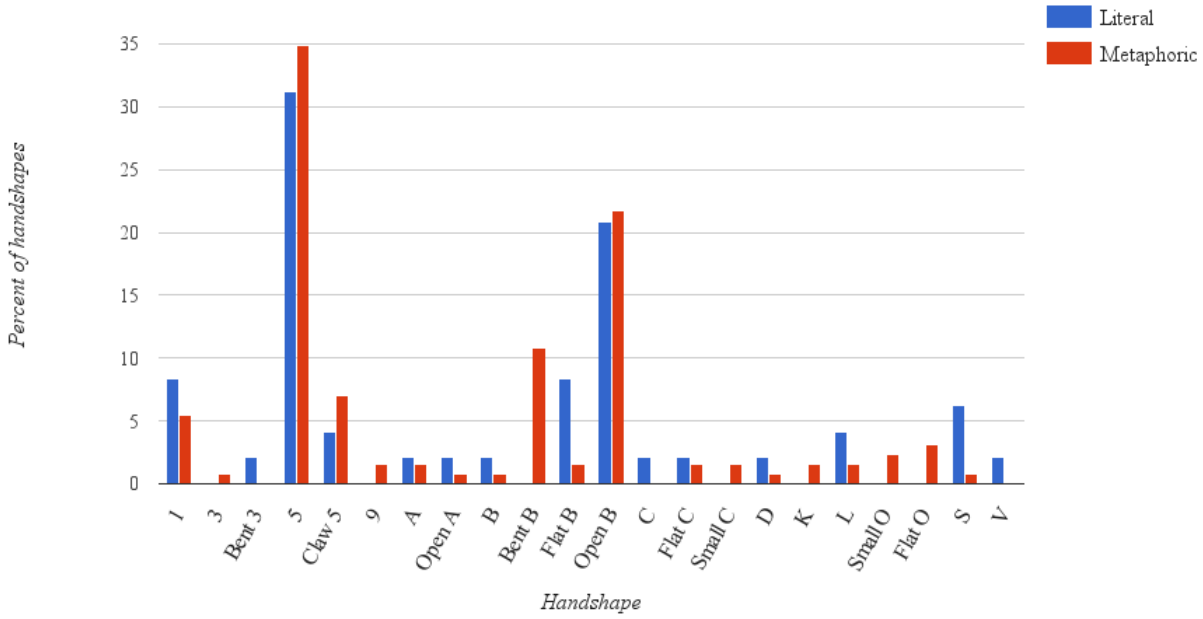
2.b.iv. *Additional annotations*

Following the transcription and annotation of all the metaphoric and literal videos in ELAN, it became apparent through initial data analysis that consolidating some of the annotations into broader categories would be useful.

Gesture image schema. As discussed in 2.b.i., the seven main gesture manner annotations can be grouped into a set of four image schemas (Path, Cycle, Complex Path, Iteration). Hence, an *image schema type* category was added, such that all *straight* and *arc* gestures were categorized as Path and *chop* as Iteration. *Circle* and *rotate* gestures were categorized as Cycle; and *zig-zag,* and *wavy* as Complex Path. The two gestures with *internal motion* of the fingers were also categorized as Iteration due to the repeated motion of the fingers.

Handshape. There were a total of twenty-three different handshapes annotated; handshape frequency (as percentages of total handshapes) is reported in Figure 20.

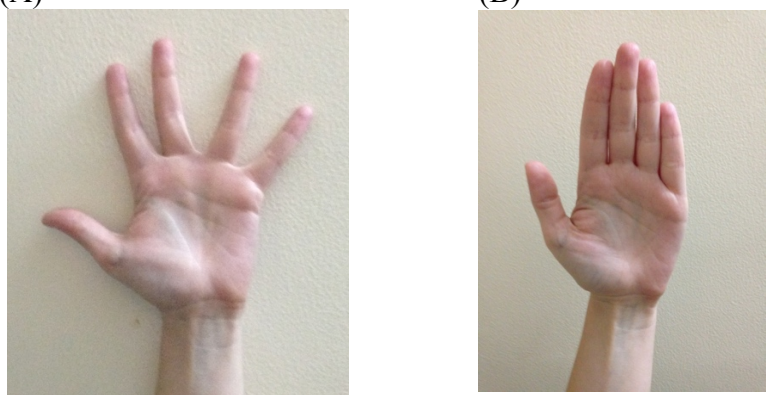
Figure 20. Distribution of handshape types between metaphoric and literal data.



As is apparent from Figure 20, a few handshapes – 5 and Open B (Figure 21) – are very common in both the literal and metaphoric data sets; together they total 52.08% of the literal handshapes and 56.59% of the metaphoric data. These two handshapes are *unmarked*, as they are the natural, default shape of the hand. It is impossible to produce a manual gesture without also producing some kind of hand shape; the 5 and Open B are the most neutral shape physically possible. Hence, they can be categorized as *unmarked*, and all other handshapes as *marked*.

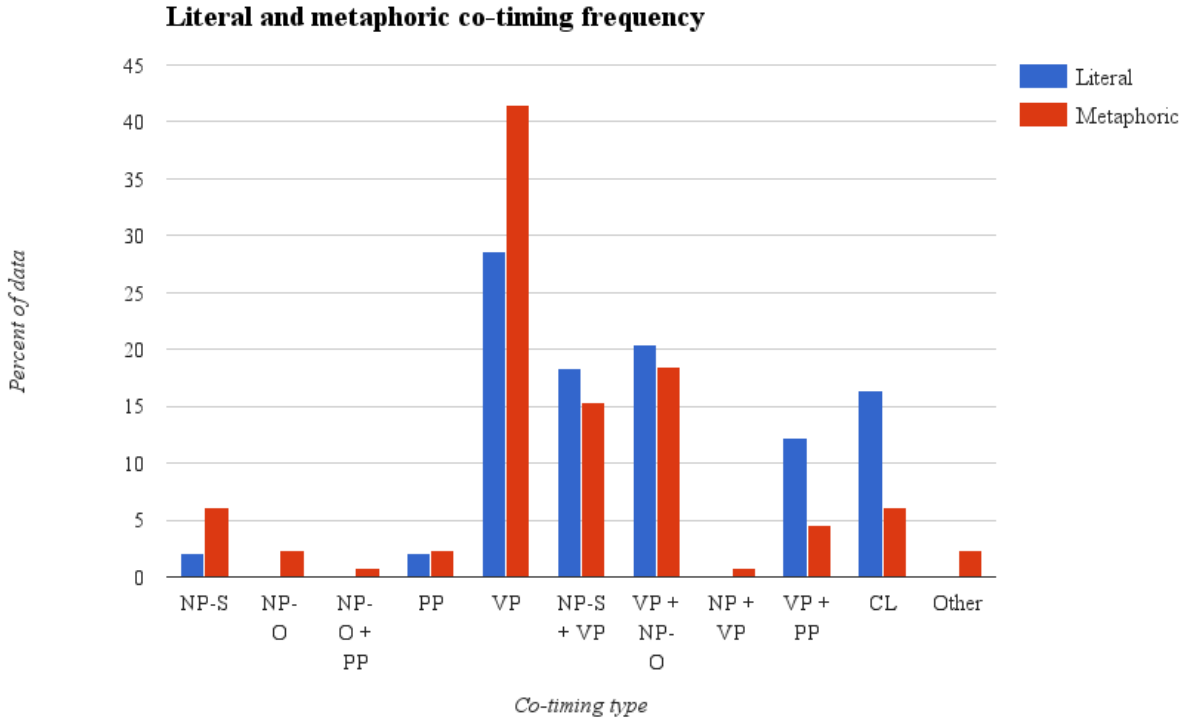
Figure 21. (A) 5 handshape; (B) Open B handshape.

(A) (B)



Cotiming. There were eleven possible types of gesture stroke co-timing, as illustrated in Figure 22.

Figure 22. Distribution of co-timing frequencies.



In the first four categories, the gesture stroke co-occurs with only a nominal (NP-S: subject NP; NP-O: object NP; PP: prepositional NP). In the fifth category – the majority of the data – the stroke co-occurs with only the verb phrase (VP). In the next five categories, the stroke co-occurs with both a nominal and verbal element: either some combination of the prior categories, or across a whole transitive clause (CL). Finally, a few gestures did not fall into any of these categories (Other). From Figure 22 we can observe that most of the gesture data is co-timed with the VP only; a small fraction co-occurs with a nominal; and a sizable remainder over a

combination of a verb and its constituents. This forms three natural *co-timing categories*: *NP Only*, *VP Only*, and *NP+VP*.

Linguistic image schema. In order to compare the image schematic structure of the gestures to that of their corresponding co-speech, it is necessary to similarly characterize the image schemas of the linguistic forms. While all of the verbs are considered manner and path verbs, *manner* here is a complex category in that it may include *how* the action was performed (e.g., "walk", "fly", "dance"); the *compulsion* of motion (e.g., "spread", "drive"); the *enablement* of motion (e.g., "walk something through") the *blocking* of motion (e.g., "stumble"); and the character of the motion itself (e.g., "shake", "rotate"). Manner verbs which reflect how the action is performed may also be used to evoke the prototypical rate of that manner (e.g., "run", "fly", "crawl"), or the shape traced by that motion (e.g., "rotate"). Similarly, some *path* verbs may only include a *path* (e.g., "drop", "rise"); others specify a *trajector-landmark* relationship (e.g., "come"). From these observations we recognize some of the image schemas previously described: Path and Cycle. We add Trajector-Landmark, and three primary force-dynamic image schemas related to caused motion: Enable, Block, and Force[54] (Johnson, 1987; Talmy, 2000). We can also add Speed and a generic self-Motion schema, to characterize manner verbs such as "run" and "walk".

Further consideration of the linguistic data reveals that speakers frequently use prepositions or particle verbs, which here are comprised of a manner verb plus a particle derived from a preposition. The image schematic structure of prepositions and verbal particles has been widely studied (e.g., Brugman, 1981; Lakoff, 1987 for foundational works), and it is beyond the scope of this work to fully discuss the details of each utterance here. However, we can broadly observe that in the current context of translational motion, these satellites can be largely understood in terms of location and path-related schemas (Location, Path, Trajector-Landmark), as would be expected given the current focus on the Location Event Structure Metaphor.

Given this basic[55] set of image schemas, I annotated each verb and satellite in its particular speech context for *verb image schema* and *satellite image schema*; my guiding framework was to choose the most-specific schema. For example, given that Trajector-Landmark inherits from Path, Trajector-Landmark is the more specific schema. Similarly, because Caused Motion is a type of Motion, verbs evoking a type of Caused Motion such as Enable would be categorized as Enable as it is more specific than generic Motion.

3.a. *Results*

To review in brief, the goal of this study is to evaluate and formalize the image schematic structure present in metaphoric gestures. In doing so, we will consider the image schemas in metaphoric gestures and compare them to the image schemas in the accompanying metaphoric co-speech; we will further compare these metaphoric multimodal utterances against similar literal utterances. We expect that co-expressive utterances should represent similar image schemas in speech and gesture. We also predict that co-expressive gestures accompanying the same linguistic utterances should be similar across speakers. In this section, I first provide

[54] "Force" is also referred to as the Compulsion schema.
[55] and admittedly oversimplified

descriptive summary statistics for the speech and gesture annotations separately, and then consider the above questions in the Discussion, section 4.

3.b. *Speech results*

Lemma. There were 85 metaphoric and 39 literal manner verbs in the data set, compared to 45 metaphoric and 10 literal path verbs. This reflects the overall trend in English as a satellite-framed language to encode manner of motion in verbs and path in satellites (Talmy 1991, 2000). Table 3 reports frequency counts for each lemma; note that some lemmas were only found for one metaphor category but not the other, with more verbs only appearing in metaphoric data. This is in part due to the fact that fewer overall literal data points were collected, but also reflective of the fact that verbs were selected for searching on the basis of metaphoric use frequency. As described in Chapter 3, use frequency varies between literal and metaphoric contexts (see also Stickles and Dodge, 2016). Given that the revised lemma list was based on metaphoric frequency rather than overall frequency, it is likely that metaphoric uses will dominate in a database search over literal ones.

Table 3. Lemma frequency measures.

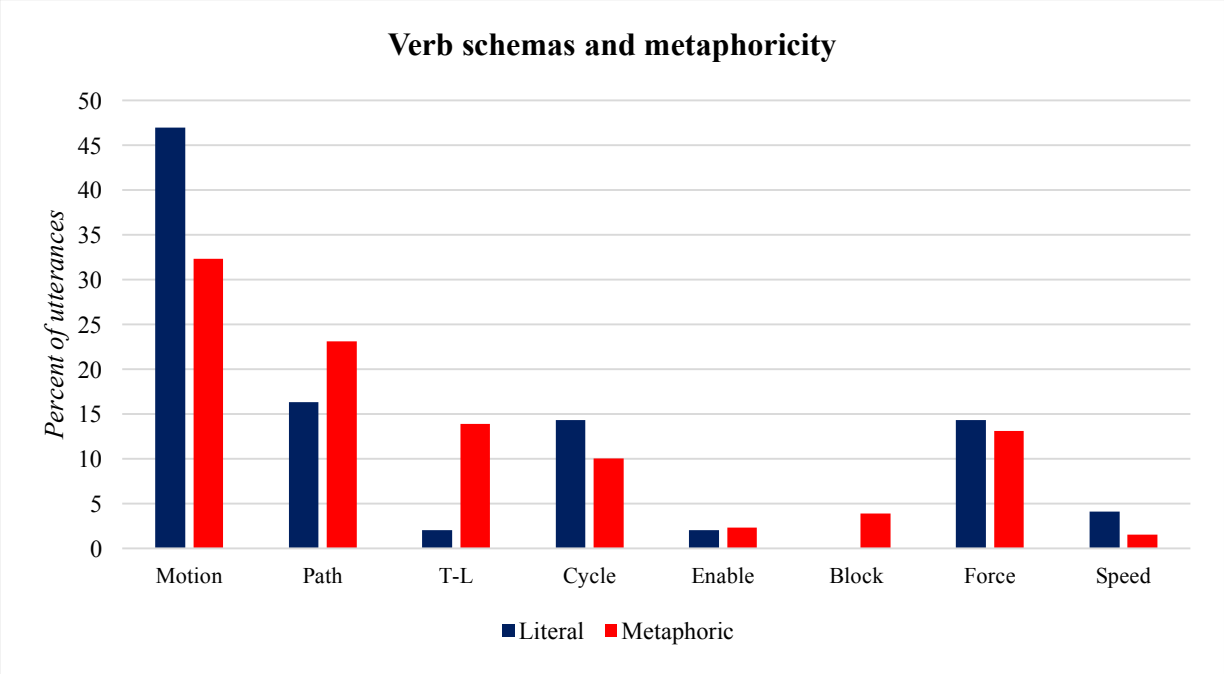
	Manner verbs	
lemma	**metaphoric**	**literal**
crawl	1	2
dance	3	1
drift	14	6
drive	6	1
float	2	2
fly	3	2
jump	1	1
leap	1	1
revolve	11	2
roll	4	1
rotate	2	6
run	4	6
shake	1	1
sway	4	2
walk	4	5
creep	6	0
rock	1	0
sail	1	0
slide	7	0
slip	1	0
spread	1	0
stumble	5	0

	Path verbs	
lemma	**metaphoric**	**literal**
advance	1	1
come	3	0
cross	2	1
climb	4	1
dip	0	2
dive	1	0
drop	12	1
enter	9	0
escape	1	0
exit	2	0
fall	3	3
plunge	5	0
rise	2	0
sink	0	1
Total	**45**	**10**

tumble	1	0
wind	1	0
Total	**85**	**39**

Linguistic image schema. Given the differences in lemma distribution between metaphoric and literal utterances, it is expected that the verb schemas will be distributed differently as well, as shown in Figure 23. However, the difference is not significant (Fisher's Exact Test, $p = 0.09$).

Figure 23. Distribution of verb schemas in metaphoric and literal utterances. "T-L" stands for Trajector-Landmark.



We can also consider a similar comparison for verbal satellite image schema, which is also not significantly different between utterance type (Fisher's Exact Test, $p = 0.17$).

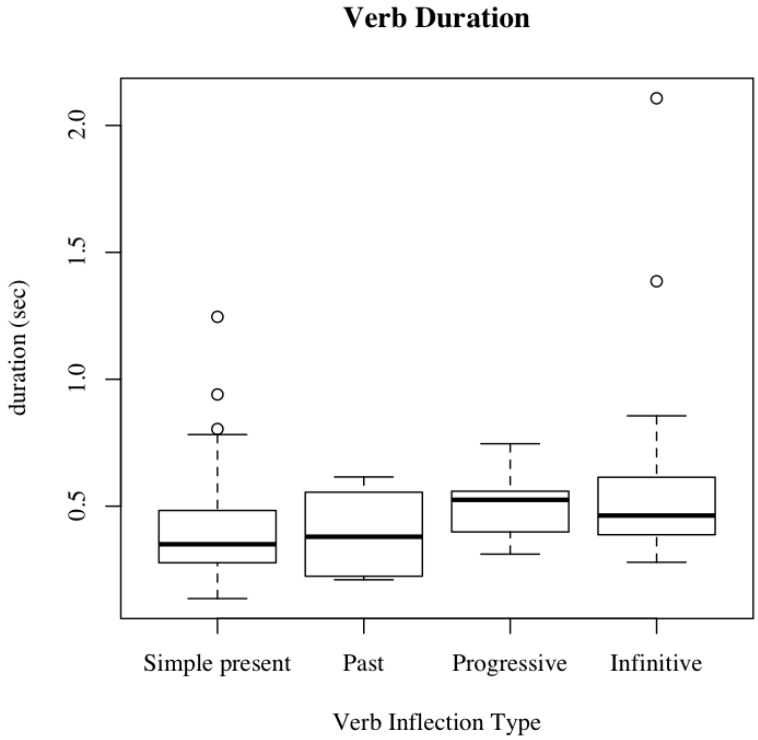
Syntax and structure. Verb inflection is summarized in Table 4. Most of the verbs were in the simple present, with the infinitive the second most common type. Infinitival forms were always accompanied by one or more auxiliary/modal verbs or pseudo-auxiliary verbs (such as "**continue** to rotate" or "**starts** to run"). There was no difference in distribution of verb inflection types between metaphoric and literal data (two-sided Fisher's Exact Test, $p = 0.71$).

Table 4. Distribution of verbal inflection.

Verb Inflection Type	**Metaphoric**	**Literal**
Simple present	84	30
Past	5	1
Progressive	12	3
Infinitive	29	15

Duration of verb utterance did not significantly differ between metaphoric and literal data (Kruskal-Wallis rank-sum test ($df = 1$) = 1.5393, $p > 0.21$). It did, however, differ between inflection categories (Kruskal-Wallis rank-sum test ($df = 3$) = 23.08, $p < 0.0001$), as illustrated in Figure 24.

Figure 24. Distribution of verb utterance duration.



Speech duration of simple present (median = 0.35 sec) and past tense verbs (median = 0.38 sec) was shorter than infinitive (median = 0.46 sec) and progressive verb duration (median = 0.53 sec), due to an effect of syllable length.

Of the metaphoric utterances, 59 were transitive, 68 were intransitive active sentences, and 3 were intransitive passives. There was no difference between the proportion of transitive and intransitive active sentences ($p > 0.05$). In contrast significantly fewer literal utterances were transitive (12) than intransitive, 37 ($p < 0.05$). Notably, 60.08% of metaphoric utterances and 59.18% of literal utterances contained a satellite (preposition or particle). However, of the utterances with a satellite, more metaphoric utterances had a verb particle (60.76%) whereas more literal utterances had a preposition (75.86%). So, while there is a distributional difference in valency (transitive/intransitive) between the metaphoric and literal verbs, this is in a sense counterbalanced by the difference in satellite distributions. The higher preposition rate in the literal utterances means that while there is a greater intransitive frequency, the utterances still have nominal discourse referents in the predicate, although realized in different syntactic positions. Complementing this is the fact that while metaphoric utterances are more frequently transitive, the high particle verb rate means that the direct objects are still packaged in a satellite-framed predicate. Indeed, a full 60% of direct objects in metaphoric utterances are in verbal

predicates with particles, whereas only 25% of those in literal utterances are. Overall, the data suggests a high rate of satellite use, irrespective of metaphoricity.

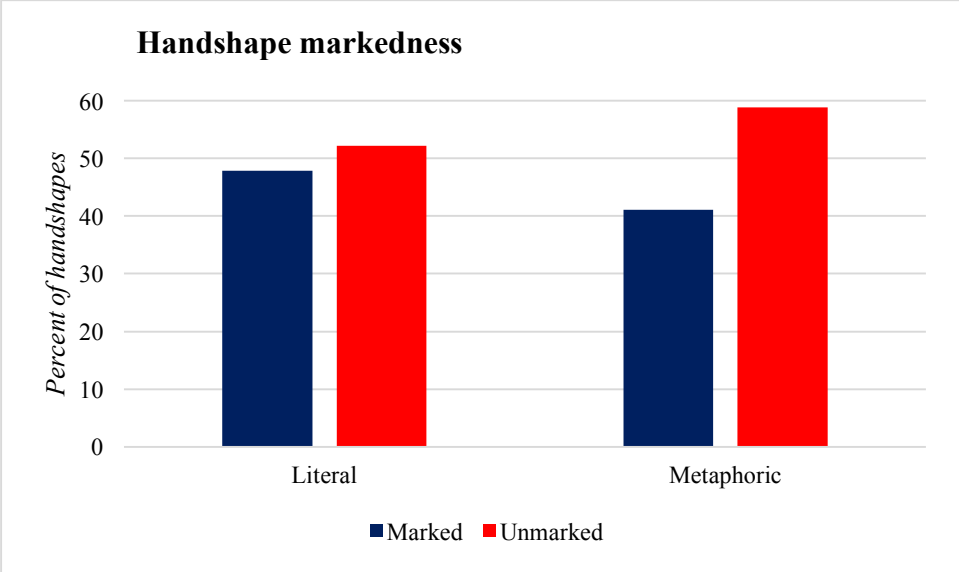
In sum, we observe relatively few linguistic differences between the metaphoric and literal utterances, beyond variation in lemma frequency. The similarities in syntactic form demonstrate that in this data set, speakers use verbs similarly in literal and metaphoric contexts. Furthermore, the lack of a difference between metaphoric and literal usages suggests that any differences in gesture we may observe are not confounded by co-varying differences in linguistic use.

3.c. *Gesture results*

In this section, I discuss results for the gesture annotation data, including form, duration, and image schematic structure. The interaction of gesture and speech will be covered separately in section 3.d.

Form: Handshape. For the purposes of this study I focused on gesture handshapes. As explicated in section 2.b.iv, handshape types were further categorized as *marked* or *unmarked* due to the distribution of handshape frequencies (see Figure 20 in 2.b.iv). As shown in Figure 25, significantly more metaphoric handshapes were unmarked than marked (one-tailed Binomial Exact Test, $p = 0.03$), while there was no difference in the literal data. However, there was not a significant difference between the literal and metaphoric data (Fisher's Exact Test, $p = 0.49$).

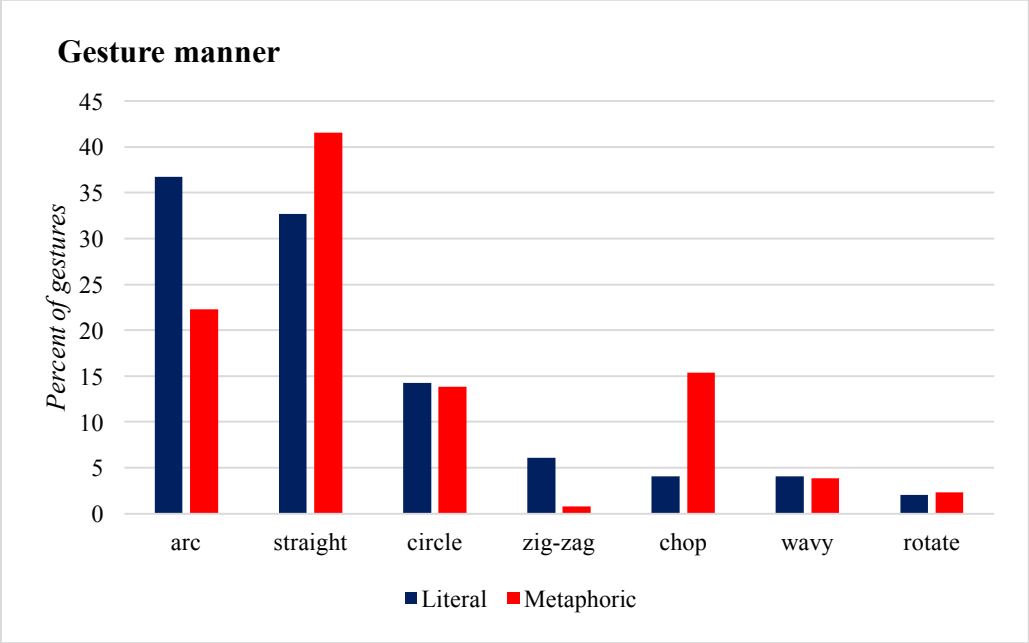
Figure 25. Handshape type frequency.



Form: Manner and Image schema. Distribution of manner categories is reported in Figure 26. *Arc* and *straight* were the most common manner categories in both the literal and metaphoric data, and both types of utterances had about the same number of *circle*, *wavy*, and *rotate* gestures. However, there were differing proportions of the *zig-zag* and *chop* gestures. Overall, there was a marginal difference in manner category distribution (Fisher's Exact Test, $p = 0.053$). To determine which comparisons were significant, post-hoc BM-corrected pairwise one-tailed proportion tests were applied to the *arc*, *straight*, and *chop* data. Others were not compared due

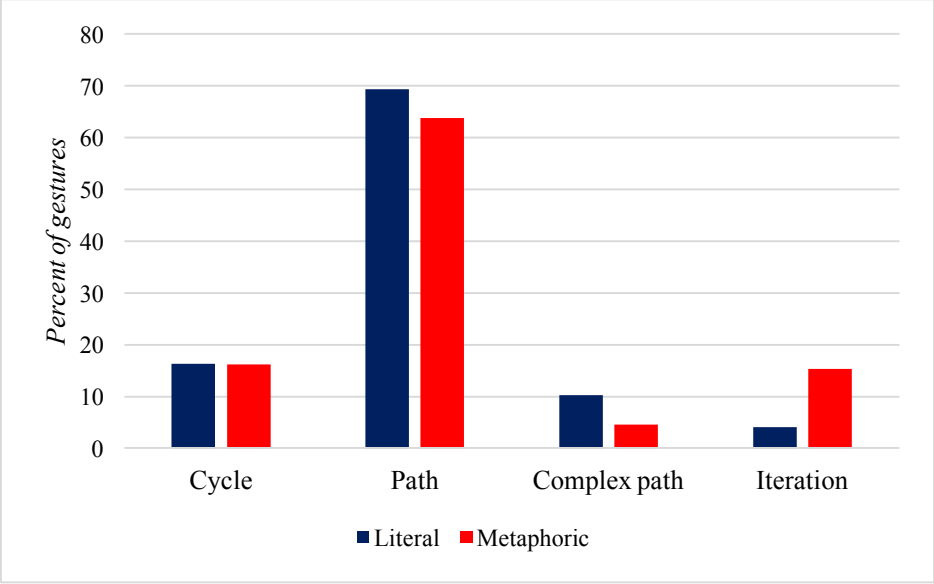
to small sample size and/or having clearly similar proportions. The *arc* and *chop* results were marginally significant, $\chi^2(df = 1) = 3.12$, corrected $p = 0.058$; and $\chi^2(df = 1) = 3.23$, corrected $p = 0.058$. The *straight* results were not significant, $\chi^2(df = 1) = 0.84$, corrected $p > 0.05$. Hence, we conclude that literal data was more likely to include *arc* gestures than the metaphoric data, and the metaphoric data more likely to include *chop* gestures, with no other differences in the distribution of gesture manner.

Figure 26. Gesture manner frequency.



We can further analyze gesture manner by considering the distribution of the basic image schemas, Path, Complex Path, Cycle, and Iteration, as illustrated in Figure 27.

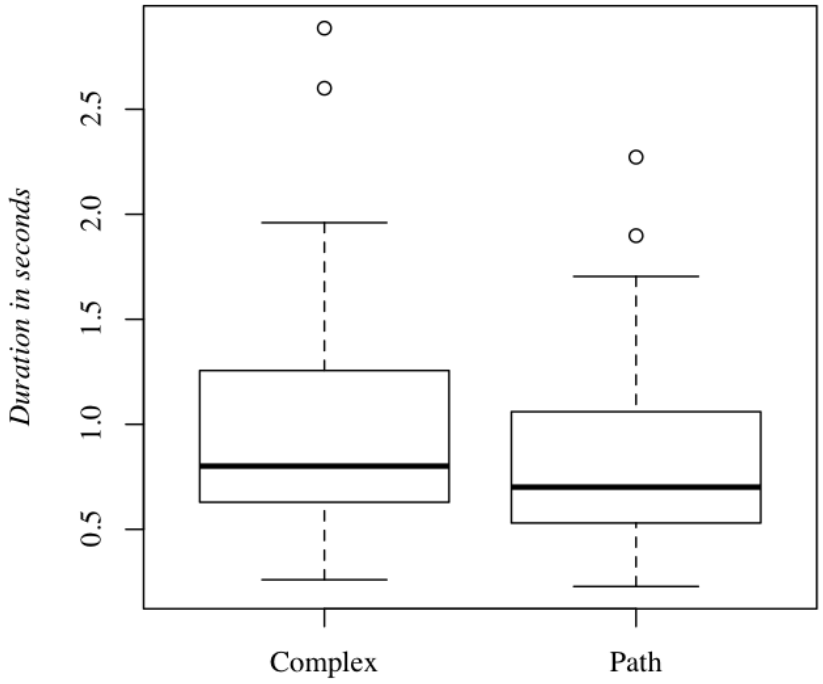
Figure 27. Distribution of gesture image schemas.



As is apparent from Figure 27, there is no significant difference between the literal and metaphoric data in terms of gesture image schema (Fisher's Exact Test, $p > 0.05$). Furthermore, it is worth noting that differentiating *arc* from *straight* in annotation was sometimes difficult, and annotators would discuss whether a very slight curve in a gesture necessarily made it an 'arc', or if the curve needed to be more pronounced to classify it as such. So, it is possible that the difference in the *arc* gestures is overstated in the data, especially since the literal data was annotated later separately from the metaphoric data and some systemic error may have been introduced due to this. When *arc* and *straight* gestures are combined together into the Path schema category, there is no difference between the literal and metaphoric data ($\chi^2(df = 1) = 0.03$, $p > 0.05$).

Stroke duration. An entire gesture phrase consists of a preparatory onset, stroke, and retraction; however, they also may include holds, which can considerably add to the duration of a gesture phrase. The holds serve to maintain the temporal synchronicity between speech and gesture, while the stroke is the part of the gesture that holds meaning (McNeill, 1992). Thus, following Duncan (2008) and Parrill et al. (2013) gesture stroke duration was measured rather than the duration of the complete gesture phrase. The mean and median stroke duration were 0.86 seconds and 0.74 seconds respectively. Stroke duration as a function of gesture image schema is shown in Figure 28 for Path vs. the other types of image schemas. Cycle, Complex Path, and Iteration were combined into a single category as a measure of increased complexity, following the distinction made in Parrill et al. (2013). Stroke duration was significantly different between the two categories, (Kruskal-Wallis χ^2 ($df = 1$) = 4.49, $p = 0.03$), with median complex gesture stroke duration = 0.80 seconds and median Path gesture stroke duration = 0.70 seconds.

Figure 28. Stroke duration as a function of gesture image schema.



3.d. *Interaction of gesture and speech results*

Co-timing. While gesture phrases are often co-timed at the clause level (McNeill, 2008), the stroke carries the meaning of the gesture. Given the emphasis here on the image schematic structure of the gesture, which is conveyed by the form of a referential gesture's stroke, gesture co-timing was calculated by annotating which parts of the utterance the stroke overlapped with. Results are shown in Figure 29. The figure is divided into four parts, as indicated by the vertical dotted lines. The left part of the graph shows frequency of nominal co-timing, the middle section verbal only, and the third shows frequency of combined verb+argument co-timing. The rightmost section comprises those strokes which did not fit into any of the above categories. Figure 30 reduces each section down into a co-timing type. Metaphoric and literal utterances differ significantly by co-timing type (Fisher's Exact Test, $p < 0.05$). A comparison of the VP + Argument proportions shows that significantly more of the literal utterances are co-timed with both the VP and one or more arguments (corrected test for equality of proportions, $\chi^2(df = 1) = 6.95$, $p < 0.01$).

Figure 29. Stroke co-timing.

