

Stage Two Report:
A Statistical Language Model for the
Generation of Figurative Language

by
Stephen McGregor

An outline of a thesis to be submitted to the University of London
for the degree of Doctor of Philosophy

Department of Electronic Engineering
Queen Mary, University of London
United Kingdom

September 2015

Abstract

MIMO (Multiple Input Multiple Output) technology has been regarded as a practical approach to increase the wireless channel capacity and reliability.

Abstract goes here...

Table of Contents

Abstract	i
Table of Contents	ii
1 Introduction	2
1.1 A Hypothesis	5
1.2 Contributions to the Field	6
1.3 Structure of The Dissertation	8
2 Background	10
2.1 Computers, Creativity, and Making Meaning	10
2.2 Concepts	13
2.3 Words	13
2.4 Word Counting	14
2.5 Minds and Spaces	16
2.6 Words and Concepts	18
2.7 Computers and Language	19
3 Theoretical Foundation	22
3.1 Concepts in Context	22
3.2 A Literal, Dynamic Language Model	23
3.3 The Geometry of Metaphor	24

4	Methodology	26
4.1	A Lexical Base Model	26
4.2	Taxonomy Recapitulation	27
4.3	Analogy Completion	28
4.4	Metaphor Generation	28
5	Implementation and Results	29
5.1	The Base Model and Projections	29
5.2	Taxonomies	30
5.3	Analogies	31
6	Evaluation	32
6.1	Taxonomy and Analogy	32
6.2	Metaphor	33
7	Conclusion	34
	References	35

Chapter 1

Introduction

In Italo Calvino's (1983/1985) whimsical novel of ideas *Mr. Palomar*, the titular character reflects upon the mating rituals of turtles. Encased as they are in an unfeeling shell, with only a filament of sensing flesh with which to prod the world, Palomar concludes that for turtles sensuality must take on an almost entirely cerebral aspect: "The poverty of their sensorial stimuli perhaps drives them to a concentrated, intense mental life, leads them to a crystalline inner awareness," (p. 18).

The chelonian condition imagined by Palomar is, at least in a certain metaphoric respect, the endemic situation of a statistical language model. The environment of such a model is at best a sparse simulacrum of existence, a trickle of information that is only about anything in a sense completely external to the model itself. If language is characterised by an aboutness that anchors itself in acts, intentions, and experiences in the world, then the blind traversal of corpora severs the semantic anchor lines and allows a model to curl into itself. The residue of this process, an accumulation of symbols trailed by kite tails of numbers, can really only be directly interpreted as an echo of language, revealing tendencies but demanding clever interpretive reconstruction in order to be of any practical use.

This thesis sets out to explore the question of if and how a statistical language model

can be involved in the generation of new meaning. That such models can at least be interpreted in a way that compellingly exposes semantic and syntactic relationships has been demonstrated, and the trend in ongoing work in this area is to continue to develop mechanisms for recapitulating datasets while allowing the features of the models themselves to drift further away from semantic transparency. The positive outcome is a battery of computationally tractable and sometimes mathematically elaborate mechanisms for processing natural language, some of which appear to have real practical applications in tasks such as information retrieval and machine translation. There is a risk, however, of a proliferation of models that are very good at solving bespoke micro-problems but reveal very little about the nature of language.

One of the objectives of this work, then, is to explore the relationship between the abstract geometry of relatively high-dimensional language models and the concrete geometry of the world that language must be about. ? places the geometric representation of the world at the foundation of mental existence, classing it in terms of fundamental *intuitions* rather than supervenient *concepts*. Already a parallel between language and conceptualisation is emerging here: concepts are about space, and also somehow composed of space, in the way that words are the world and also in the world.

When words become points, the language itself, consisting of those words and the relationships between them, becomes a spatial entity, but in an abstract space where there is not even a nominal distinction between space and information. It is no longer a space in the essential sense of Kant, but rather a generative space, a space that is somehow an index to that other space which is co-extensive with reality. In this kind of space, it is the space itself that does the work: it is the space that comes to characterise the features of each point, and, when points interact, they interact by virtue of nothing other than their positions. When a statistical language model is interpreted, it is therefore the internal space of the web of relationships working as a dynamic system unto itself that are being explored.

This thesis seeks to ground these philosophical arguments in computational exper-

imentation. The methodology described and pursued in the following pages finds its motivation in theoretical considerations of language and cognition, but its techniques are rooted in the contemporary approach to statistical language modelling. The unique contribution of the model that will be described here, though, is to offer a robust mechanism for the prolific contextualisation of the lexical information inherent in the network of statistics that delineate its spatial aspect. In this regard, what is being modelled is the relationship between language and concepts as they come about in the world, in the unfolding, ready-to-hand, non-sentential character of cognition.

With that said, no strong claim will be made here about having resolved the hard questions surrounding the mind. There is no suggestion that the model described in this thesis is somehow simulating cognition; indeed, the mechanism of a computational model such as this, substantiated by nothing more than the drizzle of data that penetrates its shell, is specifically different than the way a cognitive agent evolves into a situation of deep, tight, multifarious entanglement with the world. Furthermore, it is essential not simply to dismiss the challenges that face any attempt at a philosophically robust description of cognition. Indeed, these philosophical problems will become a guiding light for this work, with the hope that a good model that is honest about the level of abstraction at which it is applicable may reveal some interesting aspect of the elaborate whole.

The strong claim which will be made is that the model described here can participate autonomously in the discovery of new and useful meaning, and it is in this regard that this work finds its roots in the multifaceted field of computational creativity. When ? states that “only the act of meaning can anticipate reality” (p. 76), he suggests that beyond a mere syntactic encoding of events in the world, language is tied up in a process of connecting mental existence to being in the world. Creativity, construed as meaning making, is taken as a broad target for a project that attempts to model the semantics inherent in the relationship between words and concepts. The work presented here does not aspire to the kind of phenomenological richness that is at the heart of Wittgenstein’s

later philosophy, but it has been designed to generate informational structures that might be interpreted as models of internal representations with the interactional dynamics necessary for generating significant new meaning.

So, while there may be a popular perception that figurative language is in some regard more creative than plainly propositional language, the choice of metaphor as a target is actually motivated by the way in which figurative language exposes the functionality of words. In the course of constructing a metaphor, a linguistic agent is seeing meaning as an affordance for some communicative action: just as a shoe might become a hammer, or a chair a weapon, the concept of a butcher comes to stand in for a certain type of surgeon in a certain context. It is this opportunism inherent in language which the project presented here will attempt to model.

