WORDS IN CRISIS

A relational perspective of emergent meanings and roles in text

Adina Nerghes

WORDS IN CRISIS:
A RELATIONAL PERSPECTIVE OF EMERGENT MEANINGS AND ROLES IN TEXT

ADINA NERGHES

Reading Committee:

prof. dr. Marcel Veenswijk (chair)	VU Amsterdam
prof. dr. M. N. C. (Noelle) Aarts	Wageningen University/University of Amsterdam
prof. dr. L. A. (Loet) Leydesdorff	University of Amsterdam
dr. Anita M. J. van Hoof	VU Amsterdam
dr. Wouter de Nooy	University of Amsterdam

This research has been funded by the Netherlands Organization for Scientific Research (NWO) under the NWO-Aspasia 015-007-047 grant.

Cover design and layout:	Adina Nerghes
Printed by:	CPI – Koninklijke Wöhrmann B.V.
ISBN:	978-94-6203-660-4

VRIJE UNIVERSITEIT

Words in Crisis:
A relational perspective of emergent meanings and roles in text

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Sociale Wetenschappen
op dinsdag 29 maart 2016 om 9.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Adina Nerghes
geboren te Boekarest, Roemenië

promotor:	prof. dr. P. Groenewegen
copromotoren:	dr. I. R. Hellsten
	dr. Y. T. A. Taminiau

Contents

"Whatever you can do, or dream you can, begin it. Boldness has Genius, Power and Magic in it. Begin it now." - Goethe

Dedicated to my husband, Ju-Sung

Acknowledgments

The pages of this dissertation mark the end of a four-year journey full of rich and inspiring experiences. I owe my sincere gratitude to many people who accompanied me through this process and contributed greatly to my professional and personal growth. Directly or indirectly, professionally or socially, a great many people have contributed to my PhD journey.

I owe my foremost and deepest gratitude to my supervisors Prof. dr. Peter Groenewegen, dr. Iina Hellsten, and dr. Yvette Taminiau who lead me through this experience with patience and knowledge. Our many inspiring discussions, productive co-authoring, prompt email responses late at night, and your many comments on my writings have greatly contributed to the completion of this dissertation, but more importantly, to my development as a scientist. Thank you.

I am grateful to all my dissertation committee members for the time they invested in reading my work! The work of the reading committee is paramount to graduating with success and I would like to thank each and every one of them for their vote of confidence in my work.

The Organizations Sciences department at the VU has been my home for the past four years and I would like to acknowledge all those that I am grateful to for creating a relaxed yet inspiring working atmosphere. First, a special thank you to Elles and Welmoed for their vital support over the years. Many thanks to Sandra, Marieke, Bas, David, Julie F, Christine M, Nicoletta, Jeroen, Greetje, Annemiek, Ida, Sierk, Kees, Irma, Joanna, Priyanka, Christine T, Peter, Dirk and all those that joined and left the ORG department during the past four years.

I would also like to gratefully acknowledge the financial support of the Netherlands Organization for Scientific Research (NWO) under the NWO-Aspasia 015-007-047 grant.

There is very little in this life that can compare to the company of those who can easily share in one's pain and pleasure. Special thanks go to my colleagues

and friends Amanda, Michiel, Julie B, and Shan. A chat by the coffee machine, a 10-minute outdoor break (Michiel, you know what I'm talking about ☺), or a drink after work have made the process of completing this dissertation a lot more enjoyable. I am also grateful to my dear friends Mariana and Hanneke. In your own special ways, you have supplied me with just the intellectual and emotional support I needed to keep believing in myself and to always move forward. Hanneke, among many other things, you have given me a first taste of what it means to be a researcher. Mariana, I will always cherish the many after-work evenings of laughter and beer in my small student room. Thank you for always being there and for never failing to make me laugh. There are some people in life that make you laugh a little louder, smile a little bigger, and live just a little bit better. Elena, Luminita, Lidia, Raluca, and Leea thank you for providing the occasional and much needed distractions.

No statement of gratitude is ever complete without an acknowledgment of those people who had first rights in shaping one's thoughts and character. Without my parents's unwavering confidence in my choices, this journey would have been far more difficult; and I share the joy and satisfaction of completing this work with them. Mom, I know you wished I would follow in your steps but I am thankful you encouraged me to find my own way. In the end, I will still be able to write dr. in front of my name! Dad, the endless hours you spent trying to get through to my non-math oriented mind are starting to pay off and Juice is continuing the tutelage you started. Thank you both for your support and trust in me.

Last, but not least, my beloved husband, Juice. I'd like to think we would have met somehow, somewhere ... but until we can explore alternate realities, it was my PhD journey that brought me to you. Having you by my side at the end of these four years makes this moment so much more meaningful and I have so many things to thank you for: thank you for leaving your Pittsburgh life behind and joining me here, where *"the whole place is pickled ... "* (you know the rest ☺); thank you for always knowing the right time to challenge me, to comfort me, or to push me outside my comfort zone; thank you for never failing to cheer me up, for being so amazingly patient, and most of all for your

unwavering love. You are one of the smartest people I know and having the opportunity to discuss research ideas with you, to collaborate on papers, or to debate philosophical and ethical issues enriches my life tremendously. But first and foremost, you are my partner and my best friend, and whatever journey I embark on I know you will always be there, supporting me every step of the way. I love you and I am looking forward to our next adventure ☺ !

1 Introduction

"...there is no self-understanding that is not mediated by signs, symbols, and texts;" (Ricoeur, 1991)

Can we infer rich information from 'big text data'? And how can we use text-analytical methods to infer such rich information from large text collections with different characteristics? These are some of the questions that guide the aims and outcomes of the research presented in this dissertation.

What motivates these questions is the 'linguistic turn' in the social sciences and humanities. In the past decades, the increasing availability of textual information opened new venues for large-scale research in the social sciences, fostering an increase in attention given to text. The ease with which large volumes of data capturing social communication can be stored, accessed, and collected has risen to match the ambitions of social scientists in understanding behavior, structures, values or norms (Lazer et al., 2009). In particular, the numerous text documents generated daily by social actors across the world are a great wealth of knowledge and a rich source of social information. As Benoit (2011) argues, texts remain one of the most promising, but at the same time, one of the least explored sources of systematic information about our social world.

Much of the data available for social science research exists in the form of text, or can acquire a textual form. This is not to say that all social communications can be reduced to a textual form, as communication in the form of body language and gestures cannot be reduced to a textual form without difficulty. But rather, in the context of this dissertation, texts or textual data refer to "systematically collected material consisting of written, printed, or electronically published words, typically either purposefully written or transcribed from speech. "(Benoit, 2011, p. 526) A few examples of what represents textual data

are organizational press releases, reports issued by various organizations and actors, news items, memos, journal articles, open responses to questionnaires, etc., but also transcripts of public speeches, conversations, interviews, etc. Such texts, whether published online or digitized, can be stored over long periods of time at very low costs in digital archives. This is to say that the information issued by social actors (in its textual form) creates a foot print or trail that enables researchers to analyze historical episodes as well as contemporary ones in a dynamic manner.

While the availability of texts easily amenable to systematic analysis represents one of the most important advantages associated with the increased interest in textual data, texts also contain significant information about the orientations and beliefs of the actors generating them, which cannot be inferred from nonverbal forms of behavior.[1] As such, texts may contain valuable information about their authors or speakers and the ways in which meanings are attributed. Ultimately, text documents are 'vehicles for meaning' and manifestations of discourse (Chalaby, 1996). This is not to say that texts are synonymous to discourse, but rather that texts are partial manifestations of a discourse. In other words, the texts issued by a social actor (individuals and/or organization), in their entirety, construct the discourse of that social actor. As it will become clear in the following paragraphs, these texts, produced in social interaction, embed positions and interpretations of the social actors generating them, which in turn have social effects that shape and construct our social reality.

Texts, both in written or spoken form, are fundamentally interactive (Halliday, 1978). As Nystrand and Himley (1984) explain, "when readers understand a text, an exchange of meaning has taken place. Writers have succeeded in speaking to readers." (p. 198) The interactive trait of text stems from the fundamental characteristic of language production as a social practice. To produce language is to engage in discourse, and in doing so the writers or speakers

[1]For example, in a parliament setting, texts and speeches can serve as a more genuine account of the true political preferences of an actor than voting behavior. As highly strategic political acts, voting patterns are subject to strict party discipline in most contexts, and hence unreliable indicators of an actor's preference on the classical left-right policy scales (Benoit, 2011).

have to adapt their text to the social context and the audience (readers) they address. For example, writing a scientific text will involve the use of different linguistic choices and strategies than when writing a newspaper article. In this sense, the social context in which texts are produced and consumed defines *what* topics and issues may be discussed, and to some extent, *how* these topics and issues are discussed. For example, a text produced in the session of the parliament will (most likely) not discuss issues regarding the success of the latest Hollywood blockbuster, and to some degree, the linguistic choices in such a text will differ from those of a text produced by the reactions of Facebook friends to someone's vacation photographs. Thus, texts are socially constructed and express socially shared meanings in specific contexts.

But text documents are not just collections of words and ideas, chosen and organized according to constraints imposed by social contexts or audiences. The words used in texts are not rigid units of factual information transmission, they encode meanings, opinions, interpretations, positions, sentiment, etc. Because texts are produced by social actors, they will inherently contain important information about the positions of these actors on the topics and issues they write about. Employing pre-existing linguistic resources, the actors' process of generating text (written or verbal) requires a selection of linguistic practices (e.g., specific metaphors, adjectives, verbs etc.) from multiple possibilities. Even when writing or talking about the simplest of events or phenomenons, many different descriptions can be provided (Lynch & Woolgar, 1988), and thus what is selected and included in a text depends on the orientation, the interests, and intentions of a speaker or writer (Potter et al., 1990). Words and sentences "will change their meaning according to their use and the positions held by those who use them."(Ball, 2012, p.17) By using and responding to "words and sentences as semantically significant" (Rouse, 2007, p. 535) in spoken or written form, social actors engage in discursive practices through which issues are defined, interpretations are given, and positions are taken. Texts, then, present alternative descriptions and categorizations of events and issues, and create different 'versions' of such issues and events through the different discursive constructions and linguistic choices they embed.

However, while containing the discursive practices and linguistic choices of social actors, texts "do not just describe things; they do things."(Potter & Wetherell, 1987, p.6) This constructivist perspective reminds us that much of what we know about the world comes from discursive constructions and versions (Potter et al., 1990). Our knowledge of current and past world events, advances in science, or even elements of culture are gained, formed, and passed on through written and verbal text. In this sense then, texts and the discursive manifestations they contain can shape and construct our social reality (Condor & Antaki, 1997). Giving meanings to particular experiences and practices, texts can alter perceptions and world views by influencing the ways in which we make sense of the world (Fairclough, 1992; van Dijk, 1997a). Thus, the ways in which we speak or write are produced in social interaction and consequently have social effects (Fairclough, 2001).

The linguistic choices and discursive constructions of texts offer alternative descriptions and categorizations, but they also (more often than not) contain evaluations of events and issues that may shape perspectives. Take for instance the example of the recent influx of Syrian people into Eastern and Western European countries. Texts describing this particular event, in the form of media reports or even statements of various state officials, employ different linguistic choices describing and, at the same time, evaluating this event. Texts referring to this event as a 'refugee crisis' vs. a 'migrant crisis' offer a description of this event but also an evaluation. The use of 'migrant' or 'refugee', as a label for the Syrian people coming in Europe, involves an evaluation on whether these people made a conscious choice to leave their country due to economic hardship or whether they are forced to leave their country because they are at risk of persecution. Such evaluations, also known as frames, may shape people's interpretation of that story by making certain perspectives more salient (Hallahan, 1999; Iyengar, 1987; Pan & Kosicki, 1993).

The above example demonstrates that through various (intentional or unintentional) linguistic strategies employed by social actors when producing texts, issues are framed, dramatized, emphasized, packaged, etc., which may influence perceptions and attitudes towards the issues discussed. Therefore, words

and the texts containing them may become tools of influence, deception, and manipulation in the 'hands' of the writer or speaker, who may try to change the perceptions of their audiences or to impose a dominant narrative.

Such discursive strategies, are not part of the manifest content of texts, which is transmitted through explicit vocabularies, but they are part of the latent content, which denotes the implicit meaning in text or its use in social interaction (Merton, 1957; Phillips et al., 2008). Whereas the manifest content of text is reflected by the frequency of certain words or the number of words in a sentence, the latent content of texts is reflected by more subtle manifestations. Investigating the latent content of texts requires both adequate methods of analysis and interpretative frameworks that take into account the socio-cultural context in which the texts were produced (Pollach, 2012). Hence, given the ample availability of textual data and the valuable social information it contains, the question is no longer whether or not their content is worth investigating, but rather which approach is the most insightful for a given research goal.

1.1 The Analysis of Texts in the Social Sciences

Text analysis as a research strategy permeates many of the fields in the social sciences, and the range of methods employed in the analysis of texts is extensive (Bernard & Ryan, 1998). Since the end of the 1990s, the study of texts and discourse became increasingly popular in anthropology, linguistics, literary studies, sociology, cognitive and social psychology, communication studies, and political science (van Dijk, 2011). However, the popularity of text analysis is not confined to the social sciences. In broad terms, the analysis of textual data has been pursued in four academic fields: the social sciences, computer science, the (digital) humanities, and linguistics. Social scientists explore text aiming to reveal mechanisms according to which words influence and are influenced by human behavior (Roberts, 2000), computer scientists focus on quickly 'understanding' user input and the user-expected outcome (Grishman, 1986; McEnery, 1992), while the interests of the humanities have been rooted in text-specific criticism of texts (Chambers, 2001). However, by analyzing textual data, social scientists, computer scientists, and humanists alike are essentially entering the linguistic terrain (Pollach, 2012), which focuses on describing the structure of text. Although all these academic fields examine words, sentences, paragraphs, pages, documents, ideas, meanings, and even what is missing from the text, the analysis methods in linguistics, computer science, humanities, and the social sciences have evolved in parallel. Scientists from all four fields rarely refer to the work of the other fields (Pollach, 2012; Popping, 2000). This fragmentation stems primarily from the differing perspectives on text, the methodologies that mix and match linguistic and interpretative methods, and the varying degrees of methodological rigor (Johnston, 2002). In linguistics, texts are treated as objects of analysis themselves, in computer science texts are a set of strings of symbols that may be constrained by rules that are specific to it, while in the sociological and humanities traditions, texts are a window into human experience and culture (Bernard & Ryan, 1998; Sculley & Pasanek,

2008).

In the social sciences, specifically, text analysis does not aim to provide a description of the linguistic or structural features of texts, but rather it focuses on describing, classifying, interpreting or making inferences about social norms, values, behavior or structures based on a corpus of 'real' data (Bernard & Ryan, 1998).

However, text analysis in the social sciences is far from being a theoretically and methodologically coherent field of study. In fact, it covers a number of approaches that are informed by a wide variety of disciplines (Hardy et al., 2000). A few examples of the multitude of approaches to text and discourse analysis are: critical discourse analysis (Fairclough, 2012), systemic functional linguistics (Schleppegrell, 2012), rhetorical analysis (Leach, 2003), narrative analysis or narrativity (Thornborrow, 2012), mediated discourse analysis (Scollon & de Saint-Georges, 2012), dialogical and conversation analysis (Clayman & Gill, 2012), story-grammar analysis (Mandler, 1982), content analysis (Krippendorff, 2004; Roberts, 1989), and discourse–oriented ethnography (Smart, 2012). Despite a common interest in the study of texts and their use in social contexts, these approaches diverge on two of the most basic issues: the question of *what a text is*, and the question of what counts as the *social context* in which that text is used or created (Bhatia et al., 2008). The manner in which these questions are addressed by researchers depends on the discipline or disciplines they draw on. For instance, sociology and anthropology have encouraged analysts to view the use of language as a function of the context in which language is used, whereas linguistics has encouraged analysts to focus primarily on the text, with context relegated to the background.

Informed by different perspectives on the role of text and context, and being operationalized through different methods, what these approaches have in common is a focus on text to provide "knowledge and understanding of the phenomenon under study" (Downe-Wamboldt, 1992, p.314). The analysis of text in the social sciences is ultimately "concerned with the search for patterns within language in use" (Wetherell et al., 2001, p. 10). Researchers interpret text, they try to make sense of relevant parts through various methods, and

infer results that support their conclusions (Krippendorff, 2004). Furthermore, text analysis focuses on the characteristics of language as communication with attention to the content or contextual meaning of the text (Budd et al., 1967; Lindkvist, 1981; McTavish & Pirro, 1990; Tesch, 1990).

In sum, increased interest in textual data has encouraged fast-paced development of perspectives and methodologies, within and beyond the social sciences. However, driven by different research goals and operating under different understandings of what texts represent (and what the role and importance of the context is), social scientists, humanities scientists, computer scientists, and linguists have been working in parallel. This disjoint set of research streams has resulted into fragmented theoretical and methodological contributions. Even within the different disciplines of the social sciences, the multitude of approaches to texts and discourse analysis create a 'fuzzy' field of study (van Dijk, 1997b).

1.1.1 Qualitative and Quantitative Text Analysis

One clear distinction among approaches to text analysis in the social sciences is based on whether qualitative or quantitative analytical methods are employed. The qualitative approach to text analysis involves the identification of sections of text that are considered relevant, and often the selection of quotes considered representative examples of what the analyst is investigating. Using various tools (e.g., NVivo, Atlas-ti[2]), qualitative text researchers highlight sections of text, assign different codes to these sections, and cut, paste, sort, list, and enumerate the highlighted portions in terms of user-assigned categories. This approach generally involves human readers examining documents, manually coding content, and performing largely qualitative content analysis (Berg, 2007). While this approach offers perhaps the best in-depth inference of information, it also limits the amount of data that can be processed and, in most cases, it limits the extent of analytical precision due to coding/coder's bias. For example, in an experiment of reliability and misclassification of hand-

[2]NVivo and Atlas-ti are not exclusively qualitative text analysis tools, they also offer features for quantitative text analysis.

coded political texts, Mikhaylov et al. (2012) show that reliability and coder bias remain serious and systemic problems of the coding process, even when working with highly trained and experienced coders.

Generally, human coders are asked to read texts for "meaning" and categorize these texts according to some pre-developed scheme (Mikhaylov et al., 2012). Reliability problems with such coding processes, then, emerges because different human readers may attach different meaning to the same text. A reader's ability to grasp the meaning of text is linked to prior knowledge that plays an important role in the construction of meaning inferred from text (Franzosi, 1998). According to hermeneutics scholars, readers approach a text with prior assumptions and search for evidence confirming those assumptions in the text (Diesing, 1991). As such, reading texts and inferring meaning involves interpretation, which is prone to individual variation and is constrained by previous knowledge and assumptions of the text. These characteristics of human coders may result in unreliable classifications, serious validity, and reliability problems (Spooren & Degand, 2010; McTavish & Pirro, 1990), which become even more pronounced when the number of coding categories and texts increase (Mikhaylov et al., 2012). Due to these limitations and the rapid increase in the amount of accessible textual data, the quantitative approach has been gaining ground in recent years (Popping, 2000).

The increase in the popularity of quantitative methods to text analysis has mainly been fueled by technological developments that allowed researchers to exploit the large collections of available textual data in a systematic manner. However, quantitative text analysis has a long tradition in the works of George (1959), Holsti (1969), Krippendorff (2004), Lasswell (1948), Osgood (1959), Pool (1959), Stone et al. (1966), Weber (1985), and many others. In the initial stages, quantitative text analysis was popularized by a series of large-scale projects launched by Allied governments during World War II to analyze the content of Nazi propaganda (Krippendorff, 1980). These projects were conducted with the help of large numbers of coders, manually counting word frequencies and occurrences. The introduction of mainframe computers in the 1960s and the development of the General Inquirer, the first computerized text

analysis program (Stone et al., 1966), revolutionized the field of quantitative text analysis.

Due to the rapid evolution of computer technology, computer-aided text analysis has been making large strides in the past decade, facilitating collection, storage, and analysis of larger and larger text collections. Increased computational capacities and evolving machine learning techniques have also lead to greater precision, refinement, and robustness of results. However, most quantitative computer-aided text analysis methods focus on concepts, their frequencies and distributions of word usage across documents and/or over time. While providing useful summary statistics of large text collections, such approaches neglect meaning constructed through associations of multiple words and expressions. In this sense, frequency counts of words contained by text documents will reveal *what* topics are being talked about, but not *how* topics are being talked about. Consider a simple example of two text documents addressing the topic of the recent global financial crisis, both having 'bank' and 'assets' as the most frequent words. However, in one of the texts, 'bank' and 'assets' are consistently presented in conjunction with various adjectives such as 'troubled', 'toxic', or 'bad'. If such adjectives appear under many variations, and thus have low frequencies, they will most likely be regarded as unimportant. Low frequency words associate to high frequency words may build metaphorical expressions (or frames) or add valence (i.e., positive or negative sentiments), and thus become integral in the texts' subtle meanings. Whereas the example presented is trivial, the inability of frequency based content analysis methods to capture subtle linguistic manifestations becomes even more pronounced when more complex linguistic strategies are present in the text analyzed (e.g., when relatively common words are used both in a literal sense as well as a metaphorical one). Furthermore, focusing on frequency counts, which obscures meaning constructed through associations of multiple words, may lead to an overestimation of similarity between the texts using the same words in different ways (Carley, 1990). In other words, revealing meaning and comparing text documents (or even separate corpora) requires more than frequency counts or frequency distributions because meaning does not reside in individual words.

Rather, meaning is derived from the multiple relations of words with other words.

1.1.2 The Relational Perspective

Originating from the works of Cassirer (1944), Geertz (1973), and Eco (1979), the relational perspective on language and meaning posits that meaning is a relational phenomenon. Thus, it is not the individual words that generate meaning, but rather meaning is created through interrelated sets of words and concepts. In this perspective, a single word is considered to be bereft of meaning unless it is connected to other words (Carley, 1986b). Words and concepts derive their meanings from their multiple relations and integration into meaning networks. Hence, words are mere symbols whose meaning is dependent on their use; that is, their relations to other words and the similarities and differences between these interconnected words (Carley, 1986a,b, 1988; Gollob, 1968; Heise, 1969, 1970; Minsky, 1975). Consequently, all meaningful information in texts latently exists as a network of words and the relationships among those words (Carley, 1984; Eco, 1979; de Saussure, 1959).

This relational approach to text and meanings is also supported by the work of cognitive psychologists and psycholinguists, specifically by theories on semantic memory and organization of semantic information (e.g. Chang, 1986; Collins & Quillian, 1969; D'Arcais & Schreuder, 1987). Semantic memory, initially conceptualized by Tulving (1972), refers to our general knowledge of concepts and facts, and it is distinct from episodic memory which refers to our specific memory for personal experiences. Studies of semantic memory provide evidence that at a cognitive level, semantic information (i.e., language) is organized in the form of a complex network of words and concepts connected trough relationships of association (Chang, 1986). For instance, hearing or reading the word "dog" will trigger associations to concepts such as "animal", "barks", "likes bones" etc. Semantic memory, as Tulving (1972) argues, "is the memory necessary for the use of language." (p. 386) Thus, at a cognitive level information about the meanings of words is stored in the form of a network, and these meanings depend on their relations to other words.

These theories, on the relational aspect of language and meaning, provide evidence that language can be suitably represented as networks of interrelated concepts. More importantly, they also support an evolving argument that relations among words reflect cognitions and, in turn, trigger responses (Rice & Danowski, 1993). But, before a more detailed account of approaches to the analysis of relational meanings in texts is given in Chapter 2, I will first formulate the aims of this dissertation.

1.2 Aims

The main goal of this dissertation is to expand the methods of relational meaning analysis, by providing an approach suited for the detection of subtle discursive dynamic shifts in large collections of (temporal) textual data. Exploiting automated and semi-automated analytical tools of relational meaning extraction, the approach developed here reveals subtle dynamic shifts in discourse by recognizing the different roles of connective and popular words and concepts in text. This dissertation does not only demonstrate the utility of this approach in revealing subtle discursive manifestations, but it also provides evidence for the effectiveness of this approach when employed in the analysis of large text collections with different characteristics. To this end, the methods developed and presented here, are applied to text documents issued by central banks, media, and political actors. A comprehensive description of the various characteristics of these different actors is provided in Chapter 2. Additionally, the overarching context in which these three types of documents have been produced is that of the global financial crisis of 2008. The global and systematic characteristics of this recent event, presented at the end of Chapter 2, provide ample opportunities to investigate discourses in a dynamic environment, in which subtle, yet rapid discursive changes and adaptations occur. The selection of this particular event was made under the assumption that during such events as the global financial crisis, the discourses of social actors develop specific patterns that signal not only their actions but also changes in their positions and dispositions towards these events. Otherwise stated, destabilizing events that challenge or

threaten the status quo of these actors can generate rapid changes in their discursive practices, which signal subtle manifestations of position taking and which, in turn, may affect internal and external perceptions.

Although the primary aims of this dissertation are methodological, the empirical studies presented in Chapters 3 to 5 also contribute to our knowledge on how the events of the recent financial crisis were reflected in various social discourses and how the crisis drove discursive changes and adaptations. The results of my research offer a contribution to our understanding of the different aspects of the crisis that various social actors focus on. Also, by analyzing the individual discourses of three different social actors, I expose subtle yet imperative shifts in these discourses across the different stages of the crisis. These substantial transitions and adaptations have been uncovered at the level of discourses with distinctive characteristics, addressing different audiences, and fulfilling different communicative purposes.