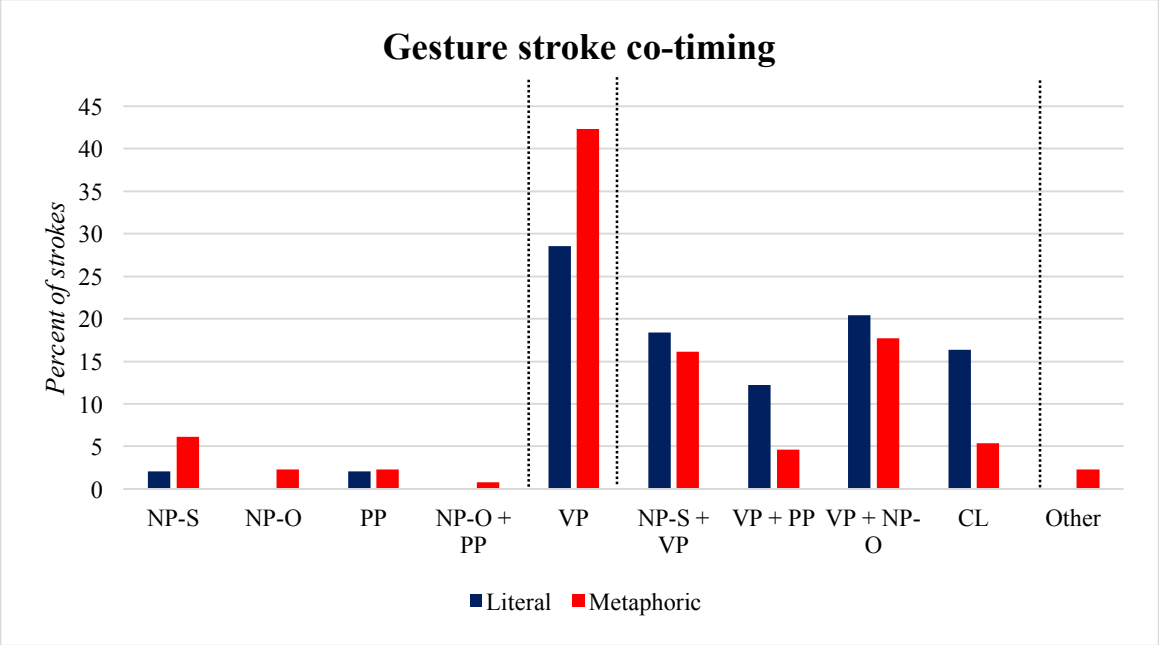
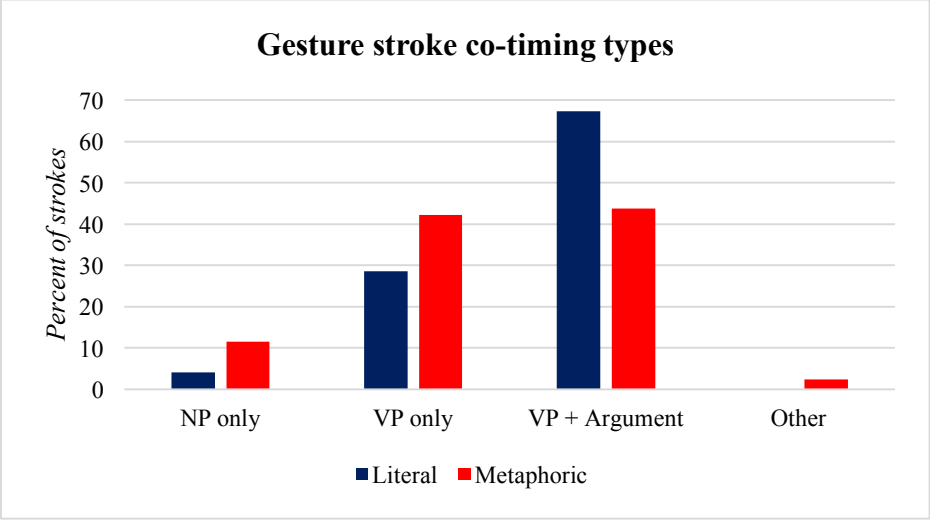


Figure 30. Stroke co-timing by type.



Due to the small sizes of the individual co-timing categories, a logistic regression was modeled on the co-timing data (Table 5) to determine which categories contributed to the difference between metaphoric and literal.

Table 5. Results of co-timing logistic regression model.

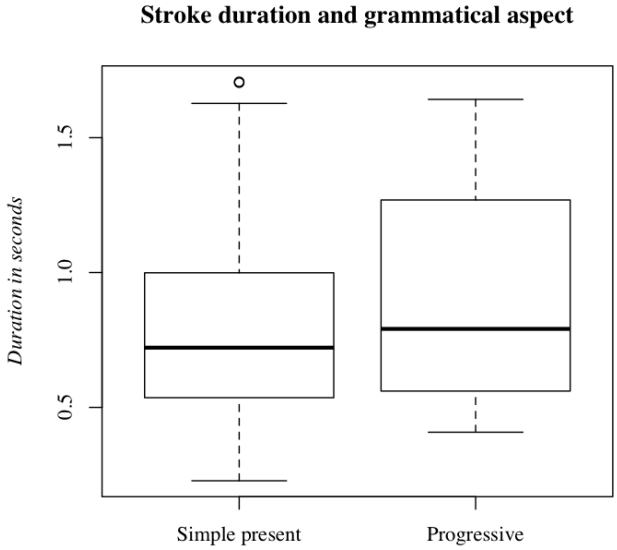
Coefficients	**Estimate**	**Std. Error**	**z value**	***p* value**
(Intercept)	0.1335314	0.5175492	0.2580072	0.7964
NP-S	-2.212973	1.180194	-1.875093	0.06078
NP-S + VP	-0.9808293	0.6531365	-1.501722	0.1332
PP	-1.232144	1.265382	-0.9737329	0.3302
VP	**-1.501807**	**0.5978859**	**-2.511863**	**0.0120 ***
VP + NP-O	-0.9664405	0.6413544	-1.506874	0.1318
VP + PP	-0.1335314	0.7753647	-0.1722175	0.8633
NP - O	-16.6996	1385.378	-0.01205418	0.9904
NP - O + PP	-16.6996	2399.545	-0.006959487	0.9944
Other	-16.6996	1385.378	-0.01205418	0.9904

Results show that one category contribute to the fit of the model: VP (significant, $p < 0.05$). By analyzing gesture stroke co-timing at these two levels of granularity we see that overall, the literal and metaphoric utterances differ in terms of stroke co-timing. Metaphoric speech is more likely to co-occur with verb phrases, whereas literal speech is more likely to co-occur with combinations of both verbal and nominal and/or prepositional elements of the utterance.

Stroke duration. Since literal gestures more often co-occur with more speech elements as discussed above, we might expect that they are also longer in duration. However, there is not an interaction of metaphoricity and stroke duration (two-tailed Mann-Whitney U = 2759, $p = 0.17$). Turning to part of speech, Parrill et al. (2013) have shown that there is an interaction between stroke duration and grammatical aspect: English speakers produce longer gestures when using the progressive as compared to the perfect. Initial comparison of progressive utterances to simple present utterances did not reveal a difference in stroke duration did not show a difference between them (Mann-Whitney U = 721.5, $p = 0.16$). However, there are several outliers with

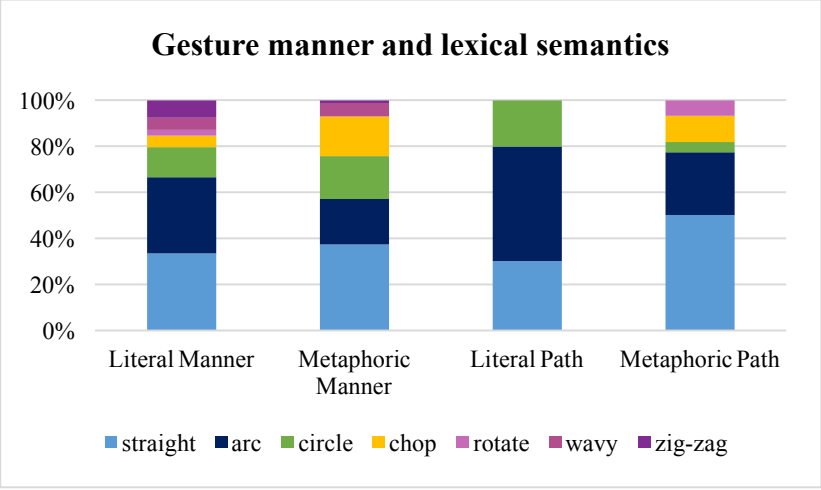
long stroke durations in the simple present data; excluding data in the upper 5% of stroke duration data reveals a greater trend towards longer duration in the progressive (Mann-Whitney U = 661.5, $p = 0.11$), as illustrated in Figure 31.

Figure 31. Interaction of grammatical aspect and stroke duration.



Semantics. Finally, we can consider the semantics of the gestures as they interact with the semantics of the speech. Figure 32 illustrates the relationship between gesture image schemas and verb type (manner vs. path) in metaphoric and literal usage. Path schemas are coded in blues; Cycle in green; and Iteration schemas in pinks. A logistic regression model was fitted to the data to test the interaction of gesture manner and verb type with metaphoricity as the dependent outcome. First combining the metaphoric and literal utterances together, there is a significant difference in gesture manner between manner and path verbs (Fisher's Exact Test, $p = 0.04$); however, there is also a difference in gesture manner between metaphoric and literal utterances (Fisher's Exact Test, $p = 0.05$). This may suggest that manner and path verbs interact with metaphoricity.

Figure 32. Distribution of gesture manner by verb type and metaphoricity.



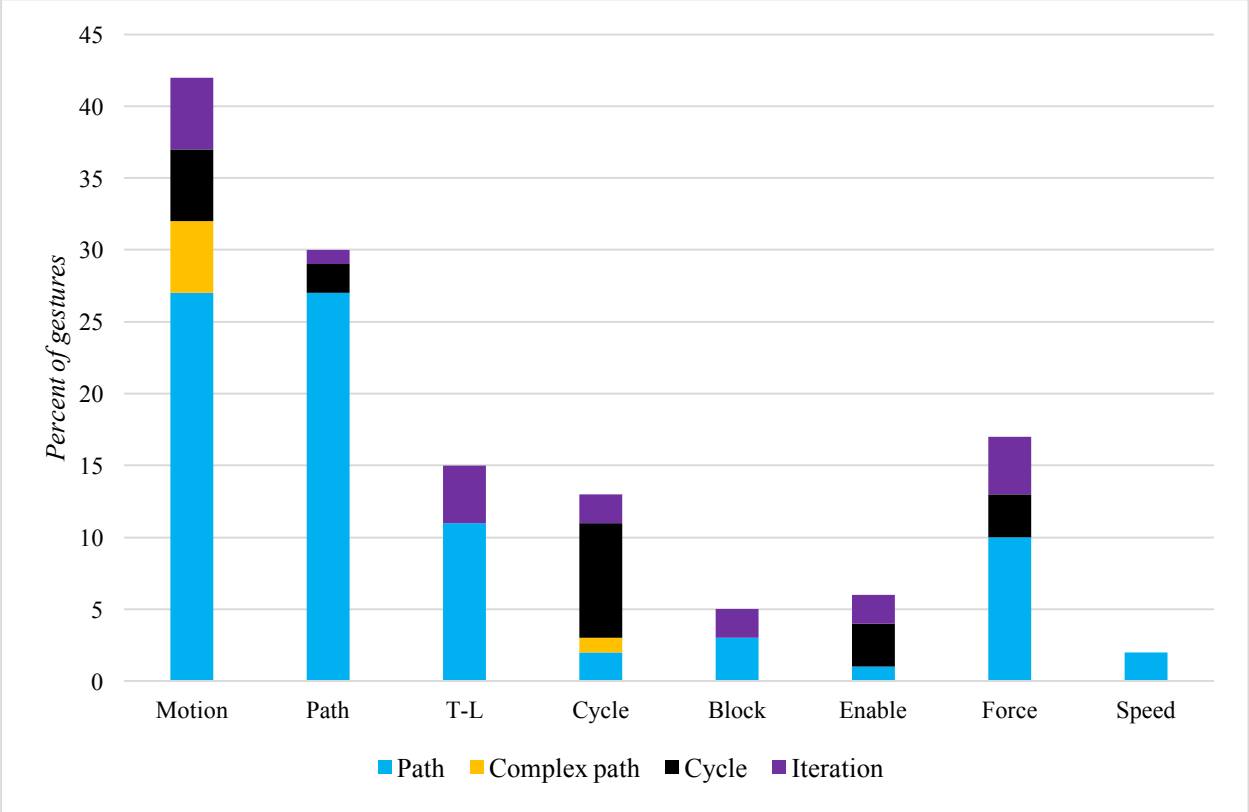
However, when we consider instead gesture image schema, there is no longer a meaningful difference between metaphoric and literal utterances (Figure 29; Fisher's Exact Test, $p = 0.11$), but still one between manner and path verbs (Fisher's Exact test, $p = 0.04$). A logistic regression was modeled on the data to test the interaction of the effects of metaphoricity and verb type on gesture image schema (Table 6).

Table 6. Results of verb type * metaphoricity logistic regression model.

Coefficients	**Estimate**	**Std. Error**	**z value**	***p* value**
(Intercept)	1.4759	0.2771	5.326	1e-07 ***
Verb type: Path	0.5782	0.5499	1.051	0.293
Metaphoricity: Literal	0.2288	0.5232	0.437	0.662
Path:Literal	-0.8967	1.0604	-0.846	0.398

Here, verb type did not significantly improve the fit of the model ($p = 0.29$). There was not an effect of metaphoricity ($p = 0.66$) or the interaction between metaphoricity and verb type ($p = 0.40$). In general, we can conclude that while utterance metaphoricity appears to interact with gesture manner, when considering instead gesture image schema it does not.
Lastly, we can consider whether there is a relationship between metaphoric verb schema and gesture image schema (Figure 33). In Figure 33, we see that most verbal image schemas are dominated by Path gestures, except for Cycle, which has predominately Cycle gestures.

Figure 33. Relationship between metaphoric verb schema and gesture image schema.



A logistic regression model (Table 7) fitted to the metaphoric data with Verb Schema as a predictor variable and Gesture Schema, using the Motion verb schema as the reference category, shows that only the Cycle verb schema significantly improved the fit of the model ($p < 0.001$).

Table 7. Results of gesture schema ~ verb schema logistic regression model for metaphoric data.

Coefficients	**Estimate**	**Std. Error**	**z value**	***p* value**
(Intercept)	2.0015	0.4765	4.201	2.66e-05***
Cycle	**-2.4715**	**0.7430**	**3.326**	**0.00088** ***
Path	0.6376	0.8734	0.730	0.46537
Trajector-Landmark	-0.3920	0.7919	-0.495	0.62053
Enable	15.5646	2284.1018	0.007	0.99456
Force	-0.4610	0.7949	-0.580	0.56190
Block	15.5646	1769.2577	0.009	0.99298
Speed	15.5646	2797.4420	0.006	0.99556

4.a. *Discussion*

In this chapter, I consider the claim that the underlying structure of metaphoric multimodal utterances is similar to that of literal multimodal utterances. To do so, I compare the following: (a) metaphoric and literal speech; (b) metaphoric and literal referential gestures; and (c) the interaction of speech and co-expressive gesture in metaphoric and literal utterances. In this section, I discuss for each of these comparisons, and argue that identifying the unifying image schematic structure of each speech/gesture package is the key to resolving unexpected results in modeling the relationship between metaphoric and literal multimodality.

4.b. *Speech results*

To analyze the linguistic component of the corpus, we focused on several aspects of speech: word choice; verb inflection; verb duration; argument structure; and verbal satellites. First, we found a difference in lemmas between metaphoric and literal utterances. This was to be expected, given that there are established differences in use frequency between metaphoric and literal usage (see Chapter 3). A deeper question, however, is whether these surface differences in lexical choice reflect more substantial structural differences. To address this, we considered verb image schema, and found a difference in distribution between the two utterance types; I consider this further in depth in section 4.d.

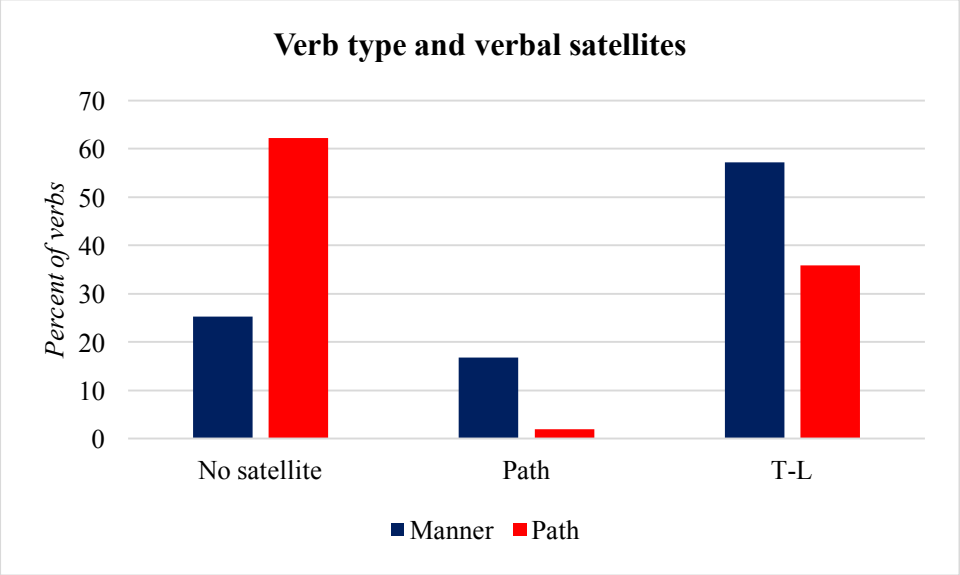
Moving to syntactic structure, we found a difference in argument structure, such that metaphoric sentences were transitive at a higher rate than literal sentences. In contrast, literal utterances were more likely to have prepositional phrases, whereas metaphoric sentences were more likely to have verbs with particles. As I argue in section 3.b., when considering argument structure as a whole, these two findings effectively complement each other: irrespective of valency, both utterance types show a high rate of satellites (prepositions or particles). It may be that differences in argument structure here relate to the lexical semantics of the verbs, given that lexemes vary as to whether they license null instantiation (e.g., David 2016). Null instantiation occurs when a frame element is evoked by an utterance, but not actually present: hence, it is "null" (Fillmore, 1986). For example, in the phrase *he climbed up*, the evoked frame includes some object by which he climbed – a ladder, or wall, or perhaps a rain-pipe – but is not present

in the sentence. Thus this object is a null instantiated element. This utterance is *intransitive,* like much of the literal data in the current corpus; but it still evokes the frame element. Since lemma distribution varies between metaphoric and literal use, null instantiation – and thus valency – may be skewed as a side-effect of that variation.

Finally, we found that metaphoric and literal utterances did not differ in terms of verbal morphology or duration. There was no difference in the distribution of tense/aspect morphology between the two. Verbal morphology did have an influence on speech duration, such that progressive verbs were of longer duration, as would be expected due to increased syllable length. Variation in speech duration was not affected by metaphoricity.

To summarize, in most respects we do not find major differences in speech. The argument structure contrasts – valency and satellite type – can be understood as an effect of the frame semantics of the specific lemmas in use. The fact that both utterance types do have high satellite rate suggests an emphasis on Path, given that satellites in this corpus convey path-related image schemas (Path or Trajector-Landmark). Furthermore, significantly more manner verbs co-occur with satellites (Fisher's Exact Test, $p < 0.0001$), as shown in Figure 34. There is no difference in satellite distribution related to metaphoricity (Fisher's Exact Test, $p = 0.08$).

Figure 34. Distribution of satellites by verb type.



The fact that a large majority of manner verbs co-occur with satellites (73.73%) while a minority of path verbs do (37.74%) tells us that when speakers are using manner verbs, they are still evoking Path-related image schemas via their use of satellites. This holds across metaphoricity.

4.c. *Gesture results*

One of the major goals of this study was to examine the relationship between the forms of metaphoric and literal gestures; I aim to show that metaphoric gestures should represent similar image schematic structure to literal gestures. However, when comparing gesture *manner* – that is to say, the manner form of the gesture – we do find differences between metaphoric and literal

gestures. Literal data was more likely to include *arcs*, whereas metaphoric data was more likely to include *chops*. From these results, we may be inclined to conclude that there are fundamental differences in structure between manner and path. However, I argue that the image schematic level, rather than form level, is the appropriate level of analysis. By breaking down a gesture into its constituent, perceptually salient features, we can identify those elements of the gesture that are semantically meaningful. Given that a gesture's meaning is available to the addressee via their shared conceptual structures with the speaker – namely, their shared image schematic representations – if we are seeking to analyze gestures as semantically meaningful, the image schema is perhaps more reliable than form alone. And indeed, when comparing metaphoric and literal gestures at the image schematic level, we find no differences between the two.

Turning to other aspects of gesture form, we observe a difference between literal and metaphoric gestures in terms of handshape: literal gestures are more likely to have marked handshapes than metaphoric gestures are, although both categories have predominantly unmarked handshapes (Open B and 5). Why would we observe a difference in handshape markedness? To understand this difference, we must first understand what types of information handshapes typically convey in iconic gestures. For this we can turn to insights from sign languages. In classifier constructions in languages such such as American Sign Language and Israeli Sign Language, the type of handshape used in the classifier conveys the size, shape, or semantic class of the incorporated noun (e.g., Supalla, 1982; Meir, 1999). For example, the F handshape can be used to represent a thin, circular object, whereas an upside-down V handshape represents a bipedal entity (Figure 35).

Figure 35. American Sign Language classifier predicates: (A) 'pipe', with an F handshape; (B) 'bipedal entity walking', with an inverted V handshape[56]

(A) PIPE 'pipe'



[56] (A) is from http://www.handspeak.com/word/list/index.php?abc=po&id=1665;
(B) is from http://www.lifeprint.com/asl101/pages-signs/c/clv.htm

(B) PERSON.WALK ‘bipedal entity walking’



Both of these are marked handshapes; in the context of classifier predicates, which rely on a strong link between form meaning to represent physical attributes of the represented nominal, it is the highly iconic nature of such handshapes that enables the link between form (two inverted fingers) and meaning (two walking legs) to occur. Furthermore, these verb classifiers have been argued to constitute noun incorporation constructions, with the salient characteristics of the noun represented by the handshape (Meir, 1999). In other words, in a complex verb classifier predicate, the handshape of the sign represents the entity; the movement parameter of the sign represents the action.

What are the implications of this relationship between handshape and the referent entity or event? As I argue in Chapter 3, in metaphoric motion the entity of the source domain is backgrounded. Although in order for a concept such as ‘prices’ to be reified (in order to be conceptualized as a physical object of motion), the size, shape, and other physical attributes of the metaphoric object are optionally underspecified. While specific elaborations of the general Location Event Structure Metaphor can certainly evoke more detailed information regarding the attributes of the metaphoric entity, those frame elements are *necessarily* evoked. For example, we can contrast a literal (1) and metaphoric (2) usage of the verb *fly*:

(1) *I fly out to Vegas on Air Force One.* [m.fly.6.26.2016a]

(2) *Twitter’s gonna fly.* (‘Twitter will react quickly to an event’) [m.fly.11.23.2014b]

In (1), the use of *fly* evokes a Flying frame; or more specifically, a Flying in a Plane frame. We can infer this because (a) the speaker (President Obama) is a human, and therefore unlikely to be evoking the Flying with Wings frame as if he were a bird or bat; and (b), he refers to *Air Force One*, the name of a particular plane that the President of the United States flies on. Hence, Obama’s use of *fly* evokes a translational motion frame that specifies the *manner* of flight: via airplane. In contrast, (2) does not evoke a Flying in a Plane or Flying with Wings frame, despite the use of the same verb. The entity in question is “Twitter”, a metonymic representation of the collective users of the social media website known for their rapid response to breaking news stories. While the individual users of Twitter are of course actual people with physical characteristics, here are they being understood as a collective Mass Entity named “Twitter” via the Mass-Multiplex image schema, wherein their individual attributes including physical ones are backgrounded. Thus, the metaphoric use of *fly* in (2) does *not* evoke a specific manner of motion of flight, and hence does not specify the physical attributes of a flying entity such as

wings. Given that handshapes typically represent entity attributes, and that metaphoric construal in these contexts does not evoke such details, it follows that the gesture will not represent detailed information in its handshape. There are simply no attributes of "Twitter" to be represented by the handshape in (2). Indeed, in this case the speaker transitions to an unmarked Open B handshape when he produces the co-expressive gesture stroke (Figure 36).

Figure 36. Use of unmarked Open B handshape. Left: prep; Center: onset; Right: stroke.



Twitter's gonna fly [m.fly.11.23.2014b]

4.d. *Gesture + speech results*

The foregrounding of the motion process and backgrounding of the moving entity is reflected in the gesture's stroke co-timing as well. 67.35% of literal gestures co-occurred with both the verb phrase and an argument (noun phrase, and/or prepositional phrase) in comparison to 43.85% of metaphoric gestures. In contrast, 28.57% of literal gestures and 42.31% of metaphoric gestures co-occur with the verb phrase only. Thus, metaphoric gestures are more likely to co-occur with the VP without a noun in comparison to the literal gestures, which are more likely to co-occur with both verbal and nominal phrases.

This combination of gestural imagistic content and linguistic content is what McNeill and Duncan (2000) term a "growth point"; it reflects the fact that speakers' mental representations are both imagistic and linguistic in nature. As such, the information and co-timing of speech and gesture will reflect both the schematized image (including metaphor-evoking image schemas) and the information packaging of the speaker's language. In English literal motion events, McNeill and Duncan find that speakers can either emphasize the manner of the event, in which case a manner gesture co-occurs with the verb; or they can downplay it by producing a path gesture synchronized with the path-conveying satellite and nominal ground.

Why, then, would English speakers produce metaphoric gestures on verbs? Given that I have argued that metaphoric gestures express path and not manner, the production of metaphoric gestures on verbs might be surprising. Since a literal gesture on a manner verb highlights manner and not path, why would speakers produce *path*-highlighting metaphoric gestures on verbs? The difference lies in the fact that McNeill and Duncan's participants were discussing literal motion events with highly salient *entities*. Not only were these literal animate entities, they were ones steeped in personality; by using animated cartoons as stimuli, they were ensuring that the characters in stimuli were full of expression and emotion. The well-known feud between

Sylvester and Tweety is driven by their fundamentally conflicting personalities which are reflected through the Canary Row stories. Thus, a gesture on the verb in the McNeill and Duncan data reflects the speaker's shift in perspective on the event from a focus on the event *entities* to the *manner of the actions* they are performing. By gesturing on the verb, the speakers are backgrounding the otherwise-foregrounded figure, and focusing on the manner of the motion event rather than the entity. In contrast, the speakers in this corpus are discussing metaphoric motion events with already highly backgrounded entities; their conceptualization of the event focuses on the motion event itself with little information as to the entity performing it. As I have argued in Chapter 3 and in this chapter, metaphoric motion backgrounds the mover entity role in the source domain and focuses on the process of motion. Hence, a gesture synchronized with the verb indicates this focus on the motion itself. Rather than reflecting an emphasis on manner, this metaphoric gesture growth point reflects the speakers' de-emphasis on the nominal elements of the event and focus instead on the action.

In addition to syntactic structure, we also observed differences between metaphoric and literal gesture with regards to the distribution of gesture manner between manner and path verbs. In particular, literal path gestures were less varied in form than metaphoric path gestures: the only manner types co-occurring with literal path verbs were *straight*, *arc*, and *circle.* However, when considering instead gesture image schemas, we no longer find a difference between metaphoric and literal data. We find instead that overall path verbs are more likely to accompanied by Path gestures than manner verbs are. However, a majority of metaphoric manner verbs are still accompanied by Path gestures. Just as we see an increase in unmarkedness in handshape form, we similarly see a preference for simple paths in metaphoric gesture manner. This appears to further reflect the emphasis on the source domain's translational motion, rather than the manner of the motion. For example, fictive motion as in (3) reflects mentally simulated motion along a path (Talmy, 2000).

(3) *the pipeline would run from Canada to the Gulf Coast* [m.run.11.15.2014a]

Whereas the pipeline itself is a static fixed object, its path from Canada to the Gulf Coast is described using the verb *run*. However, the speaker of this utterance is not implying that the pipeline is "running", but rather metaphorically construing the pipeline as moving along a path. Nonetheless, he still uses a manner verb to evoke this notion, supplementing the path with a source and goal in the adjoining prepositional phrases. As a satellite-framed language, English has a rich set of manner verbs for speakers to choose from, and fewer path verbs to do so. However, sometimes as in (3) they don't actually "need" a manner verb per se; they need a verb that evokes translational motion. As such, they can make use of manner verbs, and then evoke only those frame elements relevant to the metaphor in a partial mapping between source and target, wherein the manner of the motion does not map to the target domain. That partial mapping is reflected in the gesture's lack of manner beyond tracing a simple path, as exemplified by the gesture accompanying (3) in Figure 37.

Figure 37. Path gesture accompanying metaphoric motion verb.



the pipeline would run [from Canada to the Gulf Coast] [m.run.11.15.2014a]

Finally, we considered the interaction between gesture and speech meaning at a second level of analysis by looking at gesture image schemas and verb image schemas. Again, here it is critical to look at the meaning of the verb in specific context rather than relying on verb lemmas as a proxy for meaning, due their polysemous nature and the fact that speakers may be only evoking specific elements of the particular motion frame the lemma is associated with. We find the only meaningful difference in the co-occurrence of Cycle gestures with Cycle verbs; there are no other significant differences in the distribution of gesture image schemas between verb schema types. Referring back to Figure 30 in section 3.d, it is apparent that the other verb types are predominantly accompanied by Path gestures; the Cycle verb schema stands out with its preponderance of Cycle gestures. This would suggest that the differences between path and verb verbs in terms of gesture schemas is driven by the relationship between Cycle verbs and Cycle gestures.

In contrast to manner verbs such as *run*, *walk*, and *fly*, which specify *how* the motion is performed, or verbs like *spread*, *drive*, and *stumble* which in these contexts add *causal* elements to translational motion, the Cycle manner verbs *revolve* and *rotate* describe the *shape* of the motion's path. Namely, they both specify that the trajector is transcribing a circle. Hence, when speakers choose to use such verbs, they foreground this shape of the trajector's route. They are still describing translational motion, as the trajector does move through space, but with a focus on the circular shape of that motion. It appears that when speakers make use of the Cycle image schema, this focus on the circular aspect of the path is reflected in both speech and gesture.

4.e. *Image schematic structure*

Throughout this discussion I have emphasized the role of image schematic structure in our analysis. We have seen that apparent differences between metaphoric and literal utterances are resolved when analyzing the speech+gesture utterance package in terms of its image schematic structure(s) rather than relying on form alone. When comparing verb lemmas, we observed a significant difference. However, I have argued that lemmas are not an appropriate level of analysis due to their polysemous nature and the ways in which speakers make use of verbs to partially evoke frames, foregrounding certain elements and not evoking others. When instead

comparing both verb image schemas and those of verbal satellites, we found that metaphoric and literal utterances were comparable.

We observed similar results in terms of gesture meaning. Whereas many gesture annotation methods are moving towards understanding gesture form in terms of ever-finer detail (e.g., Brenger et al.'s (2016) 3-D motion capture heat maps of gesture stroke locations), these methods privilege form over meaning. Given that our focus is on gestures with perceptually accessible and salient meaning, our goal is to identify and categorize the form elements of these gestures that contribute to that meaning. Hence, whereas gesture manner form may initially seem to be a good measure of gesture meaning in this corpus, we find that gesture image schema produced more consistent and reliable results. Combining gesture and speech, we again found that comparing gesture image schemas with verb types showed no differences between metaphoric and literal data. Given the gesture manner vs. gesture image schema results, this level of analysis is preferable over the comparison of gesture manner. These results are all summarized in Table 8.

Table 8. Comparison of effects of utterance metaphoricity initially, and after considering image schematic structure.

Measure	**Initial results**	**Image schema results**
Speech meaning	Differences in lemma frequency	No difference in verb image schemas or satellite image schemas
Gesture meaning	Different types of gesture manner	No difference in gesture image schema
Interaction of speech and gesture meaning	Different distribution of gesture manner across verb types	No difference in distribution of gesture image schema across verb types

In sum, we find that although there are apparent differences between metaphoric and literal gestures, when using an appropriate level of analysis that reflects our natural conceptual categories – namely, image schemas – we do not find significant differences between the two in certain respects. When we do see differences – in handshape and co-timing – these differences are also reflective of our conceptual structures: the differences between literal and metaphoric usage of translational motion frames.