1.1 A Hypothesis

The premise of this this thesis is that conceptualisation can be modelled in geometric terms, not least because concepts are about a reality that is essentially geometric. Furthermore, the relationship between concepts and language is to be discovered in the way that language is in the world: language is about the world, but it is also made of the world, and the relationship between the two is as determinate as it is irretrievable. When language comes about as an affordance of communication, it is a physical opportunity for expressive action that is being directly perceived as an aspect of an environment, as a way to use and change that environment. The situation of language is inextricably spatial and functional, and so the language itself can only really properly be modelled as something that is dynamic and geometric.

Following on this, the hypothesis offered here is that figurative language can be modelled in terms of the geometrical properties of a statistical language model. In particular, this paper predicts that a dynamically generative distributional semantic model, equipped with a facility for projecting contextually informed subspaces from a base lex-

ical space, will remit clusterings of word-vectors that can be interpreted as geometric representations of conceptualisations, and the work to be described in the following chapters is predicated upon this prediction. A significant consequence of the geometric modelling of concepts is that the geometric properties of these conceptually charged word spaces will allow for the isomorphic mapping between conceptual representations as a methodology for modelling the generation of figurative language based on contextually specific information.

By way of testing this hypothesis, a statistical language model will be developed and subjected to a set of experiments. In particular, the model will be tested in terms of established approaches to the recapitulation and extension of lexical ontologies and to the completion of analogies. Furthermore, in an effort to test the pragmatic efficacy of the model, metaphoric output will be measured in a study involving human subjects and also, through a social media application of a metaphor generating bot, submitted to the tumult of public discourse. The purpose of this final measure is to test the pragmatic applicability of the model's output in a real communicative situation. It is by virtue of this last point in particular that this remains fundamentally a thesis about computational creativity, in that the trafficking of statements produced by an on-line application will be taken as affirmation that the behaviour of a system associated with the model has been interpreted as essentially creative in its generation of new, useful language.

1.2 Contributions to the Field

The concrete manifestation of the work described here is a system for generating conceptual metaphors: given linguistic input describing a basic situation in the world, for instance, "that surgeon is sloppy", the system will return a figuratively loaded description such as "that surgeon is a butcher". The mechanism involved in discovering such new relationships between conceptual entities employs dynamically interactive representations, in that the mathematical, geometric nature of the model's word-labelled repre-

sentations provides a platform for intermeshings which might lead to the emergence of representations that are dynamic on a conceptual level. As such, from the perspective of computational creativity, there is a case to be made that something like internal representations of conceptual schemes are being modelled, and it will be argued here that this model can be considered in terms of the autonomous generation of novel, compelling linguistic artefacts through a process that can be deemed creative.

A similarly concrete but broader contribution of this work is the presentation of a new language model, specifically engineered to handle the nuance of conceptualisation that the pragmatics of natural language capture so well. More than just another new model for performing a range of NLP type tasks, the goal here is to provide an endorsement for a particular type of model, namely, one that maintains features that are, in a statistical sense, interpretable as actual informational observations. As White (2009) has put it, “we need some sort of cognitive grasp” (p. 99) on the way in which statistical techniques tease a semantic model out of large scale corpora. The work presented here is designed as a transparent implementation of a theory about the contextualisation involved in mapping from words to concepts, and so the expectation is that this might at least indicate a methodological stance for achieving this cognitive desideratum.

Finally, on a more abstract level, this work is offered as an indication of one way in which computers might be used to empirically demonstrate the efficacy of philosophical arguments. In particular, the model presented here takes a theory about the nature of language and the contextual way in which language becomes deeply entangled with conceptualisation, and it applies this theory to an exercise in the design and testing of a model. One of the broad aspirations of this dissertation is to hint at a way in which future work in both computer science and philosophy can seek a harmonious balance between theory and practice. This is offered as something of an antidote to the fraught relationship between the philosophy of mind and the application of computer modelling, which has been persistently hampered by very reasonable objections to the erroneous claim that minds and computers are somehow just like each other. The counterargument

offered here is that minds and computers are fundamentally different, but there is still much to be gained in terms of both insight and application from using computers to model some of the things – language, conceptualisation, creativity – that are clearly native to the mind. The key here is to keep in mind certain truths about the physical and, on the other hand, observer dependent nature of computing, and not to lose sight of the level of abstraction at which computational models can be understood.

1.3 Structure of The Dissertation

In addition to this introductory chapter, this thesis contains six more chapters. The second chapter will present a review of relevant literature, both in terms of historical theoretical material and contemporary empirical work in the several fields related to this project. This literature review will be followed by an exposition of the theoretical foundation of the project itself, describing the conceptual basis for the experimental work that will subsequently be described.

The fourth chapter will lay out the methodological approach to the construction of a generative language model, as well as the equations describing the mechanism for using this base model to project contextually informed subspaces. It will also describe techniques for taxonomy recapitulation and the discovery of conceptual metaphors by way of mapping between word clusters discovered in the base model’s projections. The fifth chapter will describe the implementation of the language model, including its application to tasks involving completion of established test sets and recapitulation of widely used lexical ontologies. Results of these experiments will also be presented in this chapter.

The sixth chapter will discuss the experimental results presented in the previous chapter from an analytical perspective. Additionally, this chapter will present two further mechanisms for evaluating the model’s metaphor generating capabilities in particular, in terms of a study involving human subjects and a pragmatically geared application of social media. The seventh and final chapter will provide a holistic diagnosis of the

project, returning in particular to the philosophical ambitions of this work and also considering potential future applications of the model.

Chapter 2

Background

The ambitious scope of this project requires a thorough grounding in the literature from a range of disciplines. The risk in such an inclusive undertaking is that there are many opportunities for going astray. The reward, a piece of empirical work that exemplifies the possibility of using computers to do theoretically significant research, seems to justify the risk, but only if the final product is well placed in the context of existing work in the area. The following chapter will delineate the territory that this thesis seeks to explore, by way of a survey of the considerable canon of theoretical work relating to the problems of minds, concepts, and language as well as the increasingly substantial body of practical work on computational linguistics and its application to the modelling of creative language generation.