1.3 Structure

This introductory chapter has, so far, outlined the surge of attention given to text analysis in the social sciences, triggered by the increasing availability of textual data and advancements in computer technology. This chapter has shortly introduced the perspective I adopt in regards to meanings in text documents and it has also formulated the aims of this dissertation. In the following chapter, I present the relational text-analytic methods on which I expand, I give an account of the global financial crisis as the overarching contextual event under which the texts I analyze are produced, and I present a description of characteristics for the different actors generating the data sets collected and analyzed. Lastly, I conclude Chapter 2 with a comprehensive description of each of these data sets. The remainder of this dissertation is composed of three chapters presenting empirical studies (Chapters 3 to 5) and a concluding chapter (6) in which I reflect on the the utility and limitations of the methods and approaches employed and developed, and on the practical and scientific implications of the empirical findings.

2 Methods, Data, and Context

This chapter lays the methodological and contextual groundwork applicable to the rest of the dissertation. As the methodologies focus on a structural perspective of discourse, this section first introduces two analytical methods for capturing the relational aspect of words and meanings, namely semantic network analysis and topic modeling. Next, for each of these methods, I discuss the structural space approach, which expands the two text analytic methods and allows for in-depth analysis of text and highlights orthogonal dimensions of text. This chapter concludes with a description of the global financial crisis, the overarching context binding the three empirical chapters, and in addition, the distinct discourses related to the crisis that are explored using the above methods.

2.1 Relational Approaches to Text Analysis

As postulated in the previous chapter, in this dissertation I approach meanings and texts from a distinct perspective, namely the symbol-constructivist approach of relational meaning. Supported by (postmodern) linguists and philosophers (Cassirer, 1944; Eco, 1979; Fauconnier, 1994; Geertz, 1973; Polanyi, 1962; de Saussure, 1959), as well as by theories on semantic memory and organization of semantic information (Bobrow & Collins, 1975; Chang, 1986; Collins & Quillian, 1969; D'Arcais & Schreuder, 1987; Mandler, 1984; Tulving, 1972), this perspective argues that meanings are relational. This is to say that both at the cognitive level, as well as in texts, words and concepts gain meaning through their integration into networks of other words. Thus, meaning of words emerge from the relations, similarities, and differences with other words. Hence, text documents can be understood as a structure that can be formalized as a system of interrelated words. Consequently, analyzing text documents guided by this relational perspective on meaning entails moving beyond content analytic

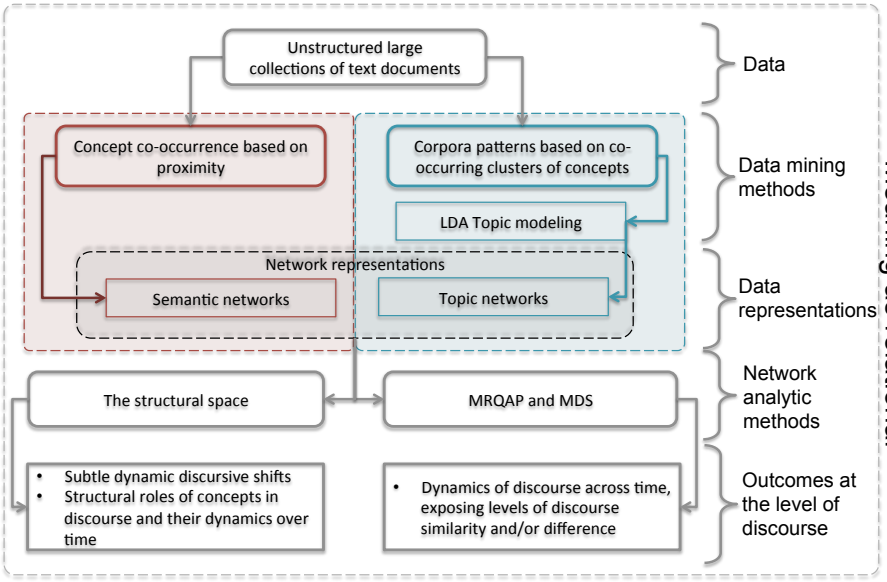


Figure 2.1: Methodology overview

methods, towards the extraction of not only words and concepts, but also the complex relations that connect them.

Following this perspective on language, radicalized by postmodern linguists and philosophers, the two text-analytic methods I employ and expand upon in the three empirical studies (Chapters 3 to 5) focus on the extraction of information about relationships between words from texts, rather than on the words themselves. As presented in Figure 2.1, the core methods employed are semantic network analysis and latent Dirichlet allocation (LDA) topic models. Designed to capture the relational aspect of language, these two analytical methods bridge the gap between qualitative and quantitative text analysis. Firstly, taking advantage of recent technological and methodological advances, these two methods are useful in analyzing vast collections of unstructured text documents. The various automated and semi-automated text processing tools developed in the past decades, are able to extract useful summary statistics (e.g., word frequencies, frequency distributions, etc.) and capture the relational character of language and meaning, without limiting the size of text collections that can be analyzed. Tools like FullText.exe (Leydesdorff, 1995,

2013), AutoMap (in conjunction with ORA) (Diesner, 2012a; Carley et al., 2013a), and ConText (Diesner et al., 2013), to name a few, operate beyond the traditional quantitative content analyses approaches by enabling automated and semi-automated extraction of relations[1]. Allowing for such an inclusive analysis of texts, these methods provide a richer understanding of textual data and the latent meanings embedded in such data.

Secondly, these tools and methods allow the analysts to stay close to the text throughout the processing steps, to make coding decisions based on knowledge of the corpora, and to explicate the results in an interpretative, qualitative manner. Although tools and methods of relational meaning extraction are highly beneficial in extracting relations among words, interpreting the resulting complex networks of interrelated words warrants familiarity with the social context in which the texts were generated and an understanding of the distinct characteristics of the social actors producing them. As postulated earlier in this dissertations, meanings are constructed in social interaction and they are constrained by specific social contexts (Chalaby, 1996; Fairclough, 2001; Pollach, 2012). These characteristics of meanings, together with the inherent versatility of language, allow writers and speakers to produce meanings specific to their goals and their social contexts. As such, these (semi-)automated methods do not eliminate the need for careful thought by researchers nor do they remove the necessity of in-depth knowledge of the context in which the texts were produced (Grimmer & Stewart, 2013). Moreover, as it will become evident throughout the remainder of this dissertations, both semantic network analysis and topic modeling facilitate integration of in-depth knowledge of text and context in the analysis of data and interpretation of results.

[1]For more details on these tools see: FullText.exe: `http://www.leydesdorff.net/software/fulltext/`; AutoMap: `http://www.casos.cs.cmu.edu/projects/automap/`; ORA: `http://www.casos.cs.cmu.edu/projects/ora/`; ConText: `http://context.lis.illinois.edu/download.php`

2.2 Semantic Networks

As one of the areas of research that has gained popularity in recent years, semantic network analysis is becoming its own research paradigm as well as a method of analysis. Referred to as maps (Carley, 1997b), semantic networks (Monge & Contractor, 2001; Popping, 2003; Lehmann, 1992), networks of concepts (Popping, 2000), or networks of words (Danowski, 1993), this approach to the analysis of text assumes language can be modeled as networks of words and the relations between them (Sowa, 1992). Translating pre-selected text into networks of concepts and the links between them, where a concept can be a word or a phrase (Popping, 2003), semantic network analysis captures the relational characteristics of meaning.

The distinctive feature of semantic networks is the ability to expose a picture of the web of meaning contained within text collections. Additionally, the network structures emerging from such an approach to the analysis of text, foster knowledge discovery because the network positions of concepts often highlight information that may be difficult to discern using other text-analytical methods (Cucchiarelli et al., 2012; Jin et al., 2012). Further attesting to the flexibility and adaptability of this method are the multitude of studies addressing a variety of topics and contexts. A few examples can be found in the works of Leydesdorff and Welbers on co-words in context (Leydesdorff & Welbers, 2011); Doerfel and Barnett on the structure of the International Communication Association (ICA)(Doerfel & Barnett, 1999); Shim, Park and Wilding on nuclear energy policy frames (Shim et al., 2015); Danowski on mapping publics of a business (Danowski, 2012); Carley and Kaufer on symbols and symbolic activity (Carley & Kaufer, 1993); Grebitus and Bruhn on the perception of food quality (Grebitus & Bruhn, 2008); Kleinnijenhuis et al. on the effects of news on voting behavior (Kleinnijenhuis et al., 2007); Kim on internet discourses of Korean supporters of Hwang Woo Suk (Kim, 2011); Leydesdorff and Hellsten on the role of language in tracking the way stem cell research is represented in various contexts (Leydesdorff & Hellsten, 2005); and Leydesdorff and Hellsten on metaphors and diaphors in scientific controversies (Leydesdorff & Hellsten,

2006).

The approaches to semantic network analysis vary based on the ways in which relations between words are extracted. For instance, the approach developed by Leydesdroff (1989) extracts semantic network relations based on cosine similarity, which is the normalized co-occurrence of concepts across documents. The focus of this method is then not on the absolute co-occurrence of a pair of words, but rather on the distinctiveness of their co-occurrence. Carley's map analysis approach (Carley, 1993) creates networks of concepts based on proximity co-occurrence and it also accounts for the direction of word associations. In this approach, the relations (links) between concepts in the semantic networks are based on co-occurrence but they are also directional, based on the positions of concepts relative to one another.

The approach to semantic network link extraction employed in the empirical chapters of this dissertation is based on the method of Rice & Danowski (1993)[2]. Supported by the work of Collins & Quillian (1969) and Chang (1986) on semantic memory and the hierarchical association of words at a cognitive level, this (undirected) proximity co-occurrence based method, generates semantic networks that represent the inherent meaning in texts. Relations between concepts are determined as they occur within a window of n words that moves sequentially through the text, one word at a time. For example, if two words co-occur within the specified window size, a link (or semantic network edge) will be formed. The window size determines the range of text words in which connections will be made between words within the window (Diesner, 2012b) and it can be two words, a sentence, a clause, a document etc. Because the links in these networks are based on co-occurrences, they can be unweighted as well as weighted. An unweighted link in a semantic network represents the existence of a relation (e.g., two words co-occurred in the specified window), while a weighted link also shows the intensity of that relation (e.g., how often two words co-occurred in the specified window). Throughout the research presented here, the value of strength for each link in the semantic

[2]For similar approaches see also the work of Sowa (1992) and the later work of Carley (1997b)

networks analyzed is determined by the frequency of co-occurrence (Wasserman & Faust, 1994). Thus, I have employed weighted links in the semantic networks. Furthermore, a concept or semantic network node can be a single word or an n-gram. N-grams are coded by replacing the spaces between words with an underscore (Carley et al., 2013a). An example of such a conversion is 'interest rate' being coded as 'interest_rate'. This procedure is used to identify the most common multi-word expressions in text documents and transforming them into what I will further refer to as concepts.

In sum, semantic network analysis becomes an invaluable and time-efficient tool for exposing patterns of large text corpora. This approach reveals key aspects of the relations between words within as well as across documents, and ultimately global shifts in discourse above and beyond what direct text analysis would reveal. Semantic networks provide insights into how language serves as a framework for representing and communicating information. However, although language can be suitably represented as a network of co-occurring words (Borge-Holthoefer & Arenas, 2010), semantic networks are often large and complex and exhibit highly intricate network structures (Bales & Johnson, 2006; Steyvers & Tenenbaum, 2005; Postma et al., 2000). The complexity of large semantic networks arises not only from the size of the corpora, but also from an array of global and local features, which in turn emerge from the structure of links between the concepts. To this end, in this dissertation I develop and employ the structural space approach, which allows the researcher to assess the content of semantic networks through a combination of popularity and connectivity features of semantic network nodes (i.e., concepts). This innovative approach, allows for the analysis of both in-depth and orthogonally informative dimensions of meanings.

2.2.1 The Structural Space of Semantic Networks

The structural space approach assesses subtle dynamic shifts in discourse through the structural positions of semantic network nodes. This approach combines two classic social network analysis structural measures, degree centrality (i.e., popularity) and betweenness centrality (i.e., connectivity) of concepts, to create

four structural roles for network nodes. The idea of structural roles in social networks has been explored through various approaches over the years. A few examples would be structural holes (Burt, 2009), equivalence (Burt, 1978, 1990; Boyd & Everett, 1999; Borgatti & Everett, 1992), blockmodels (White et al., 1976; Anderson et al., 1992), and role structure (Boorman & White, 1976). However, the identification of structural roles through the combination of structural measures has not been widely explored.

One such effort comes from Carley and Kaufer (1993), and it combines density, conductivity and consensus to explore connectivity in semantic networks. The paper of Huang et al. (2014) proposes a combination of multiple strongly correlated social network analysis (SNA) metrics to evaluate only those top ranked nodes in undirected binary networks. For visualisations purposes, NodeXL offers the possibility of plotting nodes based on their actual centrality scores but without identifying different roles (Hansen et al., 2011). The distinctive feature of the structural space approach is the identification of four structural roles based on the combination of two structural measures, and thus it is not merely focused on high ranking nodes. Building on the manner in which popular and connecting concepts play different roles in the structure and dynamics of semantic networks, this approach combines the popularity (i.e., total degree centrality) and connectivity (i.e., betweenness centrality) dimensions. This combination allows for the identification of four structural roles.

The degree centrality of a node in a network reflects the number of other nodes to whom the focal node is tied (Freeman, 1979)(or, in the case of weighted networks, the sum of the weights of all the links a node has), and thus measures the involvement of a node in its local network. Nodes with low total degree centrality are potentially more peripheral to the network (Iacobucci et al., 1996), unless they are connected to popular others. In semantic networks, total degree centrality may represent the 'importance' of a concept or its key concept status. A key concept with high degree centrality is able to activate many other key concepts; thus, it functions as a hot topic's central key concept (Diesner, 2012b). On the other hand, betweenness centrality is the sum of the proportions of the shortest paths a node lies on for every pair of nodes (out of all

shortest paths for each pair). More broadly, betweenness centrality represents the frequency with which a particular node is on the geodesic path between any other two nodes in the network (Danowski, 2012). The betweenness centrality of a concept within a semantic network is a direct indicator of its influence (Hill & Carley, 1999; Hooper et al., 2012; Wasserman & Faust, 1994). A key concept with high betweenness centrality controls access to other key concepts in the network (Brandes & Corman, 2003; Grebitus & Bruhn, 2008; Henderson et al., 1998; Hulst, 2008), and thus serves as a gatekeeper between different domains (Gloor & Krauss, 2009). The combination of these two measures, positions semantic network concepts within this structural role space, and thus, uncovers subtle structural properties of concepts and a set of changes in discourse over time. To avoid further repetition, the structural space approach will be described in more detail in the following Chapter, Chapter 3.

2.2.2 Network Analytic Methods

MRQAP

In the studies presented in Chapter 3 and Chapter 4, I also perform QAP (Quadratic Assignment Procedure) correlations and multiple regressions (MRQAP) to explore temporal dynamics of semantic networks. These methods compare one or more networks using edges and their weights as data points while controlling for their dependencies such as auto-correlation within the network structure (Krackhardt, 1987). These methods have been widely used in social network research (Dekker et al., 2003, 2007; Kilduff & Krackhardt, 1994) and also applied to research in knowledge and semantic networks (Corman et al., 2002; Broekel & Boschma, 2012; Xiang et al., 2009).

MRQAP is essentially multiple regression predicting the edge weights (including non-edges) of one network from one or more other networks. Typically, the networks (both the dependent and independent ones) are transformed into adjacency matrices so they contain edge weights as well as zeros for non-edges. These matrices are then elongated into single vectors such that the positions in each vector correspond to the same matrix cell positions. These vectors then

serve as the dependent and independent variables in the multiple regression, and an estimated regression coefficient (one for each predictor network) indicates the extent to which an independent network's edge (weights) contribute to the corresponding edge weight in the dependent network.

While the regression coefficients from an MRQAP are identical to those of a least squares regression, their significance scores (i.e., p-values) are derived by comparing the estimates against their null distributions obtained from applying the same regression model to a large sample of permutations ($m = 1000$) of the node structure (i.e., node relabelings) thereby controlling for autocorrelation (Krackhardt, 1987).

Multidimensional Scaling

To further explore the correlation analysis performed through QAP, I use multidimensional scaling (MDS) to depict differences and similarities among multiple semantic networks. Hence, MDS representations of network correlations can be used to plot compelling illustrations of the temporal dynamics of semantic networks. In MDS, also known as Principal Coordinate Analysis, an eigen-decomposition reduces the dimensionality of a matrix of distances such that a pairwise distance between each pair of data points in the reduced space is roughly proportional to the original distance between the pair. In the empirical chapters presented in this dissertations, I employ MDS by transforming distances into similarities, so that highly correlated networks appear visually closer in the 2D space, while distant points indicate relatively lower correlation in the semantic networks.

2.3 Topic Modeling

The second relational text-analytic method employed in the last empirical study presented in this dissertation (Chapter 5) is the latent Dirichlet allocation (LDA) topic models. Topic models are a class of automated text analysis tools that seek to identify, extract, and characterize the various (latent) topics contained by collections of texts. Based on the same assumption regarding the

relational aspect of meaning as semantic networks, topic models define a coherent topic as a set of word clusters. More specifically, topics are identified based on word co-occurrence patterns across a corpus of text documents, where a cluster of words that co-occur frequently across a number of documents constitute a topic. Using contextual clues, topic models connect words with similar meanings and differentiate between uses of words with multiple meanings. At a more technical level, topic modeling is based on the idea that documents are collections of topics, where a topic represents a probability distribution over words. Each topic is separately meaningful, offering a probability distribution over words which produces a consistent cluster of correlated terms (Blei et al., 2003; Griffiths & Steyvers, 2002, 2003, 2004; Hofmann, 1999, 2001). First described and implemented in the context of natural language processing, topic models use algorithms designed to browse and summarize large archives of texts. The latent Dirichlet allocation (LDA) (Blei et al., 2003), specifically, is a three-level hierarchical Bayesian model used to retrieve information contained in large collections of texts.

When fitting the LDA topic model to a collection of text documents, the analyst needs to specify the number of topics to be identified, as well as the number of words best fitting each topic. In general, selecting the most appropriate number of topics for a specific corpora implies exploration of different solutions. The output of the LDA models then represents a specified number of best fit words for each topic. For instance, in Chapter 5, where I apply LDA models to speeches given in the European Parliament, the 15 topic solution outputs ten words for each topic. These clusters of words representing predominant topics can be further labeled by the analyst to summarize the topic they represent. To further assess the similarity or distinctiveness of these topics, topic by topic networks can be generated from these labels and clusters of words. In these topic by topic networks, the links represent concepts shared by the topics. Thus, if the same word is a member of two or more topics, a link will be formed between these topics. Consequently, the values of each link in these topic networks represents the number of shared words between topics. Once these networks are generated, they can be further analyzed through standard

network measures. The analysis can also be enhanced through the addition of node (i.e., topics) attributes, for instance frequency based on the sum of frequencies of the words that belong to each topic. As shown in Chapter 5, these topic networks are useful in exploring the structure of shared concepts among topics, revealing those topics that are highly similar, as well as those topics that are distinct. When analyzing the structure of these topic networks, the structural space approach proved equally valuable in revealing the similarity and distinctiveness of each topic, as well as more general characteristics of discourse (e.g., repetitive content).

2.3.1 The Structural Space of Topic Networks

Initially developed and employed in the analysis of semantic networks, the structural space approach draws on a combination of centrality measures of semantic network nodes to uncover subtle structural properties of discourse. As detailed in previous sections (see Section 2.2.1) and in Chapter 5, this approach is based on the combination of two structural measures: total degree centrality and betweenness centrality. This combination positions the concepts within this structural role space, identifying four distinct structural roles of network nodes. However, because the links in topic networks represent shared-concepts among topics, the meanings of degree centrality and betweenness centrality need to be reconsidered, and hence the characteristics of each structural role in the context of these topic networks. First, degree centrality in the topic networks represents the number of concepts a topic shares with other topics, and thus it becomes a similarity measure. A topic with high degree centrality has higher similarity to one or more topics, than a topic with low degree centrality. On the other hand, betweenness centrality in the topic networks, denotes the connectivity potential of the topic. A topic with high degree betweenness is a gateway to other parts of the topic network, while a topic with low betweenness centrality does not fulfill a connective role. In the particular case of these topic networks, then, the structural roles will characterize topics based on similarity (rather than popularity), and connectivity potential. Hence, this approach will reveal the level of topic coherence or distinctiveness across text corpora, and

the connectivity potential of various topics. A more comprehensive description of each structural role identified by the structural space approach in the case of topic networks is given in Chapter 5, Section 5.4.1. Overall, by characterizing different part of the corpora under study, this approach can reveal which parts of the text (topics or concepts) deserve further attention, and the overall meaning and latent agenda of texts.

In Table 2.1 on the following page, I present an summary of how I employed each of the methods described above, the types of analyses performed, and the type of textual data analyzed in each of the empirical chapters. The table also gives an indication of the total number of texts included in the analysis of each chapter. Before elaborating on each of the data sets analyzed (in Section 2.5 on page 39), in the next section I expand on the context in which these texts were generated, namely the global financial crisis, and I elaborate on the three different discourses contained by the text documents I analyze. For each of these discourses (i.e., central bank discourse, media discourse, and political discourse) I provide a description of their inherent characteristics and the specific characteristics of the actors producing them. Furthermore, I present the expectations of how the meanings and content of these texts and discourses are affected by the crisis and the potential impact of these texts and their content on perceptions of the crisis.

Table 2.1: Summary of analyses and methods for empirical chapters

	Data	**Analyses**	**Methods**
Chapter 3	Press releases	Longitudinal semantic network analysis of key terminology; Structural roles of key concepts and their longitudinal dynamics	Structural space of semantic networks; MRQAP; MDS of network correlations
		Texts: 3013; Sentences: 39622; Words: 719581; Avg. Words Per Sentence: 17.61	
Chapter 4	News items	Metaphor family identification; Longitudinal analysis of metaphor family dynamics; Structural roles of metaphors and their dynamics in discourse	Structural space of semantic networks; MRQAP; MDS of network correlations
		Texts: 2817; Sentences: 87805; Words: 1971563; Avg. Words Per Sentence: 23	
Chapter 5	Speech transcripts	Identification of dominant topics; Structure of topics; Distinctiveness vs. similarity of topics; Structural roles of topics	LDA topic models; Structural space of topic networks
		Texts: 3955; Sentences: 31021; Words: 845619; Avg. Words Per Sentence: 27.21	

2.4 Context: The Global Financial Crisis

> "Crises become part of the normal organizational process, purging parts of the system that are outdated and inappropriate and creating avenues for development and change." (Murphy, 1996)

The global financial crisis that emerged in August 2007 in the United States of America resulted in the collapse of large financial institutions, the bailout of banks by national governments, and downturns in stock markets around the world (Riaz, 2009). The near collapse of the American financial system in 2008 wiped out more than $11 trillion in household wealth, left nearly 28 million people jobless, increased poverty rates to 15 percent of the population (46.2 million people), and forced the foreclosures of as many as 13 million households in the United States (Financial Crisis Inquiry Commission, 2011; Podpiera & Otker-Robe, 2013). Some of the harder-to-quantify impacts of the crisis, are the consequence of extended unemployment, reduced opportunity and increased government presence in the economy.

Impacting all areas of society, the financial crisis has been described by economists as the worst financial crisis since the Great Depression of the 1930s (Pendery, 2009). The severity and rapidity with which this crisis has spread across the globe can be considered of epidemic proportions, which has led to speculations on the origins and the subsequent processes that unfolded (Bordo, 2008; Caballero et al., 2008; Caballero & Simsek, 2009; Cavanagh & Mader, 2004; Congleton, 2009; Diamond & Rajan, 2009; Obstfeld & Rogoff, 2009; Riaz, 2009). Generally, authors talk about three events that contributed to the development of excesses of liquidity, leverage, risk taking, and greed, which in sum gave rise to what we now call the global economic crisis: (1) the US saving rate decline and its dependence on external sources, (2) globalization and (3) the global trade and finance (Caballero et al., 2008). These events are considered to be triggers of the crisis and, at the same time, mechanisms that perpetuated its rapid spread.

The first event mentioned in the literature is a two-fold historical event that has the US and Asia as main actors. This started with the global imbalance created by (1) the decline of saving rate in the US and its dependence on external resources (Bordo, 2008). The financial system in the US changed in a short time from a traditional retail banking to a highly leveraged banking system that was funded mostly through securization of assets. This development has facilitated the growth of a global imbalance on the financial markets. At the same time, another contributor has been the Asian crisis of 1997-2000, which has also added greatly to the excessive leverage and large capital flows (Bordo, 2008). The dependence of the US economy on Asian markets and the lack of precautionary measures in reaction to the meltdowns of the Asian crisis of 1997 have further contributed to the destabilization of the US financial markets and in turn to the destabilization of the global markets. The second and third processes are closely linked to each other, namely (2) globalization and (3) global trade and finance. These processes are also directly related to the fast development of information and communication technology, without which the global trade and globalization would not have been possible (Haass & Litan, 1998).