5. *Conclusions*

It is important to recognize here that I am not claiming that *all* information in the gesture can be broken down into a single image schema (or complex schema, as in the case of the Complex Path). Rather, this study represents a first attempt to identify some of the salient and dominant characteristics of metaphoric gestures, with the aims of uncovering systematicity in their use patterns. It is certainly true that the data in this study is more complex than is discussed here. For example, consider the example in Figure 38:

Figure 38. *Arc* gesture evoking a metaphoric boundary crossing.



when we [enter into] a trade agreement [p.enter.4.15.2015a]

In this example, the speaker uses a path verb, "enter" and a preposition "into", which combines the image schemas of Containment and Trajector-Landmark. The speaker produces an *arc* gesture, which comprises both a forward trajectory and a curvilinear path that traces an arc up and then down. The *arc* here reflects the semantics of both "enter" and "into": In his speech, President Obama refers to "we", meaning the United States; as its President, he is metonymic for the country and hence he positions the United States at his own body by starting the gesture close to his chest. He then moves his hand away from his body, indicating the United States metaphorically entering the trade agreement. Thus, the United States is conceptualized as a trajector entering into a bounded region (the locus of the trade agreement); this evokes the concept of "boundary crossing". When one crosses a boundary, one often has to step over a threshold or otherwise somehow cross a barrier. This boundary-crossing is reflected in the arc component of his gesture.

For this reason, this gesture would traditionally be categorized as a "manner" rather than "path" gesture (e.g., Kita and Özyürek, 2003; Kita et al., 2007). It has been argued that the information included to represent such a motion event (i.e., with the arc manner or without it) is a reflection of how the speaker conceptualizes the event, as affected by their language's syntactic packaging of event structure information (Kita and Özyürek, 2003). In other words, Obama's gesture is a meaningful reflection of his understanding of the 'entering trade agreement event', as influenced by his native English. It is notable that even though he uses a path verb in his speech, his gesture still typifies that of a speaker of a manner-oriented, satellite-framed language; this ability to package manner and path in a single gesture has been argued to be an effect of the way languages like English organize path and manner syntactically. In contrast, speakers of path-framed languages tend to produce the manner and path in separate rather than conflated gestures (McNeill, 2000; McNeill and Duncan, 2000).

However, my primary goal in this study is not to understand how event structure or syntactic structure affects gesture *form*, i.e., the specific articulatory movements created by the arm (or head, shoulders, eyebrows, etc.). Rather, it is to understand how conceptual structures interact with the *information* conveyed *via* the form of the gesture. In fact, it has been argued that these manner vs. path differences in gesture production are the result of online syntactic production rather than habitual conceptual event structure schemas (Kita et al., 2007). My purpose here is to understand how a speaker's *conceptual* structures – namely, the metaphors and the image schemas that comprise their source and target domains – interact with the gestures they produce.

For this reason it is important to focus on those specific elements of the gesture which reflect those concepts. In this chapter I have demonstrated that the image schematic information conveyed by the gesture's trajectory is the salient level of analysis in understanding the meaning of referential metaphoric gestures. It is true that the arc *form* of Obama's gesture, which reflects some nuanced element of metaphoric construal of trade agreements, conveys more information than a straight path gesture would. However, we can generalize across such variations in gesture to recognize that a straight gesture with the same source and goal loci in space, accompanied by the same co-speech, would still evoke the same primary Location Event Structure Metaphor. Ultimately, that level of analysis is our current object of inquiry. By generalizing across these variations in form in our focus on image schematic structure, we can identify the re-occurring patterns across many instances of metaphoric gestures.

While the current study has demonstrated the systematic relationships between image schemas in *co-expressive* gestures, it does have several limitations which bear acknowledgement. First, this corpus comprises co-expressive data only. This stems from an issue inherent to all studies that rely on searching databases of closed-captioned video. In order to find gestures evoking metaphoric motion, I had to rely on searching the closed-captions for instances of metaphoric language. Unfortunately, it is not yet possible to search the other way around, i.e. searching for gestures directly. As such, this corpus only includes cases in which the gesture is accompanying speech which is also metaphoric. Therefore our analysis does not consider gestures which reflect *different* metaphors than the co-speech, or gestures accompanying *non-metaphoric* co-speech (Cienki and Müller, 2008b). There are also limitations introduced by the *genre* of data in this corpus. A reader familiar with American culture will recognize some of the speakers in this chapter's examples: they are television personalities such as news anchors and commentators, talk show hosts, and politicians. These are individuals well-versed with the visual modality and know how to engage with the television audience; some of them are trained in how to gesture. In other words, they are not representative of your average American English speaker. To address these shortcomings of the current study, in the next chapter I discuss an experiment designed to elicit metaphoric gestures in semi-naturalistic discourse.

Chapter 5: An experimental approach to multimodal metaphoric utterances

1. *Introduction*

In the previous chapter, I explored the relationship between co-expressive metaphoric gesture and metaphoric speech. This compared metaphoric gestures to literal gestures by studying a corpus developed from a database of American television and established that metaphoric gestures convey the image schematic information from the metaphor's source domain. Briefly, image schemas are the experientially-based conceptual structures that represent our perceptual and motor processes (Johnson, 1987; see Chapter 1, section 3 for an overview). In Chapter 4 I discussed Cienki's (2005) work in demonstrating that referential gestures represent perceptually-available image schemas, such as a cyclic gesture that represents the Cycle image schema or a path gesture that represents the Straight image schema. I extended this work to the specific conceptual structures represented by metaphoric gestures and showed that we can best analyze the meaning of metaphoric gestures in terms of their image schematic content. However, we also recognized some limitations of the corpus methodology: (1) the data is constrained to co-expressive gestures, in which the metaphor is expressed in both the speech and gesture; (2) the gestures are produced by individuals trained in nonverbal communication, such as actors, talk show hosts, and politicians; (3) the data is limited to a pre-determined set of motion verbs.

To address these limitations, we can further explore the nature of similar multimodal metaphoric utterances with a complementary methodology, in which gestures are elicited in more naturalistic contexts. Experimental gesture studies, in which participants' gesturing is elicited – usually without actually instructing participants to gesture – has been used in a wide variety of contexts. For example, McNeill (1992) established the use of the popular *Canary Row* Sylvester and Tweety cartoons to elicit iconic gestures: participants watch cartoon videos and retell the stories, which elicit representational gestures as the storyteller participant describes the cartoon characters running in and out of buildings, climbing rain pipes, falling down stairs, and so on. Typically, these studies are run in dyadic contexts, where participants are paired and take turns as the "storyteller" and "listener"; frequently they are recruited together, such that the participants know each other. This is believed to improve the reliability of the data, as participants are more at ease with one another and their stories will be more naturalistic (Park-Doob and Stec, p.c.). Although cartoons are a common method, the stimulus varies depending on the study focus; for example, participants may simply be instructed to retell a personal story or describe an image in detail. This paradigm has been extended to studying everything from viewpoint in gesture and eye gaze (Parrill, 2009; Sweetser and Stec, 2016), non-representational interactive gestures (Bavelas et al., 1992), gesturing in second language acquisition (Gullberg, 1998; Stam, 2006), to grammatical aspect (Parrill et al., 2013). These represent only a small sample of the types of studies that make use of a stimulus to elicit gestural data.

As of yet only a small number of studies have explicitly elicited metaphoric gestures in naturalistic discourse contexts. Cienki (1998) videotaped students discussing morality and ethical issues as part of a larger project on morality; this resulted in the production of metaphoric gestures, due to the frequency with which issues of "goodness" are described in metaphoric contexts. For example, Cienki (1998) found the basic metaphor GOOD IS UP/BAD IS DOWN frequently occurred in the students' gestures, even when they were not producing the metaphor

in their language. A more targeted study of metaphoric gesture by Cooperrider and Núñez (2009) was designed to elicit spatiotemporal gestures by having participants study and then describe a large poster representing the time-course of the history of the universe. As a result, they were able to propose a five-way typology of American English spatiotemporal transverse gestures, and furthermore concluded that this quasi-experimental design holds promise in studying metaphoric gestures in well-defined semantic domains.

Despite the early promise of Cooperrider and Núñez's (2009) study, little has followed in terms of experimental gesture work. As they comment, this paradigm may be useful in studying "well-circumscribed" domains (Cooperrider and Núñez, 2009:198); the key issue here may the notion of a "well-circumscribed" domain. Time is perhaps particularly well-suited to this type of study, as the TIME IS SPACE metaphor and its variants can be represented visually via such cultural artifacts as timelines, clocks, and simply sequentially ordering images depicting the stages of an event. The visual accessibility of TIME IS SPACE has been well-exploited in metaphor studies beyond gesture, such as priming effects (e.g., Boroditsky, 2000) and fictive motion (Matlock, 2004). In Cooperrider and Núñez's (2009) study, they took advantage of this aspect of TIME IS SPACE in their use of the timeline poster, which positioned the birth of the universe with the Big Bang on the left side of the poster and the present day on the right. In contrast, it may be more difficult to make targeted elicitations of other metaphoric mappings that cannot be as easily represented without using metaphoric language itself as the stimulus.

Cooperrider and Núñez suggest that this quasi-experimental paradigm could be extended to "music, bodily experience, emotions, mental experience, and much else besides" (2009:198). Despite their optimism, the only following similar study also focused on spatiotemporal metaphor. Casasanto and Jasmin (2012) focused on eliciting both transverse and sagittal spatiotemporal metaphors, both by explicitly instructing participants to produce gestures and by using the familiar "storytelling" format. In their "storytelling" experiment, participants produced spontaneous gestures by retelling short stories both with and without metaphoric language (e.g., "before", "long ago" vs. "earlier", "many years ago"), finding that transverse gestures co-occurred with sequential language and sagittal gesture with deictic language.

The current study constitutes a first attempt to extend this paradigm to domains beyond conceptualizations of time. Given our focus throughout the dissertation on the Location Event Structure Metaphor, here we will investigate the spontaneous production of gestures representing the source domain of the LESM. In particular, this study will prompt participants to discuss emotions, one of the domains mentioned by Cooperrider and Núñez (2009); it will also elicit other types of states, such as quantity. We will focus on *changing* states, such as increasing happiness, decreasing prices, or gaining popularity. This focus derives from the CHANGE OF STATE IS CHANGE OF LOCATION variant of the LESM. Participants will be given short story vignettes to read and then re-tell to a conversational partner. Each story will convey some type of state change, such as changing stock market prices; half of the stories will have metaphoric language (prices *rose*) and the other half will have matched non-metaphoric language (prices *increased*). Given that *change of state* is conceptualized as a *change of location*, we can hypothesize that use of this metaphor will result in gestures representing *translational motion*. As an entity changes from one state to another, it is understood as changing from one location to

another – in other words, it is undergoing translational motion. Hence, we expect that this translational motion will be co-produced in gestures accompanying *change of state* speech.

Beyond some kind of translational motion, what other gesture features do we predict? As Cienki (1998) observes, speakers can produce the *same* metaphor in gesture and speech (co-expressive gesture); *different* metaphors in gesture and speech (complementary gesture); or a metaphor in *one modality* and not the other (complementary). As such, we expect participants may produce any of the above combinations. Based on the results of chapter 4, change-of-state LESM gestures will be synchronized with the linguistic locus of the metaphor: the VP. However, while this should hold true for *co-expressive* gestures, *complementary* gestures may differ in their co-timing as the speech will not have the same focus on motion. However, it is difficult to predict the *rate* of complementary vs. co-expressive gestures; it is possible participants will produce co-expressive gestures more frequently, given that maintaining a single metaphor is a lower cognitive load and therefore easier to conceptualize and produce. On the other hand, participants may be more likely to produce different information in speech and gesture when thinking in two different metaphors, as gesturing can lower cognitive load and help to explore different conceptualization strategies (Goldin-Meadow, 2005; Alibali et al., 2000; Kita, 2000). Finally, participants should produce a high rate of *unmarked* handshapes, as the metaphoric entities are backgrounded in the LESM.

Turning to speech, metaphoric speech should be heavily verbal, given the focus on state change; manner verbs will be accompanied by satellites (prepositions or particles), reflecting the emphasis on translational motion in the change-of-state LESM. When participants are re-telling stimuli with metaphoric language, they are likely to be primed by these metaphors and hence the overall metaphoric production rate should be higher. Furthermore, participants re-telling non-metaphoric stories should produce a greater variety of metaphors than those telling stories with metaphoric language, given that they are *not* being biased by the stimuli towards using LESM-evoking language. However, given that all the stimuli are state-change related, we expect that overall storytellers should predominantly make use of event structure metaphors.

2a. *Methodology*

Development of stimuli entailed writing short stories balanced for verb type (manner vs. path), aspect (perfective vs. imperfective), and metaphoricity (metaphoric vs. non-metaphoric). Metaphoricity was balanced by substituting key phrases throughout a metaphoric story with non-metaphoric language. To control for variation in emotional affect, the stories were normed for readability and change in mood. The study was run by myself and a research assistant, Emily VanLoo, on the UC Berkeley campus. Participants were undergraduate students who either received course credit or monetary compensation. Study trials were video-recorded with the participants' consent.

2.b. *Stimuli*

Sixteen short stories (average word count = 58.75) were developed for the purposes of this study. Each vignette was designed to relay a common life experience that would be either personally familiar to college student study participants, such as struggling in a college class or interviewing

for a job; or culturally familiar, such as a politician running for office or a rock band trying to gain popularity. In addition to being culturally appropriate, these life experiences were also chosen as they all entail some form of state change. This state change may be experienced by the individuals described in the story as they experience different emotions, or describe the changing state of abstract concepts such as prices or grades changing in value. Given the pervasive nature of the Event Structure Metaphor family in English, focusing on state change provides a rich target domain. Furthermore, it serves as a counterpoint to prior quasi-experimental studies investigating metaphoric gesture production, which specifically elicited spatiotemporal gestures (Cooperrider and Núñez, 2009; Casasanto and Jasmin, 2012). In contrast, this study covers a broader range of target domains – emotion, quantity, action – which all are commonly conceptualized in terms of metaphors related to event structure.

For each story, one metaphoric and one non-metaphoric variant were used, following similar methodology common to the experimental metaphor literature (e.g., Matlock, 2004; Parrill et al., 2013). In the metaphoric variants, three sentences described a change of state in terms of change of location. These metaphors were evoked by either manner or path verbs; each sentence used a different verb, but each story used either manner verbs only or path verbs only. For example, one story used the verbs *rise, skyrocket,* and *plummet* to describe changes in the stock market. All of these verbs highlight the upward or downward path of motion, and were used to evoke either INCREASE IN QUANTITY IS UPWARD MOTION (*rise*, *skyrocket*) or DECREASE IN QUANTITY IS DOWNWARD MOTION (*plummet*). Hence, all the salient metaphoric language in the story centered around the construal of changes in quantity as changes in vertical location, which is the dynamic variant of the primary metaphor QUANTITY IS VERTICALITY. In the corresponding non-metaphoric variant these expressions are replaced with a non-metaphoric phrase of similar intensity and meaning. The non-metaphoric variants corresponding to *rise*, *skyrocket*, and *plummet* were *increasing, rapidly improving,* and *drastically decreasing.* To further illustrate, in another pair of stories the metaphoric expression *sped into* was matched with the non-metaphoric phrase *quickly began*. By including the adverb *quickly*, this ensured that the relevant aspectual information of the metaphoric *sped into* was similarly conveyed by the non-metaphoric variant – i.e., that the action was begun at a fast rate. The complete set of stimuli is given in Appendix A.

The metaphoric variants were counter-balanced for verb type (manner vs. path) and all variants counter-balanced for grammatical aspect (past perfective vs. progressive past) to account for potential variation in gesture stroke duration (Parrill et al., 2013). This resulted in a 2 (metaphoric vs. non-metaphoric) x 2 (perfective vs. progressive) design overall and a 2 (manner vs. path) x 1 (metaphoric) design within the metaphoric variants (Table 1). Therefore, there were a total of eight stories written in the past perfective and eight in the past progressive. Four stories used metaphoric manner in the past perfective and four used metaphoric path in the past perfective; they were matched with the eight non-metaphoric past perfective stories. Four stories used metaphoric manner in the past progressive and four used metaphoric path in the past progressive; they were matched with the eight non-metaphoric past progressive stories.

Table 1. Stimuli balance design.

	Past Perfective		Past Progressive	
Non-metaphoric	8		8	
Metaphoric	4 manner	4 path	4 manner	4 path

A line-by-line example story is below in (1); text of all stories is provided in Appendix A. The non-metaphoric/metaphoric variants are noted in brackets. Note that in the study stimuli were presented in a regular paragraph format, rather than the line-by-line format below.

(1) a. Sunny was nominated for prom queen.
b. Her friends said her chances were [improving/rising].
c. Everyone would vote for her.
d. Her hopes were [dramatically increasing/soaring].
e. Then she tripped and broke her nose.
f. Her mood was [becoming negative/sinking].
g. She walked into the prom with a swollen, bruised face.

To ensure uniformity, each story follows the same formula (Table 2). After an introductory sentence in simple past, the stories alternate between invariant sentences (same for both metaphoric and non-metaphoric variations, in perfective aspect) and variant sentences (which contain either metaphoric or non-metaphoric language; the metaphoric language could be either manner or path verbs; and either perfective or progressive/imperfective aspect). The final invariant sentence concludes with an ambiguous ending. This ambiguous ending intentionally misdirected participants to believe the focus of the study was on how they interpreted the stories' conclusions, rather then on gesture.

Table 2. Stimuli format.

Sentence	**Content status**	**Metaphoric status**	**Aspectual status**
a	Introductory sentence	Non-metaphoric	Simple past
b	Variant	Metaphoric or non-metaphoric	Past perfective or Past progressive
c	Invariant	Non-metaphoric	Past perfective
d	Variant	Metaphoric or non-metaphoric	Past perfective or Past progressive
e	Invariant; plot twist	Non-metaphoric	Past perfective
f	Variant	Metaphoric or non-metaphoric	Past perfective or Past progressive
g	Invariant; ambiguous conclusion	Non-metaphoric	Past perfective

In Table 3, I illustrate the story format in Table 2 as it is implemented in the metaphoric variant of (1). The example in (1) uses progressive aspect and path verbs in its variant sentences. The target domains are quantity (1b) and emotion (1d, 1f). In this vignette, all three metaphoric expressions construe change of state as motion along a vertical scale. In (1b), the word *rising* conceptualizes an increase in probability as an upward movement, as an instance of the entailment MORE IS UP. In (1d), positive changes in emotional state are similarly construed via the word *soaring*. Notably, while *soar* may also convey manner, in this usage the prototypical path of *soar* rather than manner is relevant. In (1f), the downward entailment LESS IS DOWN describes a negative change in emotion state is is evoked by *sinking*.

Table 3. Stimuli format of an example metaphoric story stimulus.

Sentence	Sentence	Metaphoric status	Aspectual status
a	Sunny was nominated for prom queen.	Non-metaphoric	Simple past
b	Her friends said her chances were rising.	Metaphoric path	Past progressive
c	Everyone would vote for her.	Non-metaphoric	Past perfective
d	Her hopes were soaring.	Metaphoric path	Past progressive
e	Then she tripped and broke her nose.	Non-metaphoric	Past perfective
f	Her mood was sinking.	Metaphoric path	Past progressive
g	She walked into the prom with a swollen, bruised face.	Non-metaphoric	Past perfective

In (1), the event is one likely to be familiar to participants in this study, who were all of college age and therefore recently completed high school (i.e., the time of life when prom dances take place). Furthermore, most people have experienced an embarrassing moment. The inclusion of such relatable events helps to ensure participants are emotionally engaged in the story; heightened emotional arousal produces increased engagement on the part of both the storyteller and audience.

2.c. *Norming Study*

Prior to the gesture production study, a norming study was run on Amazon Mechanical Turk (mturk.com) to ensure consistency across the stimuli. 36 native speakers of American English (F= 20, M=16) ages 18-60 years old (mean = 35; 1 declined to state) were recruited online via the Amazon Mechanical Turk interface and consented to participate in the study. They were compensated $2.00 for their participation. The study survey itself was conducted via an online Google Forms survey interface with one stimulus presented per page. Participants were instructed as follows: “Read the following short paragraphs and answer the following questions to the best of your ability. Some questions you may need to guess the answers to; do not do any outside research. There are no wrong answers; please make your best guess.” They were then presented with each of the 16 stimuli texts; half of the participants read the metaphoric variants and half read the non-metaphoric variants. After reading a text, participants answered two norming questions on a 7-point Likert scale:

(a) Please rate how much the mood of the story changed over the course of the paragraph.
No changes in mood 1 2 3 4 5 6 7 Extreme changes in mood

(b) Please rate how easy this paragraph was to understand.
Very hard to understand 1 2 3 4 5 6 7 Very easy to understand

The summarized results of the norming study are presented in Table 4. The goal of the norming study was to ensure that stimuli were consistent both in terms of (a) range of emotional affect and ease of reading comprehension (b).

Table 4. Norming study results.

Stimulus	(a) Average mood change	(b) Average ease of reading
1. non-metaphoric	5.67	6.72
1. metaphoric	6.11	6.28
2. non-metaphoric	5.5	6
2. metaphoric	4.78	6.78
3. non-metaphoric	5.67	5.94
3. metaphoric	4.94	6.22
4. non-metaphoric	3.56	6.39
4. metaphoric	4	5.67
5. non-metaphoric	4.33	6.61
5. metaphoric	5.61	6.28
6. non-metaphoric	5	6.05
6. metaphoric	5.06	6.22
7. non-metaphoric	4.93	6.12
7. metaphoric	4.82	6.39
8. non-metaphoric	4.83	6.33
8. metaphoric	4.78	6.17
9. non-metaphoric	5.72	5.89
9. metaphoric	4.78	6.17
10. non-metaphoric	5	6.39
10. metaphoric	5.39	6.06
11. non-metaphoric	6	5.83
11. metaphoric	5.61	6.11
12. non-metaphoric	4.94	6.28
12. metaphoric	4.83	5.78
13. non-metaphoric	6.33	6.28
13. metaphoric	6.11	6.5
14. non-metaphoric	5.44	6.44
14. metaphoric	5.67	6.28
15. non-metaphoric	3.56	6.44
15. metaphoric	4.33	6
16. non-metaphoric	4.61	6.56
16. metaphoric	4.83	6.17
Non-metaphoric average	**5.07**	**6.27**
Metaphoric average	**5.10**	**6.20**
Overall average	**5.09**	**6.23**

Overall, the non-metaphoric and metaphoric variants were rated as being quite similar both in terms of emotional affect (average mood rating = 5.07 vs. 5.10) and readability (average ease of reading rating = 6.27 vs. 6.20). Given lack of order effects and the fact that participants and did

not see both variants of any given story, it appears participants evaluated each story independently.

Based on these results, two stimuli (stories 4 and 9) were chosen to be used as training stimuli, rather than test stimuli, for the production study (all the stories including the training stimuli are listed in Appendix A). These were chosen as they were towards the low end of the stimuli for average mood rating, with slightly greater differences between the two variants in comparison to the other stimuli.

2.d. *Participants*

Fifty-two (F= 25; M= 26; 1 declined to state) adult native speakers of American English ages 18-31 (mean = 20) participated in the experiment. All participants were recruited from the UC Berkeley campus community and were reimbursed either with psychology course credit or $5. After a participant was recruited, they were asked to find a friend who met the study qualifications to participate in the study with them. All participants consented to participation in the study including video and audio recording. They filled out a separate form to consent to the use of their video and audio data in various research presentation formats.

2.e. *Design and procedure*

Subjects participated in conversational dyadic pairs formed of two friends. Given that it has been shown conversational participants gesture more when they know each other, this requirement ensures that participants will be more at ease with one another and more likely to produce gestures. Pairs were seated facing one another, with two video cameras recording, one facing the pair and one set off to the side (see Figure 1). Participants were told by the researcher that they were taking part in a study on memory and storytelling. They were not told their gestures would be studied, in order to avoid influencing the participants' gestures by making them self-aware of their gesturing. After explaining the study topic and obtaining consent, the researcher explained the study procedures and conducted a practice trial using one of the training stimuli. Following the practice trial, the researcher exited the room and participants completed the test stimuli trials. After completion of the test trials, the researcher debriefed participants and explained the study would be focusing on participants' gestures. Consent was then again obtained, with participants given the opportunity to withdraw from the study and have their data immediately deleted. All participants declined the offer and re-affirmed consent following debriefing.

Figure 1. Dyadic participant configuration: Participant on the left is the current "storyteller" and participant on the right is the "listener". Story stimuli are stored in notebooks hung on the wall to encourage participants to keep their hands free while the story is being told.



[S1_11Y_A_10.9.15_EMS]

Each trial consisted of the following procedure: One participant, termed the "storyteller", reads a story stimulus, provided in a notebook. Participants were instructed to read the story thoroughly and to re-tell it to the best of their ability, but not to memorize it or attempt to re-tell it word-for-word. Then the storyteller recounts the story to the other participant (the "listener"). Both the storyteller and listener then fill out a short survey of two questions regarding the story. The questions were intended as distractors; in debriefing this was confirmed as most participants indicated they thought their answers to the survey questions, which focused on the ambiguous story conclusions, were the point of the study. Following the survey questions, the roles switch and the former listener takes a turn as the storyteller. The initial storyteller and listener roles were randomly assigned.

Participants completed a total of eight test stories, plus one of the training stories at the start. Each set of eight stories was randomly assigned and ordered. Each set included four metaphoric and four non-metaphoric stories. Twenty-six total dyads were run, for a total of 208 stories. Each story was told a total of 13 times, including both variants. Half of the metaphoric and half of the non-metaphoric stories were told seven times each, and the other half of each group six times; hence, overall there was an even number of metaphoric and non-metaphoric stories.

2.f. *Data coding*

Video and audio data from the head-on camera were used for purposes of analysis. When necessary, corresponding data from the secondary offset camera were consulted by the analyst for purposes of clarification. Recordings of each trial were split up in Quicktime (Apple Inc.,

Version 10.4) such that each story telling comprised its own separate file, for a total of 208 video files. The training trials were not analyzed, resulting in a total of 195 test videos.

Data analysis was performed in ELAN (Brugman and Russel, 2004) and focused on the events of the variant sentences in the stimuli. Those points in the story represent potential loci of salient metaphoric information, given that they comprise changes of state. In addition, focusing only on the story content serves to normalize the data across participants, as some embellished their re-tellings of the stories to include information not present in the stimuli.

Speech was first analyzed separately. Speech corresponding to the events of the variant sentences in the stimuli were identified. For example, in the story above (1), the analyst would focus on identifying content related to Sunny's changing chances of winning prom queen (1b), and her changing feelings regarding those chances (1d, 1f). Those three sentences – (1b), (1d), and (1f) – comprise those points in the story where a change of state is explicitly described. Once this speech was identified, the relevant speech was transcribed, time-aligned by word, and all usage of metaphoric language was coded. Instances of metaphoric language were coded by identifying source domain language, primarily in this data set motion verbs. In cases where the source domain language was a motion verb, the verb was tagged as a manner, path, or conflated verb if possible. It was also noted if the storyteller's word choice matched that of the story – they used the same verb or phrase as appeared in the story – or if they used different phrasing. Finally, if the relevant speech was not metaphoric, it was coded as non-metaphoric.

Co-speech gesture was analyzed following the conventions discussed in Chapter 4. Gestures were first categorized as primarily falling into one of the main conventionalized gesture types: iconic (narrative-referential non-metaphoric); metaphoric (narrative-referential metaphoric); or non-narrative (pragmatic and beat gestures). This categorization was performed in conjunction with the co-speech content, in order to determine whether or not the gesture was narrative-referential. Again, analysts focused on the portions of the story related to the variant sentences in the stories. While all gestures in the salient sections of video were categorized, only the narrative-relevant metaphoric gestures were analyzed in full. Hence, there were more overall produced metaphoric gestures than narrative-relevant metaphoric gestures.

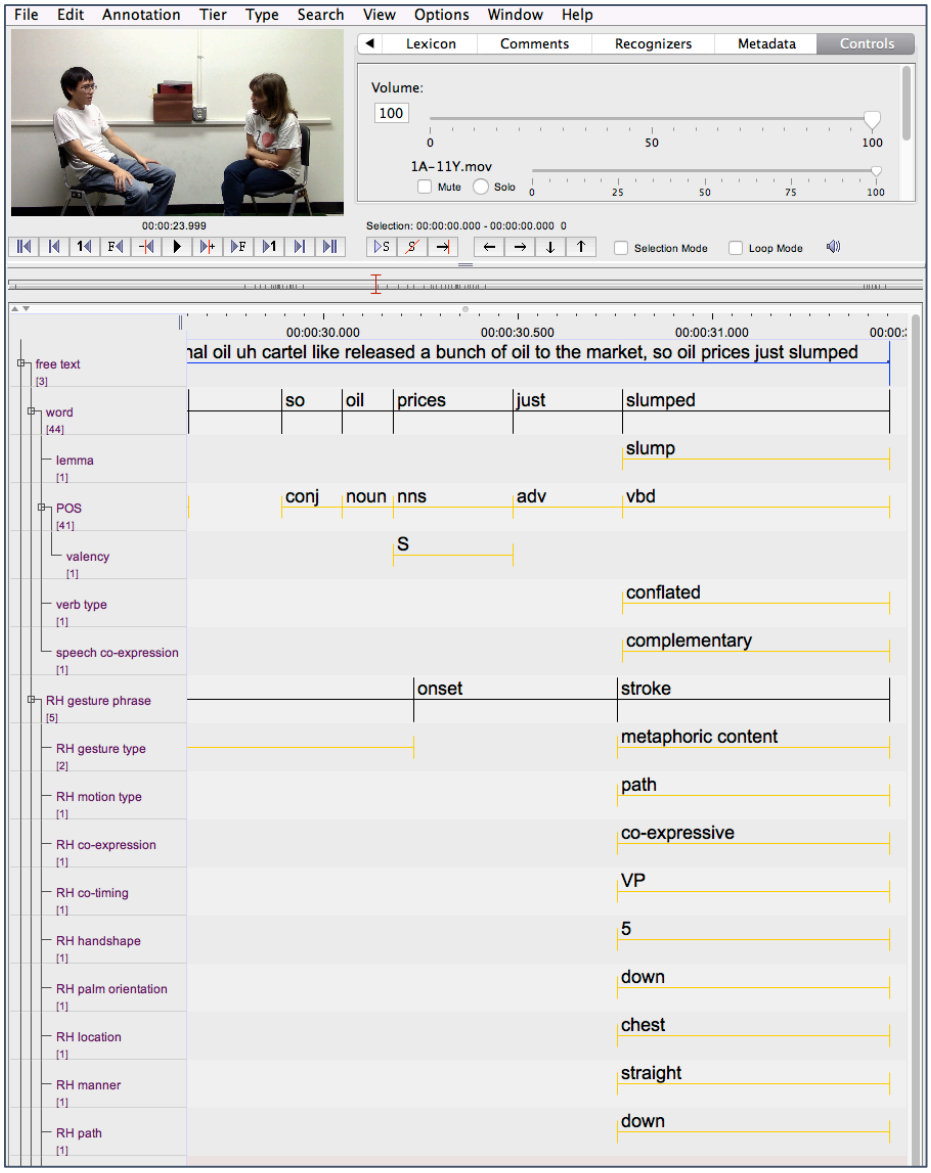
Following categorization, the stroke phase of relevant metaphoric gestures in the video clip were analyzed for form: handshape; location; palm orientation; manner; and path. In addition to gesture form, the syntactic structure of the speech co-timed with the gesture were coded. The metaphoric content of the gesture was analyzed as either *co-expressive* – conveying the same metaphor as the speech – or *complementary* – conveying a different metaphor than the co-speech (or any metaphor at all, in the case of non-metaphoric co-speech). Figures 2 and 3 provide an example of a complete analysis (Figure 3) for one salient utterance (Figure 2). Co-timing of gestures is indicated by brackets around the co-speech. Note that because here both hands are synchronized, only the right hand is annotated.

Figure 2. Example utterance.



so oil prices just [slumped] [S1_11Y_A_10.9.15_EMS]

Figure 3. Analysis of Figure 2 utterance in ELAN.



In the above example analysis, the storyteller retells a story in which the price of oil decreases due to the actions of an oil cartel (story 15). This story uses metaphoric manner verbs in the perfective: *shaken, stumbled, shook.* The storyteller produces one instance of narrative-relevant metaphoric speech: *oil prices just* ***slumped***. Here, the changing price of oil is conceptualized in terms of the metaphor QUANTITY IS VERTICALITY and its entailment REDUCTION IN QUANTITY IS DOWNWARD MOVEMENT. Furthermore, the particular verb *slump* in this context provides both path (downward movement) and manner ("slumping") of the change in price, so it is a *conflated* manner/path verb. This verb was not in the original story stimulus, so it is coded as speech that contrasts with the stimulus text. He produces two metaphoric gestures: first, the large quantity (*a bunch*) of oil released to the market is is conveyed by two hands sweeping laterally outward co-timed with *a bunch*, indicating a large expanse and conveying the metaphor QUANTITY IS SIZE, here showing that a large quantity of oil is a large area of the gesture space.

The second metaphoric gesture, co-timed with the verb phrase *slumped*, relates to the decrease in oil prices. The downward movement of the hands from chest to waist location is *co-expressive*, as it conveys the same metaphor as the co-produced speech. The downward movement evokes REDUCTION IN QUANTITY IS DOWNWARD MOVEMENT, and the range of movement conveys the extent of the price reduction. Here, the first metaphoric gesture would *not* be categorized as narrative-relevant because it is about the amount of oil released to the market, which is (a) not content from one of the variant sentences in the story and (b) not related to a relevant change of state. In contrast, the second metaphoric *is* narrative-relevant, as it is co-timed with relevant co-speech and conveys information regarding a relevant change of state.