2.1 Computers, Creativity, and Making Meaning

At its heart, this is a project about how meaning comes about in a reality which, as scientists, we accept as being fundamentally materialist. The question of the place of the mind – the instrument of meaning making – in relation to the realm of matter has been arguably the unifying theme of modern Western Philosophy, dating back to the

cogito of ? and serving as the broad target of subsequent forays into empiricism (??), idealism (??), phenomenology (??), and the verity of metaphysics itself (?). In the case of the last philosopher in particular, the project described here is concerned with what Wittgenstein has described as “the act of meaning”, that process by which being in the world as a communicative agent somehow results in the emergence of intention from the particulate morass, an evident *creatio ex nihilo*.

The recent trend in philosophy, not to mention in empirically fastidious fields such as cognitive science and psychology, has been towards a resolute materialist reductionism, to such an extent that ? reports that in the current cognitive scientific milieu, “even the word ‘Cartesian’ is often used as a term of abuse,” (p. 12). Those thinkers who do indulge in dualist theories of mind – and here Chalmers (1996) and Fodor (2008) are two of the most notable examples – do so by means of subtle argumentation and elaborate analogy.¹ By and large, though, there is a broad acceptance that the mind must somehow supervene upon physical reality, ranging from the strictly eliminativist neurophilosophy of ? to the emergentist enactivism of ? and ?. As ? has put it in his biologically grounded account of the emergence of goal-directed behaviour, “Such concepts as information, function, purpose, meaning, intention, significance, consciousness, and value are intrinsically defined by their fundamental incompleteness,” (p. 23)

So the theoretical question which underlies the empirical research presented in this thesis regards the way in which meaning emerges in a universe of particles and signals. This question is related to ?’s (?) pursuit of a *theory of strongly semantic information*—doubly so, in that, like Floridi, we will seek a quantitative approach to modelling the relationship between semantics and data, in the information theoretical tradition of ?. An early attempt to import technical insight from signal processing into the study of meaning can be found in ?’s (?), who use Shannon-type metrics as the basis for

¹Chalmers’s version of the “philosophical zombies” argument has been particularly persistent here, though parallels between this and the earlier “swampman” scenario from ? and the self-manifesting lion of ? have been arguably underexplored in the literature. The fact that thought experiments have proved so prodigious a source of ideas about the mind is in itself noteworthy, and is addressed in a certain respect in ?’s (?) case against “intuition pumps”.

quantifying the inferential properties associated with the semantic content of sentences, followed by ?, who described the formation of meaningful concepts in terms of the development of internal semantic structures that evolve to indicatively correspond with quantifiable informational situations in an environment. Subsequent forays in a *situation logic* designed to model semantic information content in a way which is simultaneously measurable and context specific (?) have led to a

Returning to the question of the *epistemic gap* between mind and matter (??), ? has proposed an information-theoretic metric for assessing the degree of consciousness exhibited by a system, an idea which ? has pursued to the extent of proposing a panpsychism by which any system in the universe has a quantifiable degree of consciousness.

In the spirit of these quantitative approaches to meaning, the work presented in this thesis seeks a data processing approach to modelling semantics—an approach which is computationally tractable but at the same time avoids committing to the case that the mind is in a computer in anything other than the most metaphoric of senses. Furthermore, when it comes to the emergence of meaning, creativity is a kind of watchword to keep us faithful to the question of how meaning

With regard to this question concerning the emergence of meaning, creativity serves as a kind of watchword, a beacon to keep us faithful to the issue of how artefacts can be about ideas in the

we are concerned, fundamentally, with the aboutness of mental states, and likewise conscious of the pitfalls of trying to cash these states out as computations.

The study of computational creativity in particular neatly aligns the

with the philosophical dilemma of the relationship between minds and computation as a broad

Searle

At stake here is, in at least small part, a definition of *creativity* itself, and ? has pointed out that this term must be considered “essentially contested,”

Nonetheless, the approach towards creativity

and if we can discover a reasonably convincing computational approach to modelling creativity, then we will likewise have discovered the ways in which

2.2 Concepts

that concerns itself with the emergence of meaning in the universe needs to likewise be grounded in a concrete sense of what constitutes a concept, as concepts are

?, who holds that concepts are fundamentally composable and, at least in some instances, atomic mental entities

what is a concept? ? has endorsed a kind of *atomism* in

(cf ?, for a case against conceptual schema in general)

The work presented in this thesis embraces the insights of Relevance Theory, which broadly postulates that

? has described this

- Lakoff - Davidson

2.3 Words

Language is the point at which the abstractions surveyed in the last section make contact with palpable data: a mainstay of the so-called Cognitive Revolution, kicked off by Chomsky’s (1959) pointed takedown of Skinner’s (1957) attempt to apply overt psycho-

logical behaviouralism to the study of language, has been the idea that language, more than simple signalling, can only be understood in terms of reflecting mental content. Indeed, there are some thinkers who would go so far as to argue, as Dennett (1996) seems to, that having language is a necessary condition for having what is thought of as a mind—Fodor’s (n.d.) notion of a language of thought, as described above, could likewise be construed as tantamount to such a stance. The notion of language as either an impediment or an incipient to pure thought was essential to Russell (1905, 1931), and was also a primary theme for Wittgenstein’s (1922) earlier work.

? advocated the idea that the signs constituting a language are involved in a physically grounded process that is perpetually unfolding in tight entanglement with the world. While Peirce’s work was fated to academic obscurity until the later 20th Century, the theme of the environmentally situated character of language would return to prominence with the later pragmatics of ?.²

2.4 Word Counting

Finally, the technical methodology of this project is grounded in recent and ongoing success in statistical approaches to language modelling. The tradition of word-counting in order to predict sequences in language races its roots back to the fastidious work of ?, who tabulated co-occurrences of characters in Pushkin’s *Eugene Onegin* by hand. The research described in this thesis is situated within the paradigm of *distributional semantics* which has its roots in Harris’s (1954) work examining “meaning as a function of distribution,” (p. 155). The guiding principle of this area of research is the idea that “words that occur in the same contexts tend to have similar meaning,” (Pantel, 2005, p. 126). Other early work in the field included

²The use of *pragmatics* to describe the work of Peirce and some of his contemporaries as compared to the latter efforts of Grice and his disciples is probably a coincidence of intent rather than an indicator of scholastic continuum, but there is nonetheless an incidental correspondence to be analysed—see Pietarinen (2004) for a treatment.

who were generally confined to theoretical output due to historical constraints on computational power.