Although previous financial and economic crises have shown similar evolution patterns (Reinhart & Rogoff, 2008), the financial crisis of 2007-2008 has spread at a faster pace than any other (Sheng, 2010), attesting to the complexity of the 21st century financial system but also to the interconnectedness of states and financial markets across the globe. The prevalence of the national states and national economies has been shadowed by the "global flows of capital, goods, services, technology, communication, and information." (Castells, 2007, p.303). This increased mobility of resources and information was possible in part due to the new infrastructure provided by the evolution of information and communication technologies, and in part by the deregulations and liberalization practices employed by national states and international organizations. The 'new economy' as discussed by Castells (2000) allows for the 'globally integrated financial markets' to manage transactions of capital within seconds with no distance boundaries. Increasingly electronic, financial markets, using cutting edge computer applications, are open to millions of simultaneous in-

vestors, having instantaneous access to the same information no matter where they are and being able to perform multimillion dollars transactions in mere seconds. Hence, (2) globalization and (3) global trade and finance are highly connected issues that can be seen in the light of a relation, where globalization lead to increased global trade and finance, and at the same time the increase in global trade and finance has fueled globalization.

Although much has been written about the triggers and causes of the crisis, and opinions often diverged (e.g., Bordo, 2008; Crotty, 2009; Diamond & Rajan, 2009), the rippling consequences of the crisis are irrefutable. The crisis has had severe consequences not only for the U.S. but also for countries all around the globe, destabilizing economies and impacting all aspects of social life. The collapse and/or rescue of major banks, the freezing of inter-bank liquidity, and the impact on stock markets, production systems, national economies, and workforces revealed key developments of the unfolding crisis. National and supra-national governmental organizations confronted a worldwide recessionary spiral. The unprecedented global connectivity of finance, production, and consumption meant that financial stability was threatened on a global scale. The global nature of the financial crisis has made clear that financially integrated markets, while offering many benefits, can also pose significant risks, with large real economic consequences.

The magnitude of the crisis and speed with which it spread across the globe resulted in great levels of anxiety and uncertainty among social actors of all kinds. "Experienced as 'turning points', crises elicit new narratives, signal the obsolescence of the status quo in markets and policy regimes, and inject deep uncertainty into agents' decision calculus." (Nelson & Katzenstein, 2014, p.362) In a crisis situation, when the complexity of the environment rises dramatically (Caballero & Simsek, 2009), social actors are confronted with non-routine decision situations, unforeseen sources of uncertainty, confusing and incomplete accounts of potential threats, and increased pressure to make appropriate decisions. Due to these high levels of uncertainty, social actors seek information looking for answers to their dilemmas but, at the same time, they become more susceptible to information from others. The general demand for

information created by a societal crisis (Coombs & Holladay, 1996) severely pressures the discourses of those regulatory and policy-making organizations expected to promptly intervene. But, at the same time, the discourses of these organizations have the potential to become highly influential in shaping how the crisis is made sense of and acted upon because social actors, trying to make sense of the unfolding events of the crisis, scrutinize these discourses even more in search for information and answers to their dilemmas.

Furthermore, the levels of uncertainty experienced by social actors during a crisis may also be driven by the media and the ways in which they convey information regarding the unfolding events. Through the meanings they confer to issues, the media can bridge or widen the gaps between social actors, playing an important role in the containment or escalation of the crisis (Mazzoleni & Schulz, 1999; Ball-Rokeach, 1985).

These interrelated processes of influence at the level of discourse render the the global financial crisis as a complex and valuable context that offers ample opportunities for text-analytic explorations. Under these multifaceted and intricate dynamics of the crisis, it is natural to assume that rapid, yet subtle shifts, changes, and adaptations of discourses may occur. Aiming to explore and expose these subtle discursive manifestations and to demonstrate the effectiveness of the text-analytic methods I employ and develop, in the subsequent empirical chapters, I focus on three distinct types of social actors and their discourses as manifested in text documents they produce. In Chapter 3 the analysis focuses on two central banks, The European Central Bank (ECB) and the Federal Reserve System (Fed), in Chapter 4 texts published by three media outlets (The Financial Times, The New York Times, and the Sun) are analyzed, and in Chapter 5 I analyze transcripts of speeches given in the European Parliament (EP). The distinctiveness of these three discourses and their specific characteristics offer opportunities to explore the flexibility and malleability of the text-analytical methods I employ and expand. In the following section, I elaborate on the individual characteristics of these actors and their discourses. Starting from the general characteristics of such discourses, I also discuss their importance and salience in times of societal crisis and the expected outcomes

of my analysis. The methodological aims of this dissertation are suitably complemented by valuable empirical findings, setting a stepping stone toward a better understanding of how the events of the recent financial crisis were reflected in various social discourses and how the crisis drove discursive changes and adaptations.

2.4.1 Three Distinct Discourses

Central Bank Discourse

Organizational text documents, such as the ones issued by central banks, are often manifestations of highly formalized discourse, which leads to redundant, structured, and even predictable (Irvine, 1979) language use. The structured manner in which central banks produce their text documents is generally governed by rules or conventions meant to standardize their content, for instance, through the use of particular sets of words or fixed-text sequences. These stylistic and structural features of texts produced by central banks, combined with the highly specialized information contained by these texts, pose important challenges for classic text-analytical approaches. When analyzing large collections of such texts through frequency counts or frequency distributions, for example, the analyst is confronted with repetitive top key concepts, indicative of the obvious and perhaps uninformative standardized communication practices of these organizations, as I will demonstrate in Chapter 3. Thus, investigating texts issued by central banks require suitable methods, capable of exploring their content beyond the core elements of formal discourse stylistic form. To this end, I demonstrate the effectiveness of the relational approach to meanings in text in combination with the structural space approach in overcoming these challenges to the analysis of large corpora of formal texts.

The importance of central bank communications has been recognized by an increasing number of researchers in recent years (Ehrmann & Fratzscher, 2007; Friedman, 2002; Kohn & Sack, 2003; Rosa & Verga, 2005; Sturm & de Haan, 2011). Many of their studies, investigating central bank discourses, have at least three findings in common: (1) central bank communications regarding

economic projections and monetary policy developments have significant impacts on the financial markets, (2) these effects are even greater when the communication channel is more formal, and (3) the more prominent the position of the communicator, the stronger the financial market reaction (Connolly & Kohler, 2004; Kohn & Sack, 2003; Chirinko & Curran, 2006; Reinhart & Sack, 2006). Thus, the communications of central banks are of great importance in times of market precariousness and financial crises. As watchdogs for price stability and economic growth, central banks use communication as tools in maintaining market stability by steering perceptions and guiding actions.

Central bank discourses fulfill a dual function. On the one hand, the general public, central bank watchers, financial media and market participants pay considerable attention to statements of central banks and their informational content because a "central bank may have, or may be believed to have, superior information on the economic outlook" (Blinder et al., 2008, p.10). These communicative acts are an important and direct source of information on future policy decisions (Jansen & de Haan, 2005a; Blinder et al., 2008), objectives of monetary policy strategy (Blinder et al., 2008), and expected future rates (Bernanke et al., 2004). On the other hand, communication has become a key tool for central banks in general, and can be used to guide private sector expectations and/or reduce noise in the financial markets (de Haan, 2008). Hence, central bank discourses may serve as a coordination device for the beliefs of financial market agents (Amato et al., 2002).

The two central banks I investigate in this dissertation, the European Central Bank (ECB) and the United States' Federal Reserve System (Fed), determine the monetary policy for two of the largest currency areas and two of the largest economies in the world. Established by the Treaty of Amsterdam in 1 June 1998 (European Union, 1997), the ECB is the formal successor of the European Monetary Institute. As one of the seven institutions of the European Union, the ECB is the central bank for the euro and administers the monetary policy of the 17 EU member states (using the Euro), which constitute the Eurozone. The Fed has a longer history than that of the ECB. Its establishment on December 23, 1913 was a response to a series of financial panics, particularly

the panic of 1907. The duties of the modern Fed include conducting the nation's monetary policy, supervising and regulating banking institutions, maintaining the stability of the financial system, and providing financial services to depository institutions, the U.S. government, and foreign official institutions (The Federal Reserve Board, 2014). Among the central banks, the Fed has a unique structure due to a mix of private and public elements which serve the interests of the general public and the interests of private bankers.

These two banks have been at the epicenter of the financial crisis, and have been involved in the efforts to combat the impact of the crisis and aid financial recovery. Although the two banks serve similar roles in their respective currency areas, their main objectives differ in that the goal of the ECB is to maintain price stability, while the Fed promotes maximum employment and moderate long-term interest rates (Belke & Klose, 2010). In terms of their behavior and decisions in the early stages of the financial crisis, the ECB has moved its policy rate much less frequently than the Fed (Sahuc & Smets, 2008). While the Fed started lowering rates already in August 2007, the ECB did not lower the interest rate until October 2008. (Belke & Klose, 2010). Throughout the different stages of the crisis, the ECB focused on fighting consumer price inflation at the cost of some output losses, while the Fed promoted output and put a smaller weight on suppressing consumer price inflation (Belke & Klose, 2010).

While previous research established a link between communications of the ECB and the Fed and their respective impacts on the financial markets (Jansen & de Haan, 2005a,b; de Haan, 2008; Hayo & Neuenkirch, 2010; Hayo et al., 2014), my focus was directed at uncovering the subtle shifts and adaptations of their discourse at the different stages of the crisis. I investigated the discursive practices of the ECB and the Fed, the dynamics of these practices, and the roles of the main objectives vs. crisis-oriented terminology in their discourses.

Media Discourse

In a general sense, news items (e.g., newspaper articles, reportages, news briefs, broadcasts, etc.) are stories presenting an event or a topic. However, to give

a story a newsworthy angle, journalists often use rhetorical devices to invoke images, increase salience of a point, and increase vividness of a report (Pan & Kosicki, 1993). In doing so, they suggest a particular perspective on the event or topic reported. For example, by referring to the financial crisis as the 'global crisis', the news media indirectly place the crisis beyond the direct control or responsibility of any actors or organizations. Hence, lexical choices constitute an important aspect of news discourse and they are frequently made in compliance with structural rules (Pan & Kosicki, 1993), often signifying the presence of a particular frame. Frames select certain features of a story and they exclude others (Iyengar, 1987). Through framing media may restrict or define a story's meaning and thereby shape people's interpretation of that story (Hallahan, 1999). Entman (1993) defines framing as a way "...to select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation." (p. 52) Thus, by highlighting certain characteristics of an issue and hiding others, framing reflects the emphasis of the author (i.e., journalist).

Among some of the most powerful framing devices is the use of metaphors (Gamson & Modigliani, 1989). Metaphors can shape the meaning audiences assign to news (Williams et al., 2011) and may also drive the effects that these interpretations carry into behavioral forms of decision making (Williams, 2013). The different theories developed over the years (Black, 1962; Johnson, 1981; Lakoff & Johnson, 1980; Ortony, 1993) approach metaphors as discussing a concept in terms of a different concept, thus transferring meaning from one concept to another. Divided into substitution and interaction theories, these theories differ by "locating metaphor either at the level of language and words as opposed to thought and context" and by "emphasising the role of metaphors as either reflecting some already existing similarities as opposed to also creating similarities between things or ideas" (Hellsten, 2002, p. 17).

Recent metaphor research has shed a different light on the social and communicative roles of metaphors and their effects on our understanding of public issues (Chilton & Ilyin, 1993; Hellsten, 2002; Lakoff & Johnson, 2003). Lakoff &

Johnson (1980) have theorized that a significant part of our everyday language is structured metaphorically, and thus we often use metaphors to understand one idea in terms of a different, more familiar idea. Metaphors have the ability to offer common grounds between discourses (Chilton & Ilyin, 1993) or to function as boundary objects (Star & Griesemer, 1989) that are at the same time flexible enough to allow several interpretations in different social contexts but also to carry a relatively fixed set of associations. Ultimately , metaphors play an important role in defining the way we perceive the world and, thus, the way we think and act (Lakoff & Johnson, 1980).

The use of metaphors to frame economic affairs of the financial crisis shapes consumers' economic perceptions in ways that hold wider significance in the global financial markets (Williams, 2013). As tools meant to either popularize or condense complex issues, or to translate highly specialized discourses, metaphors guide our perceptions and interpretations of reality and help us to frame our visions and goals, "playing a central role in the construction of social and political reality" (Lakoff & Johnson, 2003, p. 159). As such, the use of metaphors in news has the potential to influence meanings readers associate with the issues reported (Williams et al., 2011), which in turn can manifest changes in behavior and decision making (Williams, 2013). While metaphors in crisis communication have received increased attention (Bounegru & Forceville, 2011; López & Llopis, 2010; Peckham, 2013; Tourish & Hargie, 2012), the study of metaphors in debates relevant to the latest financial crisis is still in its incipient stages, even though the Metaphor Observatory[3] discusses the financial crisis as the trigger for "one of the largest metaphor spikes in recent history."

Part of the media's interest in crises comes from the fact that crises are dramatic, they are newsworthy. Kleinnijenhuis et al. (2013) found that "the news about the crisis became crisis news itself" (p. 287), reinforcing the idea that the way in which journalists report the events of the financial crisis has a major impact on the escalation of the crisis. Consequently, media reporting will not only define, but it will make salient the conditions of crises (Heath, 2010). A crisis is an event for which people seek causes and make attributions (Coombs

[3]See `www.metaphorobservatory.com`

& Holladay, 1996) . In other words, people seek information about the crisis, evaluate the causes, and the organizational responsibility for the crisis based on media coverage (An & Gower, 2009). People's perceptions of events are dependent on what information or feature is salient in the environment (Heider, 1958). In this sense, the media are highly influential in shaping the 'mood' of a population, which in turn may have political consequences associated with political preferences, attitudes, and ultimetly, voting behavior (e.g., DellaVigna & Kaplan, 2007; Gerber et al., 2009; Pinkleton et al., 1998; Quarles, 1979).

Although research on metaphors has a longstanding tradition (e.g., Chilton & Ilyin, 1993; Gamson & Modigliani, 1989; Hellsten, 2002; Lakoff & Johnson, 1980, 2003; Ortony, 1993; Tourish & Hargie, 2012; Williams et al., 2011), the identification of metaphors in large collections of texts remains a challenge. In Chapter 4, I propose a method of metaphor identification based on the extraction of words in text and their relationships. Additionally, I extend this particular method to identify metaphor families in news reporting. Metaphor families represent groups of related metaphors and have the potential to strengthen (un)intended images on the topic under discussion, and have been only seldom studied in metaphor research. The temporal expansion, evolution, and subtle changes in the use of a metaphor family are further explored through semantic networks and the structural space approach.

Political Discourse

Political language, political discourse and political text are vague terms, not a homogeneous genre (Hernández-Guerra, 2013). No universal, inherent characteristics to all types of political texts and discourse exist (van Dijk, 1998). Rather, the content of political texts is structured based on the context in which they are generated and the authors's intentions. Various types of political texts may present schematic structures, structures either imposed by legally binding requirements (e.g., Openings and Closings of official sessions of parliament) or by conventional or strategic practices. Hence, political texts, generally contain lexical choices determined by normative official criteria of decorum, but also lexical choices meant to increase their effectiveness and persuasion capacity.

Political speeches given in parliament, for instance, are expected to be held in relatively formal style of address and dialogue (van Dijk, 1998), but they will (more often than not) also include rhetorical strategies meant to emphasize or de-emphasize political attitudes and opinions, garner support, manipulate opinion, create political consent, or legitimate political power. Hence, the structure and content of political texts is contingent upon the context in which they are created and delivered, and on the political goals of their authors. Inferring meaningful information from collections of political texts, then, requires analytical methods able to parse and extrapolate beyond their normative content, into the embedded latent meanings.

Generally, researchers focusing on political texts, make use of the coding schemes developed by the Manifesto Research Group (MRG)[4] (e.g., Benoit et al., 2009; Franzmann & Kaiser, 2006; Gabel & Huber, 2000). However, the coding process employed in these schemes involves a degree of manual coding and/or human reading of the full texts, the use of pre-developed techniques of transforming text documents into numerical data (Gabel & Huber, 2000), or the involvement of expert coders. These approaches to the analysis of political texts, although valuable, involve time consuming coding techniques, coder bias, and limitations in terms of the size of the corpora analyzed.

Exploring transcripts of speeches given in the 7^{th} European Parliament, I propose an automated approach to the analysis of large collections of political texts, capable of exposing dominant discursive practices, their inherent structure, and subtle manifestations of policy positions. The fundamental features of this approach stem from the perspective on language meaning as relational, and it employs fully automated extraction of dominant clusters of words representing dominant topics. Moreover, I demonstrate the efficacy of the structural space approach in revealing the interrelatedness of dominant topics in political texts.

The importance and power of political texts have been vastly recognized and researched (e.g., Campbell & Jamieson, 1990; Gray & Griffin, 2014; Harris, 1991; Maynard, 1994; Seidel, 1988; Zupnik, 1994). In the form of policy

[4] `https://manifestoproject.wzb.eu/`

documents, written statements, or speeches, politicians transmit and shape political information, political beliefs and political opinions. "Political texts are the concrete by-product of strategic political activity and have a widely recognized potential to reveal important information about the policy positions of their authors." (Laver et al., 2003, p. 311) Through various language strategies, such as the use of metaphors (Bosman, 1987; Lakoff, 1995; Musolff, 2004), framing and reframing (Entman, 1993; Lakoff, 2004), emphasis, or repetition (Wilson, 2015), politicians are able to accentuate certain semantic features of a given utterance, while at the same time obscuring others. In other words, in political arenas, language is used to bring a certain perspective to the fore and compete over the establishment of dominant perspectives. Understanding the discursive spaces in which political action takes place and rhetorical language strategies employed, has the potential to expand our understanding of the shared perceptions of values that defines political associations. Comprehensive and inclusive approaches to political text analysis provide accounts of processes, inherent to the political realm, through which authority, legitimacy, and consensus are negotiated.

In the specific case of the European Parliament (EP), the speeches of its members offer a window into the disputes taking place in this complex discursive space, in which the ideological and cultural diversity of its members may foster critical disparities regarding political positions taken on important issues such as economic policy. As the only directly-elected institution of the European Union (EU) and one of the most powerful and influential of the EU bodies (Hix, 2011), the levels of competition and/or agreement among the Members of the European Parliament (MEPs) can directly impact the legislative process of the EP. In the context of the financial crisis, for instance, the EP became a key actor in the EU's efforts to combat the rippling effects of the crisis within the Eurozone. Through the large number of socio-economic legislative proposals adopted in a relatively short time, the EP transformed the landscape of the European financial system and the manner in which institutions do business (Broin, 2012). These direct impacts of the EP policy decisions onto European financial markets, spanning a wide range of areas and issues, reinforce the need

for scholarly research to make systematic efforts to better understand the interactions and tensions within the EP's political and discursive space. To this end, in the study presented in Chapter 5, I bring a contribution by exposing the dominant discursive practices of the two largest political groups of the EP on issues pertaining to the Eurozone financial crisis.

In the context of the global financial crisis, the three discourses described in the previous paragraphs, are expected to focus on different aspects of this crisis, in accordance with the social roles of the actors enacting these discourses. In an oversimplified example, in financial arenas, where central banks operate, the current crisis is about financial losses; in political arenas, such as the one in which the European Parliament is embedded, the crisis is about legitimacy issues; while in the media arena the crisis is about novel, dramatic, and newsworthy events. Financial, political, and media discourses also exhibit different degrees of complexity specific to the social contexts from which they emerge. Thus, by systematically investigating these three discourses, and by revealing the crisis aspects they capture and their temporal dynamics, this dissertation sets a stepping stone towards a comprehensive understanding of the effects of major societal crisis at the level of discourses. Uncovering and understanding the different discursive strategies and manifestations employed by these various actors, has the potential to provide a more inclusive depiction of some of the mechanisms that contribute to the amplification or containment of a crisis. With this dissertation, and the text-analytical approaches it provides, I hope to encourage further research to adopt more inclusive perspectives and concurrently investigate multiple discursive sites relevant to their topic of interest.

To conclude these two chapters (i.e., Chapters 1 and 2), that have set the stage for the three empirical studies presented in the subsequent chapters, in the following section, I describe the data sets collected for each study in detail.

2.5 Data

The data sets analyzed in Chapters 3 to 5 are collections of text documents collected from web databases or digital archives. A summary of the sources,

number and type of text documents, the time period they span, and the sites from where they were collected is given in Table 2.2. Overall, in this dissertation I analyze 9782 text documents representing newspaper articles, press releases, and speech transcriptions.

Table 2.2: Data collection for empirical chapters

Chapter	**Source**	**Texts**	**Type**	**Longitudinal Time Span**
3	European Central Bank	825	Press releases	January 2006 to December 2013[1,2]
	Federal Reserve System	2185	Press releases	
4	Financial Times	2212	News items	2006 to 2011[3]
	New York Times	437	News items	
	Sun	168	News items	
5	European Parliament: EPP group	2499	Speech transcripts	June, 14 2009 to June, 30 2014[4]
	European Parliament: S&D group	1456	Speech transcripts	

[1]Collected from `www.ecb.europa.eu/press/pr/date/2006/html/index.en.html`

[2]Collected from `www.federalreserve.gov/newsevents/press/all/2006all.htm`

[3]Collected from `www.lexisnexis.com/`

[4]Collected from `www.talkofeurope.eu`

The first data set analyzed and presented in Chapter 3 contains a total of 3010 press releases issued by the European Central Bank (ECB) (825 documents) and the Federal Reserve System (2185 documents). These press releases have been published by the two organizations between January 2006 and December 2013 and have been collected from their respective web archives[5]. The two corpora collected present important differences. In particular, the Fed cor-

[5]The ECB and Fed web archives can be accessed at `https://www.ecb.europa.eu/press/pr/date/2006/html/index.en.html` and `http://www.federalreserve.gov/newsevents/press/all/2006all.htm`.

pora contains a higher volume of text (i.e., total word count) than that of the ECB. The difference in word counts per document between the two organizations is significant, and points to a striking difference in the communication styles between the U.S. and European organizations. While the Fed publishes frequent yet briefer press releases, the ECB's press releases are slightly longer and are published less frequently. Hence, the higher volume of the Fed data confers larger semantic networks of prominent concepts. However, given the complex structure of language, the semantic relationships emerging in the semantic networks generated were more nuanced than a mere function of word counts. Further details regarding these two data sets and their summary statistics can be found in Table 3.1 on page 53.

Once collected, the press releases of each organization have been divided into four time periods, each spanning a period of two years. This data aggregation was motivated by the aim of assessing the discursive shifts of the ECB and Fed at the different phases of the financial crisis. The first sub-sample covers the period just prior to the financial crisis: January 2006 until December 2007. I label this period *pre-crisis*. The second sub-sample (*crisis*) includes press releases issued between January 2008 and December 2009. The third sub-sample contains press releases issued between January 2010 and December 2011, and represents the *post-crisis* period. Lastly, the fourth sub-sample includes press releases issued between January 2012 and December 2013, further referred to as *recovery*.

In Chapter 4, the data set used to investigate the evolution of the *toxic* metaphor family, consists of news articles from three newspapers: *The Financial Times (FT)*, *The New York Times (NYT)* and *The Sun (Sun)*. The selection of these newspapers was motivated by the fact that they publish very different content and thus address different types of audiences. The *FT* is a highly specialized financial reporting newspaper; the *NYT* is the most popular daily newspaper in the United States, publishing a broad variety of topics; and the *Sun* is the largest circulation daily tabloid in the United Kingdom. For each of these newspapers, the LexisNexis database was searched with the keyword *toxic* with no start date but with an end date of December 31, 2011.

By searching the data base with no start date, I aimed to discover the first use of the 'toxic' metaphor in the three newspapers in regard to financial issues, which dates back to 2004, and it was published by NYT on February 13 (see quote in Section 4.3.1 on page 94). The results of the search query were further manually selected to ensure that only articles on financial topics were included in the resulting corpora. After removing duplicates, a total of 2,817 articles remained, divided as follows: 168 articles published by the *Sun*, 437 articles published by the *NYT*, and 2212 published by the *FT*.

In a similar fashion to the data aggregation used in Chapter 3, the articles collected from each newspaper have been separated into three sub-sets that I further refer to as the *pre-crisis* period (2006−2007), the *crisis* period (2008−2009), and the *post-crisis* period (2010−2011). The only exception to this is the *pre-crisis* sub-set for *Sun* that used the *toxic* metaphor only once in 2007, and it was not included in the analysis.

In the last empirical chapter (Chapter 5), in which I investigate the speeches given by members of the two largest political groups of the European Parliament (the EPP and the S&D), I have collected all the speeches given by the members of these two political groups between the 14th of June 2009 and the 30th of June 2014 from the Talk of Europe[6] project database (van Aggelen & Hollink, 2015)[7]. The interval selected represents the mandate of the of the 7^{th} European Parliament, which has been elected in June 2009. However, it must be noted that because the 7^{th} European Parliament started its legislative activities in June 2009, the speeches collected for the year 2009 span approximately six months of data. This is also the case for the year 2014, when the 7^{th} European Parliament was replaced by the 8^{th} Parliament at the end of June.

To assess the ways in which members of the two political groups position themselves in relation to the financial crisis, and what discursive practices they employed in regards to the crisis, only those speeches containing "financial crisis" and/or "economic crisis" have been collected. After duplicate removal, a total of 2499 speeches given my members of the EPP and 1456 speeches

[6]See `http://www.talkofeurope.eu` and `http://linkedpolitics.ops.few.vu.nl/` for more details on this database

[7]In Appendix B I present two examples of queries used for the collection of the speeches.