Trials in which the storyteller did not produce metaphoric language and/or metaphoric gesture during re-telling of the salient story events were identified and discarded. (Note this does not mean the storyteller didn't use metaphor at all in their stories; rather, they didn't produce metaphoric language during the relevant points in the story.) In cases where participants held onto the story stimuli notebooks during storytelling, trials were discarded if the storyteller never gestured. If the storyteller gestured despite holding the notebook (such as holding the notebook in the non-dominant hand in the lap, while gesturing with the dominant hand), the trial was not discarded. 23 trials were discarded due to holding the notebook and 51 due to lack of metaphoric data; this reduced the total trials to 192 from 358.

3.a. *Results and discussion*

In this section, I first evaluate the experimental design by determining its efficacy in prompting linguistic and gestural metaphors, and the types of data it produces. To evaluate the experimental design, we should consider both the *quantity* of information produced by storytellers and the *type* of information conveyed. I describe the overall rate of production of metaphoric information in speech and gesture, and then discuss in detail the speech and gesture data separately, followed by the interaction between the two. Throughout I compare data produced by storytellers retelling metaphoric stories to data produced by those with non-metaphoric story stimuli.

3.b. *Overall results: Metaphoric production*

Given the short length of the story stimuli, speakers could produce at most 3 plot points per story. Over 26 trials with 7 stories per trial, this results in 546 possible data points. 23 stories were discarded due to participant error, reducing the possible data points to 523. All of these trials were discarded because participants held onto the notebooks (which held the story stimuli) rather than placing them in a box as instructed. This interfered with their gesture production, since their hands were occupied by the notebooks. Irrespective of metaphoric content, storytellers produced a total of 358 plot points, or 68.45% of the plot points. Of these, 56.42% (202) were annotated as containing metaphoric content. Storytellers produced a mean of 2.25 (median = 2.33) plot points per story and 1.14 (median = 1) metaphoric data points, or 55.04% of produced plot points included a metaphor. Storytellers' overall data production rate was quite broad, ranging from 25% to 100% of possible plot points (median = 75%). Their metaphoric production rate varied as broadly as possible, as 0% to 100% of the plot points they produced included metaphoric data (median = 54.55%).

39.11% of plot points (79) had both a linguistic metaphor in the speech and gestural metaphor (GM); 43.56% (88) had a linguistic metaphor without a co-speech GM; and 17.33% (35) conveyed a metaphor in gesture only. There was no difference in the number of co-speech linguistic metaphors produced with and without accompanying GMs (two-sided Exact Binomial Test, $p > 0.05$), but significantly more GMs were produced with a linguistic metaphor than without (one-sided Exact Binomial test, $p < 0.0001$).

How successful were the stimuli at producing metaphoric content? The gesture production rate can be compared to that in Experiment 2 in Casasanto and Jasmin (2012). They report that 28 speakers, each retelling two stories of slightly longer length (word count = 50-100), produced a total of 53 spatiotemporal metaphoric gestures, or 53 gestures in 56 stories. However, their dataset constitutes all metaphoric gestures produced in conjunction with temporal speech, whereas the current one only includes those gestures (a) produced during a relevant plot point and (b) conveying information salient to the plot point. For example, some speakers produced gestures indicating the spatial loci of story characters (e.g., one speaker positioned two candidates running against each other in an election on either side of his body) or sequential lateral gestures indicating the passage of time in the narrative, but those gestures were not included in this data.

Interestingly, Casasanto and Jasmin (2012) report a similar rate of gesture production for both non-metaphoric and metaphoric speech, whereas here gesture production is much higher during metaphoric speech. Hence, while it is not clear if the overall gesture production rate is higher here, in their study they had a higher success rate with regards to non-metaphoric speech. However, in their study, storytellers were told to retell the stories as close to verbatim as possible; thus, while speakers were not producing metaphoric *language*, they were still free to *think* metaphorically – and as demonstrated by their gesture production results, participants were still conceptualizing time using spatiotemporal metaphor. Thus it is not clear if their gesture production rate during non-metaphoric speech is in a sense artificially high; in other words, it could be that storytellers may have produced metaphoric speech given the opportunity, reflecting their current cognitive state. In contrast, in this study participants were expressly told to retell the story in their own words and discouraged from verbatim re-tellings in order to avoid this issue.

That said, storytellers *were* more likely to produce metaphoric data when the stimuli story was metaphoric (67 stories) than non-metaphoric (41 stories), and produced more metaphoric plot points when retelling metaphoric stories (72.83%) than when retelling non-metaphoric stories (36.76%), X^2 (1, N = 358) = 46.87, $p < 0.0001$. Although speakers were more likely to use metaphor in response to metaphoric stories, there was not an ordering effect, suggesting that storytellers were not influenced to use more metaphoric language as the study trial progressed (Kruskal-Wallis rank-sum test (df = 6, N = 358) = 0.6242, $p > 0.05$).

3.d.i. *Results: Speech*

Of the 202 data points with metaphoric information, 167 included linguistic metaphors. Storytellers were more likely to produce metaphors in their language when retelling metaphoric stories, reflecting the overall trend discussed above (one-sided Fisher's Exact Test, $p < 0.001$). Order did not affect linguistic metaphors production rate (two-sided Fisher's Exact Test, $p > 0.05$).

Of 167 linguistic metaphors, 62 *matched* the language used in the story – meaning the storyteller used the same lemma as in the story, although the exact wordform may have differed – and 105 *mismatched*, such that the storyteller used a different metaphor-evoking phrase. Significantly more LMs were mismatches (one-sided Exact Binomial test, $p < 0.001$). One concern with this design is that storytellers may be more likely to match stories early on as they become comfortable with the storytelling format, but order did not affect linguistic metaphor match/mismatch (two-sided Fisher's Exact Test, $p > 0.05$). Hence, storytellers' "creativity rate" – how likely they were to use a word not in the story text – didn't change over time. Of the non-metaphoric speech co-occurring with metaphoric gesture, 15 matched and 19 mismatched (*n.s.*, one-sided Exact Binomial test). Speakers were more likely to *match* their speech to the stimulus when re-telling metaphoric stimuli (Kruskal-Wallis rank-sum test (df = 1, N = 358) = 15.93, $p <$ 0.0001).

3.d.ii. *Syntax*

Utterances were coded as "verbal" if the part of speech of the primary wordform annotated as metaphoric or the wordform of the co-timed speech (in the case of non-metaphoric language) was a verb or adverb; they were coded as "nominal" if the word was a noun, pronoun, or adjective (see Tables 5 and 6). Results were largely verbal (96.8% of linguistic metaphors and 88.24% of non-metaphoric speech). This reflects the emphasis in the stimuli on state change and accords with findings that metaphoric state change in English tends to be verbal (Stickles and Dodge, 2016).

Table 5. Distribution of nominal/verbal and aspect by metaphoricity.

POS	Metaphoric speech	Non-metaphoric speech	Total
Nominal	7	4	**11**
Verbal	160	30	**190**
Simple present or past tense	102	13	**115**
Progressive	42	14	**56**
Infinitive	9	2	**11**

Table 6. Distribution of part of speech by metaphoricity.

POS	Metaphoric speech	Non-metaphoric speech	Total
Adjective	0	2	**2**
Plural noun	2	1	**3**
Noun	5	0	**5**
Pronoun	0	1	**1**
Adverb	7	1	**8**
Simple present verb	27	6	**33**
Past tense verb	75	7	**82**
Progressive present verb	42	14	**56**
Infinitive verb	9	2	**11**

3.d.iii. *Semantics*

Of the 167 linguistic metaphors, 68.27% (114) were motion verbs; of those, 41 were classified as path; 37 as manner; 27 as generic motion; and 14 as conflated path/manner. An additional 23 LMs were possession-related verbs ("have", "get"), while the remainder fell into miscellaneous categories. Table 7 lists the lemma frequencies for both matching and mismatching linguistic metaphors; lemma frequency is illustrated in Figure 4. "Match" and "mismatch" here are defined as the verb lemma matching or not matching the stimulus text, irrespective of morphological variation. Therefore, if the story text contained the phrase *prices were rising* and the storyteller said *prices were rising* or *prices rose*, their speech would be annotated as "matching". However, if the storyteller said *prices went up*, it would be annotated as a "mismatch", because they used different language than the story text. As Figure 4 shows, storytellers producing mismatched linguistic metaphors used many of the same words as can be found in the stories, which were chosen as they are frequently used metaphorically. They also commonly chose to use high-frequency light verbs ("go", "get"), which reflect the two basic event structure metaphors (respectively, the Location Event Structure and Object Event Structure Metaphors). Overall, speakers produced 54 types of lemmas.

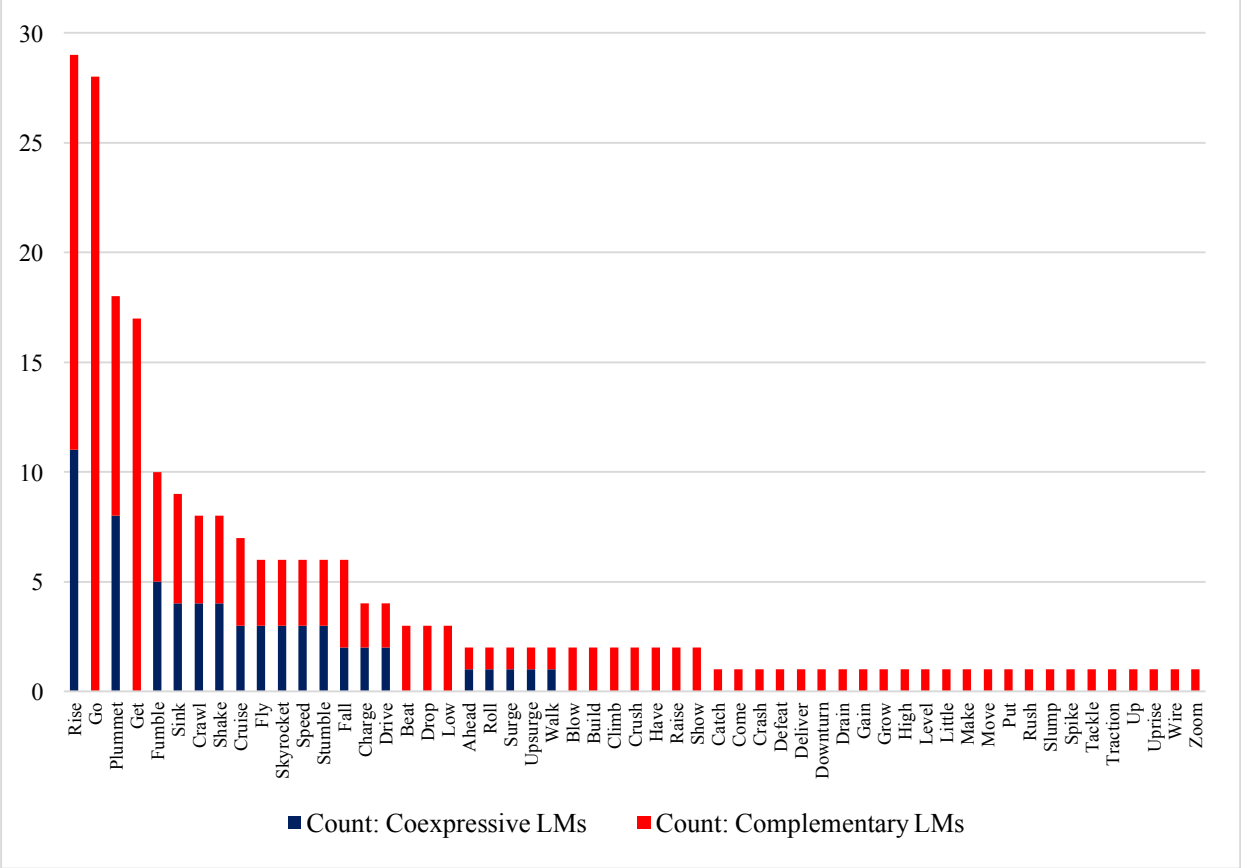
Table 7. Distribution of *match* and *mismatch* lemmas; N/A indicates a *match* is not possible as the lemma was not used in story stimuli.

Lemma	Matching linguistic metaphors	Mismatched linguistic metaphors	Total
Rise	11	18	29

Go	N/A	28	28
Plummet	8	10	18
Get	N/A	17	17
Fumble	5	5	10
Sink	4	5	9
Crawl	4	4	8
Shake	4	4	8
Cruise	3	4	7
Fly	3	3	6
Skyrocket	3	3	6
Speed	3	3	6
Stumble	3	3	6
Fall	2	4	6
Charge	2	2	4
Drive	2	2	4
Beat	N/A	3	3
Drop	0	3	3
Low	N/A	3	3
Ahead	1	1	2
Roll	1	1	2
Surge	1	1	2
Upsurge	1	1	2
Walk	1	1	2
Blow	N/A	2	2
Build	N/A	2	2
Climb	0	2	2
Crush	N/A	2	2
Have	N/A	2	2
Raise	N/A	2	2
Show	N/A	2	2
Catch	N/A	1	1
Come	N/A	1	1
Crash	N/A	1	1
Defeat	N/A	1	1
Deliver	N/A	1	1
Downturn	N/A	1	1
Drain	N/A	1	1
Gain	N/A	1	1

Grow	N/A	1	1
High	N/A	1	1
Level	N/A	1	1
Little	N/A	1	1
Make	N/A	1	1
Move	N/A	1	1
Put	N/A	1	1
Rush	N/A	1	1
Slump	N/A	1	1
Spike	N/A	1	1
Tackle	N/A	1	1
Traction	N/A	1	1
Up	N/A	1	1
Uprise	N/A	1	1
Wire	N/A	1	1
Zoom	N/A	1	1

Figure 4. Lemma frequency by co-expressivity.



The long tail of single-instance mismatched linguistic metaphors in Figure 4 further demonstrates the broad range of lexemes used by storytellers.

Speakers favored LESM-related metaphors in their speech. Both manner and path metaphoric stimuli elicited variants on the LESM, as did the non-metaphoric stimuli. Some typical complementary utterances – in which the speaker produced an utterance of their own rather than repeating the text – are provided in in the following examples. Source domain language is highlighted in italics.

The examples in (2) all make use of Manner of Motion verbs. Manner was very frequently used to refer to *rate* or *ease* of action. (2 a, b, c) all refer to the speed of the action; they make use of the LESM entailment RATE OF PROGRESS IS RATE OF MOTION. They all also make use of verbal satellites: *into, by, through,* and *ahead*. The use of these verb-accompanying particles is a further indication of their use of the LESM, as the prepositions all indicate various types of metaphoric spatial relationships between the figures (the individuals in the stories) and the grounds (the material, questions, or fellow candidates). Hence, these particles emphasize the *translational* nature of the motion; while the verb provides the *rate* or *ease* of motion, the particle provides the *forward* motion. Together they produce variations on PROGRESS IS FORWARD MOTION.

(2) a. she decides, like, to *rush into*, like, harder material [S11_12Y_A_10.29.15_ELV]
b. he *zoomed by* the questions [S16_8Y_B_11.10.15_ELV]
c. he *blew through* all the questions [S5_8X_B_10.20.15_ELV]
d. she was *cruising ahead* in the polls [S5_10Y_B_10.20.15_ELV]

In contrast to (2), the sentences in (3) all primarily make use of Path of Motion. The great majority of path-evoking utterances involve motion along a vertical path, as in (3 a-d). (3e) is very unusual in that it uses a path verb *came* that does not evoke verticality. Unlike the manner verbs in (2), when the path occurs in the verb as in (3a, b), it can occur without a satellite, as the verb itself provides the path of the motion. Forms of the generic motion verb 'go' as in (3 c, d) in conjunction with a path-bearing satellite such as a preposition (*up*) or adverb (*downhill*) also fall into this category; although the verb only conveys the notion of translational motion, the focus is on the directionality conveyed by the satellite. With the exception of (3e), path-focused utterances evoke Verticality metaphors, such as QUANTITY IS VERTICALITY as in (3a-c) or GOODNESS IS VERTICALITY as in (3d).

(3) a. so oil prices just *slumped* [S1_11Y_A_10.9.15_EMS]
b. the ratings for her opponent *rose* [S25_10X_B_12.3.15_ELV]
c. his test grades are *going up* steadily [S17_5Y_A_11.19.15_ELV]
d. and it *went downhill* from there [S10_1X_A_10.28.15_ELV]
e. his anxiety *came back* [S15_5Y_A_11.06.15_ELV]

While the majority of metaphoric data contained metaphoric language in the verb, occasionally speakers would make use of other parts of speech. (4a) uses a rather archaic noun *uprise*, whereas (4e) uses a highly conventional noun *level*. Whereas *uprise* makes use of a noun to indicate a changing location, *level* is more typical in that it indicates a steady point on a vertical state. Similar to (4e), the speaker in (4b) uses a noun *drain* to indicate a level on a vertical scale.

In (4c-e), the metaphoric speech occurs in adjectives. (4 c, e) evoke a verticality-based metaphor, whereas (4d) uses a non-event structure metaphor, QUANTITY IS SIZE, to describe the degree of popularity. (4d) and (4e) are particularly interesting in that the speakers use multiple metaphor-evoking content words in their speech. In (4e), the words *low* and *level* both use the same metaphor, QUANTITY IS VERTICALITY; whereas *level* only evokes the notion of vertical location, *low* specifies the relative location on that vertical scale. In contrast, (4d) makes use of two different metaphors with the same target domain – degree of popularity – to discuss two different aspects of that domain. The adjective *little*, evoking the non-LESM metaphor QUANTITY IS SIZE, describes the current state of the popularity; it does not imply a sense of changing state. In contrast, the verb *falls* evokes the LESM-derived DECREASE IN QUANTITY IS DOWNWARD MOTION (an entailment of QUANTITY IS VERTICALITY), indicating the changing state of the already-weak popularity.

(4) a. and now they're, like, on the *uprise* [S14_14X_B_11.06.15_ELV]
b. so her confidence is in the *drain* [S8_16Y_B_10.21.15_EMS]
c. her confidence is really *low* [S14_16X_B_11.06.15_ELV]
d. what *little* popularity they have *falls* [S15_14Y_B_11.06.15_ELV]
e. a kind of *low* popularity, um, *level* [S7_14X_A_10.21.15_ELV]

Finally, speakers also made use of non-LESM metaphoric language, as shown in (5).

(5) a. her confidence *grew* [S1_1X_B_10.9.15_EMS]
b. and then her confidence *builds* [S11_16Y_A_10.29.15_ELV]
c. she started to kinda *get* nervous [S12_7Y_B_11.03.15_ELV]
d. he *beat* the other candidates [S5_8X_B_10.20.15_ELV]
e. he was *crushed* [S15_9X_A_11.06.15_ELV]

These examples represent other common metaphors for state change. In (5a), the speaker makes use of the metaphor QUANTITY IS SIZE; this entails that CHANGE OF QUANTITY IS CHANGE OF SIZE. By making use of this perspectivized process variant of QUANTITY IS SIZE – INCREASE IN QUANTITY IS INCREASE IN SIZE – the speaker conceptualizes a person's confidence as a living entity which grows as it develops. (5b) effectively evokes the same metaphor, but with a lexical item that evokes specifics as to the nature of the metaphoric entity. By understanding confidence as a building, rather than a living organism as in (5a), the speaker still makes use of CHANGE OF QUANTITY IS CHANGE OF SIZE, but instead conceptualizes confidence as a structure that is "built up" over time. In contrast with (5a-b), (5c) is not about change in quantity, but rather change of emotion. The lexeme *get* evokes the dual to the LESM, the Object Event Structure Metaphor (OESM). In the OESM, states are understood as objects which can be acquired, held, lost, exchanged, and so on. In (5c), the individual "acquires" the state of nervousness.

(5d) illustrates the use of a non-state change metaphor. Whereas the original metaphoric story prompt describes this scene in terms of change of state (*he **flew past** the other candidates*), storytellers often focused instead on the end state of the event rather than the process itself. As such, they chose to use metaphors for this particular target domain – political elections – that describe the relationship between the winning and losing candidate. In (5d), this is realized as the

common ELECTIONS ARE PHYSICAL COMBAT, a special case of COMPETITIONS ARE PHYSICAL COMBAT.

(5e) is unusual in this category as it is not truly a non-LESM metaphor; rather, it comprises a blending of the LESM and another metaphor, EMOTIONAL HARM IS PHYSICAL HARM. The lexeme *crushed* evokes the frame of Physical Harm and hence the metaphor EMOTIONAL HARM IS PHYSICAL HARM. However, it is also the case that an inference of the Physical Harm frame is that a significantly-harmed entity is impaired in some fashion; *crush* in particular implies that the entity is so damaged it cannot move. This impediment to motion can also suggest an inability to change states; for example, *crushing poverty* is both metaphorically physical harmful and an impediment to an individual's ability to "pull" themselves out of the metaphoric poverty location – a change-of-state variant of the LESM. Thus, in (5e) a person who is emotionally "crushed" has suffered such harm that they cannot metaphorically move out of their damaged state.

Given the variety of metaphors illustrated above, we can further consider how metaphoric production was influenced by the different types of stimuli. The metaphors were grouped together into metaphor "families"; metaphors which are entailed by the same generic metaphor were grouped together. For example, both MORE IS UP and LESS IS DOWN were categorized under QUANTITY IS VERTICALITY. Table 8 summarizes the distribution of metaphor groups between stimuli types. Note that the table includes metaphors which were realized only gesturally, for comparison purposes.

Table 8. Distribution of metaphors between stimuli types.

Speech metaphor	**Manner stimuli**	**Path stimuli**	**Non-metaphoric stimuli**	**Total**
BODY IS A CONTAINER FOR EMOTIONS	0	0	0	**0**
COMMUNICATION IS OBJECT TRANSFER	0	2	0	**2**
EMOTION IS VERTICALITY	7	0	3	**10**
EMOTIONAL STATES ARE PHYSICAL STATES	0	6	1	**7**
KNOWING IS SEEING	0	0	2	**2**
Other	1	1	0	**2**
POLITICAL COMPETITIONS ARE PHYSICAL COMBAT	1	0	3	**4**
PROGRESS IS FORWARD MOTION	33	1	1	**35**
QUANTITY IS VERTICALITY	8	39	13	**60**
STATES ARE LOCATIONS	4	2	2	**8**
STATES ARE OBJECTS	4	7	4	**15**
Total	**58**	**58**	**29**	**145**

As Table 8 shows, LESM variants predominate: PROGRESS IS FORWARD MOTION, QUANTITY IS VERTICALITY, and its variant DEGREE OF EMOTION IS VERTICALITY combined comprise a full 105 out of 145 utterances. However, we also see some additional trends: EMOTIONAL STATES ARE PHYSICAL STATES predominantly co-occurs with path stimuli, whereas STATES ARE OBJECTS occurs throughout. Non-event structure metaphors POLITICAL COMPETITIONS ARE PHYSICAL COMBAT and KNOWING IS SEEING are more likely to be produced when the stimulus is non-metaphoric.

Given the preponderance of LESM-evoking speech, we can further consider how the LESM is distributed between the three types of stimuli (Table 9). The Verticality variants were combined into one category. There is significant difference in the distribution of the three types of LESM variants between the stimuli types (two-sided Fisher's Exact test, $p < 0.0001$). Stimuli with manner verbs were highly more likely to produce speech that used the PROGRESS IS FORWARD MOTION metaphor without evoking the Verticality frame, whereas stimuli with path verbs were highly more likely to elicit speech that used metaphors incorporating Verticality, particularly Motion Along a Vertical Path. Notably, non-metaphoric stimuli also preferentially evoked verticality-related metaphors over non-vertical LESM-evoking language. Speakers rarely produced LESM speech that did not incorporate motion. This stands to reason, as storytellers were discussing events of state change rather than stasis; non-motion STATES ARE LOCATIONS metaphors describe current states of being (e.g., *living in poverty*) as opposed to changing states (e.g., *falling into poverty*).

Table 9. Distribution of LESM metaphors.

LESM speech metaphors	**Manner stimuli**	**Path stimuli**	**Non-metaphoric stimuli**	**Total**
PROGRESS IS FORWARD MOTION	33	1	1	**35**
VERTICALITY METAPHORS	15	39	16	**70**
STATES ARE LOCATIONS	4	2	2	**8**

As shown in the examples in (2) and (3) above, participants produced a variety of motion verbs. However, as was also seen in Chapter 4, their use of metaphoric motion was heavily path-predominant. This can be seen in a tendency towards including path-bearing satellites in their utterances when the verb does not include path. 61 out of 111 motion verb-bearing utterances included a satellite (34 particles; 27 prepositions). 88.89% of generic motion verbs (n = 24) co-occurred with a satellite, as did 51.35% of manner verbs (n = 19). In contrast, only 27.27% of path verbs (n = 9) co-occurred with a satellite. Interestingly, none of the 14 conflated verbs did. In general, this indicates a heighted preference for including path information when the verb does not inherently evoke a path. In comparison, the conflated and path verbs were not typically modified by adverbs specifying manner of motion. All told, 81.08% (90 out of 111) of the motion verb utterances included some kind of path information, whether in the verb or satellite; 29.73% were path verbs (n = 33), showing that participants nonetheless produced more manner-bearing verbs overall.

3.e.i. *Results: Gesture*

Participants' metaphoric gesture production rate ranged from 0 to 3 gestures per story. Overall, the mean number of total gestures produced per storyteller was 2.49, with a median of 2. Excluding the seven participants who did not gesture at all during salient events, the mean gesture production rate was 2.92 total. Gesture duration also varied widely, from 0.16 to 3.47 seconds; the median and mean duration were respectively 0.58 and 0.70 seconds. Gesture duration did not vary based on the part of speech of the co-speech (Kruskal-Wallis rank-sum test ($df = 9$) = 5.79, $p = 0.76$). However, it does vary based on the image schema evoked by the gesture (Kruskal-Wallis rank-sum test ($df = 3$) = 33.75, $p < 0.0001$). Mean duration of cyclic gestures was 1.19 seconds, whereas discrete gestures averaged 0.53 seconds.

Slightly more gestures were judged to be co-expressive with speech (34) than complementary (24); however, this difference is not significant (Binomial Exact Test, $p = 0.24$).

3.e.ii. *Results: Interaction with speech*

Participants were no more likely to produce gestures when re-telling metaphoric stories than non-metaphoric stories (Kruskal-Wallis rank-sum test ($df = 1$) = 0.32765, $p = 0.57$). However, they *were* more likely to produce a gesture in co-occurrence with a linguistic metaphor (Fisher's Exact Test, $p < 0.0001$). When participants produced both a linguistic metaphor and gesture metaphor, they were no more likely to produce a *co-expressive* gesture than a *complementary* one (Binomial Exact Test, $p = 0.15$). The relationship between the linguistic metaphor and the story stimulus did not affect gesture production; participants were no more likely to produce a gesture when the speech matched the text than when the speech and text differed (Fisher's Exact Test, $p = 0.39$). Similarly, co-expressiveness of the gesture – whether or the not the gesture conveyed the same information as the speech – was not influenced by the speech matching the story text (Fisher's Exact Test, $p = 0.13$). Gesture production was also not influenced by the verb type of the metaphoric stories (Kruskal-Wallis rank-sum test ($df = 1$) = 0.36435, $p = 0.55$); nor was gesture co-expression (Kruskal-Wallis rank-sum test ($df = 1$) = 1.025, $p = 0.31$). In sum, the only significant influence on gesture production was the presence or absence of a metaphor in speech. Participants were more likely to gesture metaphorically when producing a metaphor in their speech as well.

When speakers did *not* include a metaphor in their speech but did in their gesture, their speech tended to still be verbal. Table 10 lists the non-metaphoric lemmas of the speech co-occurring with metaphoric gestures.

Table 10. Non-metaphoric lemmas co-occurring with metaphoric gestures.

Lemma	**# co-timed gestures**
Decrease	6
Do	5
Increase	5
Feel	3
Improve	3
Diminish	2
Ace	1
Be	1
Calamitous	1
Hope	1
Program	1
Quick	1
Teach	1

All but one of the lemmas that occur more than once – *decrease*, *do*, *increase*, *improve*, *diminish* – are non-metaphoric verbs evoking state change. The only non-state change lemma to co-occur with metaphoric gestures multiple times is *feel*, which may be due to the emphasis on the domain of emotion in the stimulus set.

Gesture production co-timing in this study reflected the findings of the corpus study in Chapter 4, as summarized in Tables 11 and 12. As shown in Table 11, gestures co-occurred with a variety of parts of speech; however, they are predominantly synchronized with verb phrases (73 out of 111 total).

Table 11. Gesture co-timing.

Part of speech of co-timed speech	**Count**
Clause	5
Object NP	4
Subject NP	9
Subject + VP	7
PP	4
VP	73
VP + Object NP	1
VP + PP	2
Other	6

Combining the parts of speech into nominal and verbal categories (Table 12), gestures are significantly more likely to co-occur with verb phrases than nominal phrases, both noun and verb phrases, or other parts of the utterance (typically, silence before an utterance or during a pause). This mirrors the co-timing results of Chapter 4, which found that metaphoric gestures were more likely to be co-timed with the verb phrase alone.

Table 12. Co-timing categories.

Category of co-timed speech	**Count**
Nominal	16
Nominal + Verbal	16
Verbal only	73
Other	6

3.e.iv. *Results: Gesture metaphors*

Gestures were annotated for metaphor according to the accompanying speech content. Gestures were determined to be metaphoric if the co-timed speech described the target domain of a metaphor, such as a change of state. This was evaluated irrespective of the metaphoricity of the speech; gestures could be annotated as metaphoric if they co-occurred with metaphoric language, or if they co-occurred with non-metaphoric langauge. For example, upward-moving gestures accompanying *prices rose* or accompanying *prices increased* would both be annotated as MORE IS UP, as the vertical trajectory of the gesture provides the source domain of the metaphor (Upward Motion) and the speech *prices* provides the target domain of the metaphor (Quantity). In the case of *prices increased*, the speech *increased* is non-metaphoric, but it describes a change of state (the change in price). Therefore the upward gesture is understood as referring to the change in prices, and hence is metaphoric. In the case of *prices rose*, the gesture is *co-expressive*, as both it and the speech convey MORE IS UP; in the case of *prices increased*, the gesture is *complementary*, as it conveys additional information (the metaphor) not contained in the speech.

Following annotation, the gesture metaphors were grouped together into metaphor families, as was done for the speech. Again, related metaphors, like MORE IS UP and LESS IS DOWN, were combined under their non-perspectivized variant, QUANTITY IS VERTICALITY. Table 13 reports the number of metaphor families expressed gesturally in manner stimuli, path stimuli and non-metaphoric stimuli contexts. Note that the table lists all the metaphors that were produced in speech for comparison purposes, but not all linguistic metaphors were realized gesturally.

Table 13. Metaphoric gesture counts as distributed by stimuli type.

Gesture metaphor	**Manner stimuli**	**Path stimuli**	**Non-metaphoric stimuli**	**Total**
BODY IS A CONTAINER FOR THE EMOTIONS	3	1	0	4
COMMUNICATION IS OBJECT TRANSFER	1	0	0	1
EMOTION IS VERTICALITY	0	11	9	20
EMOTIONAL STATES ARE PHYSICAL STATES	0	1	0	1
KNOWING IS SEEING	0	0	0	0
Other	1	0	0	1
POLITICAL COMPETITIONS ARE PHYSICAL COMBAT	0	0	0	0
PROGRESS IS FORWARD MOTION	18	6	8	32
QUANTITY IS VERTICALITY	6	14	18	38
STATES ARE LOCATIONS	5	6	5	16
STATES ARE OBJECTS	0	1	0	1

Table 14 summarizes the distribution of Location Event Structure Metaphor-evoking gestures. As with the spoken data, PROGRESS IS FORWARD MOTION was more likely to occur with Manner stimuli and QUANTITY IS VERTICALITY was more likely to occur with the Path and non-metaphoric stimuli. QUANTITY IS VERTICALITY was more frequent than PROGRESS IS FORWARD MOTION, and STATES ARE LOCATIONS relatively infrequent. Again the non-metaphoric stimuli evoked QUANTITY IS VERTICALITY more often than PROGRESS IS FORWARD MOTION. The fact that the gesture LESM results mirror the spoken LESM results suggests that storytellers were thinking using the metaphors and not just repeating the metaphoric language in the stimuli. If the storytellers did not make use of the metaphors at a conceptual level, we would not expect to see the metaphors produced in their gestures as well. Furthermore, storytellers frequently produced metaphoric gestures when re-telling non-metaphoric stories. This suggests that they were thinking metaphorically even when not primed by metaphoric stimuli to do so.

Table 14. Distribution of LESM gesture metaphors.

LESM gesture metaphors	**Manner stimuli**	**Path stimuli**	**Non-metaphoric stimuli**	**Total**
PROGRESS IS FORWARD MOTION	18	6	8	32
QUANTITY IS VERTICALITY	6	25	27	58
STATES ARE LOCATIONS	5	6	5	16
Total	29	37	40	106

Data in Figures 5-7 provide representative examples of complementary Location Event Structure Metaphor gestures. These gestures were typically produced when the storyteller was not making use of the LESM in his or her speech.