3

Notwithstanding ?’s (?) formidable case that the bulk of a language cannot be captured in numbers quantifying the relative situation of words, the fact remains that, when it comes to computational approaches to language, a matrix capturing the nature of some concordance or another remains one of the main games in town. Two post-millennial developments have facilitated the rapid development of research into computational methods employing word counting approaches: the accretion of large-scale, widely available digitalised textual data, and the advent of computers with the processing power to efficiently handle data of this scale.

which characterise a natural language could possibly be found in a matrix of statistics about the way that words co-occur with one another across some corpus, regardless of how comprehensive the data may be.

On the other hand, in as much as computational approaches to language modelling are concerned, co-occurrence statistics are the main game in town.

At its heart, this is a project about the potential for symbol manipulating machines to model creativity, and in particular the creative use of language. Computational Creativity is a field that, as Wiggins (2006) puts it, uses “computational means and methods” to study “behaviour exhibited by natural and artificial systems, which would be deemed creative if exhibited by humans,” (p. 210). At stake here is the problem of what creativity actually is, and, as ? have said, the concept of creativity must be *essentially contested*. Nonetheless, a basic position will be taken here that at the root of creativity is, in the spirit of Wittgenstein (1953/1967), an act of meaning-making, by which a new

³I note, in passing, that scholars in the field often allude to ?’s (?) quip “you shall know a word by the company it keeps,” (p. –) as seminal to distributional semantics. I suggest that this turn of speech has actually been taken out of the context of Firth’s intent, which was to consider the role of idioms in particular in a comprehensive programme of linguistic study at various levels of abstraction, a project which is probably actually very much at odds with purely statistical approaches to language.

way of conceptually representing something in the world comes about.

The prevalent take on computational creativity has focused on an AI approach involving the traversal of state spaces for potential creative artefacts, a methodology with deep roots in Boden (1990). Philosophical problems with this essentially symbolic mode of generation have been illustrated by McGregor, Wiggins, and Purver (2014), though: the dependency on preconditioned representations remits little in the way of actual creative behaviour on the part of the symbol manipulating agent. Here the hard question of evaluation comes up, and the evaluative challenges peculiar to computational creativity have been addressed by Ritchie (2007), Colton and Wiggins (2012), and Jordanous (2012), to name a few. At the core of this project is the proposal that, for an agent to be perceived as genuinely and autonomously instigating meaning-making, there has to be at least a nominal notion of the trafficking of dynamic, interactive representations within the agent’s cognitive framework. It is here that computational creativity will serve as a compass to keep this project on course: by returning to the question of whether the model outlined here can be described as an agent that is behaving creatively, the pragmatic issue of the role of contextualisation in the actual use of language will stay central to the thesis.

In addition to a general survey of the field, this project’s particular commitment to conceptualisation and metaphor invites a survey of the appreciable work done in the computational production of metaphor and analogy. With roots in the conceptual blending theory of ?, this area has been explored more recently by the likes of Veale and Hao (2007) and ?. Additional attention is due to work in poetry generation coming from ?, ?, ?, ?, and others.

2.5 Minds and Spaces

? cautions against considering language as an “organising system” (p. 11) for conceptual schemes. In the process of understanding a sentence, a contextually specific theory of

meaning is devised on the spot, so to speak, that permits “an acceptable theory of belief” (p. 18) regarding the sentence’s author. This means, though, that the conceptual organisation supposedly inherent in language is in fact a fleeting artefact that temporarily props up a momentary propositional stance, and language is just a mechanism for groping towards some sort of understanding between communicants regarding positions on the relationship between meaning and belief. A consequence of this theoretical insight is the temptation to resort to a view such as that taken by ?, who holds that language is “a cognition-enhancing animal-built structure,” (p. 370).

The theoretical premise of this project is that a structural, spatial, geometric language model can successfully capture significant aspects of conceptualisation. In this regard, the work is grounded in the theories of ? and ?. Gärdenfors in particular postulates a kind of cognitive middle ground where conceptual spaces, characterised by coherent dimensions which map to features of concepts, stand between low level stimuli and high level symbolic representations. Widdows, on the other hand, is more concerned with the description of a language model where a geometry of words facilitates the mechanisms of logic. In both cases, these authors provide insight into the way that a specifically geometric view of representations allows for the dynamic interaction of symbols by way of using language as a cognitively productive system.

This project’s commitment to building a system which remits evidence of dynamic representational structures is admittedly reminiscent of ?, who holds that concepts are fundamentally compositional mental entities. Any nativist forebodings engendered here are diffused somewhat by ?’s (?) insight that conceptual extension cannot strictly be found “in the head” (p. 170). In this regard, work in embodied cognitive science, rooted in the enactivism of ?, provides the cognitive foundation for a spatially charged theory of concepts. So ? goes on to offer a definition of information as “the intentional relation of the system to its milieu,” (p. 59), a view which attempts to capture the tight relationship between the spatial situation of the world and the symbolic representation of that situation (?, is another notable proponent of this approach). Similarly, the Hei-

deggarian account of ?, with reference to ?, postulates an *action-oriented representation* where the very structural nature of representations are “deeply dependent on the specific context of action,” (p. 196). In the end the theory of language embraced here will be characterised by ?’s (?) notion of *affordances*: just as with physical objects, words present themselves as opportunities for communication, and it is in the ready-to-hand appropriation of meaning that the genesis of metaphor is to be discovered.

2.6 Words and Concepts

In the spirit of ?, this thesis is predicated on the idea that the signs constituting a language are involved in a physically grounded process that is perpetually unfolding in tight entanglement with the world. As mentioned in the previous section, this stance leads on to the view that the compositional components of a language offer themselves as affordances for communication in the context of a particular environmental situation. It is by virtue of this opportunistic seizing of linguistic units as operational elements in a system of contextualised conceptual representation that language becomes so readily figurative. And it is here the work connects with the theories of metaphor produced by ? and ?, who see figurative language as interactions between isomorphic conceptual structures of one sort or another.

The account of metaphor as a simple process of intensional projection, such as that offered by ?, is confounded by ?, who makes a controversial but ultimately persuasive case for the idea that so-called figurative language must in fact be interpreted literally because the pragmatically meaningful component of metaphor transpires on a level which is non-sentential. ? fleshes out this claim, drawing on the inherently *imagistic* quality of metaphor to offer an account of how the conceptual economy inherent in figurative expression involves components that cannot possibly be captured propositionally. Similarly, ? draws out a distinction between the *meaning* and the *intimation* of a metaphor, concluding that it is this second, no-propositional quality that gives figurative language

its distinct character.