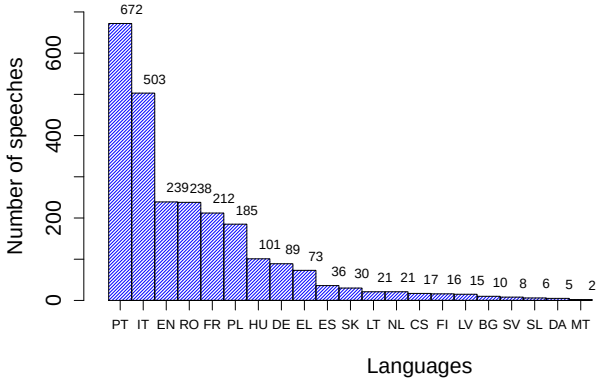
given by members of the S&D have been included in the analysis. While all the speeches collected are in English, a large number of theses speeches are translations from one of the 24 official and working languages of the European Union.

The multilingual character of the European Union, and implicitly of the European Parliament must be addressed because EP documents are published in all the official languages of the EU and every Member of the EP has the right to speak in the official language of their choice. Thus, one of the challenges this study faced was the loss of meaning through translations. MEPs are able to address the EP in their (official EU) language of choice, and these speeches are later translated. Due to this multilingual character of the EP, all legislative speech occurs in translation. In Figure 2.2 I show a breakdown of the number of speeches and the language in which they were given. The EPP members gave speeches in 21 different languages, with the dominant languages being Portuguese and Italian (Figure 2.2a). On the other hand, the S&D members gave speeches in 22 different languages and the most popular languages were English and Romanian (Figure 2.2b).

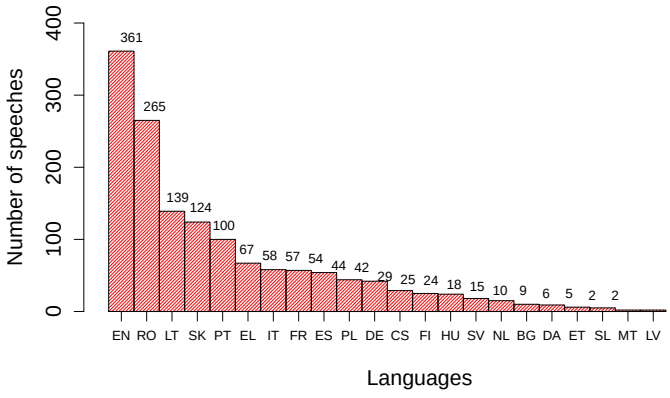
Hence, all of the EP's business occurs in multiple languages and therefore in translation. Even though so much of international politics occurs in translation, scholars have not paid significant attention to the effects of translation when using computer-based content analysis. Recognizing the multilingual character of the EP, and some of the implicit limitations imposed by this data, I argue that analyzing the speeches of the EP political groups (in translation) remains an important source of information for discourse analysts and political scientists alike.

In the following chapters, the three empirical studies are presented. Given the shared overarching aims and methods employed, theses chapters contain an acknowledged repetition of descriptions, methodology, and context. Also, because these chapters are based on published papers (Chapters 3 and 4) or on working papers in preparation for journal submission (Chapter 5), they are written using the plural 'we' form. In Appendix A, I present a detailed account of contributions brought by my co-authors in each of theses three empirical

chapters, and in Table 2.3 on the following page I summarize the scientific output relevant for each of the empirical chapters presented next.



(a) EPP speeches per language



(b) S&D speeches per language

Figure 2.2: Number of speeches per language.

Table 2.3: Scientific output for empirical chapters

Chapters	**Outputs**
Chapter 3: Mapping discursive dynamics of the financial crisis	This chapter was published in its current form (Nerghes et al., 2015b) and is related to the following: 1. Nerghes, A., Lee, J.-S., Groenewegen, P., and Hellsten, I. (2014). The shifting discourse of the European Central Bank: Exploring structural space in semantic networks. In Yetongnon, K. and Dipanda, A., (ed.), *Proceedings of SITIS*, pages 447-455, Marrakesh, Morocco. IEEE Computer Society; 2. Nerghes, A., Groenewegen, P. and Hellsten, I. (2014). Shifting discourses of the European Central Bank and the Federal Reserve System: Exploring structural space in semantic networks. Presented at the *EUSN Conference*, Barcelona, Spain; 3. Nerghes, A. and Groenewegen, P. (2014). Market Player or Regulator? A Semantic Network Analysis of the Shifting Roles of the ECB During the Financial Crisis. Presented at the *INSNA Sunbelt Conference*, St. Pete Beach, FL, USA; 4. Nerghes, A. (2013). EU Regulators: A Structure and Content Analysis. Presented at the *INSNA Sunbelt Conference*, Hamburg, Germany.

Chapter 4: A 'toxic' crisis	This chapter has been published in its current form (Nerghes et al., 2015a) and is related to the following: 1. Nerghes, Adina and Hellsten, I. (2012). A "toxic" crisis: Metaphorizing the financial crisis. Presented at the *Organizational Discourse: Processes, Practices and Performance Conference*, Amsterdam, The Netherlands.
Chapter 5: Europe Talks	This chapter is a working paper in preparation for journal submission, and is related to the following: 1. Nerghes, A., Groenewegen, P. and Hellsten, I., Taminiau, Y. (2015). The 7th European Parliament and the Eurozone financial crisis: cooperation or competition? Presented at the *Network Theory and Methods Workshop*, Amsterdam. The Netherlands; 2. Nerghes, A., Groenewegen, P. and Hellsten, I. (2015). Europe Talks: An analysis of discursive practices, their structural functions and the left-right political ideology spectrum in the European Parliament. Presented at the *INSNA Sunbelt Conference*, Brighton, UK.

3 Mapping Discursive Dynamics of the Financial Crisis: A Structural Perspective of Concept Roles in Semantic Networks

Abstract

Background/Purpose: Convenient access to vast and untapped collections of documents generated by organizations is a highly valuable resource for research. These documents (e.g., press releases) are a window into organizational strategies, communication patterns, and organizational behavior. However, the analysis of large document corpora requires appropriate automated methods for text mining and analysis that are able to take into account the redundant and predictable nature of formalized discourse.

Methods: We use a combination of semantic network analysis and network centrality measures to overcome these particular challenges and to explore the dynamic structural space of concepts in formalized documents pertaining to the recent financial crisis.

Data: For our analyses, we collect the press releases of the European Central Bank (ECB) and the United States Federal Reserve System (Fed) issued between 2006 and 2013 in order to examine their semantic networks before, during, and after the recent financial crisis. Their press releases are notably impactful in their influence on other financial institutions and society at large, especially during times of financial volatility.

Results: The structural space created from joint centrality metrics reveals salient shifts in the discursive practices of the ECB and Fed. In particular, the Fed exhibits greater attentiveness to the financial crisis especially during the crisis itself, while the ECB's attention is delayed and

Chapter based on Nerghes et al. (2014b) and Nerghes et al. (2015b)

increasing steadily. Furthermore, we show both the Fed's and the ECB's discourse transitioning into a new 'hybrid' state, rather than returning to the pre-crisis status quo.

Conclusions: Examining the semantic networks of organizational text documents, we find that our analytic approach reveals important discursive shifts, which would not have been discovered under traditional text analytic approaches. We demonstrate the utility of this approach in investigating large text corpora of organizational discourse, and we anticipate our methods to be comparably valuable in the analysis of a large spectrum of formal and informal discourse.

Keywords: Text, Semantic Networks, Centrality, Discourse, Structural Space, Financial Crisis.

3.1 Introduction

The increasing availability of online textual information opens new venues for large-scale research into organizational discourse and vocabulary shifts of organizations (Loewenstein et al., 2012). In particular, numerous text documents regarding organizational activities and objectives are generated daily across the world. However, large corpora of such text documents are difficult to analyze without proper methods which are in part automated.

Another research challenge is that organizational documents are often manifestations of highly formalized discourse, which leads to redundant, structured, and even predictable (Irvine, 1979) language use. Discourse 'acts as a powerful ordering force in [and by] organizations' (Alvesson & Karreman, 2000) because meaning is negotiated in organizations, and these meanings shape organizational practices (Grant & Hardy, 2004), as well as external presentation of organizations. As a carrier of power, language – commonly labeled as discourse – has the ability to order and constitute the social world (Alvesson & Karreman, 2000). By content and structure, discourse signals consequential information to other organizations and society in general. Its timely analysis may be crucial in order to understand the dynamic character of such signals, yet this analysis is often challenging. The approach employed in this study has been

designed to deal with complex semantic networks generated from large text corpora of formal organizational discourse (i.e., press releases). More precisely, the method assesses dynamic discursive shifts in complex semantic networks, highlighting the crucial distinction between connective and popular concepts.

Different approaches to text analysis include information retrieval, lexical analysis to study word frequency distributions, pattern recognition, tagging/annotation, information extraction, data mining techniques, etc. Direct text analytic methods such as word frequencies and tf-idf scores are limited in the subtlety of their inferences, revealing an incomplete picture of the discursive dynamics within the text. Without intending to minimize the utility and effectiveness of other text-analytical approaches, we propose that semantic network analysis becomes an invaluable and time-efficient tool for exposing subtle patterns of large text corpora. In this paper, we employ semantic network analysis to dissect organizational discourse in a structured manner. Specifically, we demonstrate how this approach reveals key aspects of the relations between words within as well as across documents and ultimately global shifts in organizational discourse above and beyond what direct text analysis would reveal.

Semantic network analysis is one of the areas of research that has gained popularity in recent years. This type of analysis maps networks of concepts (i.e., a concept being a word or multiple words) in the form of networks of meaning. Successful application of semantic network analysis to a variety of corpora addressing a multitude of contexts attests to the flexibility and adaptability of this method. A few examples can be found in the works of Leydesdorff and Welbers on co-words in context (Leydesdorff & Welbers, 2011); Doerfel and Barnett on the structure of the International Communication Association (ICA)(Doerfel & Barnett, 1999); Shim, Park and Wilding on nuclear energy policy frames (Shim et al., 2015); Danowski on mapping publics of a business (Danowski, 2012); Carley and Kaufer on symbols and symbolic activity (Carley & Kaufer, 1993); Grebitus and Bruhn on the perception of food quality (Grebitus & Bruhn, 2008); Kleinnijenhuis, Hoof, Oegema and Ridder on the effects of news on voting (Kleinnijenhuis et al., 2007); Kim on internet discourses of

Korean supporters of Hwang Woo Suk (Kim, 2011); Leydesdorff and Hellsten on the role of language in tracking the way stem cell research is represented in various contexts (Leydesdorff & Hellsten, 2005); Leydesdorff and Hellsten on metaphors and diaphors in scientific controversies (Leydesdorff & Hellsten, 2006); and Nerghes, Hellsten and Groenewegen on the evolution of metaphor families in media reports (Nerghes et al., 2015a).

Although language can be suitably represented as a network of co-occurring words (Borge-Holthoefer & Arenas, 2010), semantic networks are often large and complex and exhibit highly intricate network structures (Bales & Johnson, 2006; Steyvers & Tenenbaum, 2005; Postma et al., 2000). Some posit these networks to exhibit stylized topologies such as *small-world* or *scale-free* (Borge-Holthoefer & Arenas, 2010; Ferrer I Cancho & Solé, 2001; Steyvers & Tenenbaum, 2005; Solé et al., 2010; Postma et al., 2000). Semantic networks provide insights into how language serves as a framework for representing and communicating information. The complexity of large semantic networks arises not only from the size of the corpora, but also from an array of global and local features, which in turn emerge from the structure of links between the concepts.

Secondly, our paper applies an approach for assessing dynamic shifts in formal discourse through the structural positions of semantic network nodes. The structural space method, introduced in earlier work (Nerghes et al., 2013, 2014a,b, 2015a), combines two classic social network analysis structural measures degree (i.e., popularity) and betweenness centrality (i.e., connectivity) of concepts to create four structural roles for network nodes. The idea of structural roles in semantic networks is analogous to developments in social networks in which various approaches have been explored over the years. Examples are structural holes (Burt, 2009), equivalence (Burt, 1978, 1990; Boyd & Everett, 1999; Borgatti & Everett, 1992), blockmodels (White et al., 1976; Anderson et al., 1992), and role structure (Boorman & White, 1976) which have distinct functions in social theories of structure. However, the identification of structural roles through the combination of structural measures has not been widely explored in semantic networks yet. One such effort comes from Carley

and Kaufer (Carley & Kaufer, 1993), which combines density, conductivity, and consensus to explore connectivity in semantic networks. A recent study (Shim et al., 2015) examines structural roles of concepts alternatively using raw centrality measures. The paper of Huang et al. (2014) proposes a combination of three strongly correlated social network analysis (SNA) metrics (degree, egobetweenness centrality and eigenvector centrality) to evaluate only those top ranked nodes (core nodes and bridge nodes) in undirected binary networks. The distinctive feature of our approach is the identification of four structural roles based on the combination of two structural measures, and thus it is not merely focused on node popularity.

The corpora used in this study comprises the press releases issued by the European Central Bank (henceforth ECB) and the United States' Federal Reserve System (henceforth Fed) between 2006 and 2013. The ECB and the Fed determine the monetary policy for two of the world's largest currency areas. Established by the Treaty of Amsterdam in 1 June 1998 (European Union, 1997), the ECB is the formal successor of the European Monetary Institute. As one of the seven institutions of the European Union, the ECB is the central bank for the euro and administers the monetary policy of the 17 EU member states (using the euro), which constitute the Eurozone.

The ECB distributes large volumes of information (e.g., press releases, policy deliberations, public speeches, annual reports etc.) as one of their key policy tools. Because the ECB's only formal instrument, through which they can exert an (indirect) effect on asset prices (of key importance to the economy), is the overnight interest rate, their communications become a powerful tool. These can impact developments in the financial markets (Bernanke et al., 2004; Gürkaynak et al., 2005a,b; Kohn & Sack, 2003), directly influence private sector expectations, and are used to signal interest rate changes (Jansen & de Haan, 2005a,b). The communications of the ECB also increase the predictability of interest rate decisions (de Haan, 2008), being generally considered trustworthy and understandable by the public (Rosa & Verga, 2005).

The Fed has a longer history than that of the ECB. Its establishment on December 23, 1913 was a response to a series of financial panics, particularly the

panic of 1907. The duties of the modern Fed include conducting the nation's monetary policy, supervising and regulating banking institutions, maintaining the stability of the financial system, and providing financial services to depository institutions, the U.S. government, and foreign official institutions (The Federal Reserve Board, 2014). Among the central banks, the Fed has a unique structure due to a mix of private and public elements which serve the interests of the general public and the interests of private bankers.

Just as in the case of the ECB communications, Fed's communications contain useful information about future monetary policy (Hayo & Neuenkirch, 2010). Studies have found that the Fed communications moved financial markets in the intended direction and that these communications are even more relevant to the markets during financial turmoil (Hayo et al., 2014).

Although the two organizations serve similar roles in their respective economies, one study suggests that the ECB has moved its policy rate much less frequently than the Fed and that their interest rate behavior was rather different (Sahuc & Smets, 2008). However, both organizations have had to face the effects of the recent financial crisis.

Many studies investigating central bank communications have at least three findings in common: (1) central bank communications regarding economic projections and monetary policy developments have significant impacts on the financial markets, (2) these effects are even greater when the communication channel is more formal, and (3) the more prominent the position of the communicator, the stronger the financial market reaction (Connolly & Kohler, 2004; Kohn & Sack, 2003; Chirinko & Curran, 2006; Reinhart & Sack, 2006). As such, the communications of central banks are of great importance in times of market volatility and financial crises. In this paper, we apply semi-automated text analysis (semantic network analysis) to unravel over time changes in formal discourses and thus contribute to insight in their flexibility.

The remainder of this paper is organized as follows. In the following section, we describe our data sets and the approach we are employing. Chapter 3 presents the results of our analyses, and Chapter 4 summarizes our overall findings and discusses the benefits and limitations of our approach.

3.2 Data and Methods

For this study, a total of 3010 press releases issued by the ECB (825 documents) and the Fed (2185 documents) between January 2006 and December 2013 have been collected from their web archives. For each organization, the collected press releases have been divided in four time periods each spanning a period of two years. The aggregation of data in these four periods was motivated by our aim of assessing the discursive shifts of the ECB and Fed at the different phases of the financial crisis. The first sub-sample covers the period just prior to the financial crisis: January 2006 until December 2007. We label this period *pre-crisis*. The second sub-sample (*crisis*) includes press releases issued between January 2008 and December 2009. The third sub-sample contains press releases issued between January 2010 and December 2011, and represents the *post-crisis* period. Lastly, the fourth sub-sample includes press releases issued between January 2012 and December 2013, further referred to as *recovery*. Table 3.1 presents a summary of the number of documents collected (D), the word counts (W), and the sentence counts (S) for each organization and each period as well as summary statistics: sum of the counts (Σ) and the words and sentences per document ratio averaged across the periods (i.e., $\mu(W/D)$ and $\mu(S/D)$).

Table 3.1: Press releases collected

	ECB			**Fed**		
	Documents	**Words**	**Sentences**	**Documents**	**Words**	**Sentences**
Pre-crisis	184	59544	2133	386	80844	5104
Crisis	203	66349	2286	786	140959	7464
Post-crisis	210	70914	2528	635	104762	4869
Recovery	228	88323	3274	425	93662	4236
$\sum$	825	285130	10221	2232	420227	21673
$\mu(W/D)$		343.88	12.32		193.53	10.09

We note some differences between the corpora of the Fed and ECB. In particular, the Fed data contains a higher volume of text (i.e., total word count). However, the Fed's much higher document count translates to its press releases being smaller in word size and sentence count than those of the ECB; the difference in word counts per document between the two institutions is

significant ($p < 0.001$ for a t-test and $p < 0.05$ for a Kolmogorov-Smirnov test). This points to a striking difference in the communication styles between the U.S. and European institutions and raises additional, interesting questions and hypotheses about implications of these statistics which we leave for future exploration.[1] While the higher volume of the Fed data is expected to confer larger semantic networks of prominent concepts, we expect the relationship to be more nuanced than a mere function of word count given the complex structure of language.

In this paper, we use semi-automated coding of concepts to be included in the semantic networks (Carley & Kaufer, 1993; Danowski, 2012; Diesner, 2013). Each of the data samples *(pre-crisis, crisis, post-crisis,* and *recovery)* has been pre-processed with AutoMap (Carley et al., 2013a). The pre-processing removed all the noise words (e.g., numbers as words, verbs, single letters etc.) in the data and prepared it for the generation of semantic networks. A total of eight semantic networks (one for each organization for the four phases of the crisis) were generated using the software (see Table 3.2 for the descriptive statistics of each network). The generation of networks is based on Carley's approach to coding texts as cognitive maps (Carley & Palmquist, 1992) and Danowski's approach to proximity analysis (Danowski, 1993).

Semantic networks translate text into networks of concepts and the links between them, in which a concept may be a word or a phrase (i.e., n-gram) (Popping, 2003). The links between concepts are based on co-occurence. For example, if two words co-occur within the specified window size, a link (or semantic network edge) will be formed. The window size determines the range of text words in which connections will be made between words within the window (Diesner, 2012b). The window size used for this study was one sentence. The value of strength for each link is determined by the frequency of co-occurrence (Wasserman & Faust, 1994).

As mentioned above, a concept in our semantic networks can be a single word or an n-gram. N-grams are coded by replacing the spaces between words

[1] For example, do more frequent yet briefer press releases (such as the Fed's) affect public and organizational perception differently?

with an underscore (Carley et al., 2013a). An example of such a conversion is 'interest rate' being coded as 'interest_rate'. This procedure helps us identify the most common multi-word expressions used in text documents. Thus, when we refer to key concepts, we refer to single words as well as n-grams.

Table 3.2: Descriptive statistics of the ECB and FED semantic networks*

	ECB			
Measure	*Pre-***	*Cri-*	*Post-*	*Rec-*
Node count (n)	580	627	648	755
Sum of links	113389	131326	159091	199352
Density	0.205	0.201	0.216	0.191
	Fed			
Measure	*Pre-***	*Cri-*	*Post-*	*Rec-*
Node count (n)	827	1009	886	860
Sum of links	178782	345885	250854	260873
Density	0.147	0.155	0.152	0.150

*Each network is undirected, symmetric and valued; Only nodes with frequencies ≥ 10 have been included in the networks.
**Pre-Crisis (*Pre-*), Crisis (*Cri-*), Post-Crisis (*Post-*), Recovery (*Rec-*).

In Table 3.2, some summary statistics of the generated semantic networks are reported. The sum of links is the sum of all the values of the weighted edges/links, while the (unweighted) density expresses the proportion of non-zero (or non-null) edges to the count of possible edges in the undirected semantic network (i.e., $n/(n(n-1))$). These descriptive statistics show that even after employing a frequency threshold (≥ 10) the resulting networks are complex and relatively dense with high link counts.[2] Again, the statistics reveal some striking differences between the two institutions this time in their semantic networks. The node counts and edge weight sums increase monotonically over the periods (i.e., time) for the ECB, while these exhibit distinct peaks at the *crisis* period for the Fed. This finding is congruent with the knowledge that the United States economy was both the source of and most affected by the

[2]Most empirical social network data exhibit much lower densities.

financial crisis. That is, we expect more activity in the Fed press releases than in the ECB's during the *crisis* period. In fact, the ratio of the node and edge weights during the *crisis* against the averages across the periods are below 1.0 for ECB and above 1.0 for Fed.

The higher link weights yet lower, unweighted density of the Fed (compared to those of the ECB) appear contradictory. However, this finding indicates a higher repetition of concept co-occurrences but proportionally fewer unique dyads (i.e., pairs of linked concepts) in the Fed corpora. Whether this is due to differences in language use or policy aims between the institutions remains a research question worth exploring.

Interestingly, we find that the node counts are best predicted (statistically) by the sentence counts in Table 3.1 more so than the document or word counts which leads to a hypothesis about the structure of the language in our corpora: that unique concepts arise out of distinct sentences rather than other textual boundaries. The adjusted-R^2 for nodes predicted separately by documents, words, and sentences are respectively, 0.85, 0.87, and 0.91.

The combination of the complexity of these networks and the formal character of the documents from which they have been extracted poses a challenge for the analyst. To overcome this challenge, we propose using a structural space approach that considers total degree centrality and betweenness centrality of concepts in semantic networks, concurrently.

3.2.1 Centrality in Networks

Even after decades of social network research, the current thinking about network centrality is still mostly defined by the work of Freeman (1979) and Bonacich (1987). In 1977, Freeman developed a set of centrality measures based on betweenness (Freeman, 1977). In a follow-up article two years later, Freeman (1979) elaborates on three concepts of centrality in a social network, which have since been further developed into degree centrality, closeness centrality, and betweenness centrality. The fourth commonly used measure, eigenvector centrality, was popularized by Bonacich (1987) and is based on previous graph-theoretical research (Wei, 1952; Berge, 1958). We now define and briefly

elaborate on (total) degree centrality and betweenness centrality, the focal measures in this paper. Increasing interest in network centrality has been spurred by the growth of research on large scale networks such as the hyperlinks of the worldwide web (Barabási & Albert, 1999) and other online networks.

Total Degree Centrality

Total degree centrality is one of the most commonly used centrality measures in social network analysis (Henderson et al., 1998). The degree centrality of a node in a network reflects the number of other nodes incident to the focal node (Freeman, 1979) (or, in the case of weighted networks, the sum of the weights of all the incident links), and thus measures the involvement of a node in its local network. Nodes with low total degree centrality are potentially more peripheral to the network (Iacobucci et al., 1996) unless they are connected to popular others. In semantic networks, total degree centrality may represent the 'importance' of a concept or its key concept status. A key concept with high degree centrality is able to activate many other key concepts; thus, it functions as a hot topic's central key concept (Diesner, 2012b). Using only the local structure to calculate the degree centrality of a node, this measure alone does not take into consideration the position of the concept within the global structure of the network. Still, the distribution of network statistics, such as node centrality and even node and edge counts, can be an indication of global network properties (Barabási & Albert, 1999; Lee & Pfeffer, 2015b). In this paper, we employ the weighted version of total degree centrality.

Concept frequency is arguably a more parsimonious metric than popularity (i.e., total degree centrality). However, since we are interested in the semantic structure, focusing on popularity over frequency is appropriate. While frequency alone would reveal some of the key concepts, the semantic linkages (such structures surrounding connective concepts) would remain obscured. Still, a naïve Pearson correlation between the two metrics (weighted degree centrality and frequency) is high (ECB $r \approx 0.77$; Fed $r \approx 0.92$) for both organizations and across all of the periods. A closer inspection reveals significant variance in their relationship and that a log-linear association emerges but only for those

concepts having higher than average frequency and degree. In fact, given the heavy positive skewness of weighted degree centrality distributions (i.e., many low values), we inspect a log+1 transformation of degree centrality which reveals more comparable correlations: ≈ 0.71 and ≈ 0.77 for the ECB and Fed respectively. The ranked correlations (i.e., Kendall's τ) are even lower (≈ 0.55 and ≈ 0.70, respectively), thus using concept frequency in lieu of weighted degree centrality would reveal only a semi-structural space, having non-trivially different interpretations than the space we employ in this paper.