In Figure 5, the storyteller uses two metaphors in her speech. She conceptualizes the opponent's state in the political race as an object which can be acquired, evoking the Object Event Structure Metaphor with the word *gaining*. The viability of the political candidate's position in the race is described with the word *traction*. If there is increased traction between an entity and the surface it is moving on, the increased friction enables the entity to move faster and easier, with better control. This suggests the storyteller is *also* using the Location Event Structure Metaphor, as the political candidate's improving progress is understood as improving ability to move. If a political candidate is "gaining traction", they are acquiring the state that enables them to improve their position in the race by progressing more quickly. While the storyteller uses an entailment of PROGRESS IS FORWARD MOTION in her speech, her gesture evokes MORE IS UP (QUANTITY IS VERTICALITY) with an upward flick of her wrist. Combined, the speech and gesture provide information about multiple aspects of the event. The speech describes the state of the politician as both an object and location, and evokes ENABLING PROGRESS IS ENABLING MOTION. Her gesture, meanwhile, illustrates the increase in that ability by evoking MORE IS UP.

Figure 5. QUANTITY IS VERTICALITY gesture



so her opponent started gaining [more] traction [S24_10Y_B_12.2.15_ELV]

In Figure 6, the storyteller does not make use of a metaphor in her speech. She re-uses the text of the stimuli, *severely diminished*; her gesture traces a straight line downward. This evokes the metaphor LESS IS DOWN; in particular, EMOTION EXPERIENCE IS VERTICALITY and DECREASE IN EMOTIONAL STATE IS DOWNWARD MOVEMENT. Her production of the gesture suggests that while her speech was non-metaphoric, she was still conceptualizing the decrease in confidence metaphorically, in terms of verticality.

Figure 6. EMOTION IS VERTICALITY; QUANTITY IS VERTICALITY gesture

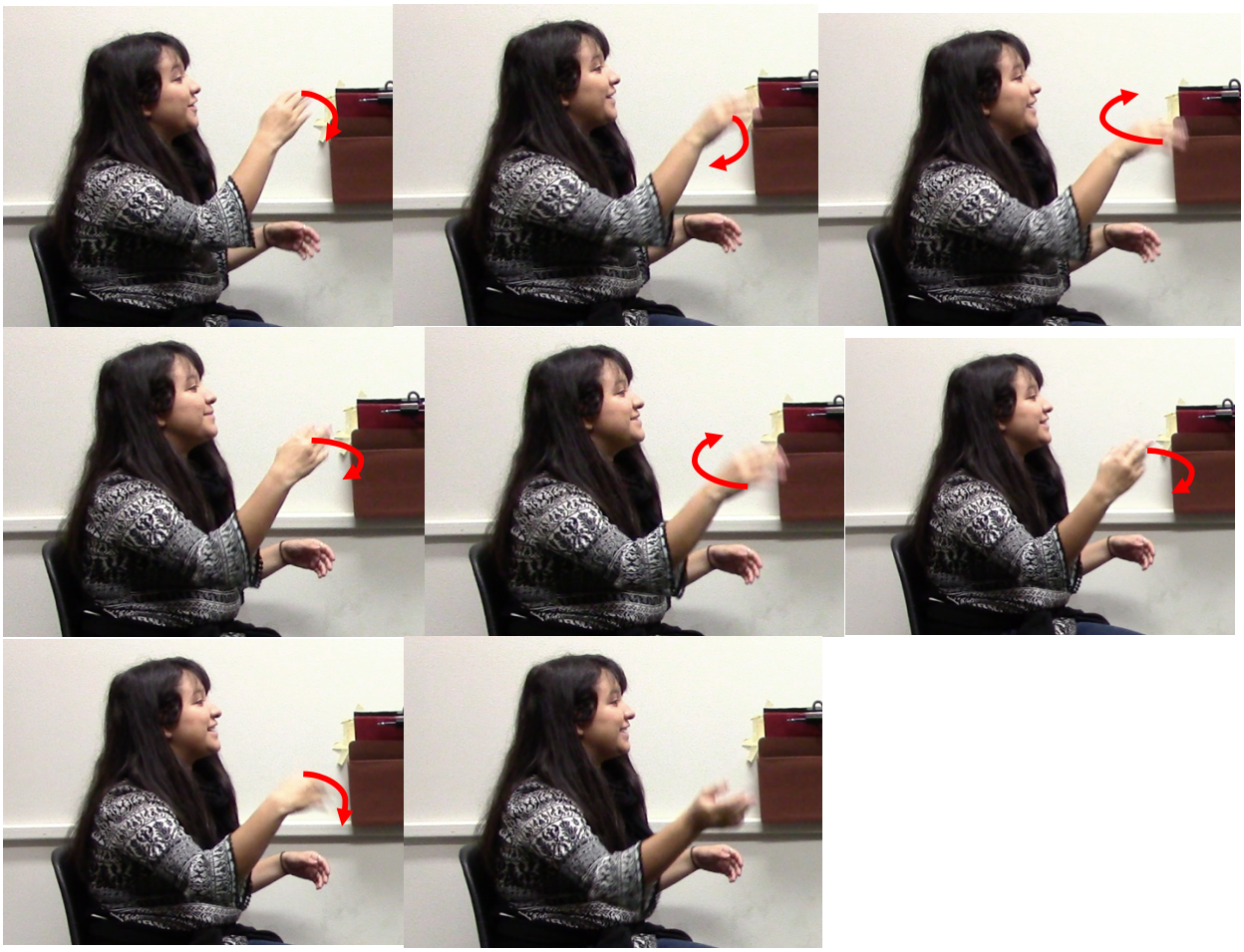


so her confidence um so it [severely] diminished [S26_16X_A_12.8.15_ELV]

Figure 7 illustrates a common cyclic metaphoric gesture. Cyclic gestures are often associated with ongoing events and have been analyzed as gestural correlates of progressive or continuous aspect (e.g., Hinnell, 2013; Parrill et al., 2013)[57]. Here the storyteller's speech describes the improvement in grades as *going up steadily*. This evokes both GOOD IS UP and MORE IS UP, as good grades are higher and a test score is generally a larger number. *Steadily* evokes a variant of STATE OF PROGRESS IS STATE OF MOTION: a steady motion is consistent and unwavering, so steady progress is consistent and continuous. Her gesture, co-timed with *steadily*, circles forward in a static position three times. It does *not* move forward along the sagittal plane but depicts continuous, cyclic movement, gesturally evoking CONTINUOUS PROGRESS IS CONTINUOUS MOTION. It is important to note that the trajectory of the circle moves forward rather than backward; given that THE FUTURE IS AHEAD and THE PAST IS BEHIND, a forward-moving gesture suggests PROGRESS IS FORWARD MOTION. Thus although the hand does not move forward away from the storyteller, its forward motion within the circle still evokes the central metaphor of the Location Event Structure Metaphor.

[57] Cyclic motion in the movement parameter is also common cross-linguistically in signed languages as a morpheme marking continuous or progressive aspect (e.g., Klima and Belugi, 1979).

Figure 7. CONTINUOUS PROGRESS IS CONTINUOUS MOTION gesture



his test grades are going up [steadily] [S17_5Y_A_11.19.15_ELV]

Figures 8-10 illustrate some of the non-LESM metaphoric gestures produced by storytellers. While most of the gestures were analyzed as evoking the Location Event Structure Metaphor as expected based on the story content, storytellers did produce non-LESM gestures.

Figure 8 is an example of common non-LESM metaphors, combined with a common LESM metaphor. The storyteller uses the Object Event Structure Metaphor in his speech as evoked by *getting*: the character in the story acquires the state of excitement. In his gesture, he touches his chest and moves his hand slightly upward as it opens from a Flat O grip handshape to a loosely cupped C handshape with an upward palm orientation. This complex gesture can be broken down into three metaphors. First, by touching his own body, his body becomes metonymic of the character's body, which contains her emotions; this evokes the BODY IS A CONTAINER FOR THE EMOTIONS metaphor, in which EMOTIONS ARE OBJECTS. Next, the upward movement of the gesture evokes the LESM variant INCREASE IN EMOTION EXPERIENCE IS UPWARD MOVEMENT: the upward-moving gesture reflects the increase in the character's excitement. Third, the changing handshape from a closed to open hand shows that as the character's excitement increases (the hand moving upward), it becomes more discernible. The contained emotion "object" held in the hand becomes visible as the hand opens up to reveal its contents, which is an example of the primary metaphor KNOWING IS SEEING.

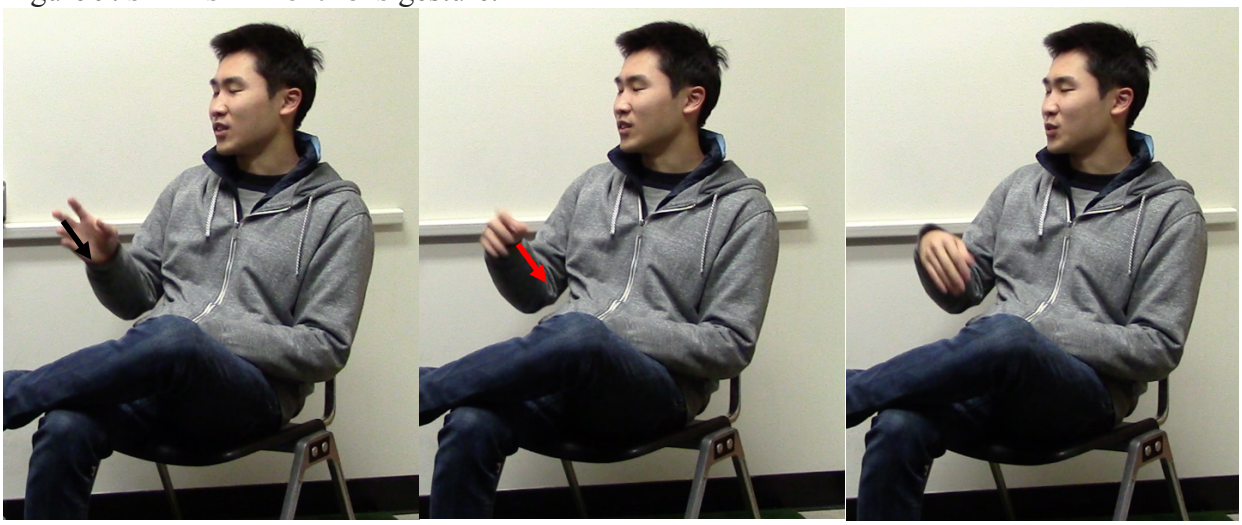
Figure 8. Multiple metaphors conveyed in one gesture.



so she was [getting] really excited [S16_1Y_A_11.10.15_ELV]

While Figure 8 illustrated a complex set of interacting metaphors, Figure 9 shows how non-LESM metaphors can co-occur in speech and gesture. In the storyteller's speech, the character's worries are *going away*. The state of worry is understood as an entity which is moving away from the character. This is an entailment of the Object Event Structure Metaphor: NO LONGER EXPERIENCING A STATE IS NO LONGER POSSESSING AN OBJECT. His gesture similarly "pushes" an object away from his body. His body is metonymic for that of the character in the story as he metaphorically moves the state of worry away from his self. However, the gesture adds causal information; in his speech, the worries are only "going away". By pushing the metaphoric object away from himself, he adds the force dynamics of causation: INTENTIONALLY NO LONGER EXPERIENCING A STATE IS PUSHING AWAY AN OBJECT. This also makes use of the entailment of the Object Event Structure Metaphor that NEGATIVE STATES ARE UNDESIRABLE OBJECTS: if something is an undesirable object, we want to get it away from us.

Figure 9. STATES ARE OBJECTS gesture.



yeah so his [worries] were going away [S26_5Y_B_12.8.15_ELV]

Third, Figure 10 illustrates another common metaphor for emotional experience, EMOTIONAL STATES ARE PHYSICAL STATES. The storyteller makes use of the Object Event Structure Metaphor with her speech *get nervous*: BEGINNING TO EXPERIENCE A STATE IS ACQUIRING AN OBJECT. Her flat horizontal hands alternate moving up and down, suggesting an unsteady surface. This evokes EMOTIONAL STATES ARE PHYSICAL STATES: EMOTIONAL UNREST IS PHYSICAL UNREST.

Figure 10. EMOTIONAL STATES ARE PHYSICAL STATES gesture



she started to kinda get [nervous] [S12_7Y_B_11.03.15_ELV]

Most of the non-LESM gestures were produced while storytellers were describing changes in emotional state, as illustrated above by Figures 8-10. Each of these is an example of a common non-LESM metaphor for the experience of emotion. Emotions are commonly understood as objects or physical states (Lakoff and Johnson, 1980; 1999). It may be that non-LESM construals of emotion are particularly well-suited for gesture due to the physicality of emotional experience; emotions are closely associated with bodily sensations. This gives rise to body-related emotion metaphors, which are readily represented in gestural form. Turning to handshape, we see a similar correlation with emotion-related metaphors.

Participants predominantly produced more unmarked (5 or Open B) handshapes than marked handshapes (Binomial Exact Test, $p = 0.0019$), a result similar to that of Chapter 4, which also found a preference for unmarked handshapes. However, when considering metaphor categories, an interaction between handshape and metaphor appears. Excluding cases where a metaphor family was only produced once, handshape markedness significantly differs between metaphor categories (Fisher's Exact Test, $p = 0.034$). Distribution of handshape types by category is shown in Table 15.

Table 15. Handshape markedness by metaphor category.

Gesture metaphor	Marked handshape	Unmarked handshape
BODY IS A CONTAINER FOR THE EMOTIONS	2	2
EMOTION IS VERTICALITY	12	8
PROGRESS IS FORWARD MOTION	7	29
QUANTITY IS VERTICALITY	14	24
STATES ARE LOCATIONS	5	7
Total	**40**	**70**

While PROGRESS IS FORWARD MOTION and QUANTITY IS VERTICALITY favor unmarked handshapes – as seen in Chapter 4 – other metaphors vary in their handshape type. BODY IS A CONTAINER FOR THE EMOTIONS and STATES ARE LOCATIONS are fairly evenly split between handshape types, whereas DEGREE OF EMOTION IS VERTICALITY actually has more marked handshapes. Curiously, this metaphor is a subcase variant of QUANTITY IS VERTICALITY. As I argued in Chapter 4, the high rate of unmarked handshapes reflects the backgrounded nature of the entity in the LESM, which is borne out in the results for the PROGRESS IS FORWARD MOTION and QUANTITY IS VERTICALITY data here. However, DEGREE OF EMOTION IS VERTICALITY has a very different type of Entity in the target domain than the other LESM metaphors here. Whereas Quantity has a purely abstract Entity, and the Action target domain evoked by Progress may have an Entity frame role of varying degrees of physicality, the Entity of the Emotion frame is not only an Animate Entity, but prototypically a Person. Furthermore, we experience emotions *physically* – the heat of anger and passion, the pain of depression – which leads to such metaphors as EMOTIONAL STATES ARE PHYSICAL STATES (Lakoff and Johnson, 1980; 1999). Thus, when discussing emotional states, the physical embodied nature of the emotion experiencer is a central aspect of the frame, unlike many other types of states. As the Entity role in this metaphor is not backgrounded in the same way as other LESM variants, the gesture handshape is more likely to represent some physical attributes of that entity, as in Figure 8 above. The change from one marked handshape to another – the Flat O to the loose C – indicated that the hand was "holding" an object, which became visible as the hand opened. By gesturing on himself, the storyteller was showing where the emotion is located – inside the body – reflecting the physical nature of the emotion experience.

3.f. *Results: Summary*

Overall, we did not observe an interaction between gesture production and the different types of story stimulus. Gesture production results generally accorded with those of the previous chapter: data strongly favored unmarked handshapes; cyclic gestures had longer durations than path gestures; and gestures were more likely to be co-timed with the VP. However, handshape type did vary by metaphor category. Gestures were more likely to be produced in conjunction with a linguistic metaphor, but otherwise gesture production was not influenced by the story stimulus or speech.

Speakers produced a wide variety of lexemes in their retellings of the salient story events, but favored verbs for state change in both their metaphoric and non-metaphoric utterances.

Metaphors in both speech and gesture were predominantly LESM variants, but a minority of metaphors either used the OESM or were not related to event structure.

4. *Conclusions*

In this study, we demonstrated the viability of this quasi-experimental paradigm in eliciting both novel linguistic metaphors and spontaneously produced gestural metaphors. Unlike prior studies in eliciting metaphoric gesture, the conceptual domains in focus centered around changes of state in an event structure context, rather than spatiotemporal metaphors. Whereas the spatiotemporal metaphors in prior studies were typically evoked by a small set of frequently-occurring lexemes, participants in this study made use of a wide variety of lemmas in evoking metaphoric language regarding state change. Nonetheless, the source domain word choice of storytellers largely revolved around a set of frames with related semantics, irrespective of whether they re-used the words in the stimuli or used non-matching language. Participants' stories made use of frames related to Motion and related spatial domains, particularly Verticality, reflecting the LESM metaphor family as discussed in Chapters 2 and 3.

However, participants did occasionally produce non-LESM linguistic metaphors. While the LESM is a central and pervasive metaphor family, there are nonetheless other common and salient metaphors, such as the OESM dual, available to English speakers when conceptualizing state change. The heavy tendency in this data towards the LESM may reflect in part the frequency and pervasive nature of the LESM, but also most likely reflects the overall priming effect of the stimuli. All participants re-told both metaphoric and non-metaphoric stories; hence, even while re-telling a non-metaphoric story, they may still have been primed to produce LESM utterances.

It is important to observe that participants' gestures were largely uninfluenced by the particular story stimulus, unlike their speech – which often matched the text of the story. This is a strong indication that storytellers were not only making use of the metaphor on a surface linguistic level – that is to say, repeating the metaphoric words but not actually *thinking* in terms of the metaphor – but rather were actually conceptualizing the target domain in terms of the source domain at a deeper cognitive level. The fact that that gestures were uninfluenced by the story stimulus, unlike their speech, shows that they were not just parroting the linguistic form of the metaphors or only making use of the metaphors that they read in the stories. Rather, they produced novel and different metaphors in their gestures. This shows that storytellers were *thinking* metaphorically, *independent* of the the priming effects of the stories. The phenomenon of metaphoric gestures has been taken as evidence for conceptual metaphor theory (e.g., Gibbs, 2011). This provides further evidence that speakers are indeed making use of these mental representations irrespective of shallow online priming effects. For example, as they talk about changes of state, they are making use of a system of conceptual metaphors which allow them to understand changes of state as changes of location. These domain-general cognitive structures are realized in their gestures, irrespective of either the stimuli they were reacting to or the linguistic content of their speech.

Whereas speakers often produced *co-expressive* gestures, in which the gesture and speech represented the same metaphor, they also produced *complementary* gestures, where the speech

and gesture conveyed different information. This is evidence that speakers are making use of multiple conceptualization strategies in understanding the same target domain. They also produced certain metaphors in one modality, but not the other. To an extent this may simply be a reflection of modality effects; it is easier to iconically represent in gesture the domain of Motion than it is to represent the domain of Seeing. One could point to the eyes, metonymically evoking the notion of vision, but it is less iconic to evoke a frame via deictic metonymy than to directly represent a frame-evoking element. Additionally, the metaphor KNOWING IS SEEING, as used by a speaker in this study, was evoked by the utterances *started to show returns* and *showed improvement*. These usages focus on a metaphoric conceptualization of change in which abstracted changes – i.e., returns on an investment and improvement in performance – are physical changes in appearance which can be visually accessed. Hence, the entailment of KNOWING IS SEEING at work is EVIDENCE OF ABSTRACT CHANGE IS VISUALLY ACCESSIBLE EVIDENCE OF PHYSICAL CHANGE. This has little or no relationship to the specific nature of eyes themselves, but only relies on the function of vision rather than the physical form by which we see. Thus, this particular metaphoric construal of change is unlikely to be realized in a gestural modality.

In contrast, some metaphors were only realized gesturally, not verbally. These are also evidence of the influence of modality on metaphor realization. In particular, BODY IS A CONTAINER FOR THE EMOTIONS was only produced in gesture. This is unlikely to be due to some difficulty in producing the metaphor linguistically. As Lakoff and Johnson (1999) discuss at length, the Container schema is central throughout metaphoric understanding of the mind and of mental experiences, including emotions. We can easily speak of happiness as something that cannot be *contained*, sadness as an object we *bury deep* within ourselves, and resentment an *infection* that *festers inside* us. However, participants in this study were strongly primed to produce LESM metaphors in their speech, perhaps at the expense of producing other metaphors for emotions. The BODY IS A CONTAINER FOR THE EMOTIONS construal of emotional experience tends to be more stasis-oriented than change-oriented, which may have also contributed to its disuse. However, the metaphor is *highly* accessible in the manual modality, given that the source domain frame of the metaphor – the body – is itself the modality that produces gestures. The body is physically present and easily accessed by the gesturer, who need only point to themselves to refer to it and evoke the metaphor's source domain.

Finally, the only significant influence on gesture production rate was linguistic metaphor production: if a speaker was making use of a metaphor in their speech, they were also more likely to produce a metaphoric gesture, *irrespective of whether or not the gesture and speech were co-expressive*.

There are a few potential explanations for this phenomenon, which I will discuss briefly but deserve further exploration. One reason may be simple individual variation: perhaps some individuals are more likely to produce metaphors in general, irrespective of the modality. Individual differences in metaphor production is an under-studied field, but there is some evidence that working memory influences metaphor production and processing (Chiappe and Chiappe, 2007), and various individual personality differences influence metaphor interpretation (Duffy and Feist, 2014). Thus, participants may have varied in how likely they were to retain

metaphors in memory during re-telling and understand the metaphors in the stories themselves. This in turn likely influenced their own metaphor production.

A second explanation may lie in the embodied nature of both gesture and metaphor, and will be discussed in the next chapter.

Chapter 6: Towards a Multimodal Embodied Construction Grammar

1. *Summary*

Throughout this dissertation, my goal has been to understand metaphor in usage, both linguistically and gesturally. Gesture provides a route to study both the cognition associated with language and the domain-generality of cognitive processes. While English speakers may be *producing* metaphoric manner verbs, are they necessarily *thinking* in terms of metaphoric manner? This is difficult to judging when looking at language alone. To answer this question, we turn to metaphoric gesture.

To that end, I first described a strategy for the formalization of metaphor which better elucidates the complex network of frames and frame elements that comprise the conceptual metaphor system. This hierarchical approach to conceptual metaphor theory, as instantiated in the MetaNet system, constitutes a major advancement in our understanding of the structure of metaphor. It enables via the MetaNet extraction engine the corpus analysis of metaphor in large bodies of text, such as the British National Corpus and the English GigaWord corpus. This new approach to metaphor in usage in turn allows for much more fine-grained statistical analyses, as demonstrated with the Location Event Structure Metaphor case study.

This detailed analysis of the Location Event Structure Metaphor in turn informed my study of gestural realization of motion metaphors. While iconic manner and path gestures have been extensively studied, little work has been done on their metaphoric counterparts. This dissertation fills this important gap by systematically investigating gestural expressions of manner and path in metaphoric contexts. I make use of both corpus and experimental approaches, and demonstrate the viability of the controlled quasi-experimental paradigm for eliciting both metaphoric speech and gesture. By establishing the possibilities of this approach, this research opens the door to further investigations of the interaction of gesture and other metaphoric domains.

The findings of this work focus on the role of image schemas in both the metaphoric source domain and the gestural expression of metaphor. I argue that image schematic meaning, rather than gesture form, constitutes the appropriate level of analysis for identifying common structures across individual gestures. Specifically, I find that path, and not manner, plays a central role in both linguistic and gestural realization of the Location Event Structure Metaphor. This is despite the privileging of manner in English lexicalization patterns. I also find that speakers are more likely to produce metaphoric gesture if they are producing metaphoric language, irrespective of gesture co-expressivity or complementarity. Metaphor tends to co-occur in speech and gesture.

Overall, I provide a usage-based approach to metaphor in language and gesture. Corpus data allows us to better understand how metaphors are actually used in speech, and which elements of metaphoric source domains are most relevant in metaphoric cognition. It also allows us to demonstrate generalizations in gesture structure, despite the idiosyncratic and holistic nature of gesture. While speech-content metaphoric gestures have been difficult to study due to their context-dependence, I have shown that they can be studied in both corpus and experimental approaches. By prompting study participants to think about semantic domains that we habitually understand in terms of metaphors, study participants produce metaphoric information

in both their speech and gesture. This demonstrates that metaphor is not only a feature of language and gesture, but a cognitive process that occurs even in the absence of linguistic priming. By studying the gestures people produce when thinking metaphorically and comparing them to those that they produce when thinking literally, we can understand differences in literal and metaphoric cognition. As I will argue in this conclusion, different conceptualization processes are at work when thinking about motion as a metaphor than when thinking about literal motion.

In the remainder of this chapter, I discuss the results of the dissertation in light of two theories relating language and cognition. The Gesture as Simulated Action framework argues that gesture is produced in conjunction with language due to the mental simulation of perceptual and motor imagery while speaking. The Thinking for Speaking Hypothesis more generally holds that cognitive processes are shaped by habitual use of language-specific patterns. Following this discussion, I review the notion of a multimodal construction and provide an Embodied Construction Grammar analysis of illustrative data from Chapters 1 and 4. I conclude that formalization of gesture representation is necessary to further our understanding of the interaction of language, gesture, and meaning.

2.a. *Theoretical implications: The Gesture as Simulated Action framework*

In the conclusion of Chapter 5 (section 4), I suggested there are at least two possible explanations for the relationship between metaphoric speech and gesture production. First, it may be that some individuals are more likely to maintain metaphoric representations in their working memory, such that metaphors they were previously exposed to (either in reading or hearing metaphoric stories, or their own prior narrations) were maintained. While they were primed to produce a particular metaphor in language by the story, these metaphors "carried over" in working memory and were then expressed gesturally. However, I did not find an effect of storyteller in gesture/speech mismatch rates. This does not preclude the possibility of individual differences in metaphor production (Chiappe and Chiappe, 2007), but suggests that more than individual differences may be at work. Therefore, I propose that we can understand these results when we interpret them in light of the Gesture as Simulated Action framework (Hostetter and Alibali, 2008) as it relates to contemporary conceptual metaphor theory.

According to the Gesture as Simulated Action framework, gesture production stems from the perceptual and imagistic mental simulations that underlie the embodied nature of cognition (Hostetter and Alibali, 2008). In this framework, gestures are produced when a gesture production activation threshold is lowered such that the mental simulation of action (whether by perceiving someone else performing an action, or performing the action ourselves) becomes a physical representation of that action in the gestural modality. Hostetter and Alibali (2008) posit that at least four different factors may contribute to the changing activation threshold; two are of relevance here. First are neural factors, such as the strength of connections between premotor and motor cortices; these variations in neural anatomy may explain individual differences in gesture production. Second is the social communicative situation. Speakers may produce more gestures in pursuit of a communicative goal, such as during instruction. In this study, storytellers were aware that they had to convey enough information for the listener to respond to comprehension questions at the end of the story. While the stories were all rated as having high readability

during the norming study, some storytellers may have felt greater communicative pressures than others, leading to a social context that engendered a higher metaphoric gesture production rate in order to better explain the story.

Third, Hostetter and Alibali (2008) identify a variety of cognitive factors. Gesture is known to interact with cognitive load. Goldin-Meadow (2005) and others have shown that gesturing lightens cognitive load and frees up resources for completing complex reasoning tasks. Hostetter and Alibali argue that conversely, maintaining a high activation threshold – inhibiting the simulated action from being produced as actual action – requires more cognitive effort. Thus, cognitively demanding tasks may result in lowering the gesture activation threshold in order to free up resources. In this study, we saw evidence that participants entertained multiple metaphors, producing different conceptualizations of the same target domain in their speech and gesture. Simultaneously maintaining multiple, sometimes contradictory, semantic domains will constitute a greater challenge than maintaining a single domain. Depending on the "difficulty" of the metaphor, metaphor comprehension may be cognitively taxing; for example, the processing of challenging metaphors increases right hemisphere activation, reflecting greater involvement of non-linguistic cognitive processes such as executive function (Prat et al., 2012). Metaphor production is similarly associated with additional activation of neural areas governing executive control (Benedek et al., 2014). Therefore, multiple metaphoric construals should produce an additional cognitive load. And so, when mentally "juggling" two or more source domains, this increased cognitive load may lead to a lower activation threshold and therefore result in gesture production.

Activating multiple source domains may allow the storyteller to understand the target domain in different ways, and thus explore alternative implications of the different conceptual structures afforded by the various metaphoric entailments. Goldin-Meadow (2003) has argued that the *speech-gesture mismatch*, in which the speech and gesture convey two different strategies or beliefs, can indicate a transitional cognitive state. In the context of learning, she has shown that mismatches index a "ready to learn" knowledge state, in which the learner is on the cusp of acquiring a new belief state. Similarly here we see gesturers in a cognitive state of multiple, alternative conceptualizations. Sometimes these can be mutually exclusive, as in when the speech makes use of the Object Event Structure Metaphor and the gesture the Location Event Structure Metaphor. Additionally, splitting metaphoric construals between the speech and gestural modalities allows the storyteller to convey different elements of the target domain by utilizing the features of each modality. For example, a storyteller might use the Location Event Structure metaphor in her speech and point to her heart to evoke THE BODY IS A CONTAINER FOR EMOTIONS gesturally. This makes use of the physical body to deictically index the source domain of the metaphor in a way that is not possible when evoking it linguistically.

Hostetter and Alibali summarize this part of their framework as follows: "a speaker's propensity to gesture is the product of (1) the amount of simulated action underlying his or her current thinking and (2) his or her current threshold or resistance to allowing this simulated action to be transferred into an overt motor plan" (2008: 505). If we consider the *embodied* nature of conceptual metaphors, the relationship between metaphor and gesture production becomes apparent. Hostetter and Alibali argue that the amount of *simulated action* underlying a person's current thinking contributes to the likelihood of gesture production because gestures are the

expression of perceptual and imagistic representations; the greater the mental simulation of perception and action, the lower the gesture threshold. According to the embodied cognition theory of metaphor, metaphors arise from these very simulations of perceptual and motor experience. Neural evidence has shown that processing of literal language and metaphoric usage of the same language activates the same premotor areas of the brain (Aziz-Zadeh and Damasio, 2008; Jirak et al., 2010). In other words, when we make use of sensorimotor information metaphorically, we still mentally simulate that sensorimotor experience.

Thus, when we are thinking metaphorically, we are increasing the current amount of simulated action in our thinking. In the Gesture as Simulated Action framework, this is one of the key factors that lowers the gesture activation threshold. This predicts that metaphoric language should lower the activation threshold and therefore, speakers using metaphoric language should also produce gestures. This is what I found in Chapter 5: the only significant factor that predicted metaphoric gesture production rate was linguistic metaphor production. This was true irrespective of whether gestures were co-expressive or complementary. In the case of multiple metaphor activation, we might expect this to produce an even lower activation threshold, due to the activation of multiple source domains and therefore greater sensorimotor activation. In theory, this should lead to a higher gesture production rate for complementary gestures in comparison to co-expressive gestures. This is difficult to accurately evaluate here, as the amount of sensorimotor simulation will vary depending on the particular activated source domain. Not all conceptual metaphors have source domains based in experiences that will translate readily to gesture. For example, MORAL DISGUST IS PHYSICAL DISGUST and UNDERSTANDING IS GRASPING both have source domains grounded in physical experience. However, the notion of Disgust does not have a direct manual correlate, and Grasping does. Thus while they both activate sensory and/or motor representations, the perceptual imagery of Disgust might not be expressed manually (although it might be realized in a disgusted facial expression). It may be that metaphoric motion is so heavily motoric that it effectively topped out storytellers' metaphoric production rates, and produced a ceiling effect irrespective of the number of metaphors a storyteller was thinking with. A study that elicits metaphors with varying degrees of motoric representation could test this prediction: storytellers primed with multiple motion-based metaphors should produce more gestures than those primed with non-motoric metaphors, as they have greater motoric activation and therefore a lower gesture activation threshold.

Additionally, if conceptual metaphors activate neural premotor regions just like comparable literal representations do, we might expect gesture production rates to be the same in matched metaphoric and literal contexts. However, Zima's corpus study (2014) found lower rates of gesture in conjunction with metaphoric expressions than with the same expressions used literally. In Chapter 5, I found that of the 202 produced metaphoric expressions, 56.44% expressed metaphoric information gesturally. Zima (2014) found metaphoric gesture production rates ranging from 37% to 83%, depending on the particular utterance. Considering the wide range of these production rates, there must be more factors influencing whether or not a gesture is produced during metaphoric cognition beyond simply the binary presence or absence of metaphor. Furthermore, there is some evidence that metaphor use does not activate neural premotor regions *just* like literal usage: sensorimotor activation decreases with increased conventionalization (Desai et al., 2013) and familiarity (Desai et al., 2011). Thus we might expect that highly conventional metaphors should actually be less likely to be accompanied by

gesture, given that the decreased sensorimotor activation levels will not lower the gesture activation threshold as much as novel metaphors will. Ironically, this runs somewhat counter to the conventionalization and entrenchment required of constructional analysis. If gestures are less likely to be produced in conjunction with conventional utterances, they are less likely to themselves become conventional. Given that studies of correlates of metaphor production at both the neural and gestural levels are still in their infancy, much remains to be done to address these contradictory predictions.