At stake here is a basic notion about the contextual nature of conceptualisation, something that ? has described in terms of the *haphazard* way in which concepts are formed in response to a situation in an environment. In the same vein, it is again ? who, from a relevance theoretical perspective, has characterised the formulation of concepts as *ad hoc*, in particular analysing the way in which meaning is appropriated for some pragmatic purpose in the course of metaphor making. In response to Carston, ? have made the case that *ad hoc* concepts cannot possibly be *atomic*, preferring instead to view concept formation as a process of inference involving some sort of clustering of information on a contextual basis.

This last theory of conceptualisation as a process of informational clustering begins to look a lot like the computational model that will be described in this thesis. By pushing these clusters into an explicitly geometric representation, the hope is that a kind of non-propositional structure will emerge, satisfying apt points regarding the nature of metaphor raised by Davidson and his acolytes and, at the same time, providing a concrete basis for the sort of conceptually anatomic mapping proposed by Lakoff and Johnson. Viewed in this light, computational linguistics, with its propensity for the construction of informational structures in a high-dimensional space, becomes the natural domain for a dynamically unfolding, context sensitive model of figurative language.

2.7 Computers and Language

At last, from a practical perspective, this work must be placed in relation to the ongoing work in computational linguistics, and in particular on the type of high dimensional, corpus based, generally unsupervised language modelling that will be at the core of the project presented here. The idea that statistical word representations can be constructed in such a way as to be dynamically interactive is inherent in the work on compositionality done by, for instance, Coecke, Sadrzadeh, and Clark (2011) and ? (see Mitchell &

Lapata, 2010, for an overview). While the model described in this thesis does not target compositionality, the idea that linguistic representations can only capture conceptualisation by being somehow dynamically interactive is implicit in distributional semantic approaches to compositionality.

In terms of vector space models of distributional semantics in general, Clark (2015) offers a good contemporary overview of work in the field. The model described here, populated as it is by literal co-occurrence statistics rather than abstract values learned through iterations over a weighted network, distinguishes itself from the ongoing trend towards models with abstract dimensions, which can trace its roots back to the matrix factorisation based techniques developed by, for instance, Deerwester, Dumais, Furnas, Landauer, and Harshman (1990) and Blei, Ng, and Jordan (2003). More recently, the neural network based approaches of, for instance, Bengio, Ducharme, Vincent, and Jauvin (2003), Collobert and Weston (2008), and Mikolov, tau Yih, and Zweig (2013) have achieved impressive results. Mikolov et al in particular, along with Pennington, Socher, and Manning (2014), who present a model involving a hybrid of weighted networks and matrix factorisation, have achieved state-of-the-art results on analogy completion tasks. In the case of these models, however, the opacity of the scalars which constitute the word-spaces means there can be no hope of capturing the contextually informed theory of conceptualisation surveyed in the previous section. Instead, to the degree that context is captured at all, it is found in the multiplicity of directions of movement offered by the dimensionality of the one irreducible space. Ultimately, it seems that the analogy test sets which these systems target really just embodied a very specific type of figurative language, and subsequent systems have, by focusing on this particular objective, built this specificity into their own procedures.

Beyond analogy, ongoing work in taxonomy completion and, more generally, the inferential capacities of statistical language models is germane to the project described in this thesis. To this end, work from Cimiano, Staab, and Tane (2003) and Snow, Jurafsky, and Ng (2006) provides a good exemplar of current directions in the field. More

recently, Santus, Lenci, Lu, and im Walde (n.d.) have presented an energy based model for determining hypernymy based on lexical statistics, while Levy, Remus, Biemann, and Dagan (2015) have offered a rebuttal to some of the ongoing work in the area, suggesting that a range of supervised learning language models actually do not recognise conceptual relationships between words, but rather simply model the probability of a term being used to describe a categorical *prototype*. In related work, Socher, Chen, Manning, and Ng (2013) have proposed a neural tensor model for completing knowledge bases, again with impressive results.

Ultimately there is a certain correlation between the project described in this thesis and the comprehensive objective of research from Baroni and Lenci (2010), who describe a generalised statistical language model designed to perform well on a wide array of language processing tasks. Elsewhere, Baroni and Zamparelli's (2010) work on composing adjective-noun pairs into singular vector representations is also relevant, as the adjective, represented on its own as a matrix, might be taken as a contextualising mechanism.

Chapter 3

Theoretical Foundation

This chapter presents the theoretical framework for a novel distributional semantic model.

3.1 Concepts in Context

The language model presented here is based on the premise that conceptualisation happens in way that is fundamentally contextualised by an immediate situation in an environment. The cognitive flexibility inherent in conceptualisations is reflected in the lexical ambiguity that is prevalent in language. To the degree that it makes sense to talk about lexical semantics, the attribution of word meaning must be seen as a kind of *a posteriori* procedure, arising from the observation of how a word is generally related to a particular concept.

By the time a distributional model is generated from a large scale linguistic corpus, the blur from use to lexicalisation and then back towards figuration is already inherent in the statistical comportment of the data—so, for instance, it is not surprising to find that a word such as *state* can be statistically associated, in terms of co-occurrence, with both words like *solid*, in the conceptual sense of PHYSICAL STATE, and words like *Ohio*,

in the sense of AMERICAN STATE.