Betweenness Centrality

Betweenness centrality (C_{B}) is the sum of the proportions of the shortest paths a node lies on for every pair of nodes (out of all shortest paths for each pair). The formulation for unweighted betweenness is:

$$C_{\mathrm{B}}(i) = \sum_{s \neq i \neq t} \frac{\sigma_{s,t}(i)}{\sigma_{s,t}}$$

where $\sigma_{s,t}$ indicates the count of shortest paths between nodes s and t. For weighted betweenness, the shortest paths are computed using the inverse of the edge weight since heavier edges should warrant greater flow (and hence higher betweenness). We employ this inversion as most of the edges between concepts are valued (i.e., weighted). More broadly, betweenness centrality represents the frequency with which a particular node is on the geodesic path between any other two nodes in the network (Danowski, 2012). As such, betweenness centrality captures one aspect of a node's position in the graph, thus taking into account the global structure of the network. The effectiveness of this measure is limited by the extent of connectedness in the network. That is, if a network contains many disconnected components, betweenness centrality becomes less of a global measure. However, this is not an issue with our semantic networks because each constitutes a single component. The betweenness centrality of a concept within a semantic network is a direct indicator of its influence (Hill & Carley, 1999; Hooper et al., 2012; Wasserman & Faust, 1994). A key concept

with high betweenness centrality controls access to other key concepts in the network (Brandes & Corman, 2003; Grebitus & Bruhn, 2008; Henderson et al., 1998; Hulst, 2008), and thus serves as a gatekeeper between different domains (Gloor & Krauss, 2009). For semantic networks, it is presumed that a node with high betweenness centrality has a higher likelihood to become activated or activate connections across domains (or topic communities).

3.2.2 Structural Roles

By combining popularity and connectivity of concepts in semantic networks, we expect to capture emerging topics within the texts and subtle shifts in formal discourse through the classification of nodes according to their structural roles.

Table 3.3: Concepts with the highest total degree centrality in each ECB network

Pre-crisis		***Crisis***		***Post-crisis***		***Recovery***	
Concept	**Deg.**	**Concept**	**Deg.**	**Concept**	**Deg.**	**Concept**	**Deg.**
ecb	8200	ecb	9800	ecb	11300	ecb	12700
european	5600	market	6200	financial	6900	bank	10400
eu	5200	central bank	6200	euro area	6300	european	9400
system	4900	eurosystem	5900	market	6200	financial	8200
eurosystem	4700	eu	5500	bank	6000	include	7500
euro area	4600	euro area	5200	include	5900	market	7400
central bank	4500	operate	4800	system	5800	monetary	7300
include	4500	national	4500	eu	5800	eu	7200
market	4300	include	4200	central bank	5300	central bank	6900
operate	4000	increase	4100	economic	5200	euro area	6900

Note: concepts are color-coded to highlight their similarity.

Because the discourses of the Fed and the ECB are highly formal and the resulting networks are complex, looking separately at 1) the top most frequent concepts, 2) the top most central concepts (see example in Table 3.3), or 3) concepts having the highest betweenness centrality (see example in Table 3.4) will not be very informative.[3] These top concepts are very similar across the four periods for both organizations and constitute the core issues under discussion.

[3]For conciseness, the above tables display only one measure per organizational data source; however, the problem of repetitive highly ranked concepts persists for betweenness centralities of the ECB networks and degree centralities of the Fed networks.

Table 3.4: Concepts with the highest total betweenness centrality in each Fed network

Pre-crisis		***Crisis***		***Post-crisis***		***Recovery***	
Concept	**Deg.**	**Concept**	**Deg.**	**Concept**	**Deg.**	**Concept**	**Deg.**
bank	77942	fed	250152	bank	4767507	bank	106529
fed	25589	bank	232961	fed	2242646	capital	89368
rate	24073	loan	96151	credit	1273791	fed	89311
financial	20868	credit	73132	financial	1244069	financial	59273
agency	12662	financial	41794	loan	675665	committee	48583
president	10151	capital	33451	agency	552549	credit	41855
credit	9809	market	27397	institution	452635	agency	39651
institution	8804	agency	26388	rate	422863	loan	32514
board	6860	institution	21378	consumer	422569	institution	17825
loan	6730	rate	15927	mortgage	364074	rate	16707

Note: concepts are color-coded to highlight their similarity.

In order to explore both in-depth and orthogonally informative dimensions of the ECB and the Fed discourses, we characterize these discourses using two distinct measures, building on the manner in which popular and connecting concepts play different roles in the structure and dynamics of semantic networks. Combining the popularity (i.e., total degree centrality) and connectivity (i.e., betweenness centrality) dimensions allows for the identification of four structural roles. This combination positions the concepts within this structural role space.

3.2.3 Correlations Among Centrality and Text Analytic Measures

While high correlations among measures can be used to identify obviously important network nodes (e.g., Huang et al., 2014), our structural space approach would be limited if degree and betweenness centralities in semantic networks (including our own) were highly correlated and it would fail to identify subtle roles of network nodes.

Research on correlations between degree and centrality measures in empirical networks and well-known topologies is mixed. Li et al. (2015) find moderate Pearson correlations between degree and betweenness centralities (average

≈0.6) in real-world networks and significant overlap in top node memberships (for nodes ranked by each degree and betweenness centrality) in Erdős-Rényi (ER) networks (Erdős & Rényi, 1959) and scale-free (SF) networks (Barabási & Albert, 1999) (> 0.95). However, this finding is limited as their simulated networks were of one size and specific densities. Also, Pearson correlations can be misleading as they can be heavily influenced by outliers (e.g., a single node having extremely high betweenness and degree centralities). Lee & Pfeffer (2015a) show that while ER, Watts-Strogatz (WS) small world networks (Watts & Strogatz, 1998), and SF networks display high average Pearson correlations (for degree and betweenness centralities) across a range of network sizes and densities, these correlations can be poor for networks with low densities or high centralization; these characteristics are found in many empirical networks.

Thus, while correlations between centrality metrics require further investigation, it is fair to argue that highly correlated degree and betweenness centrality metrics would have major implications on their joint interpretation. The semantic networks explored in this particular study exhibit only moderate correlations between degree and betweenness centrality (Kendall's $\tau \leq 0.51$) as shown in column DB of Table 3.5.[4] Thus, these measures are distinct, yet conceptually related, and can be jointly used to reveal distinct structural roles.

Table 3.5: Average correlations among text-analytic and semantic network measures

	FT	*DF*	*BF*	*DT*	*BT*	*DB*
ECB	0.704	0.553	0.496	0.383	0.391	0.517
	(0.059)	(0.025)	(0.016)	(0.031)	(0.021)	(0.045)
Fed	0.867	0.709	0.472	0.679	0.477	0.422
	(0.024)	(0.023)	(0.026)	(0.033)	(0.024)	(0.009)

These correlations are averaged across the four periods.
The standard deviations appears just below the averages.
F = frequency, T = tf-idf, D = degree centrality, and
B = betweenness centrality

[4]Kendall's τ ranked correlation is preferred here in order to mitigate the impact of extreme points (outliers).

One might consider the use of direct text analytic metrics such frequency of words (or in our case, frequency of concepts) and the tf-idf score (term frequency-inverse document frequency).[5] In Table 3.5, we also report average Kendall's τ rank correlation coefficients among concept frequency, tf-idf, and the two centrality measures; the correlations are averaged across the four periods for each of the data sets. The correlations between either text analytic measure and the centrality measures are modest, indicating that the network measures are not substitutable (with text analytic measures). In fact, the two text analytic measures exhibit the highest correlations (within each data set) and are more substitutable for one another than for the network measures. So, while a direct analysis through frequency counts or tf-idf values of the documents issued by the ECB and the Fed could potentially reveal noteworthy findings, we argue that semantic network analysis adds a new dimension to this text analysis by reveal deeper insights into discursive structures.

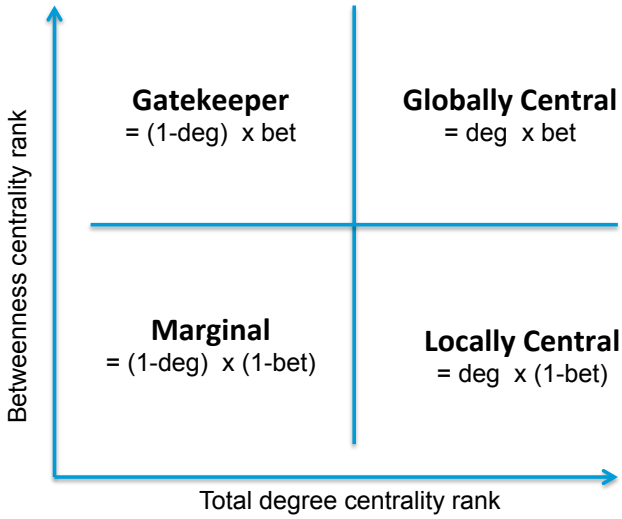


Figure 3.1: The four quadrants of the structural space

In order to connect the concepts with these structural roles, each concept

[5]Tf-idf reflects the importance of a word in a collection or corpus, and it is often used as a weighting factor in information retrieval and text mining. The tf-idf value increases proportionally to the number of times a word appears in the document, but is offset by the frequency of the word in the corpus.

in the network has been ranked based on its total degree centrality (C_{D}) and its betweenness centrality (C_{B}). For these rankings, we first consider the *set* of unique, unordered values X derived from some vector (or bag) of measures X_b. The order set of X is then:

$$X_{\text{ordered}} = \{x_1, \ldots, x_n | x_i \in X; n = |X|; x_1 < \cdots < x_n\}$$

where n is the number of unique measure values. We also define an index set J such that $x_j^{\text{ordered}} | j \in J$ is the j^{th} element of ordered set X_{ordered}. We now have a bijection $X_{\text{ordered}} \rightarrow X_{\text{rank}}$. For degree centrality, we replace X_b with the degree centrality measures C_{D} and obtain X_{ordered} which contains the unique, ordered degree centralities. For each node i, $C_{\mathrm{D}}(i)$, we obtain the degree centrality rank $C_{\mathrm{D}}^{\text{rank}}(i)$:

$$C_{\mathrm{D}}^{\text{rank}}(i) = 100 \cdot \tfrac{j}{n} | \left(x_j^{\text{ordered}} = C_{\mathrm{D}}(i) \right).$$

The rankings are rescaled (hence normalized) to the [0,100] interval so that we can easily compare rankings across semantic networks. The rankings for betweenness centrality are obtained in a similar fashion (i.e., using C_{B} for X_b).

In simpler terms, we rank the total degree centrality and betweenness centrality scores for the concepts from each time period network into a normalized range between 0 and 100. Ranking was employed because a) the networks are of different sizes and densities (hence, we want to be able compare across time periods) and b) using the raw centrality scores produces less compelling and readable visualizations due high skewness of the centrality distributions.[6]

We highlight the four structural roles as four quadrants of the four structural space in Figure 3.1. The Globally Central (GC) role includes concepts with high degree centrality and high betweenness centrality [high values of $C_{\mathrm{D}}^{\text{rank}} \times C_{\mathrm{B}}^{\text{rank}}$], where $C_{\mathrm{D}}^{\text{rank}}$ and $C_{\mathrm{B}}^{\text{rank}}$ are the normalized rankings of total degree centrality and betweenness centrality. A GC concept is a central key concept of a hot topic

[6] Alternatively, we could have employed normalized centrality scores. However, these exhibit the same skewness and still require transformation. Our approach is mathematically similar to using ranks of normalized scores.

because not only is it highly connected to other concepts (i.e., popular) but it also serves as a bridge between different parts of the network (i.e., connective).

The Locally Central (LC) role contains concepts with high degree centrality and low betweenness centrality [high values of $C_{\mathrm{D}}^{\mathrm{rank}} \times (100 - C_{\mathrm{B}}^{\mathrm{rank}})$]. LC concepts are very popular concepts that do not have a strongly connective role. In other words, an LC concept is the central key concept of a local hot topic because it is highly connected but does not serve as a bridge in the network.

The Gatekeeper (G) role incorporates concepts with low degree centrality and high betweenness centrality [high values of $(100 - C_{\mathrm{D}}^{\mathrm{rank}}) \times C_{\mathrm{B}}^{\mathrm{rank}}$]. These types of concepts are highly connective concepts that aren't very popular. A G concept is influential in the network because although it is not highly connected, it acts as a bridge in the network, linking different themes or topics. Such a concept can mark the emergence of merging themes.

Lastly, the Marginal (M) role includes concepts with low degree centrality and low betweenness centrality [high values of $(100 - C_{\mathrm{D}}^{\mathrm{rank}}) \times (100 - C_{\mathrm{B}}^{\mathrm{rank}})$]. M concepts are neither popular, nor connective but retain the potential to become emergent concepts and transition into other roles.

Figure 3.2 illustrates a scatterplot layout example for the four structural roles described above using empirical centrality ranks from one of our subsamples. The darker the red shade of the nodes, the higher embedded these nodes are in the region of the specific structural role.

Alternatively, the structural role scores could have been computed by simply adding the total degree centrality and betweenness centrality score components (i.e., the multiplicands). However, this addition produces inaccurate role mappings. For example, Globally Central (GC) concepts become classified also as Gatekeepers (G) due to their high betweenness centrality irrespective of their high total degree centrality. Similarly, Marginal (M) concepts can appear as Gatekeepers due to their extremely low total degree centrality. We find multiplication of the role components to parsimoniously produce distinct role assignments.

After identifying the four roles and the nodes that belong to each role, additional visual dimensions can be added by sizing, shaping, and/or coloring

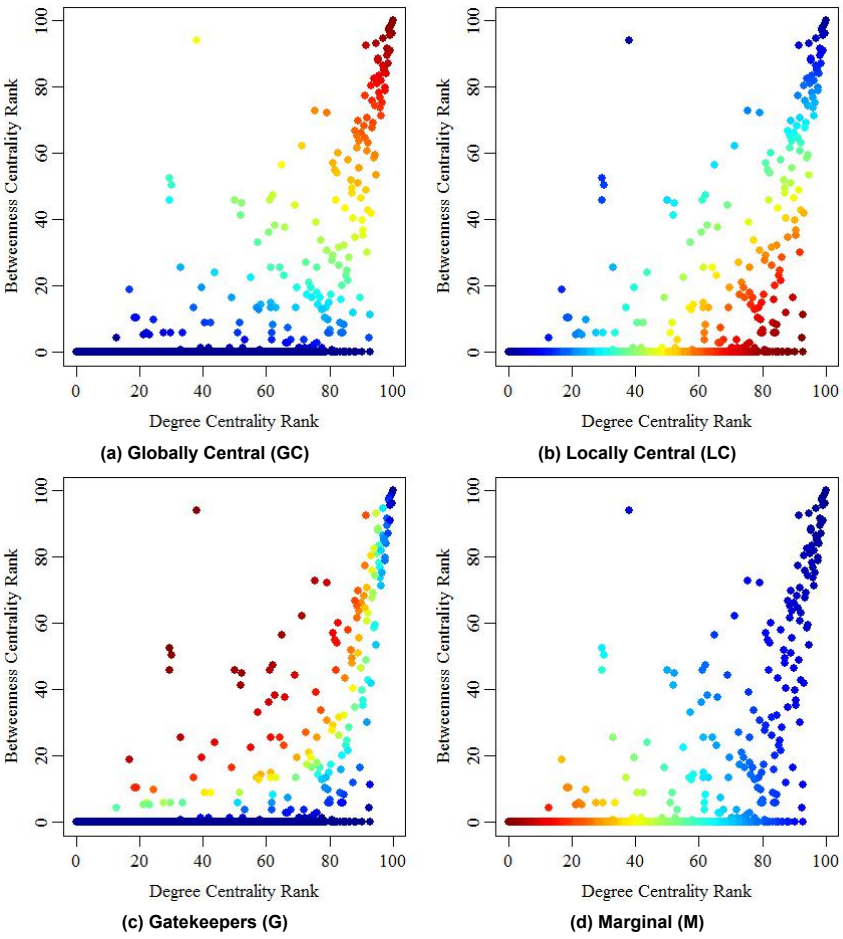


Figure 3.2: Example of structural roles

the nodes based on other measures. This approach provides other ways to explore each role individually or the structural space as a whole in terms of distinct or subtle patterns. For example, other network metrics may expose other types of structural roles.

While, in semantic networks, total degree centrality represents the popularity of a concept and betweenness centrality represents the links between two different thematic areas, the combination of these two measures has the potential to uncover more subtle structural properties of concepts and thus a set of changes in discourse over time.

3.3 Results

3.3.1 Structural Roles

For the purposes of this article, we will focus our structural space analysis on a set of selected key concepts within the discourses of the ECB and the Fed, and the dynamics of these concepts across four stages of the financial crisis. As presented in Table 3.6, the first set of concepts have been selected because they reflect the main objectives of the ECB (European Union, 1997) and the Fed (Federal Reserve System, 2005). The second set of concepts has been included in the analysis to represent crisis-oriented terminology. While six of the crisis-oriented concepts are similar for both organizations, the other four represent financial instruments employed by the ECB and the Fed during the different stages of the crisis. Below, we describe the acronyms in Table 3.6:

- *Longer Term Refinancing Operations (LTRO)* provide additional, longer-term refinancing to the financial sector;
- *Main Refinancing Operations (MRO)* serve to drive short-term interest rates, to manage the liquidity situation, and to signal the monetary policy stance in the euro area;
- *Term Asset-Backed Securities Loan Facility (TALF)* was created in 2008 to accommodate the credit needs of consumers and small businesses by

Table 3.6: Selected key concepts

ECB	**Fed**
Main objectives	
Interest rate	Interest rate
Stability	Stable price
Growth	Employment
	Unemployment
Crisis oriented concepts	
Crisis	Crisis
Debt	Debt
Inflation	Inflation
Lend	Lend
Refinancing	Refinancing
Risk	Risk
LTRO	TALF
MRO	TSLF

facilitating the issuance of asset-backed securities collateralized by loans such as student loans, auto loans, credit card loans, commercial mortgages, and loans guaranteed by the Small Business Administration;

- *Term Securities Lending Facility (TSLF)* is a special lending facility set up by the Federal Reserve in 2008 to loan Treasury securities to primary dealers for 28 days.

We will begin by reporting our findings for each organization separately by highlighting the observed variations in discourse as it develops across the different phases of the financial crisis.

First, in Figures 3.3 to 3.10, we plot the structural space positions of all the ECB concepts (with a raw frequency ≥ 10) in the semantic networks of the ECB later followed by those of the Fed. We highlight the selected key concepts by labeling them in the structural space and we add another visual dimension by coloring the nodes based on their raw frequencies of occurrence in the text data. The color spectrum ranges from dark **blue** (low frequency) to dark **red** (high frequency).

We also display edges among the focal concepts; that is, the subgraph induced by the node set comprising these concepts. The edges are visually

weighted and represent the count of co-occurrences (within the one sentence window) between these focal nodes. We note that these edges do not represent the total activity of the focal concepts but just the activity among themselves (for presentation purposes). Finally, at the top of each graph, we report 1) the number of considered nodes n (i.e., having a raw frequency ≥ 10); 2) the number of distinct edges of the subgraph of key concepts $|E|$; and 3) the sum of the edge weights of that subgraph Σw.

ECB

Figure 3.3 on the next page shows that even in the *pre-crisis* period, before the end of 2007, crisis-oriented key concepts are present in the ECB discourse, some of them having relatively high total degree centrality (i.e., 'MRO', and 'risk') and being connected to the main objectives of the ECB. The globally central (GC) position of 'risk' as well as the "on-the-fence" position of 'MRO' (which borders the locally central (LC) and the marginal (M) quadrants) could indicate that some of the ECB's attention was focused on the emerging financial crisis before the end of 2007. We also observe that unlike 'interest_rate' and 'stability', which are highly ranked GC concepts, 'growth' (one of the ECB's main objectives) is a highly ranked LC concept. This indicates that, during the *pre-crisis* period, 'growth' was a popular concept but not a very connective one.

In the *crisis* period (see Figure 3.4 on page 70), most of the crisis-related key concepts are becoming more prominent. Concepts such as 'inflation', 'loan', 'MRO', and 'debt' are ranked higher in total degree centrality and in betweenness centrality than in the previous period, suggesting they became more central and connective of different domains in the discourse of the ECB during the *crisis* period. At the same time, the betweenness and degree centralities of 'interest rate' and 'stability' noticeably decrease, suggesting once again that the main objectives of the ECB lose rhetorical ground against the full-blown financial crisis. The concept 'risk' is even higher ranked in the GC category during the *crisis* becoming one of the 'hottest' topics of the ECB discourse. We also observe the emergence of 'LTRO', a concept that was not present

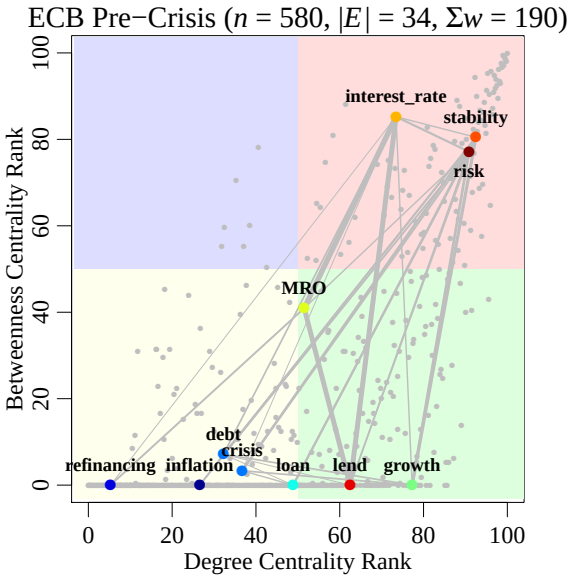


Figure 3.3: Structural space of the ECB *Pre-crisis* semantic network

in the *pre-crisis* period. 'LTRO' enters the discourse of the ECB as a very highly ranked G concept, indicating that it connected disparate topics during the *crisis*. The link weights also increase in the *crisis* period, highlighting the increased co-activity among these concepts in the ECB press releases during this period. An interesting finding is that 'crisis' remains a marginal concept during the *crisis* period. The similar marginal position of 'crisis' during the *pre-crisis* as well as the *crisis* period raises questions regarding the discursive practices employed by the ECB. Had the ECB avoided the use of such concepts to avoid creating panic among stakeholders? Or had the ECB denied or ignored the existence of the crisis? Or did they perhaps use less value laden concepts?

Figure 3.5 on page 71 plots the structural space of the *post-crisis* semantic network, showing all the main objectives of the ECB in the GC quadrant. While in the *pre-crisis* only two of the main objectives were in GC positions and in the *crisis* period the centrality of these two concepts decreased, in the *post-crisis* period all the three main objectives return to being globally central, GC. At the same time we observe that the betweenness centrality of 'crisis' increases, while the betweenness of 'debt' and 'inflation' decreases significantly.

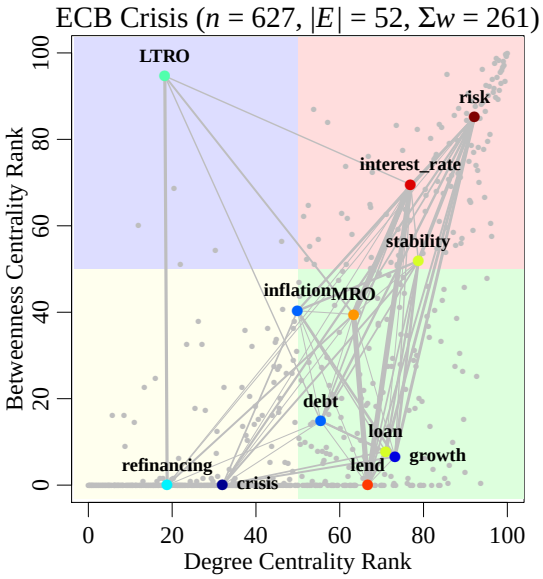


Figure 3.4: Structural space of the ECB Crisis semantic network

These positional changes suggest that the shift in the ECB discourse could be at least partly explained by their struggle to deal with the aftermath of the crisis, while at the same time refocusing on their core objectives. The link weights show increased activity for 'LTRO', 'MRO', and 'refinancing', lending further evidence to the ECB's resumed focus on the aftermath of the crisis.

Figure 3.6 plots the structural space of the *recovery* semantic network, revealing significant changes in the discourse of the ECB beyond the crisis. Compared to the previous period, 'interest_rate' is now an LC concept. This concept, signifying one of the main objectives of the ECB, maintained a GC position in all the three previous periods analyzed. While 'LTRO' suffers a drastic decrease in betweenness centrality (becoming an M concept), 'MRO' and 'refinancing' become G concepts. We also note the positional change of 'loan', moving from the LC quadrant to the GC quadrant. Based on all these structural changes, we argue that the *recovery* period exhibits a clear shift towards a discourse directed at dealing with the aftermath of the crisis. By assessing the width of the links, we see that the positional changes described above are also reflected in the co-occurence levels. While 'MRO' and 'refinanc-

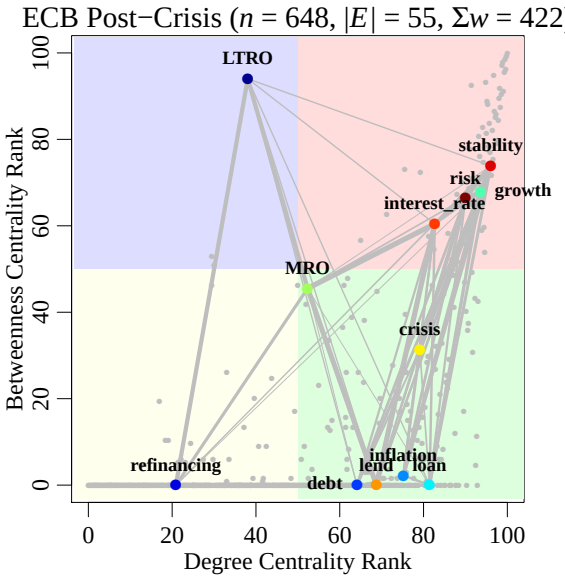


Figure 3.5: Structural space of the ECB *Post-crisis* semantic network

ing' show increased activity, 'LTRO' co-occurs less often with the other key concepts.