2.b. *Theoretical implications: Thinking for Speaking Hypothesis*

Whereas the prior section focused on applying the Gesture and Simulated Action framework to the results of this dissertation, this section will discuss implications of these results for the Thinking for Speaking Hypothesis (Slobin, 1987, 1996). This hypothesis argues that our habitual language use shapes the way we think, particularly while we are speaking. For example, because English speakers must frequently reference manner of motion when talking about motion, we should habitually attend to manner of motion; meanwhile, because Spanish grammar includes obligatory aspectual morphology, Spanish speakers should be more likely to attend to the aspectual structure of events. Furthermore, the elements of experience encoded in a language's grammatical categories will lead speakers of that language to privilege those elements linguistically. While Hebrew speakers can include aspectual information periphrastically (e.g., with adverbs), because it is optional and not grammaticalized, they are less likely to do so than Spanish speakers, who will do so *even* when it is optional. These tendencies have a snowball effect: a language with a rich manner lexicon will add even more manner verbs as its speakers will particularly attend to manner, discriminate finer distinctions in manner, and then add more words to describe those finer distinctions (Slobin, 2006).

Prior gesture studies investigating the Thinking for Speaking Hypothesis have focused on literal manner and path of motion. Kita and Özyürek (2003) demonstrated that language-specific differences in English, Japanese, and Turkish influence gestural expression of motion events. As a speakers of a satellite-framed language, English speakers were more likely to encode both the manner and path of a ball rolling into a bowling alley with a single conflated gesture, whereas as speakers of verb-framed languages, the Turkish and Japanese participants produced a higher rate of manner-only or path-only gestures. The separation of manner and path into separate gestures by the Turkish and Japanese speakers reflects their languages' separation of manner and boundary-crossing paths into separate clauses, and the combination of manner and path into a single gesture by the English speakers reflects their language's combination of the two in single clauses.

However, it is not the case that more frequent manner in language simply translates to more manner in gesture. Duncan (2001) shows that speakers of English, Spanish, and Chinese actually produce manner in gesture at about equal rates; the difference is in which contexts and combinations of other elements of the scene (e.g. figure, ground). She argues that these differences entail *joint highlighting*, in which the speech and gesture together foreground particular aspects of the scene rather than a *compensation* model in which the gesture makes up for the lack of information in the speech. She concludes that variation between languages' manner conceptualization is "variation with respect to the dynamic interplay of the various

components of motion" rather than "less versus more" manner (Duncan, 2001: 369). This argument is commensurate with work by McNeill and Duncan (2000), which showed that English speakers include manner of motion in their gestures only when it is a salient aspect of the scene and Hickmann et al. (2013), who found that English speakers gesture mostly in relation to path, but also conflate manner and path in their gestures. These findings collectively suggest that while English speakers may frequently make use of manner of motion in their speech, they are freer in their choice of manner expression in their gesture; they do so when it is perceptually salient. Hickmann et al. summarize this pattern thusly:

> "This phenomenon allows English speakers to downplay Manner and/or to highlight Path in speech by gesturing about Path instead (Brown and Gullberg 2008). Arguably, in V-languages speakers can (and frequently do) downplay Manner by simply omitting it, given the tendency to express Manner peripherally (e.g. *en nageant* 'by swimming'). The situation is different in S-languages where Manner is expressed in the main verb (e.g. *he swam*…) so that one way to shift focus away from Manner in the clause is to gesture about Path. Crucially, backgrounding Manner and/or foregrounding Path is achieved by aligning Path gestures with spoken Path elements…The choice to background or foreground Manner is presumably a pragmatic one, not necessarily guided by linguistic structures per se, but rather by the communicative situation…Aligning Path gestures with Path speech seems to be a default. In contrast, conflating MP in gesture in English…is not even the predominant pattern in adults. Foregrounding Manner and Path in gesture thus seems to be done by choice." (Hickmann et al., 2013: 149-150)

In Chapters 4 and 5, I show that English speakers typically downplay manner and foreground path gesturally, reflecting the findings by Hickmann et al. (2013). When they do include manner, it is highly salient, such as in the Cycle-evoking gestures; this accords with Duncan and McNeill's (2000) observations. Therefore, these results extend prior findings in the study of literal event construal in gesture to metaphoric event construal in gesture. Just as English speakers must frequently *speak* about literal manner but can *choose* whether or not to gesture it, English speakers can choose whether or not to gesture about metaphoric motion. Hickmann et al. (2013) suggest that this choice is a pragmatic one rather than a linguistic one, as it is determined by communicative need. In the case of metaphoric event structure, it is perhaps more appropriate to consider it a semantic "choice" in the sense that it is driven by the conceptual structure of the metaphor. The relevant elements of the source domain frame are activated by the metaphor; those elements are available for linguistic and/or gestural expression. In speech, it sometimes is the case that they are evoked by lexemes that can also evoke other semantic information. Manner verbs *can* evoke manner of motion frames, but they may be used to evoke motion frames that do *not* specify manner: in *time flies*, *fly* evokes speed of motion but not air flight. In gesture, manner is more easily optional and therefore simply omitted completely in such cases.

Importantly, focusing on metaphoric data strengthens the evidence for the Thinking for Speaking Hypothesis. A weakness of the literal motion evidence is that it relies on the argument that English speakers, who must attend to manner as speakers of a satellite-framed language, can "choose" to ignore manner gesturally when it is not salient. Thus while manner is obligatorily part of their conceptual structure due to its lexicalization in English, they can "turn off" its

expression in gesture when it is backgrounded. However, in metaphor, we have seen that the conceptualization of manner is *not* obligatory. Because metaphoric mappings between source and target domain are partial, only the relevant elements of the source domain frame are conceptually active (Grady, 1997). We saw evidence for this in Chapter 3, wherein usage of the the Location Event Structure Metaphor in the Corpus of Contemporary American English backgrounds manner information, even when manner verbs are used. Manner verbs frequently occur in conjunction with path-bearing prepositional phrases and in constructions that downplay the role of the entity – whose features are responsible for determining the manner of motion – in the event. Furthermore, I find that English speakers make frequent use of directional motion verbs in metaphoric usage, despite speaking a language that lexicalizes a preference for manner verbs. Therefore, when producing speech reflective of metaphoric reasoning – "speaking for metaphoric thinking" – English speakers habitually make use of *path* rather than manner-oriented structures.

This evidence from English metaphoric motion usage patterns, combined with the preference for path gestures in English metaphoric gestures found in Chapters 4 and 5, suggests that English speakers are typically simply not conceptualizing manner of motion when thinking with the Location Event Structure Metaphor. Thinking about metaphoric motion is not the same as thinking about literal motion. Therefore, thinking for speaking metaphorically is not the same as thinking for speaking literally. We see evidence of this in the path-predominant gestures English speakers produce when thinking, and sometimes speaking, metaphorically.

Having discussed implications of the results of Chapters 3, 4, and 5 for models of language, cognition, and gesture, I now turn to discussing how to represent them in a formal framework.

3.a. *Multimodal construction grammars*

The notion of a "multimodal construction grammar" or "multimodal grammar" has begun to receive interest in recent years. In 2012, Steen and Turner laid out a broad vision for the possibilities of a multimodal construction grammar that incorporates visual imagery, non-linguistic audio, and joint attention as well as gesture into its representation of a communicative act. However, as Schoonjans et al. describe, little work has done thus far to establish the theoretical foundation necessary for the notion of multimodal constructions:

> "La systématicité des co-occurrences, ou liens conventionnels entre éléments verbaux (de nature lexicale et/ou grammaticale) et les patterns gestuels, n'a pas fait à ce jour l'objet d'une étude empirique systématique et requiert un fondement théorique… en étude gestuelle, les analyses allant au fondement conceptual de cette relation geste-mot sont plutôt rares." (Schoonjans et al., 2016: 36)
>
> 'The systematicity of co-occurrences, or conventional links between verbal elements (of a lexical or grammatical nature) and gestural patterns, have not thus far been the object of a systematic empirical study and require a theoretical foundation… in gesture studies, analyses of the conceptual foundation of the gesture-word relationship are quite rare.' (Translation my own)

Whereas much work has been done on the interaction of speech and gesture, there are few studies that seek to empirically establish the *systematicity* of these interactions. As constructions are *familiar* form-meaning pairings that reoccur with "sufficient frequency" (Goldberg, 2006: 6), for a gesture/speech combination to be considered a "construction", the regularity of that combination needs to be demonstrated.

In their studies of German and Dutch, Schoonjans (2014) and Schoonjans et al. (2016) work towards establishing how the gestural affiliates of specific words contribute to their semantics and pragmatics. For example, in their corpus, German *einfach* 'simply' co-occurs with a palm-up open hand gesture in over 72% of utterances (Schoonjans et al., 2016). This illocutionary gesture has been analyzed as "presenting" information or evidence. They argue that this and other frequent gestural correlates of *einfach* support the conclusion that the word is an evidential particle. In another study of multimodal constructions, Zima (2014) focuses on English motion events, as I do in this dissertation. Her corpus study analyzed gestures accompanying four speech constructions: [*V(motion) in circles*], [*zigzag*], [*N spin around*], and [*all the way from X PREP Y*]. She found recurrent gestures, such as cyclic gestures in conjunction with *V(motion) in circles*, occurring with 60% to 80% of the data. Notably, she found lower rates of occurrence with metaphoric utterances than literal ones, but overall demonstrated consistent relationships between gestures and lexical affiliates. However, Zima (2014) also raises the concern that it is not clear what the threshold for constructional entrenchment is. Gestures don't co-occur with speech 100% of the time, so a demonstrated 100% rate for a speech-gesture pairing pair is highly improbable. Zima argues that these lower 60%-80% rates are not necessarily signs that no entrenchment has occurred; given that speakers are known to adapt and change linguistic constructions to conversational contexts, it stands to reason they can do so with gestural constructions as well. She argues that these rate are sufficiently high to be taken as evidence of some degree of conventionalization and entrenchment, defending the notion of a multimodal construction. In their study of gestures associated with conditional constructions, Smith and Sweetser (2015) demonstrate that different gestural strategies are reliably produced with different types of conditionals. Causal conditionals are marked with contrastive gestures indicating alternative spaces, whereas speech act conditionals are marked with elaborating gestures that do not indicate alternative spaces. Rather than focus on establishing a high rate of gesture/speech co-occurrence, they show that the reliable and predictable *distribution* of each type of gesture between the two types of constructions supports a multimodal construction analysis.

While Zima (2014), Schoonjans et al. (2016), and Smith and Sweetser (2015) all work to establish evidence in favor of treating speech/gesture combinations as constructions, they do not provide formal constructional analyses in a particular construction grammar framework (see Fried and Östman, 2004; and Hoffman and Trousdale, 2013 for examples of a variety of construction grammar approaches). While there are to the best of my knowledge no extant formalizations of multimodal construction grammar, Kok and Cienki (2016) provide an analysis of an iconic gesture to demonstrate the potential for using a Cognitive Grammar framework (e.g., Langacker, 1987; 2008a) in representing multimodal utterances. They argue that cognitive linguistic models, including Cognitive Grammar, are appropriate for such a task because of their emphasis on the relationship between language and spatial cognition in their representations. In their analysis, they make use of Cognitive Grammar's diagrammatic approach to represent the

conventional and ad-hoc aspects of a gesture and integrate them into an analysis of the semantics of the whole utterance (see section 3.c. for further discussion of their analysis).

In their empirical support for a constructional treatment of gesture, Schoonjans (2014), Zima (2014), and Schoonjans et al. (2016) all focused on specific lexical affiliates: either single lexemes such as *einfach*, or lexically-instantiated constructions such as *all the way from X PREP Y*. In contrast, in this work I have cast a broader net by analyzing gestural data co-occurring with any syntactic pattern in which a motion verb is used metaphorically. For this reason, we could consider the gestures in these multimodal utterances to be analogous to argument structure constructions (Goldberg 1995; 2006). Argument structure constructions pair form and meaning at the syntactic level; for example, the Caused Motion Construction combines the meaning of Caused Motion with the pattern Subject Verb Object Directional, as in *he kicked the ball into the room* and *he sneezed the napkin off the table* (Goldberg, 1995). Metaphoric variants of argument structure constructions additionally specify the relationship between the semantics of the source and target domains of the metaphor and the lexical slots in the construction that evoke the metaphor (Sullivan, 2007; 2013).

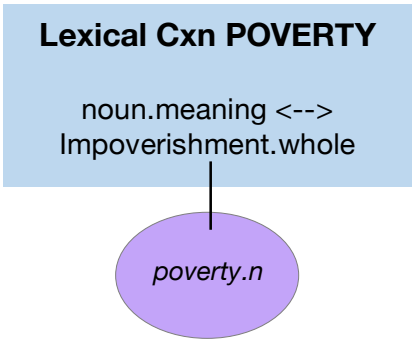
Therefore, a metaphoric gesture construction that describes the metaphoric gestures in this study does not specify the form of the gesture or a particular meaning like a lexical construction does[58]. Instead it describes which elements of the gesture evoke the various semantics of the metaphor, just as a metaphoric argument structure construction specifies which lexemes will evoke the source and target domain of the metaphor. In the case of a metaphoric gesture construction, it will only evoke the source domain. Since referential gestures are context-dependent, the target domain will be slotted in via binding to the speech content. Because I have argued that the image schemas evoked by the gesture form are the critical components to the gesture's meaning, the construction will specify that the gesture image schemas evoke the source domain of the metaphor. As we have seen previously in Chapter 3, Embodied Construction Grammar is an ideal route for representing the interaction between frame semantics, metaphor, and constructions. Due its commitment to embodied cognition and its emphasis on the representation of spatial information and force dynamic structure, its formalisms stand as the natural choice for a formal account of metaphoric gesture constructions.
In the following sections, I present a sketch of an Embodied Construction Grammar approach to integrating gesture and speech in a construction grammar account. First I briefly review the Embodied Construction Grammar framework and its application to metaphoric language, although a more detailed discussion is provided in Chapter 3. Then, I extend this approach to co-expressive and finally complementary metaphoric gestures.

3.b. *A review of metaphoric speech in ECG*

To represent gesture schematically in the ECG framework, we first begin with the representation of lexical constructions, as in Figure 1. The lexeme *poverty* evokes the lexical noun construction. The meaning of the construction is linked to the Impoverishment semantic frame.

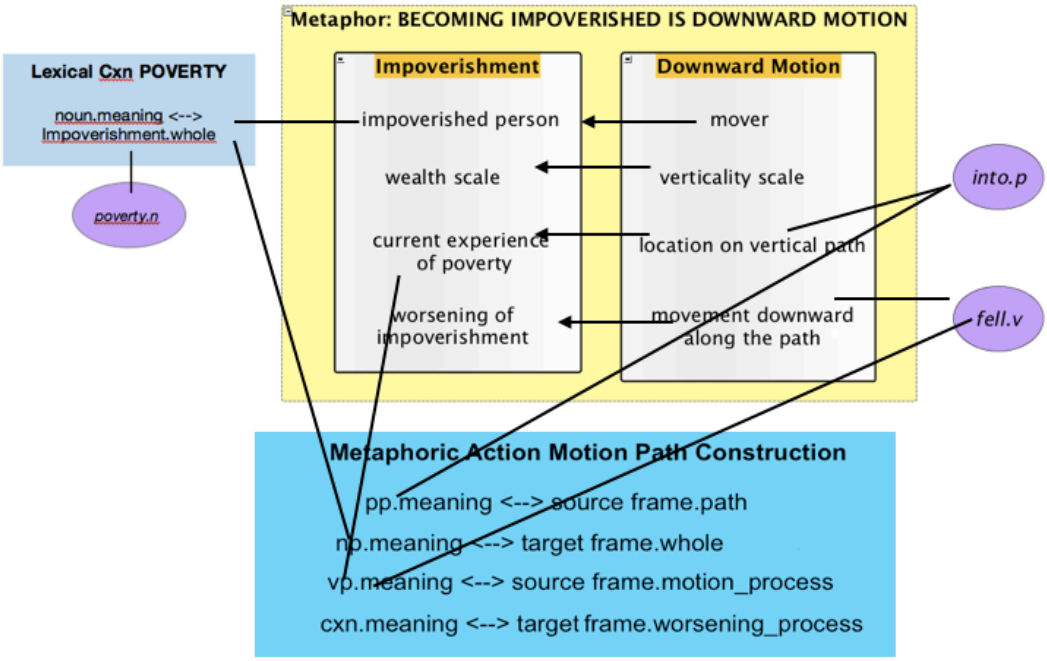
[58] In contrast, highly conventional gestures like emblems or possibly common gesture families like the palm-up open-hand could be good candidates for analysis as lexical gesture constructions.

Figure 1. Lexical construction: Poverty



This construction is incorporated into the larger analysis of *fell into poverty* from Chapter 3, reproduced here as Figure 2. Note that the lexemes *fell* and *into* would also evoke lexical constructions, but they are not represented here.

Figure 2. Metaphor construction analysis of *fell into poverty*



This analysis makes use of the metaphoric Active Motion Path Construction, again reproduced from Chapter 3 (Figure 3):

Figure 3. Metaphoric Active Motion Path Construction

```
Construction Metaphoric Active Motion Path
Subcase of Active Motion Path
Meaning: Metaphoric Motion Path
  Constraints
   pp.meaning <--> source frame.path
   np.meaning <--> target frame.whole
   vp.meaning <--> source frame.motion_process
   cxn.meaning <--> target frame.worsening_process
```

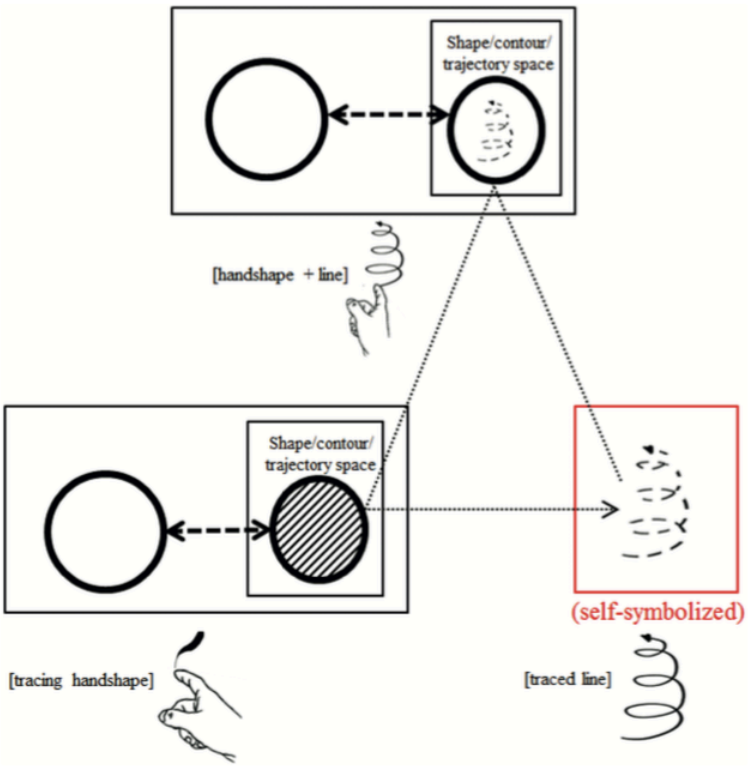
In this analysis, the meaning of each phrase evokes a frame or frame element of the source or target domains. The noun phrase *poverty* evokes the meaning of the lexeme *poverty*, which is the target frame Impoverishment. The verb phrase evokes the meaning of the lexeme *fell*, which is the process of motion in Downward Movement; and the prepositional phrase *into poverty* evokes the meaning of *into*, which is the Path from Downward Movement. *Poverty* is metaphorically the Goal of the of the Source-Path-Goal image schema in Downward Movement. Thus the meaning of the whole construction is the worsening process in Impoverishment, because the motion process in the source domain maps onto the worsening process in the target.

3.c. *Co-expressive metaphoric gestures in multimodal ECG*

Now that we have reviewed how metaphor is analyzed in an ECG framework, we can consider how we might represent gestural information in ECG. Similar to language, there is a distinction in gesture between *form* and *meaning* (McNeill, 1992; Saussure, 1916). However, Saussure argued language form is *arbitrary* – there is no natural or inherent reason why a particular form would be related to a particular meaning. In contrast, the gestures we have seen here are *iconic* – the form of the gesture represents some element of its meaning. Nonetheless, an iconic or metaphoric gesture's meaning cannot be understood through its form alone; context is required to interpret the meaning of the form. As Kendon (1980) first observed, gestures can be understood along a cline of *conventionality*; "emblems" are quite word-like in that their meaning is tightly entrenched. However, it has been questioned as to whether more idiosyncratic gestures such as iconics and metaphorics can be considered to have systematic structure, due to their context-dependence (Kok and Cienki, 2016).

Kok and Cienki argue that because they do have commonalities and recurrent features, such as certain discourse functions, they can be argued to have "a type of very schematic grammatical structure, sharing only very few formal and semantic features in common" (2016: 72). They conclude in their Cognitive Grammar approach that such gestures should not be dismissed outright from inclusion in grammar; they are just highly schematized and more variable than other types of gesture. In their analysis, rather than attempt to represent the semantics of gestures schematically, the "ad-hoc" components of an iconic tracing gesture are represented pictorially and then composed with a visual representation of the entity being depicted. In one example, the meaning of a hand tracing a spiral line is analyzed as the combination of the functional *category* of tracing gestures with the ad-hoc representation of a spiral (Figure 4).

Figure 4. Cognitive grammar analysis of a tracing gesture. (Kok and Cienki, 2016: 83)



Their analysis relies on the analog representation of the gesture in the pictures of the spiral line to convey both its form and meaning, rather than attempt to schematically represent it at a higher level of abstraction. They acknowledge that their approach leaves open the question of *how* to describe these gestures; they ask what level of abstraction would be best appropriate: "whether they are most adequately captured in terms of individual form parameters…, more holistic patterns…, or more theoretical constructs such as image or action schemas" (2016: 94). In this dissertation, I have shown that *image schemas* can be highly effective in analyzing the salient components of these gestures. While my analysis has focused on *metaphoric* referential gestures, the same arguments could easily hold for similar *iconic* gestures, as I discuss in Chapter 4. Indeed, Cienki (2005) has shown that iconic gestures have similarly accessible image schematic representations to comparable metaphoric gestures.

If we take the salient image schema(s) of the gesture as our starting point – as I have argued is the appropriate level of abstraction – then our challenge is to represent the *form* of the gesture and link that representation to the *image schemas* and their *meanings*. Recalling our analysis of the lexical construction, we note that lexical-level analysis in ECG separates form and meaning. Thus, we can represent the gesture's form separately, and make use of the tools we already have from ECG to represent meaning. For ease of simplicity, I first address the analysis of a *co-expressive* gesture, which evokes the same meaning as its co-speech (Figure 5).

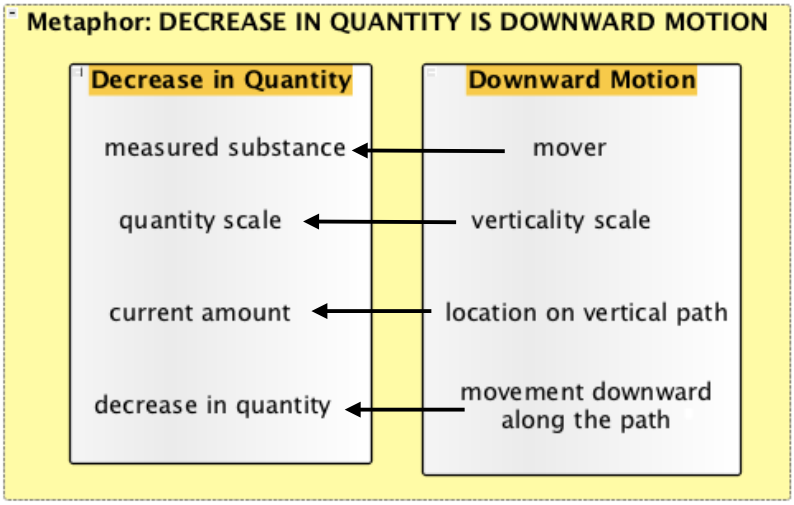
Figure 5. Co-expressive metaphoric gesture



if [oil prices drop] [p.drop.2.18.2015]

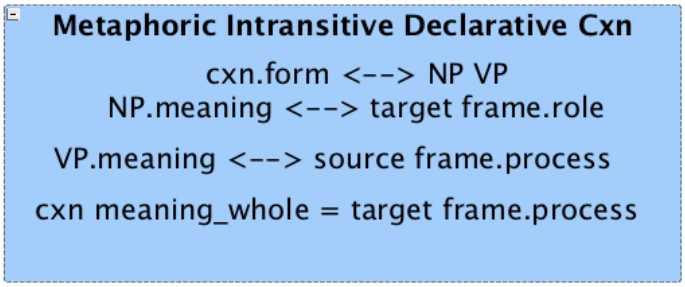
In Figure 5, the speaker produces a downward movement with his hand while uttering *oil prices drop*. Because the meaning of the gesture is contextually dependent, I present an analysis of the speech first. The speaker metaphorically construes *prices* as points on a vertical path; decreasing in quantity is therefore construed as moving downwards along the path. *Prices* evokes the target domain of Quantity and *drop* the source domain of Downward Movement; Figure 6 presents the metaphor analysis. The moving entity maps onto the measured substance (here, prices); movement downward maps onto the decrease in quantity.

Figure 6. Metaphor analysis of *prices drop*



Although embedded in a larger Conditional Construction, we analyze *prices drop* as a Metaphoric Intransitive Construction. Thus, the metaphoric intransitive inherits its structure from the Intransitive Construction; I only present the metaphoric analysis in Figure 7.

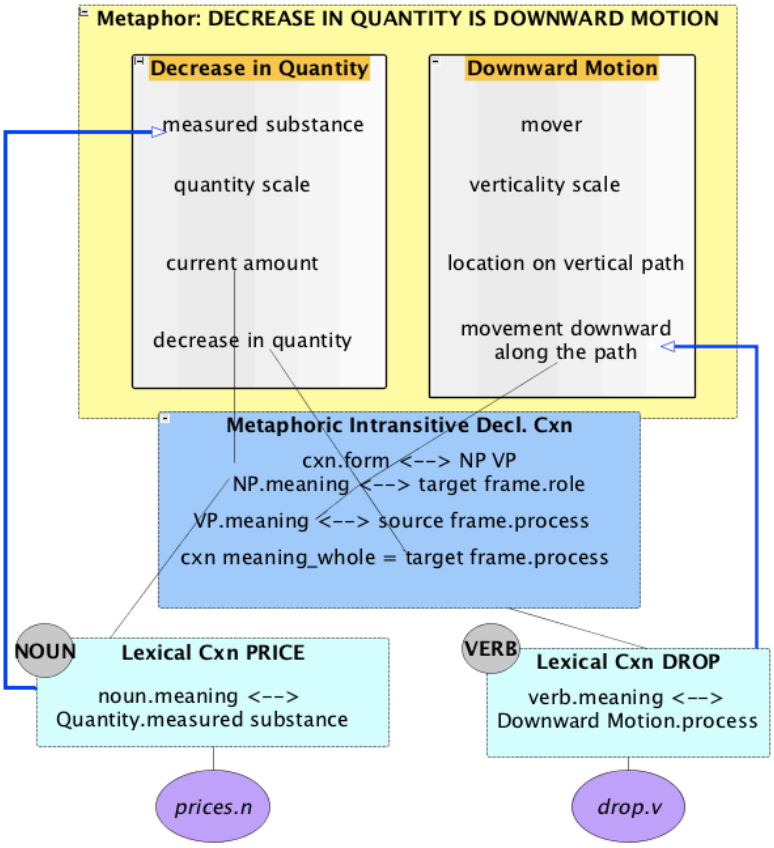
Figure 7. ECG analysis of the Metaphoric Intransitive Declarative Construction



The intransitive takes the form NP VP. As with the previous analysis of the Action Motion Construction, the noun phrase evokes the target frame and the verb phrase's meaning evokes the the source domain frame. The meaning of the construction as a whole is the target domain, reflecting its metaphoric semantics.

Now we can combine our semantic metaphor analysis with our ECG analysis for a full representation of *prices drop* (Figure 8). Lexical constructions *price* and *drop* make use of the Noun and Verb constructions and are incorporated into the NP VP form of the intransitive. The noun phrase in the Metaphoric Intransitive Construction evokes the measured substance role in the Quantity target frame. The verb phrase's meaning evokes the movement process in the source domain Downward Motion frame, and the meaning of the construction as a whole is the target domain process of decreasing quantity.

Figure 8. Metaphor and ECG analysis of *prices drop*

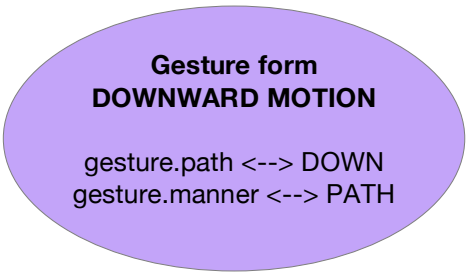


Now that we have analyzed the speech component of Figure 5, we can consider the gesture component. This gesture has the following form parameters:

Handshape: BentB
Location: Chest
Palm orientation: Down
Path: Down
Manner: Straight

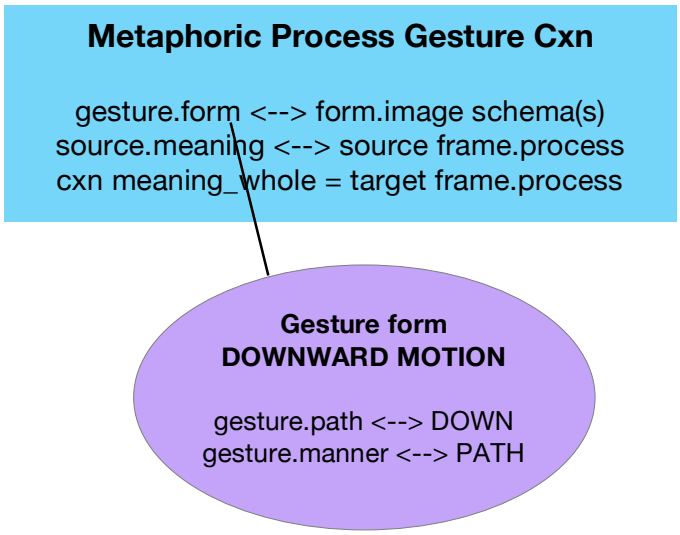
Based on the accompanying speech context, we first focus on the Path and Manner of the gesture. The straight downward gesture is *co-expressive* with the lexeme *drop* in the speech. The salient image schemas in this gesture are therefore Path and Down, as they are the prototypical schemas in the Downward Motion frame, as evoked by *drop*. We can represent gesture form just as we represent lexical form. In a complete constructional analysis, lexical form is represented by specifying the phonemic realization of the utterance; here we represent gesture form as codified by its constituent parameters (Figure 9). In the interest of space, as is typical with constructional analyses we include only the elements currently under analysis – the path and manner of the gesture. These *form* representations are linked to the image schemas they evoke.

Figure 9. Representation of gesture form and image schema.