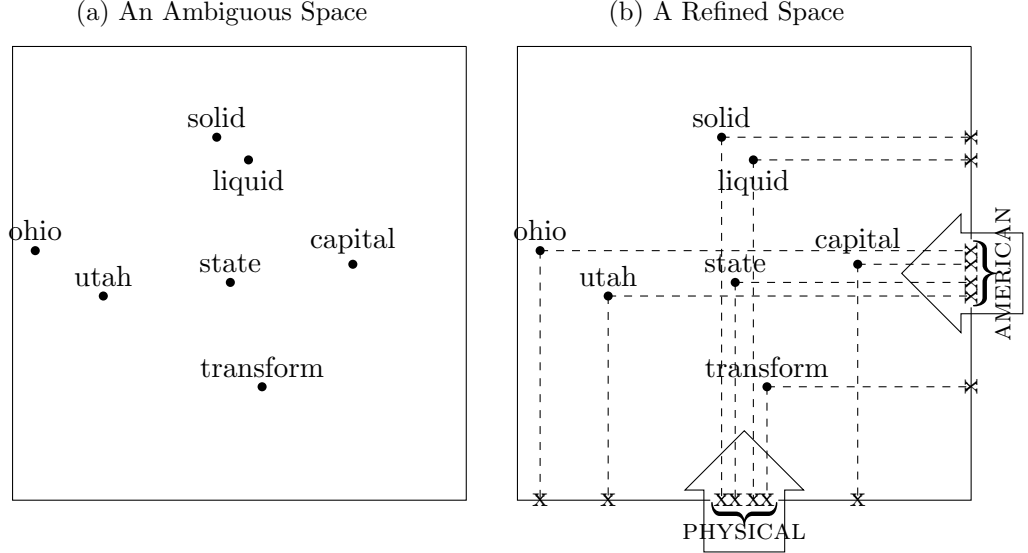


Figure 3.1: An ambiguous two-dimensional space of words relating to the term *cat* (a), and a refined space where, based on two different one-dimensional projections, two different conceptual mappings of *cat* emerge as word clusters (b).

3.2 A Literal, Dynamic Language Model

The model described here is characterised by two crucial features:

1. The dimensional features of the space are literally interpretable, in that they correspond to information theoretical statistics about the co-occurrence of words in a large scale corpus.
2. Following on this first point, the model operates by a procedure of continuously unfolding projections of subspaces, with the character of these projections determined by a continuous dimensional analysis based on input terms.

In the first instance, the maintenance of dimensions populated by statistics pulled directly from a corpus traversal differentiates this approach from recent work by the likes of Mikolov, Chen, Corrado, and Dean (2013) and Pennington et al. (2014), who have made use of neural networks and matrix factorisation to build models which perform well

on hand-crafted test sets designed to test the models’ performance on specific aspects of conceptualisation. It is precisely the statistically concrete nature of the present model which allows it to remain conceptually dynamic. Because the model’s dimensions correspond literally to co-occurrence terms, an analysis of the dimensions that contribute strongly to terms observed as occurring in a particular contextual context can be drawn out and used to project a subspace from a base model.

3.3 The Geometry of Metaphor

The contextual dynamism of concepts is implicit in the isomorphic mappings between conceptual schemes at the root of ?’s (?) conceptual metaphors. It is the internal anatomy of each conceptual scheme, produced by the environmental and consequent cultural entanglements inherent in language, that allows the alignment of conceptual schemes: there is always an implicit geometric foundation to these mappings (and this foundation is often rather more explicit, as in the case of, for example, orientation and conduit metaphors). This project seeks to realise a practical implementation of this theoretical stance, in particular by utilising the geometric qualities of the contextualised subspaces that have already been described. In particular, the al

There is once again a point of comparison with the models that have been described in the contemporary NLP literature here. The analogy completions performed by the model of, for instance, Mikolov, Chen, et al. (2013) employ a clever layer of basic linear algebra to word-vectors derived through the complex non-linear processing of corpus data. These operations are, however, dependent on a certain consistency of scale within a space. The algebraic relationship $\overrightarrow{king} - \overrightarrow{queen} = \overrightarrow{man} - \overrightarrow{woman}$, for instance, which has been discussed extensively in recent papers, depends on a comparability not only in terms of orientation but also in terms of distance between the spaces of ROYALS and PEOPLE. This is a special category of analogy within a distributional model, however, as there are demonstrably comparable conceptualisations mappable from word clusterings

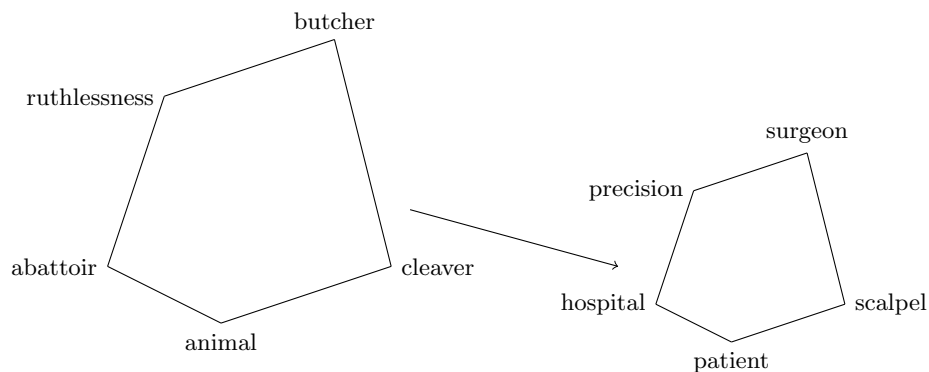


Figure 3.2: Congruences discovered in subregions of a vector space model suggest metaphoric mappings. The regions do not necessarily have to be of the same scale in order to identify a possible alignment.

of various scales.

By instead relying on the general geometric property of congruity rather than the specific scale of a space, the model described here offers a more robust approach to drawing connections between spaces. If a concept is represented as a cluster of words that emerge in a distributional semantic subspace, then the theory here is that there should be mappings between the words in two different such subspaces which, taken as a whole, substantiate a conceptual metaphor. The intuition here can be related to a more traditional transference view of metaphor: there will be some notable overlap between the co-occurrence profiles of the respective distributional spaces of, for instance, SURGERY and BUTCHERY. By identifying this overlap, it should be possible to draw out correspondences between conceptual components such as *scalpel* and *cleaver*, and this in turn allows for the postulation of an alignment between the spaces which represents a metaphor. Thus it is the geometry of the spaces which ultimately affords the opportunity for constructing a new semantic connection between conceptual domains, and the evident transference of intensions might be interpreted as an *a posteriori* logical artefact of this process.