As for the graph-level statistics, we observe that the count of nodes (i.e., n, the count of non-infrequent concepts) increases almost exactly linearly to the word counts of non-noise words in the collected documents for each period. These word counts are 28155, 30991, 33538, and 42892 (from *pre-crisis* through *recovery*) and the correlation is 0.997. However, the activity in the focal concept subgraph does not follow suit. Specifically, the edge count increases initially and then stabilizes at ∼53, and the sum of edge weights peaks at *post-crisis* and then decreases. We surmise that the ECB discourse becomes expansive with the inclusion of additional topics (not identified in this paper). Hence, a naïve analysis using simple, relative frequencies of these key concepts would only diminish their importance. On the other hand, our structural role analysis reveals that some of the concepts associated with ECB's objectives (here, 'stability' and 'growth') in fact remain prominent.

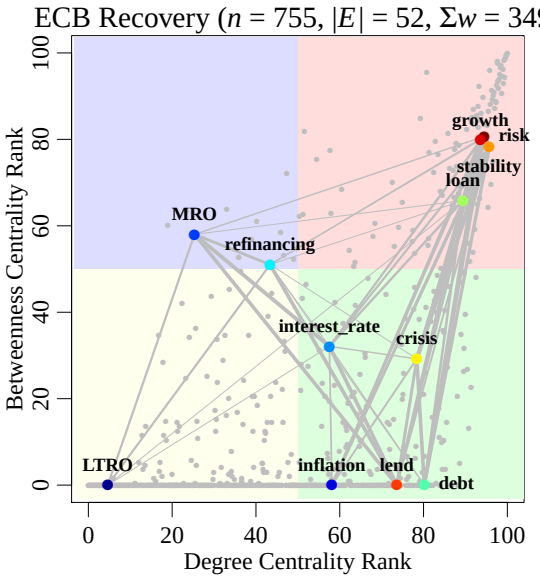


Figure 3.6: Structural space of the ECB *Recovery* semantic network

FED

In Figure 3.7 on the following page, we see that, in the *pre-crisis* period for the Fed data, eight of the 13 selected key concepts are present. Interestingly, two of Fed's main objectives are marginal concepts, thus having low popularity and low connectivity. The third concept representing Fed's main objectives, 'stable_price', is a popular concept (LC concept) during the *pre-crisis* but it does not serve as a bridge between other topics discussed. The three highest ranked GC concepts in this period are 'inflation', 'loan', and 'risk', concepts that we categorized as crisis-oriented terminology. With 'risk' being the highest ranked concept for this period and with the M and LC positions of the three main objectives of the Fed, we can conclude that the Fed's discourse is not focused on their main objectives but on what appears to be the initial signs of the financial crisis.

In the *crisis* period (see Figure 3.8 on page 74), 12 of the 13 selected key concepts are present. Similar to the previous period analyzed, not all of the main objectives of the Fed are present, and while 'employment' maintains its marginal position, 'stable_price' moves from the LC quadrant into the GC

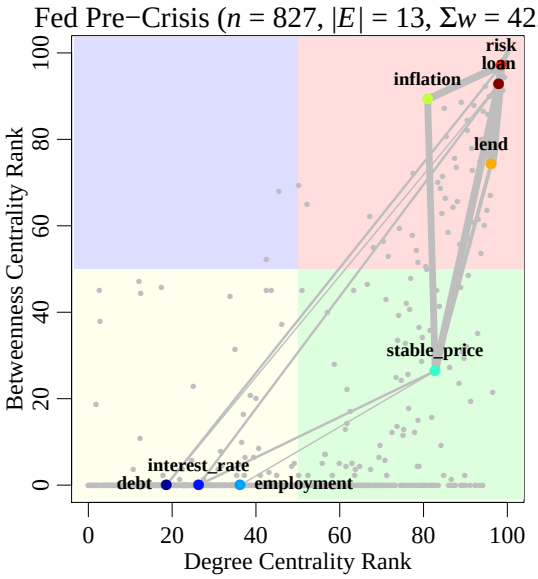


Figure 3.7: Structural space of the Fed *Pre-crisis* semantic network

quadrant, becoming a more popular and a more connective concept during this period. Unlike the *pre-crisis* period which had 'risk' as the top GC concept, in the *crisis* period the highest ranked GC concept is 'loan' (concept that was already a GC concept in the *pre-crisis* period). We also observe that 'TALF' (Term Asset-Backed Securities Loan Facility) and 'TSLF' (Term Securities Lending Facility) enter the discourse of the Fed, with 'TALF' being a highly popular and connective concept, while 'TSLF' is mostly a popular concept. The *crisis* period clearly reveals the discourse of the Fed focusing on the events of the financial crisis, while at the same time being focused on maintaining price stability. Just as in the *pre-crisis* period, 'unemployment', the fourth main objective of the Fed, is not present. Based on the node count (n) of the *crisis* and the *pre-crisis* periods, we can conclude that the Fed's discourse expands in the *crisis* period. In other words, the Fed has been using more unique concepts than in the previous period. This period also exhibits the most subgraph activity ($\sum w$), indicating the increased co-ocurrence of these concepts.

Figure 3.9 on page 75 depicts the *post-crisis* period in which 12 of the key

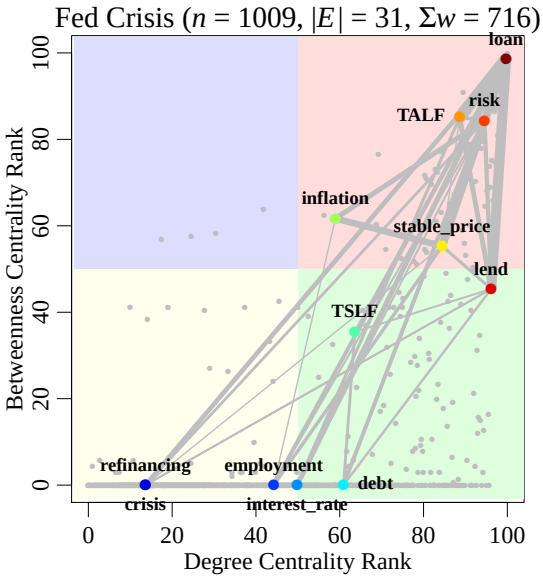


Figure 3.8: Structural space of the Fed *Crisis* semantic network

concepts are present. During this period, we note that all four of the Fed's main objectives are present, with 'employment' shifting from its M position to a GC position. Also, while 'stable_price' becomes an even more popular concept, 'unemployment' enters the discourse of the Fed directly as a popular concept (LC). It is also worth mentioning the new position of 'crisis', an emerging concept in the *crisis* period, which has now become a connective concept (G). The degree centrality and betweenness centrality of 'TALF' decrease, placing this concept into the M quadrant.

The *recovery* period presented in Figure 3.10 on the next page reveals all four main objectives of the Fed as popular but not connective (LC) concepts, while the three prominent GC concepts ('loan', 'risk' and 'inflation') are all crisis-related terms. Based on these findings, we can argue that while the Fed's focus was shifting towards their main objectives, the aftermath of the crisis was still evident in their discourse.

The graph-level statistics for the Fed also reveal an interesting trajectory. While the *crisis* semantic network has the highest node count (n) and the highest sub-graph activity ($\sum w$), the *post-crisis* network shows the highest number

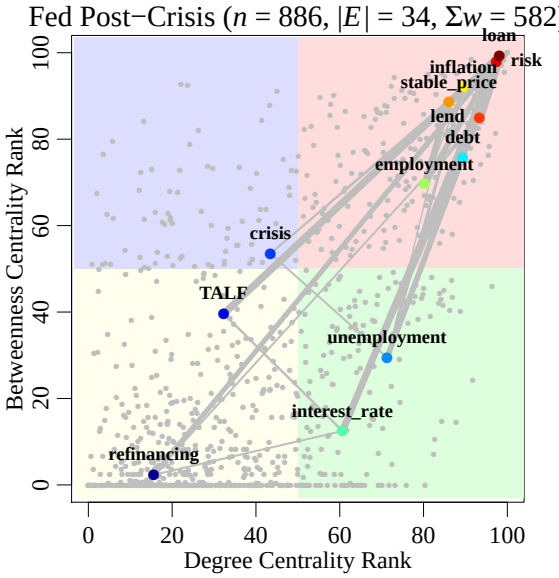


Figure 3.9: Structural space of the Fed *Post-crisis* semantic network

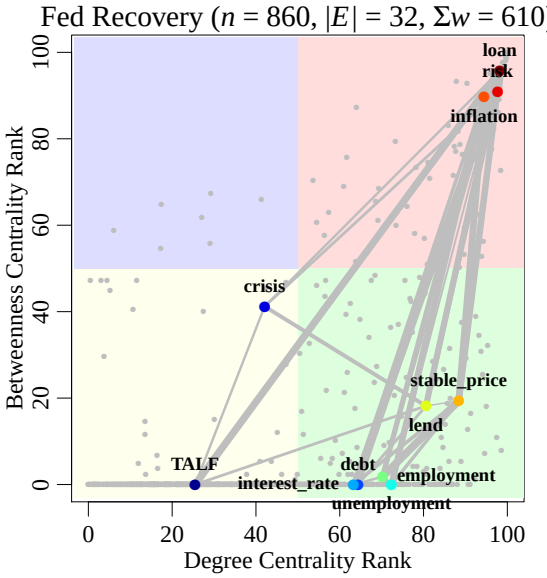


Figure 3.10: Structural space of the Fed *Recovery* semantic network

of links ($|E|$) between the focal concepts. In contrast with our findings for the ECB, where the expansion of their discourse happens gradually over the four periods, the Fed's discourse expands during the *crisis* period but contracts in the following periods. In other words, while the ECB's use of unique concepts grows throughout the four different stages, the Fed's use of unique words expands rapidly between the *pre-crisis* and the *crisis* and decreases thereafter. This finding is also supported by the word counts of the pre-processed, cleaned text documents (without noise words) which are 43564, 72919, 53537, and 50286 (from *pre-crisis* through *recovery*). The overall Pearson correlation between word counts and concept-node counts here is also high and significant, $r = 0.85$.

3.3.2 MRQAP

As the last part of our analysis, and in light of the findings above, we performed QAP (Quadratic Assignment Procedure) correlations and multiple regressions (MRQAP). These methods compare one or more networks using edges and their weights as data points while controlling for their dependencies such as autocorrelation within the network structure (Krackhardt, 1987). These methods have been widely used in social network research (Dekker et al., 2003, 2007; Kilduff & Krackhardt, 1994) and also applied to research in knowledge and semantic networks (Corman et al., 2002; Broekel & Boschma, 2012; Xiang et al., 2009). This type of analysis is appropriate for our networks because we are using valued data, and we can characterize each of the four periods as a function of its previous periods for each of the two organizations.

MRQAP is essentially multiple regression predicting the edge weights (including non-edges) of one network from one or more other networks. Typically, the networks (both the dependent and independent ones) are transformed into adjacency matrices so they contain edge weights as well as zeros for non-edges. These matrices are then elongated into single vectors such that the positions in each vector correspond to the same matrix cell positions. These vectors then serve as the dependent and independent variables in the multiple regression, and an estimated regression coefficient (one for each predictor network) indi-

Table 3.7: QAP correlations for the ECB semantic networks

	Pre-crisis	*Crisis*	*Post-crisis*	*Recovery*
Pre-crisis	–	0.758***	0.697***	0.593***
Crisis	–	–	0.758***	0.627***
Post-crisis	–	–	–	0.851***

*** indicates significance at $p < 0.001$

cates the extent to which an independent network's edge (weights) contribute to the corresponding edge weight in the dependent network.

While the regression coefficients from an MRQAP are identical to those of a least squares regression, their significance scores (i.e., p-values) are derived by comparing the estimates against their null distributions obtained from applying the same regression model to a large sample of permutations ($m = 1000$) of the node structure (i.e., node relabelings) thereby controlling for autocorrelation (Krackhardt, 1987). For example, the inherent popularity of certain concepts is controlled for. The permutations of the predictor matrices represent alternative "worlds" in which the structure remains the same but nodes have been repeatedly re-assigned. The same applies to the computation of a QAP Pearson correlation. The networks are first conformed by node count as the networks sizes need to be identical as required by standard regression. For this analysis the semantic networks of each organization have been conformed through intersection of the node sets, which retains only nodes that are common to all the networks.

MRQAP ECB

The Pearson correlations reported in Table 3.7 are moderate to high despite the complexity of the four semantic networks.

Their patterns show what we would nominally expect: proximal time periods bear the most resemblance while those farther apart differ the most. For example, the *pre-crisis* network's correlations diminish with more recent periods. Interestingly, the *post-crisis* and *recovery* periods exhibit more similarity to one another than any other adjacent pairs of periods, suggesting these periods are not as distinct as those other pairs and that *recovery* was likely already

Table 3.8: MRQAP coefficients for the ECB semantic networks

Dependent	Intercept	*Pre-crisis*	*Crisis*	*Post-crisis*
Recovery	0.361***	0.037ˆ	−0.071***	0.968***

Adj-R^2= 0.725; ˆ = $p < 0.1$, *** = $p < 0.001$

underway during the *post-crisis* period. Because the *pre-crisis* and *recovery* networks are the least similar, we can argue that the *recovery* network is a transition phase in the ECB discourse towards a new, possibly hybrid, state and not a resumption of the status quo of the *pre-crisis* period.

We are also interested in the extent to which the first three periods constitute the *recovery* period. Conforming the four periods' networks by the intersection of their common nodes yields (n) 382 nodes per network. In Table 3.8, we show the results of the MRQAP regression for predicting the *recovery* network from the earlier periods' networks. As suggested by correlations in Table 3.7, the *post-crisis* period is the most predictive of *recovery*. Similarly, the *pre-crisis* period adds very little to the *recovery* period; however, the coefficient remains positive indicating a contribution to the similarity. On the other hand, the negative (and significant) coefficient for *crisis*' prediction on *recovery* indicates a slight reversal in the semantic structure from that period. That is, semantic associations of high prominence in *crisis* appear less prominently in *recovery*, controlling for the effects from the other two periods. In other words, the ECB seemed more inclined to focus less on the financial crisis and more on the subsequent recovery, an observation supported by the shifts in the structural roles (Figures 3.6 to 3.9).

In light of the findings of the structural space analysis and the MRQAP, we can conclude that the *recovery* period is a different state in the discourse of the ECB. This new state in their discourse exhibits elements characteristic to the *post-crisis*, moving further away from the *crisis* period yet not reverting to the status quo of the *pre-crisis* period.

Table 3.9: QAP correlations for the Fed semantic networks

	Pre-crisis	*Crisis*	*Post-crisis*	*Recovery*
Pre-crisis	–	0.624***	0.581***	0.499***
Crisis	–	–	0.756***	0.536***
Post-crisis	–	–	–	0.745***

*** indicates significance at $p < 0.001$

Table 3.10: MRQAP coefficients for the Fed semantic networks

Dependent	Intercept	*Pre-crisis*	*Crisis*	*Post-crisis*
Recovery	0.647***	0.336***	−0.128***	0.817***

Adj-R^2= 0.568; *** = $p < 0.000$

MRQAP Fed

In Table 3.9, the correlations among the intersection conformed Fed networks are reported. The intersection node set size here is $n = 356$ nodes.

While the correlations are modestly lower here than for the ECB, their patterns mirror those of the ECB correlations. Specifically, adjacent periods exhibit higher correlations, and those pairs of periods with the greatest temporal distance have the lowest correlations.

We again employ a full MRQAP model predicting the *recovery* period from the prior three periods and report the coefficients in Table 3.10.

While the pattern of coefficients' magnitudes and valences bear some resemblance to those of the ECB MRQAP model, there are also some striking differences. The coefficient for *pre-crisis* is higher here than it is for the ECB, indicating that the structure of language (and hence the policy focus) in the *recovery* period does not depart as widely from *pre-crisis* in the Fed data as it does for the ECB data. Conversely, the coefficient for *crisis* is negatively higher here than for ECB indicating the language structure in the Fed *recovery* period departs even more from *crisis* period language.

3.3.3 Multidimensional Scaling

In Figure 3.11, we display multidimensional scalings (MDS) in 2 dimensions for both the ECB and Fed correlation matrices.[7] The correlations were converted to distances via a $1 - r$ transformation where r is a period pair-wise Pearson correlation. In MDS, an eigen-decomposition reduces the dimensionality of a matrix of distances (in our case, the extent of non-correlation) such that a pairwise distance between each pair of data points in the reduced space (in our case, 2-D) is roughly proportional to the original distance between the pair. Thus, distant points (or periods of the ECB of FED) in Figure 3.11 indicate relatively lower correlation in their semantic networks than proximal period points.

The arrows in the figure indicate the temporal progression from *pre-crisis* all the way through *recovery*. The plotted axes have been aligned across sub-figures to allow for comparison. The root mean squared errors between the original correlations (upper triangle) and the Euclidean distances of the MDS coordinates are low (0.006 and 0.008 for ECB and Fed respectively), indicating that two dimensions adequately capture the temporal differences of the conformed networks and additional dimensions would not add to the results.

The figures succinctly echo our analysis of the raw correlation matrices: that adjacent periods bear the most similarity (i.e., are closest together in the MDS). Interestingly, the trajectories for both the ECB and Fed are not merely straight lines departing from *pre-crisis*, but instead there is a curvilinear shape to them, which indicates the language structure partly returning to the status quo. However, we can maintain our earlier assertion that *recovery* is a new state given that the trajectories along one of the dimensional axis (the x-axis) proceeds monotonically away from *pre-crisis*(for both the ECB and the Fed discourses). Finally, the higher correlations of ECB are reflected in the MDS trajectories occupying a smaller area of the space.

[7] MDS is also known as Principal Coordinate Analysis.

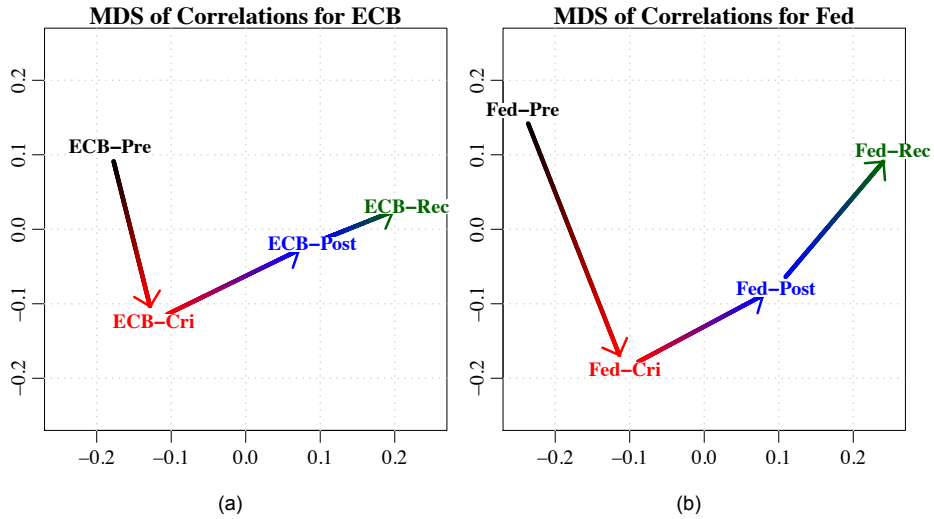


Figure 3.11: MDS of Correlations

3.3.4 Conclusion

The goals of the present article were three fold: 1) we sought to overcome two common challenges in text analysis, namely the size of the text corpora and its formal character; 2) we aimed to explore the benefits of the structural space dimensions; and 3) we wanted to investigate how the discursive practices of the ECB and the Fed have been affected by the recent financial crisis.

The structural space method employed by this study revealed substantial and imperative shifts in the discourse of both the ECB and Fed, demonstrating that it could be a valuable instrument for change detection in formal discourse. As demonstrated, focusing on just the obvious most central concepts in formal discourse does not always reveal the underlying and subtler shifts across the periods investigated. Formal organizational discourse contains repetitive top key concepts, indicative of the obvious and perhaps uninformative central topics of an organization. The structural space analysis proved to be more explanatory regarding the shifts and changes in formal discourse, by combining structural measures and looking beyond the core of the network structure. At the same time, structural roles of key concepts may be good predictors of emerging topics and the dynamics of discursive change.

In recent years, an increasing number of researchers have been focusing on the importance of central banks' communication (Ehrmann & Fratzscher, 2007; Friedman, 2002; Kohn & Sack, 2003; Rosa & Verga, 2005; Sturm & de Haan, 2011). As key policy instruments and valuable sources of information, the communications of the ECB and the Fed are of critical importance for financial market participants and for society at large. This being said, in times of crisis, the role of these central banks in guiding financial market expectations through communication is particularly important due to higher market uncertainty. While previous research established a link between communications of the ECB and the Fed and their respective impacts on the financial markets (Jansen & de Haan, 2005a,b; de Haan, 2008; Hayo & Neuenkirch, 2010; Hayo et al., 2014), our focus was directed at uncovering the shifts and adaptations of their discourse in a time of crisis and market volatility.

The structural space dimension of the selected key concepts exposed significant changes in the ECB and the Fed discourses. Below, we summarize our main findings for each organization and each of the examined periods.

ECB Pre-crisis: Our analysis showed that crisis-oriented key concepts were already present, suggesting that even before the end of 2007 the ECB's discourse shifted towards crisis terminology, and their focus may have been on the impending financial crisis.

ECB Crisis: The crisis-related key concepts detected by our method in the *pre-crisis* period became more prominent in the *crisis* period. We showed that the key concepts associated with the main objectives of the ECB lost ground in front of the effects of the full-blown financial crisis. Also, we noted the emergence of 'longer-term refinancing operations' (LTRO) as a highly ranked connective (G) concept and similarly the increase in betweenness for 'main refinancing operations' (MRO). These changes denote the focus of the ECB on refinancing operations during the *crisis* period. Interestingly, the 'crisis' concept was not a highly connective, nor a popular concept during this period. The marginal position of this concept in both the *pre-crisis* period and the *crisis* period could denote an intentional attempt of the ECB to minimize panic reactions among the stakeholders, or it could be explained by a narrow

focus of the ECB's discourse towards the overwhelming market defaults and not towards the crisis as a whole.

ECB Post-crisis: This period revealed the ECB's discourse in a distinct state, where all the main objectives of the ECB are in a prominent position, while 'crisis' itself became a more connective concept. The changes observed in this period point towards a focus of the ECB's discourse towards dealing with the aftermath of the crisis.

ECB Recovery: While in the *crisis* and the *post-crisis* periods 'LTRO' is the highest ranked G concept, during the *recovery* period it suffers a drastic decrease in betweenness centrality ranking, appearing now as a marginal concept (M). At the same time, we show 'loan' becoming more popular, and 'MRO' and 'refinancing' becoming more connective. These structural changes in the semantic structure show the shift towards dealing with the aftermath of the crisis more clearly than in the previous period.

Finally, our findings revealed that by the end of 2013, the discourse of the ECB had in no way returned to the *pre-crisis* levels, but perhaps advanced to a new state altogether. This 'new state' could be explained by the role of the ECB in dealing with the aftermath of the financial crisis. Also, during the *recovery* period, the ECB seemed to focus less on the financial crisis and more on the subsequent recovery process. This particular finding was supported by the structural space analysis as well as the MRQAP coefficients and multidimensional scaling.

Fed Pre-crisis: The *pre-crisis* period exposes the Fed discourse as having a focus on such terms as 'risk', 'loan', and 'inflation'. At the same time, we observe that the main objectives of the Fed are not central to its discourse in this period. With 'employment' and 'interest_rate' as marginal and emerging concepts and 'stable_price' as an LC (popular) concept, we can argue that the focus of the Fed seems to already be aimed at the impending financial crisis and not towards their main objectives.

Fed Crisis: During the *crisis* period, the Fed discourse expands considerably, and we also note the increased activity in the induced subgraph. The focus on crisis-oriented terminology becomes even clearer during this period,

with the same main objectives present. The prominent positions held by even more crisis-oriented concepts leads us to conclude that the Fed discourse during the *crisis* period has been mostly focused on dealing with the events of the financial markets, while at the same time being focused on maintaining price stability.

Fed Post-crisis: At this stage the Fed's discourse seems to regain a focus oriented towards their main objectives but the crisis terminology remains prominent at the same time. Interestingly, 'crisis', a concept that has been only marginal in the pervious period, now becomes a connective concept in their discourse. These findings suggest that while the focus of the Fed was still very much aimed at the effects of the crisis, they also returned to a state in which their main objectives became more popular and connective in their discourse.

Fed Recovery: During the final period analyzed, we find the Fed's discourse in an unexpected state, in which the main objective concepts are no longer in the GC quadrant, but instead they are highly ranked LC concepts. Here, these concepts are popular in the Fed's discourse, but they do not serve a connective role between topics being discussed. In light of these findings, and based on the high GC ranks of 'loan', 'risk', and 'inflation', we can postulate that Fed is acknowledging the (still) precarious state of the financial system, while at the same time dealing with the repercussions of the crisis.

The Fed data MRQAP analysis results are relatively similar to the ones of the ECB but they also reveal some noteworthy differences. While the general trend is comparable to the one of the ECB data, the *recovery* period discourse of the Fed moves even further away from the *crisis* period, but appears more similar to the *pre-crisis* period discourse. Thus, while the ECB's *recovery* period discourse enters a new state which seems to retain aspects of the *crisis* period but also returning slightly to the *pre-crisis* status quo, these changes in discourse are even more pronounced in the Fed data set.