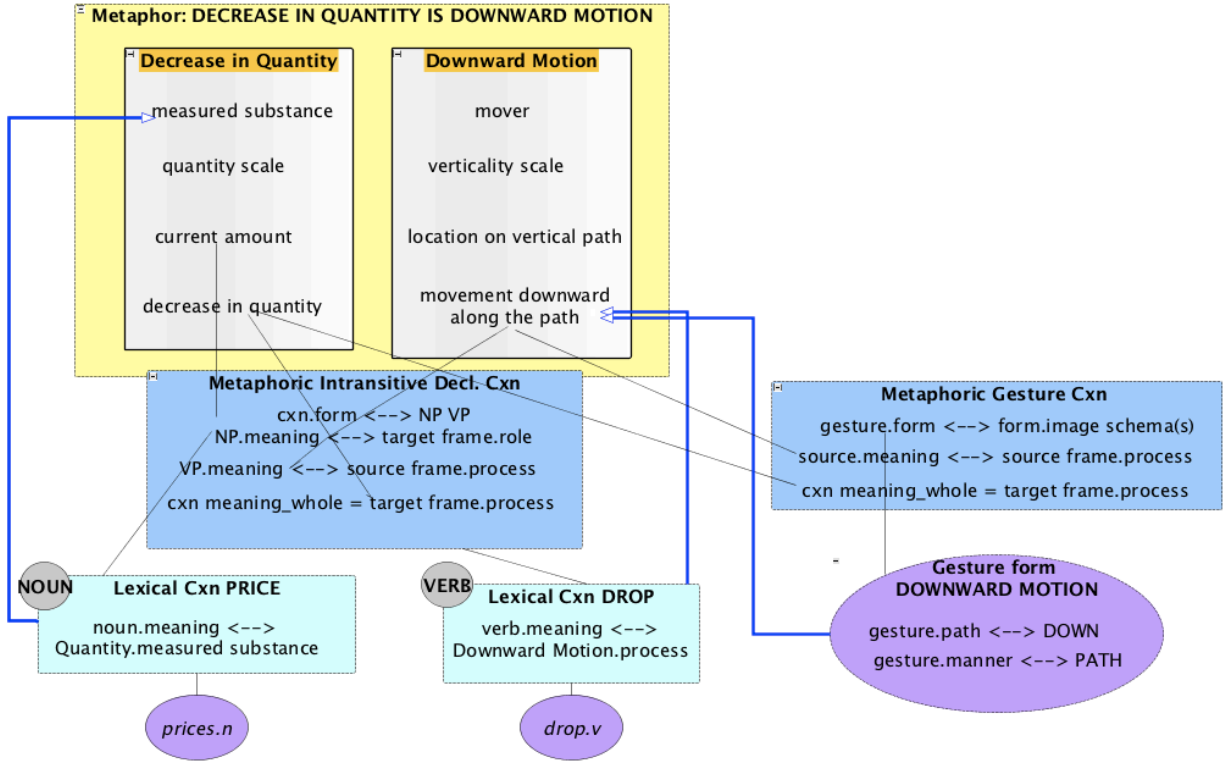
Whereas emblem gestures are similar to single words, iconic and metaphoric gestures tend to represent both entities and actions. They are therefore arguably closer in structure to predicate-level argument structure constructions than single lexical constructions. For this reason, I represent the Metaphoric Process Gesture Construction, which allows the gesture to evoke a metaphoric meaning, in a similar format to argument structure constructions such as the Metaphoric Intransitive Construction (Figure 10). The gesture evokes the process role of the metaphor's source domain by linking the form of the gesture to its image schemas, which form the relevant parts of the source domain process. In this case, the process role of the Downward Movement domain is movement downward along a path. The relevant image schemas in this frame element are Path and Down. Thus, the salient image schemas in the gesture – Path and Down – are linked to those in the source domain. The meaning of the whole gesture construction is the target domain process.

Figure 10. Metaphoric Process Gesture Construction analysis.



At this point, we have associated the form of the gesture with its meaning, and can incorporate into a representation of the multimodal utterance as a whole (Figure 11).

Figure 11. Complete metaphoric ECG analysis of the multimodal utterance.

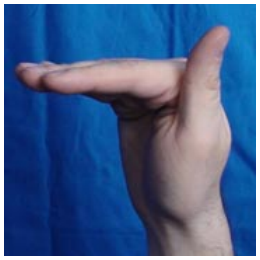


The gesture's image schemas evoke the movement downward along a path role in the Downward Motion frame via the gesture construction, which binds the gesture form and image schema meaning. The gesture's image schemas evoke the source domain process, and the meaning of the construction binds to the process of decrease in quantity in the target domain. Hence, it evokes

the metaphor DECREASE IN QUANTITY IS DOWNWARD MOTION. We understand this to be a metaphoric gesture as it is produced simultaneously with the speech *prices drop*, which also evokes the metaphor through the Metaphoric Intransitive Declarative Construction and specifies the target domain of Quantity.

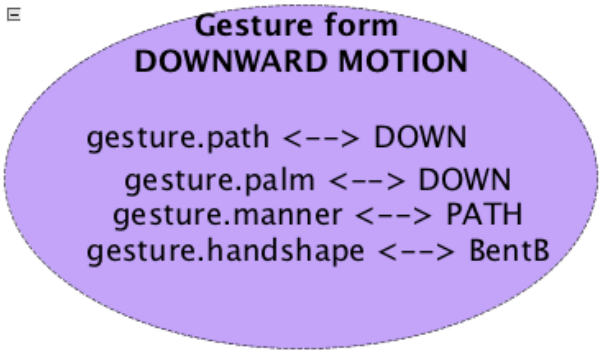
This analysis has focused on the path and manner of the gesture. However, we can also observe that the palm orientation *down* and the handshape *BentB* also evoke salient information as well. The palm orientation reinforces the Down image schema. The handshape becomes relevant when we further consider the larger speech context: the speaker actually says *if oil prices drop to a certain **level***. By introducing the metaphoric goal *level*, he focuses his utterance on the endpoint of the path rather than the trajectory itself. The BentB handshape with its 90° angle bend at the knuckles emphasizes the flatness of the plane projected by the fingers (Figure 12).

Figure 12. BentB handshape.



This handshape represents information similar to that of his lexical choice *level*. *Level* evokes the location on the vertical scale within the Downward Motion frame. The handshape, by delineating a flat horizontal plane, similarly evokes this location role. Adding in the palm orientation and handshape information, our more complete analysis of the gesture form is as follows (Figure 13).

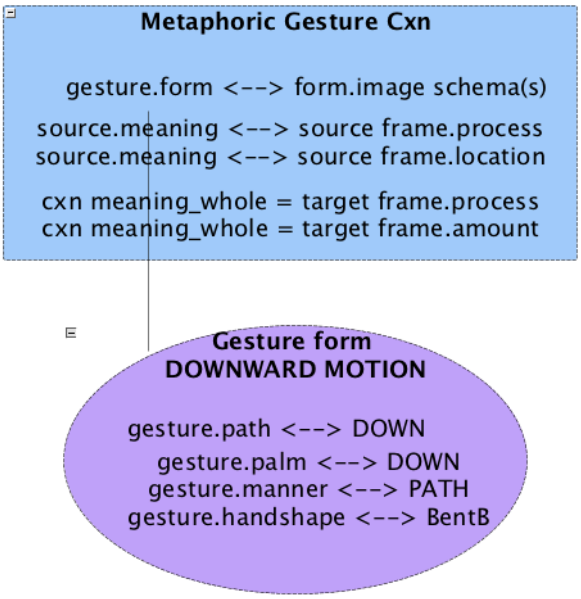
Figure 13. Re-analysis of the metaphoric gesture form.



By incorporating additional role information into our analysis of the gesture, we posit a different metaphoric gesture construction (Figure 14). The first analysis in Figure 10 was a *process* gesture construction, which evoked specifically the process role of the target domain. This re-analysis of the form evokes both process and entity information. Notably, the gesture here has a *marked* handshape. As I argue in Chapter 4, unmarked handshapes in metaphoric gestures are an indication of a de-emphasis on the metaphoric entity. In this case, the gesture does have a marked handshape. This represents the notion of a 'level' – the location entity on the path. For

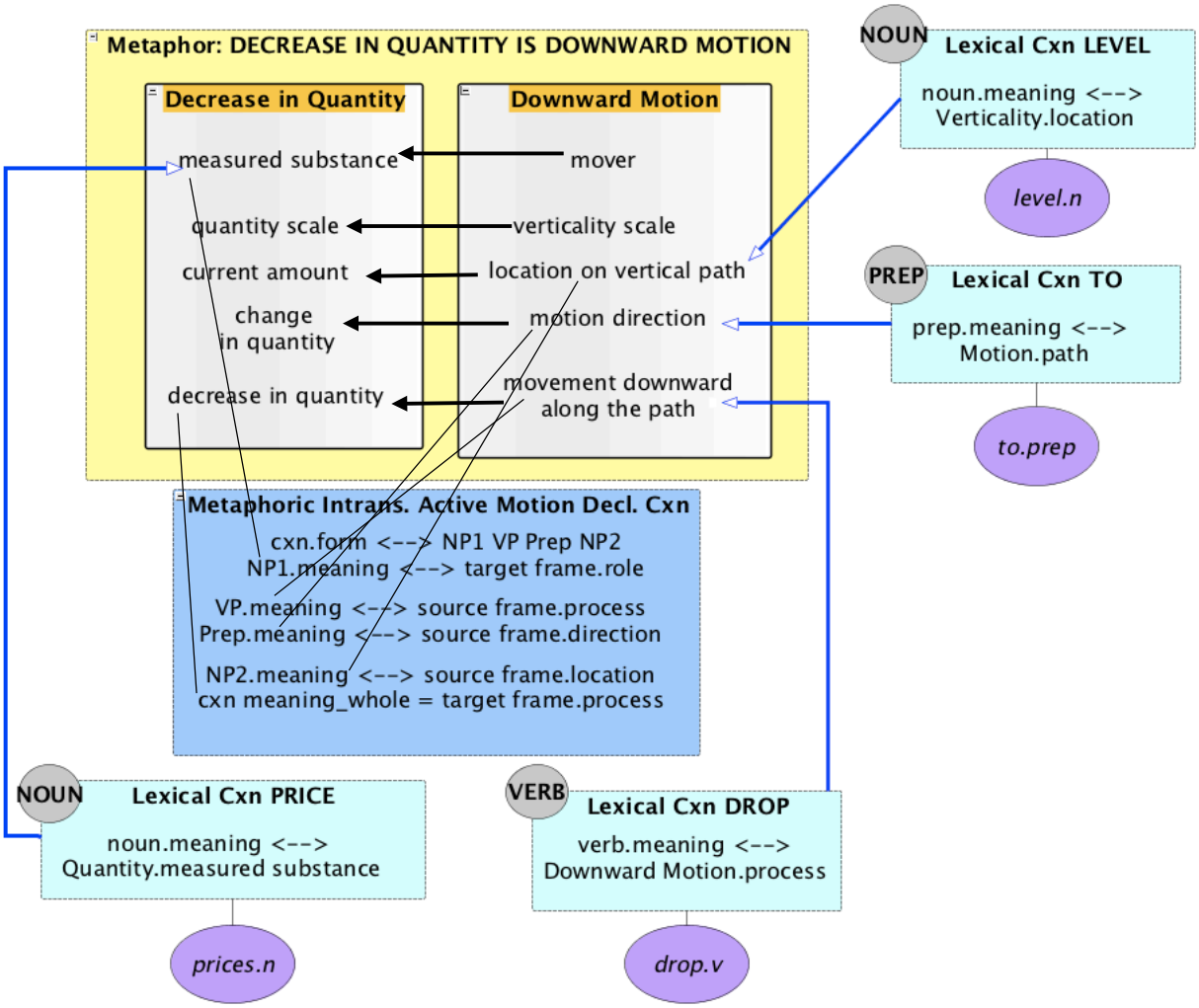
this reason, it makes use of a metaphoric gesture construction that evokes both the process and amount in the target domain.

Figure 14. Re-analysis of the metaphoric gesture construction.



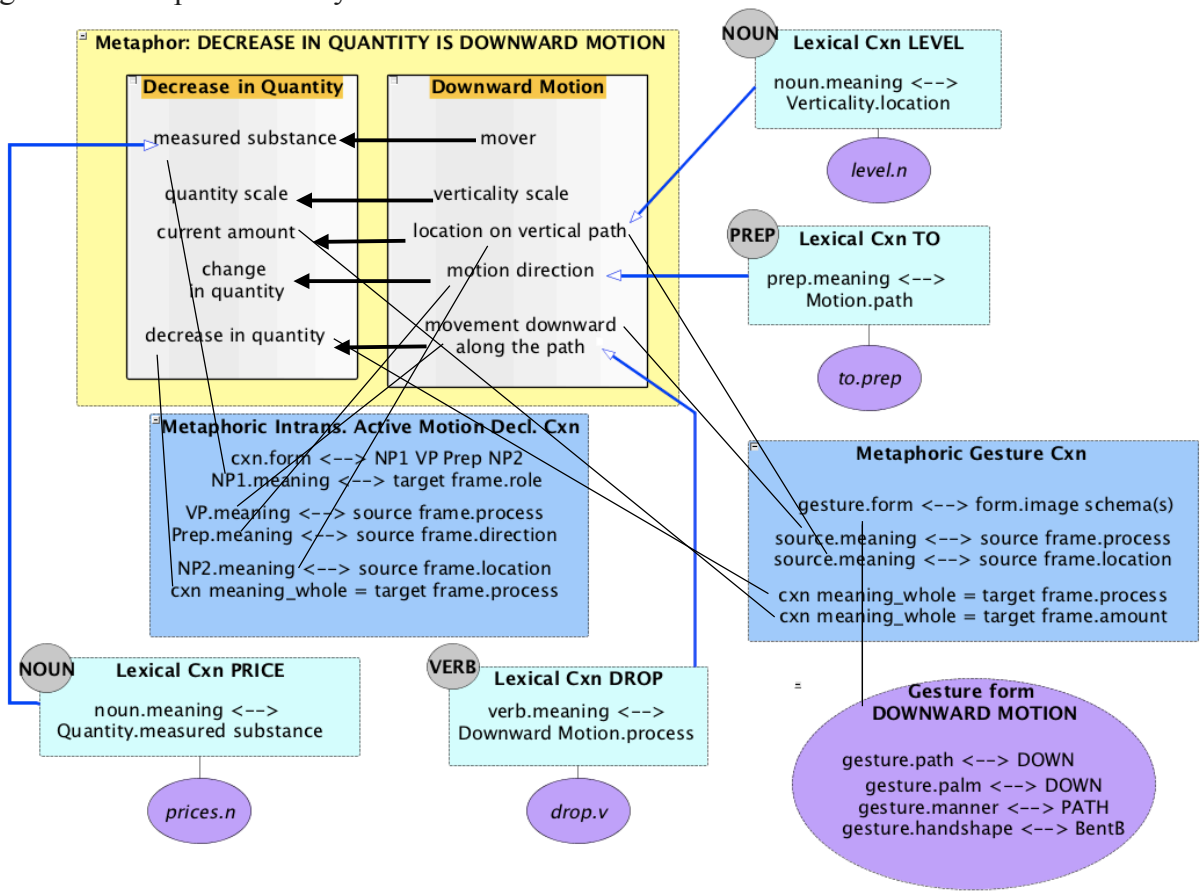
We also must incorporate the additional linguistic information into the ECG representation. I have already discussed an analysis of the Metaphoric Action Motion Path Construction, which is similar to the current construction, in Chapter 3. Therefore, I present here only a complete analysis (Figure 15). In addition to our prior analysis of *prices drop*, we add the prepositional phrase *to a certain level*, focusing on the preposition and noun. The preposition *to* evokes the direction of motion along the path and *level* the goal of the path. These map onto the the change in quantity and quantity amount, respectively. In effect, the non-metaphoric Active Motion Path Construction as instantiated in the construct *drop to a certain level* combines with the Metaphoric Intransitive Construction in *prices drop*, as unified by their shared verb *drop*. Whereas in *drop to a certain level*, the only semantic domain is Downward Motion, in *prices drop* both Quantity and Downward Motion are evoked. This creates a composite Metaphoric Intransitive Active Motion Construction in the complete utterance *prices drop to a certain level*.

Figure 15. ECG analysis of prices drop to a certain level



Now, the remaining step is to compose the gesture and speech utterances, as done previously in Figure 11. The difference is only that we are incorporating additional information from the speech and gesture into our analysis (Figure 16).

Figure 16. Complete re-analysis of the multi-modal utterance.



The analysis of the gesture form and metaphoric gesture construction (Figure 14) is incorporated into the ECG speech analysis (Figure 15) by connecting the meaning of the gesture image schemas to their respective roles in the source domain Downward Motion frame. The meaning of the construction as a whole evokes elements in the target domain Decrease in Quantity frame. This is the same approach as previously shown in Figure 11; we have only added information conveyed by the gesture form handshape and palm orientation parameters. Both the palm orientation and motion evoke downward motion, and the BentB handshape evokes the location of the moving entity.

The process of re-analysis of the utterance from Figure 5 to Figure 16 reflects the *compositionality* of the gestures (Kok and Cienki, 2016). Initially, I analyze only a small set of elements of the multimodal utterance: the speech *prices drop* and the path and manner of the gesture. These compose in the analysis in Figure 5. I then analyze the speech *to a certain level* and the palm orientation and handshape of the gesture. This additional analysis is composed with the previous one to produce the resulting analysis in Figure 16. Mittelberg (2014) argues that gesture is not what is traditionally considered to be compositional, as gesturers create new semiotic material with each typical multimodal utterance, rather than select from a lexicon or inventory as they do in typical speech. Furthermore, gestures do not have a defined grammar as language does syntax (McNeill, 1992). However, gestures can nonetheless be considered to be compositional in that they combine conventionalized features, such as *tracing*, with ad-hoc

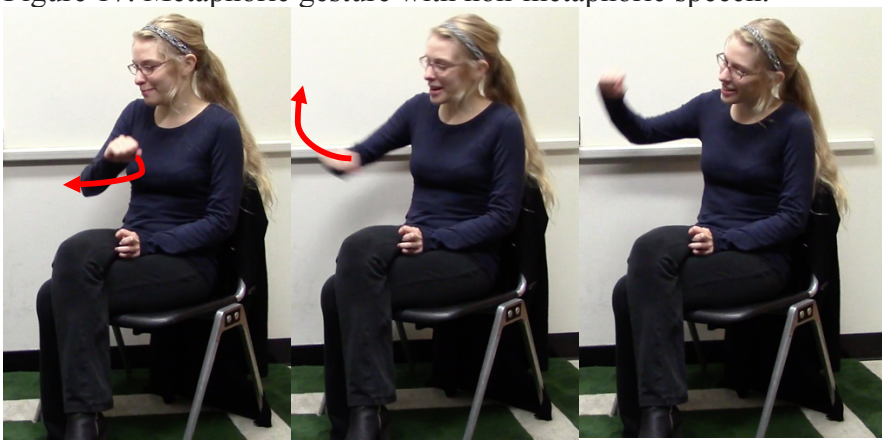
representations of specific spatial/physical characteristics (Mittelberg, 2014). For example, in their Cognitive Grammar analysis, Kok and Cienki (2016) compose the semantics of the conventional element with the ad-hoc symbolization as shown in Figure 4. The semantics of the conventional *tracing gesture* are understood to mean that the "trace" is of an entity with a spatial property such as a path or contour; the ad-hoc spiraling of the gesture specifies that contour.

I argue that the conceptual structures of the image schemas and metaphoric constructions provide the necessary conventionalization to compose the holistic and idiosyncratic nature of these gestures with their co-speech. Image schemas are abstracted, conceptual representations of sensorimotor information, whereas constructions are established patterns of form-meaning mappings. These provide a structural basis for the gesture's visual/spatial properties to bind to. To be clear, this Embodied Construction Grammar analysis is not at odds with Kok and Cienki's (2016) Cognitive Grammar analysis; rather, they are complementary and commensurate. Whereas their approach specifies in greater detail how the ad-hoc spatial representations of the gesture combine with the conventional aspects of the gesture, mine focuses more on the details of the composition of image schematic structure and speech. While they are agnostic as to the details of the relationship between form and function – indeed, they identify it as an open question – here I claim that image schemas constitute the most appropriate level of analysis, at least for this class of metaphoric gestures.

3.d. *Complementary metaphoric gestures in multimodal ECG*

To further illustrate this approach, I now turn to a case where the speaker/gesturer produces a non-metaphoric utterance in her speech and a metaphoric utterance in her gesture, thereby producing a metaphoric multimodal utterance. We first saw this utterance as Figure 1 in the beginning of Chapter 1; I reproduce it here as Figure 17. For the purposes of this analysis, I will focus on the first part of the narrator's utterance, *he feels [better]*.

Figure 17. Metaphoric gesture with non-metaphoric speech.



he feels [better] [S9_2Y_B_10.27.15_ELV]

As previously discussed in Chapter 1, in this sentence the narrator does not produce metaphoric language; she uses a (non-metaphoric) active intransitive declarative construction, shown in a more detailed analysis in Figure 18. The narrator simultaneously produces an upward arcing

gesture, co-timed with *better*. The gesture's form comprises an *arc* manner and *up* path, which evoke the Path and Up image schemas.[59]

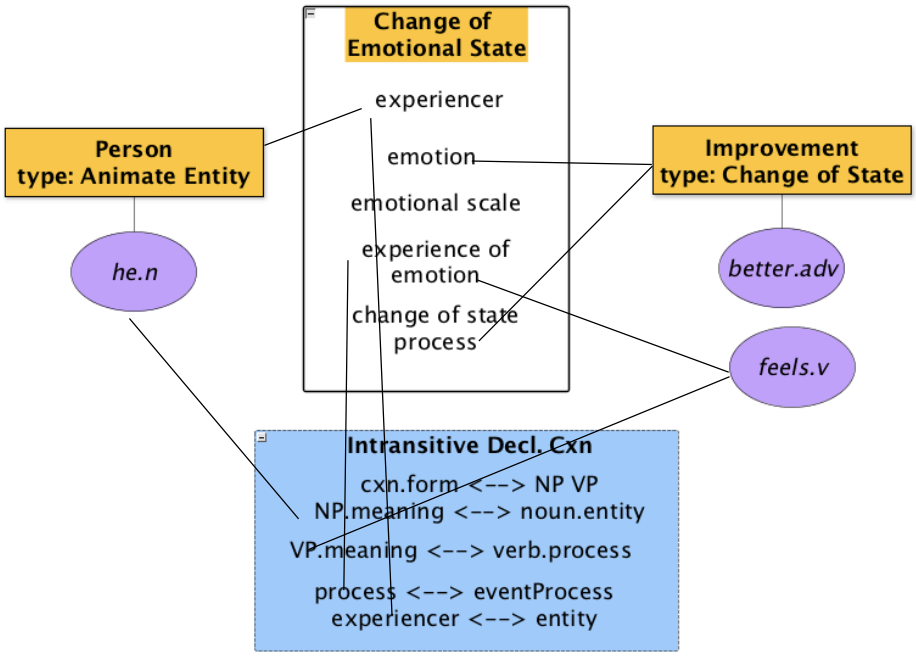
Figure 18. ECG analysis of the Active Intransitive Construction.

```
Construction Active Intransitive
Subcase of ArgumentStructure
Form: NP VP
Meaning: Intransitive
  Constraints
   np.meaning <--> noun.entity
   vp.meaning <--> verb.process
   process <--> eventProcess
   experiencer <--> entity
```

The active intransitive takes the form NP VP; the meaning of the NP is its head noun and the meaning of the VP is the verb. I represent the relationship between lexemes, construction, and frames in Figure 19. The verb evokes the process role of the Process schema, which comprises an Action Process and the Entity that performs that Process. The noun evokes an Entity schema; this binds to the agent which performs the action. In this particular sentence, the pronoun *he* evokes the Person schema, which specifies that the Entity is of type Person. As an Animate Entity, Person entities experience Emotions, the frame of which is evoked by *feel*. The Change of Emotional State frame has two processes: the change of state process and the experience of the emotion. *Feels* evokes the emotion experience. The comparative in *better* evokes the change of state variant of the Emotion frame. The adverb *better* modifies *feels* to specify the nature of his emotional state: it indicates he is undergoing state change as his state improves. Note that this is a representation of a partial analysis; for example, I do not include the constructional components that specify the comparative evoked by *better* or the modifying Verb-Adverb Construction; I also do not include lexical-level constructions. I focus on the frames and construction most relevant here to our analysis.

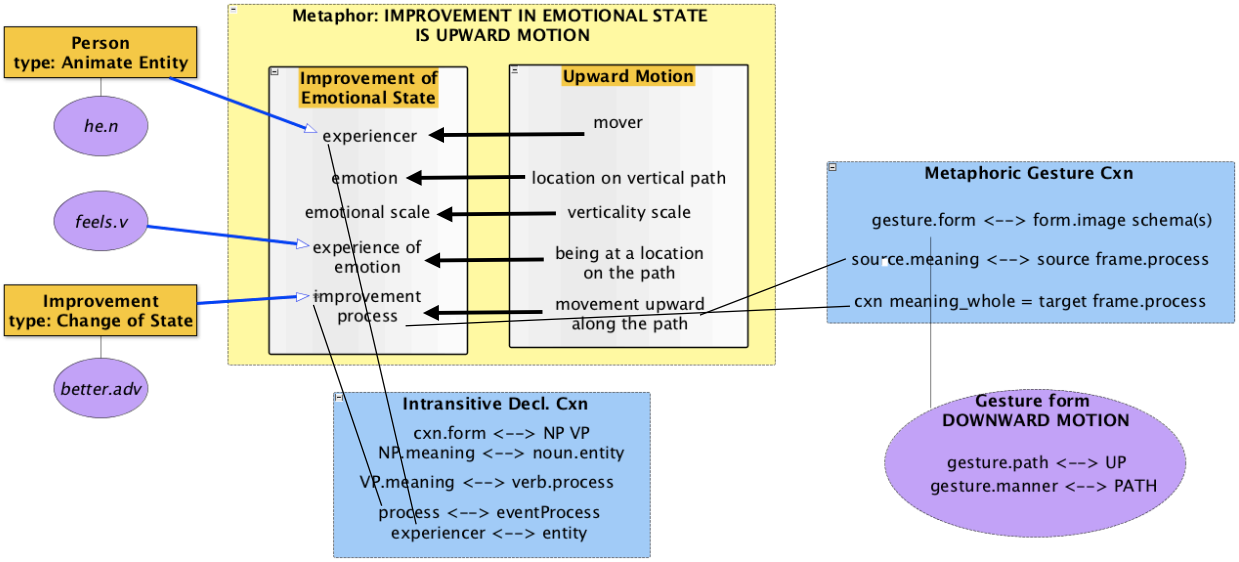
[59] The narrator also produces a fisted S handshape. I do not include the handshape in my analysis, because I have not heretofore discussed the semantics of Force image schemas or metonymy. In short, the fist handshape metonymically evokes the Force schema. This occurs via the experience of using a fisted hand to produce sufficient tension and strength to forcefully cause change to another entity's physical state (for example, *punching* or *squeezing*). For a detailed explanation of the force dynamics analysis, see Talmy (2000). The fist becomes a metonymic representation of the force behind the "punching" or "squeezing". Cienki (2013) argues this type of referential semantics derives from *mimetic schemas*, which "mimic" real-life actions. In this metaphoric context, the force evoked by the handshape maps onto the degree of the change of state. A "strong" change has a greater scalar difference between start and end states than a "weak" change (see also Sweetser, 1990). This type of force dynamics-driven metonymy and metaphor is also attested in sign languages; for example, the sign for modal CAN derives from the sign for STRONG, which similarly has S fist handshapes (Wilcox and Wilcox, 1995).

Figure 19. ECG and frame analysis of *he feels better*.



Now that we have analyzed the non-metaphoric speech, we can approach the metaphoric co-speech gesture. Recalling that the Metaphoric Gesture Construction specifies that the meaning of the construction is the target domain of the metaphor, we can link the meaning of the gesture construction to the Change of Emotional State frame evoked by the speech (Figure 20). This is accomplished via the IMPROVEMENT IN EMOTIONAL STATE IS UPWARD MOVEMENT metaphor. The upward movement of the gesture evokes the source domain and the speech fills the target domain information, as in our prior analysis of the co-expressive gesture.

Figure 20. Complete multimodal analysis.



4. *Conclusions*

This dissertation is a wide-ranging exploration of the interaction of metaphor, gesture, and grammar in usage; consequently, its diverse findings have implications for conceptual metaphor theory, theories of language and cognition, and gesture studies. With regards to conceptual metaphor theory, the formalization of metaphor representation and its instantiation in the MetaNet repository advances both metaphor analysis itself and corpus approaches to metaphor. By proposing to represent metaphors as a complex network of frames, mappings, and bindings, we advance the representation of conceptual metaphors to a level that interfaces more accurately with representations of frames and constructions in FrameNet and Embodied Construction Grammar. In turn, this enables the detailed analysis of metaphors and metaphor systems, as exemplified by the Location Event Structure Metaphor case study. This corpus-based study reveals not only how the metaphor system is evoked in language, but further illustrates the conceptual structure of the metaphor. The foregrounding of path information in linguistic realization of metaphoric motion runs counter to the privileging of manner in English lexicalization patterns. This finding lays the groundwork for the investigation of the same metaphor system in gesture.

The corpus and experimental studies find that English metaphoric motion gestures pattern similarly to the use of English metaphoric motion verbs in speech. The predominant image schematic structures conveyed by these gestures include path and direction information, which reflects prior findings in investigations of English gestural correlates of literal motion events. These results expand on the gestural event structure literature by providing the first systematic investigation of metaphoric event structure in gesture. By making use of both corpus and experimental approaches, we can study these gestures in naturalistic and controlled environments across a variety of speakers and contexts. The parallel metaphoric and literal corpora provide matched comparisons which are controlled for lexical structure, whereas the experimental data is unconstrained by the limitations of database searches. These results support the Thinking for Speaking Hypothesis by revealing the particular conceptual structures at work when reasoning with metaphoric motion, and show that thinking for speaking metaphorically is different from thinking for speaking literally. These results can be understood in terms of the Gesture as Simulated Action framework due to the embodied nature of both gesture and metaphoric cognition. The frequent co-occurrence of metaphoric language and metaphoric gesture can be best understood when both language and gesture are considered in terms of embodied cognition.

Finally, I have demonstrated the potential of representing gesture form and meaning in an Embodied Construction Grammar framework. While some initial work has shown that multimodal utterances may be considered constructions, this is the first work to propose an actual constructional analysis that formally represents the relationship between gesture form and meaning and language form and meaning. I argue that co-occurrence of language and metaphoric gesture can be considered akin to metaphoric argument structure constructions, wherein the source and target domains of the metaphor are predictably evoked by certain elements of the multimodal utterance. The analysis presented here is the first attempt at such a formalization, and undoubtedly will undergo revision in subsequent research. However, this represents an important step in a unified approach to our understanding of language and gesture as a single multimodal communicative act. In formalizing our representation of the multimodal utterance, we can refine

our understanding how form and meaning in each modality interact and combine to create a coherent whole.

References

Ahmad, K. (2005, June). Terminology in text. Paper presented at the Tuscan Word Centre Workshop. Siena, Italy.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes*, *15*(6), 593-613.

Aziz-Zadeh, L., & Damasio, A. (2008). Embodied semantics for actions: Findings from functional brain imaging. *Journal of Physiology, 102,* 35-39.

Bailey, D. (1997). *A computational model of embodiment in the acquisition of action verbs*. Ph.D. dissertation, University of California, Berkeley.

Barcelona, A. (2000). Introduction: The cognitive theory of metaphor and metonymy. In A. Barcelona (Ed.), *Metaphor and metonymy at the crossroads: A cognitive perspective* (pp. 1-28). Berlin/New York: de Gruyter.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade., A. (1992). Interactive gestures. *Discourse Processes, 15*(4), 469-489.

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A, Silvia, P. J., Dunst, B., & Neubauer, A. C. (2014). Creating metaphors: The neural basis of figurative language production. *Neuroimage*, *90*, 99-106.

Bergen, B. K. (2012). *Louder than words: The new science of how the mind makes meaning*. New York: Basic Books.

Bergen, B. K., & Chang, N. (2005). Embodied Construction Grammar in simulation-based language understanding. In J.-O. Östman & M. Fried (Eds.), *Construction grammars: cognitive grounding and theoretical extensions* (pp. 147-190). Amsterdam/Philadelphia: John Benjamins Publishing.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Boas, H. C. (2013). Cognitive construction grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 233-254). Oxford, UK: Oxford University Press.

Borin, L., Dannélls, D., Forsberg, M., Gronostaj, M. T., & Kokkinakis, D. (2009, December). Thinking green: Toward Swedish FrameNet++. Poster presented at TLT-8: The 8th International Workshop on Treebanks and Linguistic Theories. Milan, Italy.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, *75*(1), 1-28.

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science, 13*, 185-189.

Bouveret, M., & Sweetser, E. (2009). Multi-frame semantics, metaphoric extensions, and grammar. In I. Kwon, H. Pritchett, & J. Spence (Eds.), *Proceedings of the 35th Annual Meeting of the Berkeley Linguistics Society* (pp. 49-59). Berkeley, CA: Berkeley Linguistics Society.

Brenger, B., Schüller, D., Priesters, M., & Mittelberg, I. (2016, July). *3D heat maps of multimodal travel planning: Correlating prepositional and adverbial phrases with locating and routing gestures.* Paper presented at the 7th meeting of the International Society of Gesture Studies, Paris, France.

The British National Corpus (2007). Version 3 (BNC XML Edition). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of Manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, *30*(2), 225-251.

Brugman, C. M. (1988). *The story of over: Polysemy, semantics, and the structure of the lexicon*. M.A. Thesis. University of California, Berkeley.

Brugman, H., & Russel, A. (2004). Annotating multimedia/ Multi-modal resources with ELAN. In M. Lino, M. Xavier, F. Ferreira, R. Costa, & R. Silva (Eds.), *Proceedings of the 4th International Conference on Language Resources and Language Evaluation (LREC 2004)* (pp. 2065-2068). Paris, France: European Language Resources Association.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods 41*, 977-990.

Calbris, G. (1990). *The semiotics of French gestures*. Bloomington, IN: Indiana University Press.

Calbris, G. (2003). From cutting an object to a clear cut analysis: Gesture as the representation of a preconceptual schema linking concrete actions to abstract notions. *Gesture 3*, 19-46.

Casasanto, D. (2008, August). Conceptual affiliates of metaphorical gestures. Paper presented at Conference on Language, Communication, & Cognition. Brighton, UK. Available from http://www.casasanto.com/papers/Casasanto_ConceptualAffiliatesGestures.pdf.