Chapter 4

Methodology

4.1 A Lexical Base Model

The general model is trained by a single traversal of a large scale corpus. Token by token, co-occurrence frequencies are tabulated for each word type, and the dimensions of the model are populated with information theoretical statistics measuring the mutual information associated with the observation of a context word occurring within a window of a certain number of words from a vocabulary word. The result is a matrix with rows corresponding to word-vectors and columns corresponding to co-occurrence terms, such that an element $M_{w,c}$ measures the mutual information inherent in the co-occurrence of term c with word w , where N is the total count of word tokens in the corpus, $n_{w,c}$ is the frequency with which c occurs in the context window of w , n_w is the total frequency of word w , n_c is the total count of times that c is observed in the context of another word, and a is a smoothing constant:

$$M_{w,c} = \log_2 \left(\frac{n_{w,c} \times N}{n_w \times (n_c + a)} + 1 \right) \quad (4.1)$$

Once this base space is established, a dimensional analysis of input terms provides

the basis for the projection of contextually informed subspaces. So, for instance, given an input of terms w_1 , w_2 , and w_3 , the model analyses the corresponding word-vectors \vec{w}_1 , \vec{w}_2 , and \vec{w}_3 to determine the salient co-occurrence dimensions shared between all three terms, which is to say, those universally non-zero dimensions with the highest mean value. These dimensions are then used as the basis for the generation of a subspace, with the expectation that conceptually relevant terms will emerge in this new, significantly lower-dimensional space. In particular, conceptual constituents are identified by searching for word-vectors that cluster around a positive central point considerably removed from the origin, by searching for the word-vectors with the greatest norms, and by searching for the word-vectors that are, on average, furthest away in terms of Euclidean distance from any other word vectors.

4.2 Taxonomy Recapitulation

The conceptual projections of the lexical base model perform particularly well at extending the membership of short lists of words corresponding to constituents of some conceptual category (so, for instance, a list containing *lion*, *tiger*, and *bear* is expanded to include other exemplars of the concept WILD ANIMAL). With this in mind, a mechanism for establishing hypernymic relationships between such lists and terms denoting a general categorisation of such lists should lead to the effective reconstruction of hand-crafted lexical ontologies such as WordNet.

The proposed way forward at the time of writing is to establish a technique for analysing the co-occurrence data for singular terms and then to generate short lists of other words as candidates for list expansion operations. This will most likely involve inspecting potential groupings of the terms that emerge from a subspace for high degrees of mutual co-occurrence, and then expanding top scoring sets to discover potential groupings of terms to be considered as constituents of a certain conceptual subtree. There will also have to be a system for maintaining and then eventually comparing hypotheses

about the conceptual relationships between the denotations of word pairs, as well as a mechanism for ultimately resolving these hypotheses into a well formed ontology.

4.3 Analogy Completion

Where work from researchers such as ?, ?, and ? has focused on building distributional semantic spaces where simple arithmetical operations between word-vectors yield good results for certain types of analogies, the project presented here will focus on more robustly geometric methods to identify congruences between different contextualised projections. As a first pass at this problem, searching for word-vectors that complete basic geometric transformations within a space seems like a viable approach to this task. So, for instance, there is some transformation from the space of BUTCHERY to the space of SURGERY that preserves the overlapping co-occurrence relationships between words like *cleaver* and *scalpel* while aligning the

4.4 Metaphor Generation

The ultimate goal of metaphor generation will involve a kind of reverse engineering of the model’s approach to analogy completion: once the mechanism for performing analogical transformations between conceptually rooted word-subspaces has been refined, these transformations can be applied to a given projection based on the dimensional information inherent in the influence of some contextualising term. Once these transformations are applied, the relatively sparse transformed source space can be searched for the points that best align with the geometry of the target space.

Chapter 5

Implementation and Results

The base model has been trained on the English language text of Wikipedia, compiling statistics for a vocabulary of the 200,000 most frequent word types in the corpus and the terms that occur within a window of five words on either side of each token of the elements of the vocabulary. In the cases presented here, the 200 most salient dimensions for each input query are used to project the contextualised subspaces.

5.1 The Base Model and Projections

Tests performed so far on the base model have returned good results for some sample contextual projections from the base model. Two different types of input queries have been considered, one involving terms which together denote some contextually specific concept and another which offers a short list of conceptual exemplars. Results for four sample queries are offered in 5-A, with the top ten output words in terms of norm being reported.

Table 5-A: Contextualised Word Spaces

BY HYPERNYM		BY LIST	
<i>wild animals</i>	<i>professional occupations</i>	<i>lion,tiger,bear</i>	<i>surgeon,butcher,builder</i>
pocupines	technicians	leopard	bricklayer
deer	accountants	dhole	grocer
boar	technologists	hyena	apprenticed
rabbits	therapists	boar	wheelwright
foxes	electricians	langur	blacksmith
raccoons	physiotherapists	macaque	plumber
boars	dentists	tapir	shoemaker
goats	hygienists	chital	industrialist
squirrels	pharmacists	lion	joiner
hares	nurses	rhinoceros	cabinetmaker

5.2 Taxonomies

Some preliminary results, comparing the performance of the model described here versus the word2vec model of ?, are offered in 5-B. The linear algebraic technique of using vector addition and then rating similar terms based on cosine similarity has been applied in the case of word2vec, as described in the literature. Here *hits* refer to terms that can be found in the WordNet subtree of terms associated with the appropriate synset for each input query, *misses* refer to terms that are in WordNet but are not in the appropriate subtree, *absent* tallies terms that are not in WordNet at all, and *duplicate* indicates the count of terms that are morphemically identical to other words already returned by each model.

Table 5-B: Comparative Hypernymy

	<i>wild animals</i>		<i>visual artist</i>	
	MODEL	W2V	MODEL	W2V
HITS	40	34	16	8
MISSES	5	9	24	30
ABSENT	2	2	7	9
DUPLICATE	3	5	3	3

5.3 Analogies

Preliminary experiments have been carried out to explore the comparative geometry of different word-spaces. The approach thus far has been to measure the angles of the vertices of different constellations of word-vectors within a subspace, and then to examine how these angles line up with the angles in another space with an eye towards describing a prospective conceptual mapping. Experiments with the model's projections are under way, and results will be forthcoming. For the time being, 5-C offers a comparison between terms within the general, non-reduced space. Three mappings are offered; in each case, all possible alignments of words between each space are compared. The average difference in angle between each of the two spaces for each possible alignment is calculated, and the ranking of the alignment that is deemed, for the purposes of this test run, conceptually satisfactory as compared to all other alignments is returned. The high performance of the appropriate alignments even in the unrefined base space motivates further and more rigorous testing of mappings between projected spaces, which should remit a much higher degree of conceptual coherence.