In sum, the structural space approach has exposed key findings regarding the subtle shifts in the discursive practices of the ECB and the Fed throughout the different phases of the financial crisis. One of the most striking findings

is that unlike the ECB, the Fed has not been entirely focused on all of their main objectives with the exception of the *post-crisis* period. While the three main objectives of the ECB are present in all the four periods, all of Fed's main objectives are only present in the *post-crisis* and *recovery* periods. During the *recovery* period, all of the Fed's main objectives are present, but they are all positioned in the LC quadrant suggesting that these concepts are only locally central and thus not connective between various topics present in the Fed's discourse.

Furthermore, there is greater presence and prominence in positions of certain crisis concepts in the Fed than in the ECB. That is, more Fed crisis concepts emerged to achieve GC positions in *crisis* and *post-crisis* than the crisis concepts of the ECB. In particular, 'TALF' emerged as a GC concept (*crisis*) as did 'lend' and 'debt' (*post-crisis*), suggesting a change in the focus of the Fed between *crisis* to *post-crisis*. Even in *pre-crisis* and *recovery*, we observe the presence of most of the Fed's crisis concepts as well as GC prominence of several of them. The prominence of crisis terms in these periods, in addition to others, point to the Fed's inclination to alert consumers not only about the actual but also a potential crisis. All these findings indicate the Fed's concerns (more so than the ECB's) being oriented towards elements of the financial crisis with either expectation or caution (in the case of *pre-crisis*) or as a reaction to the present crisis (*crisis*) or in a reparative capacity throughout the aftermath (*post-crisis* and *recovery*). This observation coincides with our earlier claim that the Fed exhibits a greater attentiveness to the crisis based on the the summary network statistics alone (see Table 3.2).

Even if the focus on their main objectives differs between the ECB and the Fed, one similarity worth noting is that the 'crisis' concept does not occupy a prominent position in neither of their discourses until the *post-crisis* period. Whether this finding is influenced by an organizational strategy meant to prevent further panics on the financial markets or simply a delayed acknowledgment of the crisis from the two organizations, this finding warrants further investigation into the possibility that other, less value laden concepts have been used by the ECB and the Fed to describe the events that were affecting the

financial system.

Although the method we have employed in this study revealed important findings, one of its primary limitations stems from the general limitations of semantic network analysis. The process of transforming textual data into networks of concepts (or words) implies a series of coding choices which can greatly impact the results of the analysis. That is, the techniques used when preprocessing the raw text (e.g., removing noise words, removing numbers, etc.), the identification of nodes to be included in the network, and/or the parameters used for the creation of links (e.g., window size and/or stop unit) can strongly impact the structure of the resulting network. As such, these coding choices should be closely aligned to the objectives of the researcher and should be chosen with care.

While our method of classifying nodes into one of the four structural roles was used to highlight only a handful of key concepts, the classification may easily be broadened to identify lists of top concepts (e.g., top ten) within each of the roles. This enumeration of the roles offers a more complete depiction of the roles and their evolution.

Our naïve treatment of weighted degree centrality, while typical in network research, raises some concerns. Specifically, weights and the number of distinct ties ought to be considered separately as the same total degree centrality score of a node can arise from starkly different ego-centric structures. While the exploration of this issue is beyond the scope of this paper, we hope (and expect) that future research will improve the use of weighted degree centrality in semantic and social network analysis.

As for the complex structure of our semantic networks, some diagnostic tests reveal that they are only mildly small-world and not at all scale-free, contrary to the findings of other work. Still, further investigation (outside the scope of this paper) would be required to determine if these inconsistencies are due to the types of organizational semantic networks inferred in our study or the exact nature of semantic network extraction or simply that semantic networks can vary widely in their topologies.

As for metric comparisons with other research, our within-network correla-

tions for our two centrality measures were modest, echoing some other findings, e.g., (Valente et al., 2008; Lee & Pfeffer, 2015a) and low enough to permit the identification of outliers in structural space, particularly in the gatekeeper (G) and locally central (LC) roles. Also, the differences in the correlations between the two data sets may relate to structural differences between the discourses of the ECB and Fed and warrant further investigation.

Furthermore, the direct text analytical metrics cannot be used in lieu of the network measures due to lack of high correlation. In fact, semantic network analysis in combination with the structural space approach outperforms more direct text-analytical approaches (i.e., frequencies or tf-idf). For example, the relevant G role of 'LTRO' ('longer-term refinancing operations') remains obscured under direct text-analysis which shows high percentiles/rankings for tf-idf (0.98) and medium-high frequency (0.73) but omits its low popularity/degree (0.15) and high connectiveness/betweenness (0.98).[8] Thus, a text analytic approach would fail to uncover the highly connective role of 'LTRO' and other concepts which bridge prominent topics and themes of the ECB discourse.

Our use of centrality ranks as opposed to actual centrality scores warrants additional, future inquiry. We suspect that in order to compare them more precisely across networks of varying sizes and densities, tighter controls must be exerted. We envision the development of highly robust comparative indices that account for both the relative or ranked centrality score as well as the absolute score.

To conclude, we can argue that our approach proved beneficial for the analysis of large corpora of formal organizational discourse. We anticipate our noteworthy results to open new avenues for semantic network research dealing with formal discourses and beyond the context of the financial crisis.

[8] While the tf-idf here coincides with betweenness, the mild correlation between the two measures in Table 3.5 reveals this is not always the case.

4 A 'Toxic' Crisis: Metaphorizing the Financial Crisis

"Metaphors are the mind's eyes and society's tools."
(Nerlich & Jaspal, 2012)

Abstract

The recent financial crisis has been covered in newspapers with metaphors such as toxic assets and toxic loans. Although these groups of related metaphors (i.e., metaphor families) may strengthen the intended images on the topic under discussion, they have been only seldom studied in metaphor research. This article investigates the ways in which metaphor families fulfill a translator role for emerging terminology in financial discourses. We explore the expansion and evolution of the toxic metaphor family, revealing subtle changes of metaphor use in three newspapers over time. Our results show a transition from generic image-creating metaphors toward financial-instrument-targeted metaphors. Overall, the evidence brought by this study is a stepping-stone for further research on metaphor families.

Keywords: Metaphor, Semantic Networks, Financial Crisis, Structural Space.

4.1 Introduction

"The growing use of the language of toxicity during the past two decades may be attributable to the fact that it conveys a destructive force, a poison, or a dysfunction spreading throughout an environment, a human body, or a human system" (Goldman, 2008, p. 243).

Chapter based on Nerghes et al. (2015a)

Whereas the metaphors of choice in the late 1990s savings and loan crisis evoked fears of contagion, the 2008 financial crisis moved toward "environmental and climatic rather than epidemiological metaphors."(Smith, 2009, p. 409) A set of novel metaphorical combinations such as *toxic assets, toxic loans*, and even *toxic banks*, which frame the financial system and its main operations negatively, has increased in use in most newspapers. The repeated use of such related metaphors has the potential to strengthen the images they invoke while at the same time potentially strengthening each individual metaphor to the point of conventionalization.

Metaphors are defined as cross-domain mappings across two separate domains of experience: a source and a target domain (Lakoff & Johnson, 1980). Discussing the financial crisis in terms of *toxic loans*, for example, maps together the source domain of *toxic* and the target domain of *loans* and builds an image of *loans* as something highly negative and perhaps even lethal. The same source domain can be used for describing several related issues, such as *toxic banks* and *toxic assets*. In metaphor research, such related metaphors have seldom been studied despite the important role they play in strengthening a specific image or a frame of the issue. Metaphors may be regarded as "condensed" ways of framing issues (Snow & Benford, 1992) and as providers of specific perspectives on issues (Gamson & Modigliani, 1989). Our focus is on the framing of the financial crisis as toxic by investigating the family of *toxic* metaphors. We define a metaphor family as a set of metaphors that use the same source domain but different target domains.

In this article, we systematically map the evolution of the *toxic* metaphor family in three newspapers, *The Sun (Sun)*, *The New York Times (NYT)*, and *The Financial Times (FT)*, each representing a different level of specialization and audience, over a five-year period prior to, during, and after the financial crisis. We extend metaphor research beyond the analysis of single metaphors to a family of related metaphors. In particular, we are interested in the stage of the financial crisis that leads to the emergence and expansion of the *toxic* metaphor family. We expect the three newspapers to use toxic metaphors differently because their audiences are different.

4.1.1 Metaphors of the Financial Crisis

The different metaphor theories that have been developed over the years (Black, 1962; Johnson, 1981; Lakoff & Johnson, 1980; Ortony, 1993) approach metaphors as discussing a concept in terms of a different concept, thus transferring meaning from one concept to another. Divided into substitution and interaction theories, these theories differ by "locating metaphor either at the level of language and words as opposed to thought and context" and by "emphasising the role of metaphors as either reflecting some already existing similarities as opposed to also creating similarities between things or ideas" (Hellsten, 2002, p. 17). The conceptual metaphor theory we build upon considers metaphors as playing an important role in defining the way we perceive the world and, thus, the way we think and act (Lakoff & Johnson, 1980). According to Lakoff & Johnson (1980), the human conceptual system is metaphorically constructed, and everyday language is largely based on metaphorical ways of thinking. Conceptual metaphor refers to the understanding of one domain in terms of a different domain (Lakoff & Johnson, 2003). Lakoff & Johnson (1980) have theorized that a significant part of our everyday language is structured metaphorically, and thus we often use metaphors to understand one idea in terms of a different, more familiar idea. Metaphors are flexible in the sense that they can be mappings between two discrete concepts (e.g., He's living on borrowed time) or between a descriptor and an object, in our case toxic asset.

Recent metaphor research has shed a different light on the social and communicative roles of metaphors and their effects on our understanding of public issues (Chilton & Ilyin, 1993; Hellsten, 2002; Lakoff & Johnson, 2003). In this tradition, we considers metaphors as tools of communication (Hellsten, 2002), and we take into account their ability to offer common grounds between discourses (Chilton & Ilyin, 1993) or to function as boundary objects (Star & Griesemer, 1989) that are at the same time flexible enough to allow several interpretations in different social contexts but also to carry a relatively fixed set of associations. The concept *toxic* is engaging because via this concept, a rich web of financial issues (e.g., *toxic assets*, *toxic loans*, and even *toxic banks*)

can be translated to something that is expected to be familiar to different discourses. Associating various financial concepts in discussions about the crisis to the term *toxic* puts blame on products, organizations, and their proponents without discussing specifics and personal involvement while at the same time implying a role in the crisis. Such a translation process does not mean that *toxic* carries the same meaning in various social contexts (Zeiss & Groenewegen, 2009). On the contrary, the power of metaphors is in their flexibility in uses and interpretations. *Toxic* can be easily adapted to fit the expected worldviews of the readers of different newspapers.

Metaphors have been studied in economic discourse broadly (e.g., Alejo, 2010; Charteris-Black & Musolff, 2003; Hayes, 1997; Rhodes & Garrick, 2002), but the study of metaphors in debates relevant to the latest financial crisis is still in its incipient stages. The Metaphor Observatory [1], for example, discusses the financial crisis as the trigger for "one of the largest metaphor spikes in recent history." In recent years, attention has been paid to the use of metaphors in crisis communication by banking executives (Tourish & Hargie, 2012), variations in the use of positive and negative metaphors between Spanish and English financial texts (López & Llopis, 2010), editorial cartoons representing the global financial crisis (Bounegru & Forceville, 2011), and contagion in the general press parallel to the avian flu scare (Peckham, 2013).

Kleinnijenhuis et al. (2013) found that "the news about the crisis became crisis news itself" (p. 287), reinforcing the idea that the way in which journalists report the events of the financial crisis has a major impact on the escalation of the crisis. Journalistic discourse often contains compelling metaphors and warrants investigation. This important role of metaphors in the mass media is confirmed by the vast array of compelling results published by researchers of various fields (e.g., Berdayes & Berdayes, 1998; Hellsten, 2000; Kennedy, 2000a; Kitis & Milapides, 1997; Nerlich et al., 2000; Paris, 2002; Petersen, 2005). However, not many authors have investigated the use of metaphors in newspaper reporting of the recent global financial crisis (e.g., Bounegru & Forceville, 2011; López & Llopis, 2010; Tourish & Hargie, 2012).

[1] See `www.metaphorobservatory.com`

In journalistic discourse, metaphors are used to "popularize, concretize, or dramatize issues" (Hellsten, 2002). The use of metaphors makes issues newsworthy and interesting for audiences, and metaphors may also be used to address different audiences simultaneously (Bucchi, 1998). According to Kennedy (2000b), "metaphors are often said to be helpful in creating and dealing with what is novel" (p. 209). In other words, metaphors can be used by the media to introduce a new issue to their wider audience in terms of something more familiar (Wyatt, 2004) or engaging. In the case of the financial crisis, since 2008, many novel issues and terms have emerged [2]. This abundance of new terms has prompted the use of metaphors as translators in the media. Using metaphors to describe such terms as derivatives, *collateralized debt obligation* (CDO), or *asset-backed mortgage* translates these terms into what is perceived as more concrete, familiar, or engaging. This is not to say that the translation role of metaphors is restricted to mere descriptors of unfamiliar terms. On the contrary, when part of a growing metaphor family, such metaphors as *toxic* can function as poisonous or even deadly labels for each target domain associated with them. Such metaphor families suggest a specific image − a negative image in the case of *toxic* metaphors − on the issues while suppressing alternative views, thus reducing the complexity of issues.

As tools meant to either popularize or condense complex issues, or to translate highly specialized discourses, metaphors guide our perceptions and interpretations of reality and help us to frame our visions and goals, "playing a central role in the construction of social and political reality" (Lakoff & Johnson, 2003, p. 159). As such, the use of metaphors in news has the potential to influence meanings readers associate with the issues reported (Williams et al., 2011), which in turn can manifest changes in behavior and decision making (Williams, 2013).

[2] See "The Layman's Finance Crisis Glossary" at `http://news.bbc.co.uk/2/hi/uk_news/magazine/7642138.stm`

4.2 Aims and Goals

Using data collected from three newspapers, *The Sun (Sun), The Financial Times (FT)*, and *The New York Times (NYT)* between 2007 and 2011, we contribute to metaphor theory by widening the focus from conceptual metaphors to metaphor families − or hybrid word families (Thelwall & Price, 2006) that share a common source domain. So far, little is known about when such related metaphors emerge and how they develop over time. By identifying metaphors sharing the source domain *toxic* in newspapers and by revealing these metaphors' dynamics in the financial crisis debate, this study contributes to a better understanding of the ways different publics are drawn into a specific framing of the financial crisis.

To structure our search for variation in the use of *toxic*, we follow the differentiation of the metaphor family by examining the following factors:

1. The evolution of the *toxic* metaphor family across stages of the financial crisis, with attention to the following questions:
   (a) At what stage of the financial crisis did the *toxic* metaphor family emerge?
   (b) At what stage of the financial crisis did the *toxic* metaphor family show most variation in usage of unique metaphors?
2. The structural roles of the shared metaphors identified across newspapers and how they changed across the periods analyzed.

4.3 Data and Methods

4.3.1 Data Collection

The texts used in this study come from three newspapers: *The Financial Times (FT)*, *The New York Times (NYT)* and *The Sun (Sun)*. These three newspapers publish very different content and thus address different types of audiences.

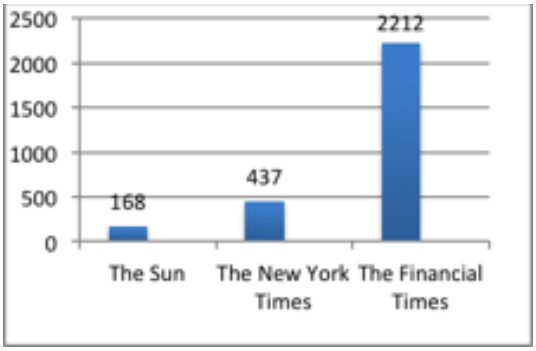


Figure 4.1: Number of articles collected for each newspaper

FT is a highly specialized financial reporting newspaper; *NYT* is the most popular daily newspaper in the United States, publishing a broad variety of topics; and *Sun* is the largest circulation daily tabloid in the United Kingdom. Each of these newspapers target relatively different audiences and are expected to employ the toxic metaphor family differently.

We collected the data from LexisNexis by searching with the keyword *toxic* with no start date but with an end date of December 31, 2011, in each of the three newspapers selected for inclusion in our analysis. All the articles retrieved were then manually selected, and only articles on financial topics were included in the analysis. After removing duplicates, a total of 2,817 articles remained (see Figure 4.1).

The first article using the *toxic* metaphor in regard to financial issues dates back to 2004, and it was published by *NYT* on February 13. The article discusses the debt created by the building of the Eurotunnel:

> "Michael Wilkins, a managing director at Standard & Poor's, says Eurotunnel's senior debt is investment grade, but its lowest-rated debt is in the low junk category, and there are billions in debt below that, most of it *toxic* waste." (Norris, 2004, p. C1, emphasis added)

The articles collected from each newspaper have been separated into three sets that we call the *pre-crisis* period (2006−2007), the *crisis* period (2008−2009), and the *post-crisis* period (2010−2011). The only exception to this is the *pre-*

*crisis*set for *Sun* that used the *toxic* metaphor in only one article in 2007.

4.3.2 Metaphor Identification

Repeatedly addressed in natural language processing (NLP) research, automatic identification of metaphors continues to be a challenge (Gedigian et al., 2006; Shutova, 2010) due to the complexity of language. Semiautomated methods such as part-of-speech tagging, sentence clustering, and lexical patterns are still limited because they require manual annotation or other manual coding (Birke & Sarkar, 2006; Fass, 1991; Gedigian et al., 2006; Goatly, 1997; Krishnakumaran & Zhu, 2007; Miller et al., 1990; Peters & Peters, 2000).

In this article, we make use of a method that allows for automated text processing and extraction of metaphors based on their target or source domains. Using this method, we identify metaphors that use a specific word or multiword expression as the source domain (i.e., *toxic*) present in large corpora of unstructured text documents.[3] Specifically, we will focus on metaphors where the source domain, *toxic*, precedes the target domain, as this is the most common case for this metaphor family. Research on metaphor identification based on target or source domains remains limited to this day, and only a few authors have undertaken similar efforts (Mason, 2004; Thelwall & Price, 2006; Ureña Gómez-Moreno & Faber, 2010). After preprocessing the text (removing all noise words), we generated semantic co-reference lists using a window size of two words and a stop unit of one sentence as a method of identifying metaphors. The window size determines the range in which connections are made between words (Diesner, 2012b). A window size of two will create a link between each two consecutive words within the limit of one sentence. Because these lists were generated to identify metaphors in the toxic family, they are unidirectional. The semantic co-reference lists include co-occurring concepts based on the window size and the frequency with which they occur. This part of the analysis has been used as a method of detecting the unique *toxic* metaphors used by each selected news source. We identified 25 metaphors from the *toxic* metaphor family used by *Sun*, 60 metaphors used by *NYT*, and 171

[3]Conversely, the method can be used to identify metaphors based on their target domain.

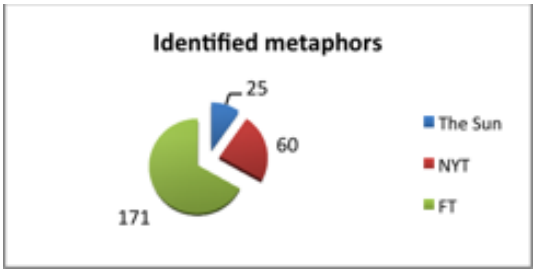


Figure 4.2: Number of metaphors identified for each newspaper

metaphors used by *FT* (see Figure 4.2). Of all the metaphors identified, only 8 are common to all the newspapers and will be further analyzed to track their dynamics.

Once the metaphors were identified, each of them was recoded in the corpora by using n-gram conversion, which creates single concepts from multiword n-grams by replacing the space between the words with an underscore (Carley et al., 2013a). An example of such conversion is *toxic asset* becoming *toxic_asset*.

4.3.3 Semantic Networks

After we identified and recoded the metaphors, we generated semantic maps using Automap (Carley et al., 2013a). Semantic networks translate selected text into networks of concepts, in which a concept can be a word or a phrase (i.e., an n-gram) (Popping, 2003), and the links between them (in this case, relations among concepts are defined by co-occurrence). The value of the strength of each link is determined by frequency of co-occurrence (Wasserman & Faust, 1994). The methods of extracting networks of concepts from texts have been referred to as maps Carley (1997a,b), semantic and communication networks (Lehmann, 1992; Monge & Contractor, 2001; Popping, 2003), networks of concepts (Popping, 2000), and networks of words (Danowski, 1993). Named differently, all these methods focus on content analysis that assumes language can be modeled as networks of words and their relations (Sowa, 1984). Unlike content analysis, our approach does not require extensive manual coding. The semantic

networks for all three newspapers were generated using a window size of eight, with a stop unit of two sentences. These choices of window size and stop unit are the most appropriate for generating semantic networks from newspaper corpora (Diesner, 2012b). Table 4.1 contains the descriptive statistics for each of the networks generated.

Network	Node count	Link count	Density
Sun Pre-Crisis semantic network	37	118	0.177
Sun Crisis semantic network	2738	55396	0.007
Sun Post-Crisis semantic network	2381	40102	0.007
NYT Pre-Crisis semantic network	2539	44852	0.006
NYT Crisis semantic network	14370	699629	0.003
NYT Post-Crisis semantic network	7012	218618	0.004
FT Pre-Crisis semantic network	3953	90668	0.006
FT Crisis semantic network	20231	1550492	0.004
FT Post-Crisis semantic network	12690	665702	0.004

Table 4.1: Descriptive statistics of each semantic network generated

4.3.4 The Structural Space Method

The resulting semantic networks were analyzed through the structural space method (Nerghes et al., 2014b), which combines total degree centrality (i.e., popularity) and betweenness centrality (i.e., connectivity) of concepts in a semantic network. The total degree centrality of a node in a network is the number of other nodes to which the focal node is tied (Freeman, 1979). In semantic networks, total degree centrality may represent the importance of a concept or its key concept status. Betweenness centrality is the frequency with which a particular node is on the geodesic path between any other two nodes in the network (Freeman, 1979).

The betweenness centrality of a concept in a semantic network is an indicator of its influence (Hill & Carley, 1999; Hooper et al., 2012). Such a concept controls access to other key concepts in the network (Brandes & Corman, 2003; Grebitus & Bruhn, 2008; Henderson et al., 1998; Hulst, 2008), serving as a gatekeeper between different domains (Gloor & Krauss, 2009).

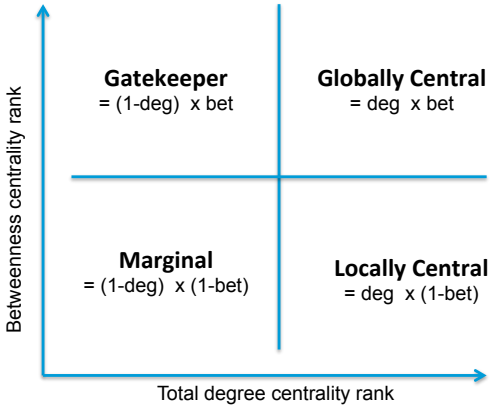


Figure 4.3: The four quadrants of the structural space

The combination of node count and link count reveals four structural roles as quadrants of the structural space (see Figure 4.3), and we use the approach to uncover more subtle structural positions of concepts and changes in discourse over time. The globally central (GC) concepts have high degree centrality and high betweenness centrality. These are very popular and highly connective concepts and become central, key concepts of a hot topic because they are highly connected to other concepts and serve as bridges between parts of the network (or topics). The locally central (LC) concepts have high degree centrality and low betweenness centrality. These concepts are very popular but do not have a strongly connective role. They serve as key concepts of a local hot topic because they are highly connected to other concepts, but they do not serve as bridges between topics. The gatekeeper (G) concepts are characterized by low degree centrality and high betweenness centrality. These concepts serve as connective concepts that are not very popular, but they are influential in the network because although they are not highly connected, they act as bridges, potentially linking themes or topics. Last, the marginal (M) concepts have low degree centrality and low betweenness centrality. These concepts are neither popular nor connective, but they can be emergent concepts. (For more details on this approach, see Nerghes et al. (2014b))

We use the structural-space method to evaluate and track the dynamics

of the eight shared metaphors (used by all three newspapers) of the *toxic* metaphor family.

4.4 Results

This section of will start by outlining the results we found for *Sun*, followed by those for *NYT*, and last by those for *FT*. The second part of this section compares the three newspapers, and the last part discusses the structural shifts in metaphor use detected with the structural-space method for each newspaper.

4.4.1 The Sun

The articles published by *Sun* are relatively short and are designed to catch readers' attention at a glance. In the 168 articles published between 2007 and 2011, we identified 25 unique metaphors from the *toxic* metaphor family. Out of the 25 unique metaphors, only one was used in the *pre-crisis* period, in an article published on June 6, 2007. In this article, *Sun* metaphorizes the word income (i.e., *toxic income*). This demonstrates that *Sun* did not use metaphors with the source domain *toxic* in the years preceding the financial crisis to characterize the emerging events of the financial markets for its readers.

During the *crisis* period (between 2008 and 2009), *Sun* used 15 metaphors of the *toxic* metaphor family. The most frequently used metaphor was *toxic debt*, which occurred 35 times in the articles published in this period.

In the *post-crisis* period (between 2010 and 2011), *Sun* used 17 metaphors of the *toxic* metaphor family. The most frequently used metaphor was *toxic loan*, which occurred 23 times.