Casasanto, D., & Jasmin, K. (2012). The hands of time: Temporal gestures in English speakers. *Cognitive Linguistics, 23*(4), 643-674.

Chiappe, D. L., & Chiappe, P. (2007). The role of working memory in metaphor production and comprehension. *Journal of Memory and Language*, *56*(2), 172-188.

Chui, K. (2011). Conceptual metaphors in gesture. *Cognitive Linguistics*, *22*(3), 437-458.

Cienki, A. (1997). Some properties and groupings of image schemas. In M. Verspoor, K. D. Lee, & E. Sweetser (Eds.), *Lexical and syntactical constructions and the construction of meaning* (pp. 3-16). Amsterdam/Philadelphia: John Benjamins Publishing.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. In J. P. Koenig (Ed.), *Discourse and cognition: Bridging the gap* (pp. 189-204). Cambridge, UK: Cambridge University Press.

Cienki, A. (2005). Image schemas and gesture. In B. Hampe (Ed.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 421-442). Berlin/New York: Mouton de Gruyter.

Cienki, A. (2008). Why study metaphor and gesture? In A. Cienki & C. Müller (Eds.), *Metaphor and gesture* (pp. 195-217). Amsterdam/Philadelphia: John Benjamins Publishing.

Cienki, A. (2013). Image schemas and mimetic schemas in cognitive linguistics and gesture studies. *Review of Cognitive Linguistics*, *11*(2), 417-432.

Cienki, A. (2016). Aspect and gesture. Symposium presented at the 7th meeting of the International Gesture Studies Society, Paris, France. July 2016.

Cienki, A., & Müller, C. (Eds.) (2008a). *Metaphor and gesture*. Amsterdam/Philadelphia: John Benjamins Publishing.

Cienki, A., & Müller, C. (2008b). Metaphor, gesture, and thought. In R. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought*. Cambridge, UK: Cambridge University Press.

Clausner, T. C., & Croft, W. (1999). Domains and image schemas. *Cognitive Linguistics*, *10*(1), 1-32.

Colletta, J. M., Kunene, R. N., Venouil, A., Kaufmann, V., & Simon, J. P. (2009). Multi-track annotation of child language and gestures. In M. Kipp, J.-C. Martin, P. Paggio, & D. Heylen, (Eds.), *Multimodal corpora: From models of natural interaction to systems and applications* (pp. 54-72). New York: Springer Berlin Heidelberg.

Comrie, B. (1976). *Aspect*. Cambridge, UK: Cambridge University Press.

Cooperrider, K., & Núñez, R. (2009). Across time, across the body: Transversal temporal gestures. *Gesture*, *9*(2), 181-206.

Croft, W. (1993). The role of domains in the interpretation of metaphors and metonymies. *Cognitive Linguistics*, *4*(4), 335-70.

Croft, W. (2001). *Radical construction grammar*. Oxford: Oxford University Press.

Croft, W. (2002). *Typology and universals*. Cambridge, UK: Cambridge University Press.

Croft, W. (2009). Connecting frames and constructions: A case study of *eat* and *feed*. *Constructions and Frames 1*(2), 7-22.

David, O. (2016). *Metaphor in the grammar of argument realization*. Ph.D. Dissertation, University of California, Berkeley.

David, O., Lakoff, G., & Stickles, E. (2016). Cascades in metaphor and grammar: A case study of metaphors in the gun debate. *Constructions and Frames, 8*(2).

Davies, M. (2008-) *The Corpus of Contemporary American English: 520 million words, 1990-present*. Available online at http://corpus.byu.edu/coca/.

Deignan, A. (2005). *Metaphor and corpus linguistics*. Amsterdam/Philadelphia: John Benjamins.

Desai, R. H., Binder, J. R., Conant, L. L., Mano, Q. R., & Seidenberg, M. S. (2011). The neural career of sensory-motor metaphors. *Journal of Cognitive Neuroscience*, *23*(9), 2376-2386.

Desai, R. H., Conant, L. L., Binder, J. R., Park, H., & Seidenberg, M. S. (2013). A piece of the action: Modulation of sensory-motor regions by action idioms and metaphors. *NeuroImage*, *83*, 882-869.

Dodge, E. (2016). A deep semantics corpus-based approach to metaphor analysis: A case study of metaphoric conceptualizations of poverty. *Constructions and Frames 8*(2).

Dodge, E., Hong, J., & Stickles, E. (2015). Metanet: Deep semantic analysis. In E. Shutova, B. B. Klebanov, & P. Lichtenstein (Eds.), *Proceedings of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies 3rd Workshop on Metaphor in NLP* (pp. 40-49).

Dodge, E. K., & Lakoff, G. (2005). Image schemas: From linguistic analysis to neural grounding. In Hampe, B. (Ed.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 57-91). Berlin/New York: Mouton de Gruyter.

Dodge, E. K., & Petruck, M. R. L. (2014). Representing caused motion in Embodied Construction Grammar. In Y. Artzi, T. Kwiatkowski, & J. Berant (Eds.), *Proceedings of the 2014 Association for Computational Linguistics Workshop on Semantic Parsing* (pp. 39-44).

Duffy, S. E., & Feist, M. (2014). Individual differences in the interpretation of ambiguous statements about time. *Cognitive Linguistics*, *25*(1), 29-54.

Duncan, S. (2001). Co-expressivity of speech and gesture: Manner of motion in Spanish, English, and Chinese. In C. Chang, M. J. Houser, Y. Kim, D. Mortensen, M. Park-Doob, &

M. Toosarvandani (Eds.), *Proceedings of the 27th Annual Berkeley Linguistic Society Meeting* (pp. 353-370). Berkeley, CA: Berkeley Linguistics Society.

Duncan, S. (2008). Gesture: Annotative practice. Available at http://mcneilllab.uchicago.edu/pdfs/susan_duncan/Annotative_practice_REV-08.pdf.

Feldman, J. (2006). *From molecule to metaphor.* Cambridge, MA: The MIT Press.

Feldman, J. A., Dodge, E. K., & Bryant, J. (2009). A neural theory of language and embodied construction grammar. In H. Narrog & B. Heine (Eds.), *The Oxford handbook of linguistic analysis* (pp. 111-138). Oxford, UK: Oxford University Press.

Feldman, J. A., & Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain and Language*, *89*(2), 385-392.

Fillmore, C. J. (1976). Frame semantics and the nature of language. In S. R. Harnad, H. D. Steklis, & J. Lancaster (Eds.), *Origins and evolution of language and speech* (pp. 20-32). Annals of the NY Academy of Sciences, Vol. 280.

Fillmore, C. J. (1982). Frame semantics. In *Linguistics in the morning calm: Selected papers from SICOL* (pp. 111-137). Seoul: Hanshin.

Fillmore, C. J. (1986). Pragmatically controlled zero anaphora. In V. Nikiforidou, M. VanClay, M. Niepokuj, & D. Feder (Eds.), *Proceedings of the 12th Annual Meeting of the Berkeley Linguistics Society* (pp. 95-107). Berkeley, CA: Berkeley Linguistics Society.

Fillmore, C. J. (1988). The mechanisms of construction grammar. In S. Axmaker, A. Jaisser, & H. Singmaster (Eds.), *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society* (pp. 35-55). Berkeley, CA: Berkeley Linguistics Society.

Fried, M., & Östman, J.-O. (2004). *Construction grammar in a cross-language perspective*. Amsterdam/Philadelphia: John Benjamins Publishing.

Gibbs, R. W. (2009). Why do some people dislike conceptual metaphor theory? *Cognitive Semiotics*, *5*(1-2), 14-36.

Gibbs, R. W. (2011). Evaluating conceptual metaphor theory. *Discourse Processes*, *48*(8), 529-562.

Gibbs, R. W., & Colston, H. (1995). The cognitive psychological reality of image schemas and their transformations. *Cognitive Linguistics*, *6*(4), 347-378.

Gibbs, R. W., Lima, P. L. C., & Francozo, E. (2004). Metaphor is grounded in embodied experience. *Journal of Pragmatics*, *36*(7), 1189-1210.

Givón, T. (2001). *Syntax: An introduction. Vol 1.* Amsterdam/Philadelphia: John Benjamins Publishing.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work*. Cambridge, UK: Oxford University Press.

Goldin-Meadow, S. (2005). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Grady, J. E. (1997). *Foundations of meaning: Primary metaphors and primary scenes*. Ph.D. Dissertation, University of California, Berkeley.

Graff, D. & Cieri, C. (2003). English Gigaword LDC2003T05. Web Download. Philadelphia: Linguistic Data Consortium.

Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*. Lund, Sweden: Lund University Press.

Hickmann, M., Hendricks, H., & Gullberg, M. (2013). Developmental perspectives on the expression of motion in speech and gesture: A comparison of French and English. In D. Bassano & M. Hickmann (Eds.), *Grammaticalization and first language acquisition* (pp. 129-155). Amsterdam/Philadelphia: John Benjamins Publishing.

Hinnell, J. A. (2013, June). TAM and gesture in North American English: A multi-modal corpus study. Paper presented at the 12th meeting of the International Cognitive Linguistics Conference, Edmonton, Canada.

Hinnell, J. (2014, July). Multimodal aspectual constructions in North American English: A corpus analysis of aspect in co-speech gesture using Little Red Hen. Paper presented at the 6th meeting of the International Society of Gesture Studies, San Diego, CA.

Hoffman, T., & Trousdale, G. (2013). *The Oxford handbook of construction grammar.* Oxford, UK: Oxford University Press.

Hong, J. (2016). Automatic metaphor detection using constructions and frames. *Constructions and Frames 8*(2).

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment as simulated action. *Psychonomic Bulletin and Review*, *15*(3), 495-514.

Israel, M. (1996). The *way* constructions grow. In A. Goldberg (Ed.), *Conceptual structure, discourse, and language* (pp. 217-230). Stanford, CA: CSLI.

Jamalian, A., & Tversky, B. (2012). Gestures alter thinking about time. In N. Miyake, D. Peebles & R. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp.503-508). Austin, TX: Cognitive Science Society.

Jirak, D., Mareike, M. M., Buccino, G., Borghi, A. M., & Binkofski, F. (2010). Grasping language – A short story on embodiment. *Consciousness and Cognition*, *19*, 711-720.

Johnson, C. (1999). *Constructional grounding: The role of interpretational overlap in lexical and constructional acquisition.* Ph.D. Dissertation, University of California, Berkeley.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, interaction, and reason.* Chicago, IL: University of Chicago Press.

Kay, P., & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: The what's X doing Y? construction. *Language*, *75*(1), 1-33.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207-227). The Hague: Mouton de Gruyter.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.

Kita, S., (2000). How representational gestures help speaking. In McNeill, D. (Ed.), *Language and gesture* (pp. 162-185). Cambridge: Cambridge University Press.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16-32.

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, *22*(8), 1212-1236.

Klima, E., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. Cambridge, UK: Cambridge University Press.

Kövecses, Z. (2008). Conceptual metaphor theory: Some criticisms and alternative proposals. *Annual Review of Cognitive Linguistics*, *6*(1), 168-184.

Kövecses, Z. (2010). *Metaphor: A practical introduction (2nd ed.)*. Oxford and New York: Oxford University Press.

Kövecses, Z. (2011). Methodological issues in conceptual metaphor theory. In S. Handl & H-J. Schmid (Eds.), *Windows to the mind: Metaphor, metonymy and conceptual blending* (pp. 23-40). Berlin/New York: Mouton de Gruyter.

Krötzsch, M., Vrandecic, D., Völkel, M., Haller, H., & R. Studer (2007). Semantic Wikipedia. *Journal of Web Semantics, 5*(4), 251-261.

Ladefoged, P., & Johnson, K. (2010). *A course in phonetics.* Boston, MA: Wadsworth.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago, IL: University of Chicago Press.

Lakoff, G. (1990). The Invariance Hypothesis: Is abstract reason based on image schemas? *Cognitive Linguistics*, *1*(1), 39-74.

Lakoff, G. (2008). The Neural theory of metaphor. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 17-38). Cambridge, UK: Cambridge University Press.

Lakoff, G. (2012). Explaining embodied cognition results. *Topics in Cognitive Science*, *4*(4), 773-785.

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

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic books.

Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor.* Chicago, IL: University of Chicago Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Langacker, R. W. (2008a). *Cognitive grammar: A basic introduction*. Oxford, UK: Oxford University Press.

Langacker, R. W. (2008b). Metaphoric gesture and cognitive linguistics. In A. Cienki & C. Müller (Eds.), *Metaphor and gesture* (pp. 249-251). Amsterdam/Philadelphia: John Benjamins Publishing.

Lederer, J. (2015a). Assessing claims of metaphorical salience through corpus data. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1255-1260). Austin, TX: Cognitive Science Society.

Lederer, J. (2015b). Exploring the metaphorical models of transgenderism, metaphor and symbol. *Metaphor and Symbol*, *30*(2), 95-117.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Lewis, T., & Stickles, E. (In press). Gestural modality and addressee perspective influence how we reason about time. *Cognitive Linguistics.*

Marcus, M. P., Marcinkiewicz, M. A., & Santorini, B. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, *19*(2), 313-330.

Mason, Z. J. (2004). CorMet: A computational, corpus-based conventional metaphor extraction system. *Computational Linguistics*, *30*(1), 23-44.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition*, *32*(8), 1389-1400.

Matlock, T., Ramscar, M., & Boroditsky, L. (2005). On the experiential link between spatial and temporal language. *Cognitive Science, 29*, 655-664.

McGlone, M. S. (2007). What is the explanatory value of a conceptual metaphor? *Language & Communication*, *27*(2), 109-126.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (1997). Growth points cross-linguistically. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 190-212). Cambridge, UK: Cambridge University Press.

McNeill, D. (2008). *Gesture and thought.* Chicago, IL: University of Chicago Press.

McNeill, D., & Duncan, S. (2000). *Growth points in thinking-for-speaking.* In D. McNeill (Ed.), *Language and gesture* (pp. 141-161). Cambridge: Cambridge University Press.

McNeill, D. & Levy, E. (1982). Conceptual representations in language activity and gesture. In R. J. Jarvella & W. Klein (Eds.), *Speech, place, and action* (pp. 271-295). New York: John Wiley and Sons.

Meir, I. (1999). Verb classifiers as noun incorporation in Israeli Sign Language. In G. Booj. & J. van Marle (Eds.), *Yearbook of morphology 1999* (pp. 299-319). Dordrecht, Netherlands: Springer.

Meir, I. (2010). Iconicity and metaphor: Constraints on metaphorical extension of iconic forms. *Language 86*(4), 865-896.

Mittelberg, I. (2014). Gestures and iconicity. In C. Müller, J. Bressem, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill & J. Bressem (Eds.), *Body – language – communication: An international handbook on multimodality in human interaction*, Vol. 2 (pp. 1712-1732). Berlin/Boston: De Gruyter Mouton.

Mittelberg, I., Hinnell, J., Beecks, C., Hassani, M., Seidl, T. (2015, July). Emergent grammar in gesture: A motion-capture analysis of image-schematic aspectual contours in North American English speaker-gesturers. Paper presented at the 13th meeting of the International Cognitive Linguistics Conference, Newcastle-upon-Tyne, UK.

Moore, K. E. (2011). Frames and the experiential basis of the Moving Time metaphor. *Constructions and Frames*, *3*(1), 80-103.

Müller, C. (1998). *Redebegleitende gesten. Kulturgeschichte – theorie – sprachvergleich.* Berlin: Berlin Verlag A. Spitz.

Müller, C. (2004). The palm-up-open-hand: A case of a gesture family? In C. Müller & R. Posner (Eds.), *The semantics and pragmatics of everyday gestures* (pp. 233-256). Berlin: Weidler Verlag.

Müller, C. (2008). *Metaphors: Dead and alive, sleeping and waking.* Chicago: University of Chicago Press.

Murata, T. (1989). Petri nets. In M. G. Singh (Ed.), *Systems and control encyclopedia: Theory, technology, applications* (pp. 3665-3670). Elmsford, New York: Pergamon Press.

Nam, S., Park, J., Kim, Y., Ham, Y., Hwang, D., & Choi, K-S. (2014). Korean FrameNet for semantic analysis. *Proceedings of the 13th International Semantic Web Conference.*

Narayanan, S. S. (1997). *Knowledge-based action representations for metaphor and aspect (KARMA)*. Ph.D. Dissertation, University of California, Berkeley.

Núñez, R. E., & Sweetser, E. (2006). With the future behind them: convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science*, *30*(3), 401-450.

Oakley, T. (2007). Image schemas. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 214–235). Oxford: Oxford University Press.

Ohara, K. H., Fujii, S., Ohori, T., Suzuki, R., Saito, H., & Ishizaki, S. (2004). The Japanese FrameNet Project: An introduction. In J. C. Fillmore, M. Pinkal, C. F. Baker, & Katrin Erk (Eds.), *The Fourth International Conference on Language Resources and Evaluation (LREC 4); Proceedings of the Satellite Workshop "Building Lexical Resources from Semantically Annotated Corpora"* (pp. 9-11).

Özyürek, A., Kita, S., Allen, S. E., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures?: Insights from crosslinguistic variations and similarities. *Gesture*, *5*(1-2), 219-240.

Parrill, F. (2008). Form, meaning, and convention: A comparison of a metaphoric gesture with an emblem. In A. Cienki & C. Müller (Eds.), *Metaphor and gesture* (pp. 195-217). Amsterdam/Philadelphia: John Benjamins Publishing.

Parrill, F. (2009). Dual viewpoint gestures. *Gesture*, *9*(3), 271-289.

Parrill, F., Bergen, B. K., & Lichtenstein, P. V. (2013). Grammatical aspect, gesture, and conceptualization: using co-speech gesture to reveal event representations. *Cognitive Linguistics*, *24*(1), 135-158.

Parrill, F., & Sweetser, E. (2004). What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture 4*(2), 197-219.

Petruck, M. R. L. (1996). Frame semantics. In J.-O. Östman, J. Blommaert, & C. Bulcaen (Eds.), *Handbook of pragmatics*. Amsterdam/Philadelphia: John Benjamins Publishing.

Petruck, M. R. L. (2013). Advances in frame semantics. In M. Fried & K. Nikiforidou (Eds.), *Advances in frame semantics* (pp. 1-12). Amsterdam/Philadelphia: John Benjamins Publishing.

Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, *22*(1), 1-39.

Prat, C. S., Mason, R. A., & Just, M. A. (2012). An fMRI investigation of analogical mapping in metaphor comprehension: The influence of context and individual cognitive capacities on processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2), 282-294.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 254-283). Cambridge: Cambridge University Press.

Ruiz de Mendoza Ibáñez, F. J., & Perez Hernandez, L. (2011). The contemporary theory of metaphor: Myths, developments and challenges. *Metaphor and Symbol*, *26*(3), 161-185.

Ruppenhofer, J., Ellsworth, M., Petruck, M. R., Johnson, C. R., & Scheffczyk, J. (2010). FrameNet II: Extended theory and practice. Berkeley, CA: International Computer Science Institute.

Ruppenhofer, J., & Michaelis, L. A. (2010). A constructional account of genre-based argument omissions. *Constructions and Frames*, *2*(2), 158-184.

Salomão, M. M. M., Torrent, T. T., & Sampaio, T. F. (2013). A linguística de corpus encontra a linguística computacional: Notícias do projeto FrameNet Brasil. *Cadernos de estudos linguísticos*, *55*(1), 7-34.

Saussure, F. de. (1983 [1916]). C. Bally & A. Sechehaye (Eds.) *Course in general linguistics*. Trans. R. Harris. La Salle, Illinois: Open Court.

Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 266-298). Cambridge, UK: Cambridge University Press.

Schoonjans, S. (2014). *Modalpartikeln als multimodale Konstruktionen. Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*. Ph.D. Dissertation, KU Leuven.

Shibatani, M. (1985). Passives and related constructions: A prototype analysis. *Language*, *61*(4), 821-848.

Slobin, D. (2006). What makes manner of motion salient. In M. Hickmann & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories* (pp. 59-82). Amsterdam & Philadelphia: John Benjamins.

Slobin, D., & Hoiting, N. (1994). Reference to movement in spoken and signed languages: Typological considerations. In S. Gahl, A. Dolbey, & C. Johnson (Eds.), *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society* (pp. 487-505). Berkeley, CA: Berkeley Linguistics Society.

Smith, I., & Sweetser, E. (2015, July). Conditional constructions, gestural space, and mental spaces. Paper presented at the 13th meeting of the International Cognitive Linguistics Conference, Newcastle-upon-Tyne, UK.

Stam, G. (2006). Thinking for speaking about motion: L1 and L2 speech and gesture. *International Review of Applied Linguistics*, *44*(2), 143-171.

Steen, G. J. (2007). *Finding metaphor in grammar and usage: A methodological analysis of theory and research*. Amsterdam/Philadelphia: John Benjamins Publishing.

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. Amsterdam/Philadelphia: John Benjamins Publishing.

Steen, F., & Turner, M. (2012). Multimodal construction grammar. In M. Borkent, B. Dancygier, & J. Hinnell (Eds.), *Language and the creative mind* Stanford, CA: CSLI Publications.

Stefanowitsch, A. (2005). The function of metaphor: Developing a corpus-based perspective. *International Journal of Corpus Linguistics*, *10*(2), 161-198.

Stefanowitsch, A., & Gries, S. T. (Eds.). (2007). *Corpus-based approaches to metaphor and metonymy*. Berlin/New York: de Gruyter.

Stickles, E., David, O., & Sweetser, E. (2016a). Grammatical constructions, frame structure, and metonymy: Their contributions to metaphor computation. In A. Healey, R. N. de Souza, P. Pešková, & M. Allen (Eds.), *Proceedings of the 11th meeting of the High Desert Linguistics Society* (pp. 317-345). Albuquerque, NM: High Desert Linguistics Society.

Stickles, E., & Dodge, E. (2016, January). Literal vs. figurative language use affects the frequency of syntactic patterns. Paper presented at the 90th Annual Meeting of the Linguistic Society of America, Washington, D.C.

Stickles, E., Dodge, E., David, O., & Hong, J. (2016b). Formalizing contemporary conceptual metaphor theory: A structured repository for metaphor analysis. *Constructions and Frames 8*(2).

Stickles, E., Dodge, E., & Hong, J. (2014, November). *A construction-driven, MetaNet-based approach to metaphor extraction and corpus analysis.* Paper presented at the 12th Meeting of Conceptual Structure, Discourse, and Language, Santa Barbara, CA.

Strömqvist, S., & Verhoeven, L. (Eds.) (2004). *Relating events in narrative, Vol. 1 and 2.* Mahwah, NJ: Lawrence Erlbaum.

Sullivan, K. S. (2006). Frame-based constraints on lexical choice in metaphor. In Z. Antić, C. B. Chang, E. Cibelli, J. Hong, M. J. Houser, C. S. Sandy, M. Toosarvandani, & Y. Yao (Eds.), *32nd Annual Meeting of the Berkeley Linguistics Society* (pp. 387-400). Berkeley, CA: Berkeley Linguistics Society.

Sullivan, K. S. (2007). *Grammar in metaphor: A construction grammar account of metaphoric language*. Ph.D. Dissertation, University of California, Berkeley.

Sullivan, K. S. (2013). *Frames and constructions in metaphoric language.* Amsterdam/Philadelphia: John Benjamins Publishing.

Supalla, T. (1992). The classifier system in American Sign Language. In C. G. Craig (Ed.), *Noun classes and categorization* (pp. 181-214). Amsterdam/Philadelphia: John Benjamins.

Sweetser, E. (1990). *From etymology to pragmatics*. Cambridge, UK: Cambridge University Press.

Sweetser, E. (1998). Regular metaphoricity in gesture: Bodily-based models of speech interaction. In *Actes du 16 congres international des linguistes*. Elsevier.

Sweetser, E. (2007). Looking at space to study mental spaces. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 201-224). Amsterdam/Philadelphia: John Benjamins Publishing.

Sweetser, E, & Sizemore, M. (2008). Personal and interpersonal gesture spaces: Functional contrasts in language and gesture. In A. Tyler, Y. Kim, & M. Takada (Eds.), *Language in the context of use: Cognitive and discourse approaches to language and language learning* (pp. 25-52). Berlin: Mouton de Gruyter.

Sweetser, E., & Stec, K. (2016). Managing multiple viewpoints with gaze. In B. Dancygier, W. Lu, & A. Verhagen (Eds.), *Viewpoint and the fabric of meaning: Form and use of viewpoint tools across languages and modalities* (pp. 237-258). Berlin/New York: de Gruyter.

Talmy, L. (1983). How language structures space. In H. L. Pick Jr. & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research, and application* (pp. 225-282). New York: Plenum Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language Typology and Syntactic Description*, *3*, 57-149.

Talmy, L. (1991). Path to realization: A typology of event conflation. In L. A. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society* (pp. 480-519). Berkeley, CA: Berkeley Linguistics Society.

Talmy, L. (2000). *Toward a cognitive semantics* (Vol. 1 and 2). Cambridge, MA: MIT press.

Taub, S. (2001). *Language in the body: Iconicity and metaphor in American Sign Language.* Cambridge, UK: Cambridge University Press.

Taub, S., & Galvan, D. (2001). Patterns of conceptual encoding in ASL motion descriptions. *Sign Language Studies*, *1*(2), 175-200.

Taub, S., & Galvan, D. (2004). The encoding of motion information in ASL. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative, Vol. 2, Typological and contextual perspectives* (pp. 191-217). Mahwah, NJ: Lawrence Erlbaum.

Taub, S., & Piñar, P. (2001). Comparing the informational content of speech/gesture and sign language. In C. Chang, M. J. Houser, Y. Kim, D. Mortensen, M. Park-Doob, & M. Toosarvandani (Eds.) *Proceedings of the 27th Annual Meeting of the Berkeley Linguistics Society* (pp. 465-473). Berkeley, CA: Berkeley Linguistics Society.

Tong, Y., & Cienki, A. (2015, July). Motion, metaphor and gesture: A comparison between referential gestures referring to concrete and abstract motion. Paper presented at the 13th meeting of the International Cognitive Linguistics Conference, Newcastle-upon-Tyne, UK.

Tong, Y., & Cienki, A. (2016, July). Referential gestures and physical actions: A form-based analysis. Paper presented at the 7th meeting of the International Gesture Studies Society, Paris, France.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes.* Oxford, UK: Oxford University Press.

TV News Archive. 2012-present. Available at https://archive.org/details/tv.

UCLA NewsScape Archive. 2005-present. Los Angeles, California. Available at http://newsscape.library.ucla.edu/.

Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, *34*(3), 459-520.

Wehling, E. (2017). Discourse management gestures: An embodiment account of gesture pragmatics in face-to-face discourse. *Gesture, 16*(2).

Wilcox, S., & Wilcox, P. (1995). The gestural expression of modality in ASL. In J. Bybee & S. Fleischmann (Eds.), *Modality in grammar and discourse* (pp. 135-162). Amsterdam/Philadelphia: John Benjamins.

Wilcox, S., & Xavier, A. N. (2013). A framework for unifying spoken language, signed language, and gesture. *Todas as Letras-Revista de Língua e Literatura*, *15*(1), 88-110.

Winter, B., Perlman, M., & Matlock, T. (2013). Using space to talk and gesture about numbers: evidence from the TV News Archive. *Gesture*, *13*(3), 377-408.

Yu, N. (2011). A decompositional approach to metaphorical compound analysis: The case of a TV commercial. *Metaphor and Symbol*, *26*(4), 243-259.

Zima, E. (2014). English multimodal constructions: A construction grammar perspective. In *Papers of the Linguistic Society of Belgium 8*. Available from http://uahost.uantwerpen.be/linguist/SBKL/sbkl2013/Zim2013.pdf.

Appendix A: Text of experimental stimuli for Chapter 5.

Note: Stories 4 and 9 were used as training stimuli.

I. Stories with metaphoric path

a. Stories with progressive aspect

1. John invested his life savings in the stock market. The price of his stock was slowly [rising/increasing]. John became very excited. The stock market was [skyrocketing/rapidly improving]! Then, the country went into a recession. The stock market index was [plummeting/drastically decreasing]. John's savings were gone.

2. Adrian was enrolled in a calculus class. His exam grades were steadily [rising/improving]. He visited the tutoring center regularly. Adrian's test worries were [falling away/lessening]. Then he took the final exam. His anxiety was [climbing/increasing]. He couldn't remember anything he'd learned.

3. Sunny was nominated for prom queen. Her friends said her chances were [rising/improving]. Everyone would vote for her. Her hopes were [soaring/dramatically increasing]. Then she tripped and broke her nose. Her mood was [sinking/becoming negative]. She walked into the prom with a swollen, bruised face.

4. Robert worked really hard at his job. He was slowly [climbing/gaining status] within the company. He was watching his life plans develop. His life was finally [advancing/developing]. Robert's wife Tina felt like Robert spent too much time at work. Robert was [escaping/resolving] his depression by working more. He continued to progress within the company.

b. Stories with perfective aspect

5. Martin was a senior in high school. His worries [fell/decreased] as he filled out college applications. His counselor said his SAT scores were good. His confidence [rose/improved]. Then he found out his top choice only accepted 10% of applicants. His hopes [sank/decreased]. Later, he received a letter from the college.

6. A local punk band named Limp Lizards played a lot of local bars. They didn't have a big enough fanbase to [climb the charts/get more views]. It seemed like they would never make it as a band. They weren't on the radio and their meager popularity [fell/decreased]. Then an agent took a chance on them and signed them on with a big record label. Their radio airtime [upsurged/quickly increased]. The Limp Lizards might make it big.

7. Manuel was a college student who applied to a very competitive business school. His emotions [toppled/were calamitous] as he looked over his application results. He didn't make it in, and he felt like a failure. He felt as if his academic career [had tumbled/was unsuccessful]. But then Manuel became inspired and applied himself to his other areas of interest. He [entered/became a part of] every club he could find. Michael felt like he found new goals and aspirations.

8. Marina was a new student and was taking French for the first time in her life. She was confused and her confidence [plummeted/became low]. Marina felt like she would never understand anything in this complicated language and considered dropping the class. Every day she came into class her anxiety levels [rose/increased]. Then, her professor told her to go find a French tutor. Her confidence slowly began to [climb/increase]. Marina continued taking the class.

II. Stories with metaphoric manner
    a. Stories with progressive aspect

9. Emily was ready for a date with Sam. Her stomach was [tumbling/aching]. She was so nervous as she waited at the café. Her heart was [racing/pounding]. Sam hadn't shown up yet. Emily's excitement was [stumbling/diminishing]. As she walked home, she happened to see an old flame from high school.

10. Eva was campaigning for re-election to her Senate seat. She was [pulling ahead/improving] in the polls. She had more campaign contributions than her opponent. Eva was [cruising to victory/likely to win]. Then she was accused of accepting bribes. Her opponent's support was [surging/increasing]. On election day, the candidates were tied.

11. Alexia had a job interview. She was [flying through/acing] the skills evaluation. The interviewer seemed impressed. Alexia was [driving home/emphasizing] her qualifications. The interviewer asked about her prior job experience. Her answers were [fumbling around/confusing]. The interviewer thanked her for her time.

12. Remy was panicking because he just remembered he had guests coming over for a dinner party tonight. He thought of his mother-in-law's condescending comments at the last dinner as he was to [rolling through/quickly preparing] the hors d'oeuvres. Remy didn't want to be embarrassed in front of his in-laws again. Remy's cooking was [sailing smoothly/progressing] when he dropped the food onto the floor. But then, he remembered that there was a wonderful Chinese restaurant around the corner who could deliver quickly. He was nervously [stumbling through/completing] an order just thirty minutes before the expected arrival of his in-laws. The doorbell eventually rang.

b. Stories with perfective aspect

13. Arnold had a big final project in his computer science class. He [jumped into/started] programming right away. He stayed up all night to work on it. His programming [charged ahead/quickly progressed]. Then his program stopped working. He [stumbled/faltered] and couldn't find the error. The project was due the next day.

14. Cole Thurman was elected for another term in office. He [flew past/defeated] the other candidates in the race. Mr. Thurman was prepared to introduce his new bills in the legislature. He was ready to [advance through/eliminate] the problems ahead. However, the media let out a secret about his affair with another woman. When he was interviewed about, he [rolled through/answered] lots of questions. He was worried about how he was being portrayed in the media.

15. Jebediah the oil tycoon had dedicated his life to finding oil. Recently, Jeb [was shaken by/learned] some terrible news. An international oil cartel released millions of barrels into the market. Oil prices had [stumbled/decreased] by a lot. It had seemed that Jeb would have to sell his business for good until the one fateful day that Jeb had struck oil. [His heart shook/ He felt morose].

16. Uyen was a dedicated elementary school teacher. She carefully [walked through/explained] the material [with/to] her fifth grade class. They earned good grades on their first math test. Maria [sped into/quickly began] tougher material. The principal warned her that the districts would fire teachers whose students failed the state math exam. The students' performance [crawled/got worse]. At the end of the school year, they took the state exams.