Table 5-C: Angular Mappings of Conceptual Regions

target source	ROYALS PEOPLE	SURGERY BUTCHERY	EMOTIONS ORIENTATIONS
rank	1 out of 24	15 out of 720	3 out of 720
mean	0.062	0.046	0.048
high	-	0.037	0.045
low	0.198	0.094	0.101
mappings	man → king woman → queen boy → prince girl → princess	butcher → surgeon cleaver → scalpel abattoir → hospital lamb → patient pig slaught. → operation ruthlessness → precision	up → happiness down → sadness out → loneliness in → togetherness atop → control beneath → incapacity

Chapter 6

Evaluation

This section for now very broadly outlines some of the evaluative objectives of this project.

but the crucial

considered here is the necessity of resorting to some real world mechanisms for assessing the model's performance, both through formal studies with human subjects and through a less structured but more public testing of the system

6.1 Taxonomy and Analogy

One problem inherent in the use of datasets, be they test sets specifically designed to examine features of computational language models or general and highly public frameworks such as WordNet or DBpedia, is the necessarily biased and incomplete process of assigning conceptual relationships to senses of words. As such, in addition to the straightforward analysis of results over test sets and reporting of precision, recall, and related statistics for ontology recapitulation, it will probably be necessary to resort to asking human subjects about the efficacy of the model's conceptual mappings.

6.2 Metaphor

Even more than with the only moderately controversial tasks of determining whether conceptual and analogical relationships are appropriately captured by the models output, the problem of assessing the creative value of metaphoric artefacts generated by the model is fraught with the

In the end, the criteria of usefulness and novelty generally taken as the basic standard for computationally creative output can serve as a good guide for the model's target, but not as much more than this. Simply asking humans whether they consider the model to be performing along these lines must be a hopelessly subjective enterprise

Chapter 7

Conclusion

References

- Baroni, M., & Lenci, A. (2010). Distributional memory: A general framework for distributional semantics. *Computational Linguistics*, 36(4).
- Baroni, M., & Zamparelli, R. (2010). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of the 2010 conference on empirical methods in natural language processing* (p. 1183-1193).
- Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (2003). A neural probabilistic language model. *Journal of Machine Learning Research*, 3, 1137-1155.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Boden, M. A. (1990). *The creative mind: Myths and mechanisms*. London: Weidenfeld and Nicolson.
- Calvino, I. (1983/1985). *Mr. palomar*. New York: Harcourt Brace Jovanovich. (Translated by William Weaver)
- Chalmers, D. J. (1996). *The conscious mind*. Oxford University Press.
- Chomsky, N. (1959). A review of b. f. skinner's verbal behavior. *Language*, 35(1), 26-58.
- Cimiano, P., Staab, S., & Tane, J. (2003). Automatic acquisition of taxonomies from text: Fca meets nlp. In *Proceedings of the ecml/pkdd workshop on adaptive text extraction and mining* (p. 10-17).
- Clark, S. (2015). Vector space models of lexical meaning. In S. Lappin & C. Fox (Eds.), *The handbook of contemporary semantic theory*. Wiley-Blackwell.
- Coecke, B., Sadrzadeh, M., & Clark, S. (2011). Mathematical foundations for a compo-

- sitional distributed model of meaning. *Linguistic Analysis*, 36(1-4), 345–384.
- Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on machine learning*.
- Colton, S., & Wiggins, G. (2012). Computational creativity: The final frontier? In *Proceedings of the 21st european conference on artificial intelligence*.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal for the American Society for Information Science*, 41(6), 391-407.
- Dennett, D. (1996). *Kind of minds*. London: Weidenfeld & Nicolson.
- Fodor, J. A. (n.d.). *The language of thought*. Cambridge, MA: Harvard University Press.
- Fodor, J. A. (2008). *Lot 2: The language of thought revised*. Oxford University Press.
- Harris, Z. (1954). Distributional structure. *Word*, 10(23), 146–162.
- Jordanous, A. K. (2012). *Evaluating computational creativity: A standardised procedure for evaluating creative systems and its application* (Unpublished doctoral dissertation). University of Sussex.
- Levy, O., Remus, S., Biemann, C., & Dagan, I. (2015). Do supervised distributional methods really learn lexical inference relations? In *Proceedings of the 2015 conference of the north american chapter of the association for computational linguistics*.
- McGregor, S., Wiggins, G., & Purver, M. (2014). Computational creativity: A philosophical approach, and an approach to philosophy. In *Proceedings of the fifth international conference on computational creativity*.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. In *Proceedings of iclr workshop*.
- Mikolov, T., tau Yih, W., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies* (p. 246-251).
- Mitchell, J., & Lapata, M. (2010). Compostion in distributional models of semantics.

- Cognitive Science*, 34(8), 1388-1439.
- Pantel, P. (2005). Inducing ontological co-occurrence vectors. In *Proceedings of the 43rd annual meeting on association for computational linguistics* (pp. 125–132). Stroudsburg, PA, USA: Association for Computational Linguistics.
- Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Conference on empirical methods in natural language processing*.
- Pietarinen, A.-V. (2004). Grice in the wake of peirce. *Pragmatics & Cognition*, 12(2), 295–315.
- Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1), 67-99.
- Russell, B. (1905). On denoting. *Mind*, 14(56), 479-493.
- Russell, B. (1931). *The scientific outlook*. London: George Allen and Unwin.
- Santus, E., Lenci, A., Lu, Q., & im Walde, S. S. (n.d.). Chasing hypernyms in vector spaces with entropy. In *Proceedings of the 14th conference of the european chapter of the association for computational linguistics*.
- Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley Publishing Group.
- Snow, R., Jurafsky, D., & Ng, A. Y. (2006). Semantic taxonomy induction from heterogeneous evidence. In *Proceedings of the 21st international conference on computational linguistics and the 44th annual meeting of the association for computational linguistics*.
- Socher, R., Chen, D., Manning, C. D., & Ng, A. Y. (2013). Reasoning with neural tensor networks for knowledge base completion. *Advances in Neural Information Processing Systems*, 26.
- Veale, T., & Hao, Y. (2007). Comprehending and generating apt metaphors: A web-driven, case-based approach to figurative language. *AAAI*, 1471-1476.
- White, G. (2009). On scholarship. In T. Bartscherer & R. Coover (Eds.), *Switching codes: Thinking through digital technology in the humanities and the arts*. University of Chicago Press.
- Wiggins, G. A. (2006). Searching for computational creativity. *New Generation Com-*

puting, 24, 209-222.

Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. London: Kegan Paul, Trench, Trubner & Co, Ltd.

Wittgenstein, L. (1953/1967). *Philosophical investigations* (3rd ed.). Oxford: Basil Blackwell. (trans. G. E. M. Anscombe)