4.4.2 The New York Times

In the 437 articles published between 2006 and 2011, we identified 60 unique metaphors from the *toxic* metaphor family. Although our analysis includes articles published only between 2006 and 2011, it is important to mention that the first metaphorical use of the word *toxic* in relation to financial issues by

NYT dates back to February 13, 2004, and discusses the debt created by the building of the Eurotunnel (quoted above in the Data Collection section).

In the *pre-crisis* period, *NYT* used 10 of the 60 metaphors identified. The first such metaphor is used in an article published on October 6, 2006, and includes the word stock (i.e., *toxic stock*). The most frequently used metaphors during this period are *toxic market* and *toxic waste*, both being used twice. In the *crisis* period, *NYT* used 45 of the 60 metaphors identified. The most frequent metaphor was *toxic asset*, used 271 times.

During the *post-crisis* period, *NYT* used 23 of the 60 metaphors identified, with the most frequent being *toxic asset*, used 24 times.

4.4.3 The Financial Times

FT publishes slightly longer and more elaborated articles. In addition, its readers are expected to be particularly interested in financial issues. It is important to mention here that our LexisNexis search revealed that *FT* first used the term *toxic* in relation to financial issues in an article published on January 27, 2004. This article talks about *"toxic levels of debt"* (Roberts, 2004, p. 24).

In the 2,212 articles published by *FT* between 2006 and 2011 and collected for this study, we identified 171 unique metaphors, the first metaphor identified being *toxic combination*, used in an article published on March 17, 2006.

During the *pre-crisis* period, *FT* used 26 of the 171 metaphors identified, the most frequently used metaphor being *toxic waste*, with nine uses. In the *crisis* period, 113 metaphors were used, with *toxic asset* the most frequent, with 1,820 uses.

In the *post-crisis* period, 87 metaphors of the 171 identified were used. Just as in the *crisis* period, the most frequently used metaphor in the *post-crisis* period was *toxic asset*, with 216 uses.

To summarize, the *crisis* period was most prolific in expanding the *toxic* metaphor family, both with novel metaphors and in the frequency with which the most popular metaphors were used (Figure 4.4) in *NYT* and *FT*. The number of unique metaphors used by *Sun* increased slightly between the *crisis*

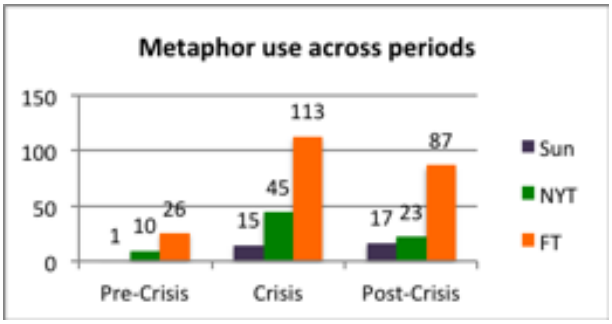


Figure 4.4: Number of metaphors used by each newspaper across periods

and the *post-crisis* period.

4.4.4 Comparing Newspapers

In terms of the number of distinct metaphors, the discourse of *FT* is the richest, having used 171 unique metaphors of the *toxic* metaphor family. This can be partially explained by the higher number of articles published by *FT* between 2006 and 2011. The high number of metaphors and the high number of articles published can be linked to the fact that *FT* is a highly specialized newspaper on financial issues. However, in terms of frequency of metaphors, *Sun's* discourse suggests more variety. As shown in Table 2, *toxic asset* is by far the most frequent metaphor in both *NYT* and *FT*, with the second most used metaphor appearing comparatively many fewer times. The top most frequent metaphors used by *Sun*, on the other hand, are much closer in frequency, suggesting more discursive diversity. Therefore, we posit that *Sun* engages a more diverse audience.

SUN		NYT		FT	
Metaphor	Freq.	Metaphor	Freq.	Metaphor	Freq.
Toxic loan	52	Toxic asset	295	Toxic asset	2036
Toxic debt	47	Toxic mortgage	72	Toxic mortgage	117
Toxic asset	25	Toxic security	40	Toxic security	106
Toxic bank	19	Toxic waste	20	Toxic loan	101
Toxic mortgage	6	Toxic loan	18	Toxic debt	96

Table 4.2: Top five most frequent metaphors per newspaper

4.4.5 Structural Roles

While the structural-space method can be used to look at, for instance, top-ranking concepts in each of the structural roles, we focus on the eight shared metaphors (Figure 4.5). In the next section, we analyze the dynamics of these shared metaphors across periods and newspapers. For each of the structural roles plotted in this section, the background colors are a rough estimation of the four structural roles, the highlighted nodes are colored by frequency (red being the highest and blue being the lowest frequency), and n represents the number of nodes in the network plotted.

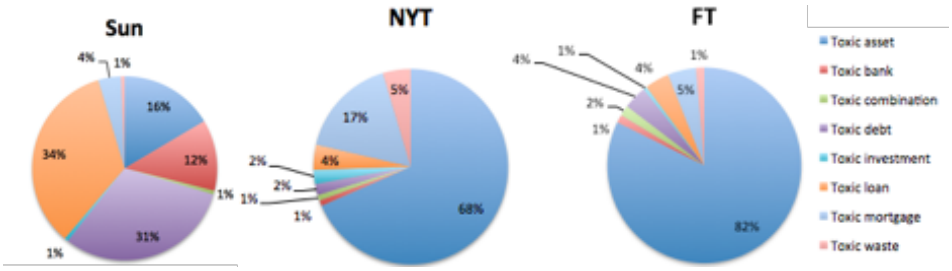


Figure 4.5: Use of selected metaphors across newspapers.

As mentioned before, in the *pre-crisis* period, *Sun* published only one article on financial topics using a *toxic* metaphor (i.e., *toxic income*). This metaphor is not one of the eight shared metaphors we have analyzed with the structural-space method because it was not used by all three newspapers.

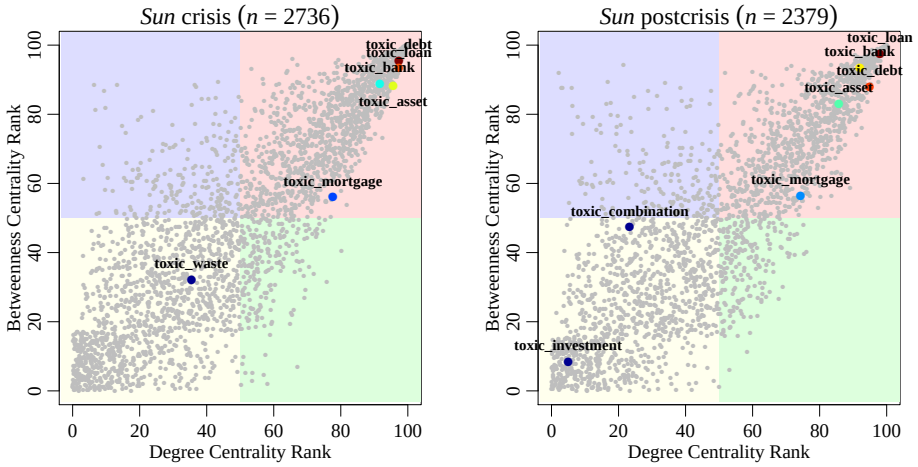


Figure 4.6: Structural-space plots for the semantic networks of the Sun.

During the *crisis* period (Figure 4.6a), *toxic debt, toxic loan, toxic bank*, and *toxic asset* rank highly on the globally central (GC) structural role. This means that these four metaphors, which were not present in the *pre-crisis* period, rapidly became popular in the *crisis* period while also becoming bridges between the topics *Sun* debated. *Toxic mortgage* was a more popular concept than a connective one, whereas *toxic waste* was a marginal and potentially emerging concept. The discourse of *Sun* shifted quickly from the *pre-crisis* state in which none of the eight metaphors were used to a crisis discourse that included six of those metaphors.

In the *post-crisis* period, *Sun* used seven of the eight common metaphors (see Figure 4.6b). Interestingly, *toxic loan, toxic bank, toxic debt*, and *toxic asset* maintained their GC position: they were popular in *Sun's* discourse and used to connect various topics. In the *crisis* period, *toxic debt* was the highest ranking GC metaphor, but in the *post-crisis* period, *toxic loan* became the highest ranking. This shows that the focus of *Sun* subtly changed. *Toxic mortgage* also maintained its position as a more popular than a connective metaphor. While *toxic waste* was no longer present in *Sun's* discourse, *toxic combination* held a borderline position between the G role and the M role. At the same time, *toxic investment* entered the discourse as an M metaphor,

ranking low on both total degree centrality and betweenness centrality.

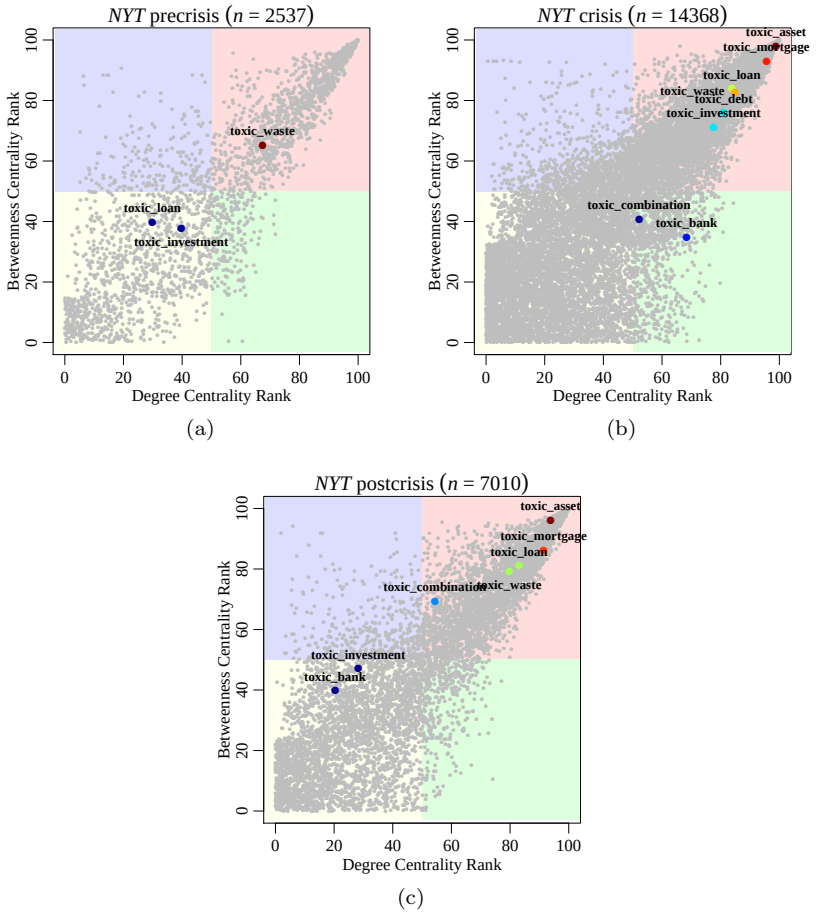


Figure 4.7: Structural space plots for the semantic networks of NYT.

For *NYT*, the structural roles of the *pre-crisis* period reveal that only three of the eight common metaphors were used: *toxic waste, toxic loan*, and *toxic investment* (see Figure 4.7a). While *toxic loan* and *toxic investment* were emerging marginal metaphors, *toxic waste* ranked higher in both total degree centrality and betweenness centrality. Thus, in the *pre-crisis* period, *toxic waste* was a popular metaphor that was also used to connect various topics under discussion in *NYT's* discourse. Based on these results, we can conclude that in the *pre-crisis* period, *NYT* was mostly focused on one of the eight

common metaphors (*toxic waste*), which is arguably a generic metaphor meant perhaps to offer a broad characterization of the emerging crisis rather than a specific characterization of particular financial instruments, as *toxic asset* does. The metaphors generally used to characterize specific financial instruments were marginal and possibly emerging in *NYT's pre-crisis* discourse.

During the *crisis* period, *NYT* used all eight common metaphors, with six of these ranking as globally central (GC) metaphors (see Figure 4.7b). *Toxic waste* was still a GC metaphor, but the fact that the top-ranked GC metaphor was *toxic asset*, followed by *toxic mortgage*, suggests a subtle shift in *NYT's* discourse toward characterizations of particular financial instruments. This is also established by the GC positions of *toxic loan, toxic debt*, and *toxic investment*. Remarkably, while *toxic investment* was a marginal metaphor entering the discourse in the *pre-crisis* period, in the *crisis* period, this metaphor became GC. The positions of *toxic combination* and *toxic bank* in the structural-space plot (Figure 4.7b) indicate that these metaphors are not highly ranked on any of the four structural roles.

In the *post-crisis* period, *NYT* used seven of the eight common metaphors (see Figure 4.7c). *Toxic debt* was no longer part of the discourse, but *toxic combination* increased in betweenness centrality to become a more connective concept than it had been in the *crisis* period. At the same time, *toxic investment* was no longer a GC metaphor, now fulfilling a more connective role. *Toxic bank* also decreased in degree centrality during this period.

The *pre-crisis* period exposed that *FT* used only five of the eight common metaphors in its discourse, with *toxic waste* and *toxic loan* being the highest ranked GC metaphors (see Figure 4.8a). Unlike in the case of the *pre-crisis* period of *NYT*, the structural position of these two metaphors suggests that *FT's* focus was twofold: on a more generic portrayal of the events, and on a characterization of specific financial instruments. During the *pre-crisis* period, *toxic combination* and *toxic mortgage* were marginally popular metaphors, but these metaphors were not highly connective ones. *Toxic investment* was a marginal metaphor in this period.

In the *crisis* period, *FT* used all eight common metaphors, with *toxic as-*

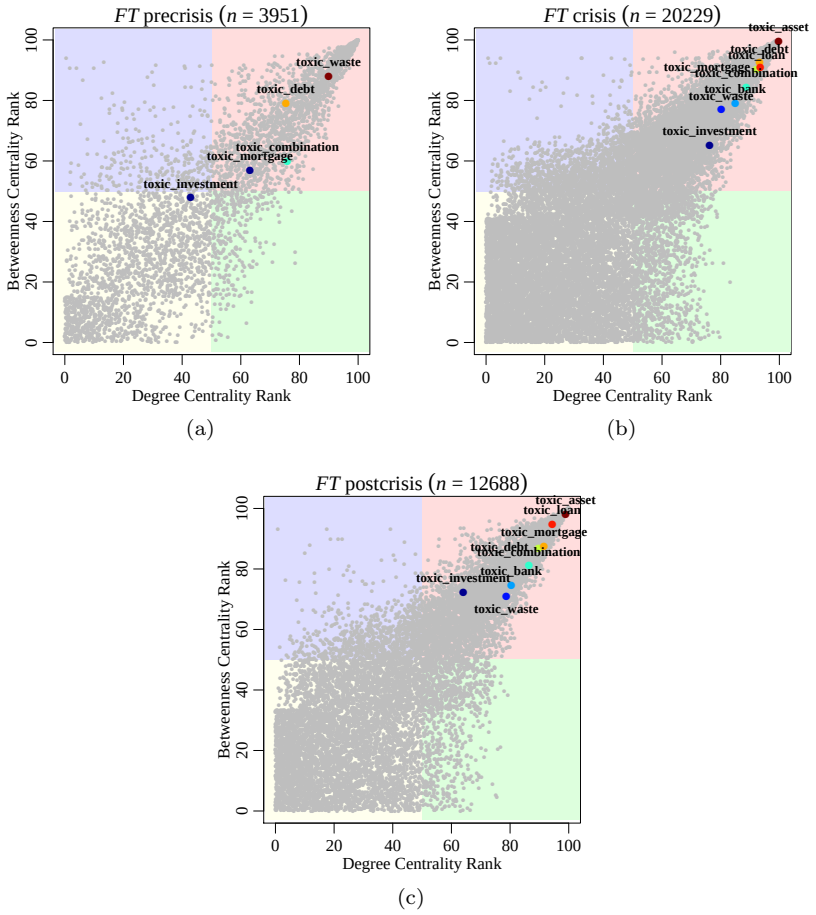


Figure 4.8: Structural space plots for the semantic networks of FT.

set ranking the highest on the GC scale (see Figure 4.8b). Two of the three metaphors that entered the *FT* discourse in the *crisis* period are related to financial instruments (i.e., *toxic asset, toxic loan*), and the third refers to banks (*toxic bank*). Although all eight metaphors can be considered GC during this period, the metaphors ranking highest in this role (*toxic assets, toxic debt, toxic loan, toxic mortgage*) suggest a clear shift toward portrayals of financial instruments.

In contrast to the *post-crisis* period for *Sun* and *NYT*, the *post-crisis* period for *FT* does not reveal significant differences to the *crisis* period for *FT* (see

Figure 4.8c). This finding suggests that while *NYT* and *Sun* adapted their discourse to a new stage (period), *FT* used the same discourse as it had in the *crisis* period.

None of the three newspapers used *toxic asset* in the *pre-crisis* period, but this metaphor became the top-ranked GC metaphor in the *crisis* period in all the newspapers. This metaphor also retained a high GC ranking in the *post-crisis* period in the three newspapers.

To summarize, the structural-space method revealed that in the *pre-crisis* period, *NYT* and *FT* used mostly generic portrayals of the emerging events by primarily using *toxic waste*, which we identified as the most GC *toxic* metaphor. The position of the *toxic waste* metaphor in the *pre-crisis* period suggests that the translation role of this metaphor is initiated through a generic, familiar association of it to financial issues, which potentially elicited a familiar negative image for audiences. Conversely, the structural-space analysis of the *crisis* period has shown that all three newspapers focused more on metaphors characterizing financial instruments than they did on the generic characterization metaphors seen in the *pre-crisis* period. The translation initiated in the *pre-crisis* period with generic metaphors evolved into the use of a richer variety of *toxic* metaphors.

With the exception of *FT*, the newspapers showed significant changes in metaphor use in the *post-crisis* period, implying a discursive shift toward a different and perhaps new state.

To further explore this particular finding, we looked beyond metaphor use to the general discourse of each newspaper and the differences that arose across periods. We treated the correlation coefficients among the semantic networks of each newspaper and each period as distances and used multidimensional scaling (MDS) to plot these values. MDS offers a way to visualize the (dis)similarities among a set of points (Gower, 1966). Figure 4.9 confirms that our findings regarding metaphor use remained valid for the general discourse of *FT* between the *crisis* and the *post-crisis* periods. Figure 4.9 also shows that although *Sun* and *NYT* discourses moved further away after the *crisis* period, *FT's post-crisis* discourse remained very close to its *crisis* discourse. Interestingly, the

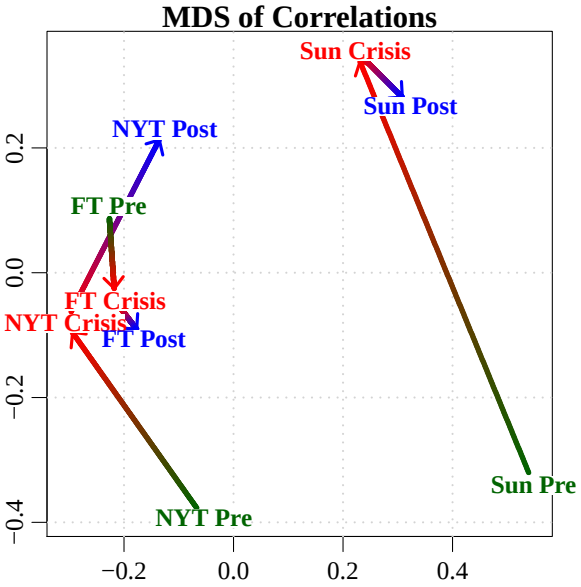


Figure 4.9: MDS representation of correlations across periods and newspapers. The periods are Pre-Crisis, Crisis, and Post-Crisis

MDS plot also shows significant similarities between the *crisis* discourse of *NYT* and *FT*, and the *post-crisis* discourse of *FT*.

4.5 Conclusion

We have analyzed financial crisis reporting in three newspapers with the aim of understanding the ways in which metaphor families fulfill a translator role for emerging new terminology in financial discourses. We have also contributed to research on the expansion and evolution of metaphor families. To map the changes as different ways of framing, we used the structural-space method to reveal subtle discursive changes of metaphor use in the three newspapers over time. Although metaphors using source domains such as *toxic* are common to discourses of all kinds, the way in which mass media, for instance, use such metaphor families to communicate financial issues and debates to their audiences plays a potentially important role in the way the public understands these issues and debates.

Our analysis showed that most of the *toxic* metaphor family variations were created in the actual *crisis* period, not in the *pre-crisis* and *post-crisis* periods. It seems that the communication needs of the *crisis* period led to the creation of new metaphors, suggesting that translating the *crisis* to the readers of these newspapers required a larger set of metaphors (i.e., a metaphor family) than the one used in the *pre-crisis* period. Our metaphor-identification method revealed *FT's toxic* metaphor family to be the richest, containing 171 unique metaphors. This can be partially explained by the high number of articles published by *FT*, but it can also be linked to the highly specialized or professional nature of this newspaper. Because *FT* specializes in reporting financial issues, it can be argued that a larger, more diversified metaphor family is required to translate these many issues to its audience. While *FT* used the largest *toxic* metaphor family, its focus was clearly on *toxic asset*, the metaphor with the highest frequency in its published articles. This also applies to *NYT*. In contrast, *Sun's* articles showed more variety in metaphor use. The top most frequent metaphors used by *Sun* were much closer in frequency, suggesting more discursive diversity.

In the second part of our analysis, we employed the structural-space method, which combines popularity and connectivity potential of concepts in semantic networks for a more comprehensive understanding of subtle dynamic discursive shifts within the investigated newspapers. Selecting the eight metaphors common to all three newspapers, we exposed shifts in the focus of individual newspapers across the three periods and differences and similarities between the newspapers. During the *pre-crisis* period, *NYT* and *FT* mostly used generic portrayals of the emerging events by using *toxic waste* as the most globally central metaphor.

In contrast, during the *crisis* period, all three newspapers focused more on metaphors characterizing financial instruments than they had on generic characterization metaphors in the *pre-crisis* period. In the *pre-crisis* period, the translation function of these metaphors was initiated with the use of *toxic waste* as the most popular and connective metaphor in both *NYT* and *FT*. This catchy, novel metaphor set the stage for the *crisis* period by eliciting negative,

familiar images for the readers of these newspapers and thus strengthening the variety of *toxic* metaphors that emerged in this period.

These findings propose that during the full-blown financial crisis, the focus of the newspapers shifted to translating the many emerging flawed financial instruments. Translating such issues to audiences with a metaphor family created a bucket into which all these instruments could be thrown, and thus they were labeled as poisonous or even deadly. This kind of translation (labeling) leaves little room for any neutral or positive associations, strengthening the negative impact of such metaphor families even further. The *toxic* metaphor family suggests a wider view on toxic capitalism (Smith, 2009), and this may have wider implications for which actions are taken to solve the crisis. *NYT* and *Sun* showed significant changes in their discourse in the *post-crisis* period, implying a discursive shift to a different and perhaps new state. The discourse of *FT* did not show significant changes between the *crisis* and the *post-crisis* periods, suggesting that by the end of 2011, *FT's* discourse had not transitioned into the *post-crisis* stage.

Because of the highly specialized character of *FT*, this particular finding raises some interesting questions regarding the messages such discourse stability sends to its audience. If the discourse of *FT* remains unchanged, should we not talk about a *post-crisis* period? Is the financial crisis not yet over?

4.6 Discussion

Rich metaphor families provide a new field of research in metaphor theory. We provide an approach through which a vast array of metaphors can be identified and analyzed in a timely manner. Our approach offers the possibility of longitudinal analysis of metaphors in semantic networks over any time frame and thus opens new possibilities for theoretical and empirical advances in metaphor-evolution research. In this sense, our analysis adapted the notion of conceptual metaphors as proposed by (Lakoff & Johnson, 1980) to a more focused and effective approach to the study of source domains, their use in discourse, and their evolution. The results show that the source domain *toxic* was applied

to increasing numbers of target domains, and at the same time, the idea of financial crisis seemed to become more conventional by the increased use of metaphors such as *toxic asset*, which was introduced into Oxford Dictionaries in 2010 (Toxic, 2012). Our findings show how such metaphors evolve from a stage in which they are used as general descriptive devices, such as "*toxic waste* of the debt markets" (Davies, 2006, p. 4, emphasis added), to the stage in which they become translating devices for unfamiliar terminology, such as "*toxic mortgage* backed securities" (Schwartz, 2010, emphasis added). While Thelwall & Price (2006) focused on extracting data about the rich family of metaphors related to the source domain of Frankenstein's monster, we have taken a step further and applied the structural-space method to the analysis of the rich family of *toxic* metaphors.

We performed automated extraction of metaphors and showed how the *toxic* metaphor family has been used differently in three news sources (*Sun, NYT*, and *FT*) in three time periods. Different from most other approaches (e.g., Birke & Sarkar, 2006; Fass, 1991; Gedigian et al., 2006; Goatly, 1997; Krishnakumaran & Zhu, 2007; Mason, 2004; Miller et al., 1990; Peters & Peters, 2000; Thelwall & Price, 2006; Ureña Gómez-Moreno & Faber, 2010; Wilks, 1978), the metaphor-identification method we used proved efficient in identifying large numbers of metaphors sharing the same source domain from large volumes of text with minimal data pre-processing and no manual coding. Based on the preferences of the analyst, the method can be adapted to extract metaphors of different lengths (e.g., by increasing the number of co-occurring concepts to include) and can also be applied to identify metaphors based on their target domain.

This type of metaphor identification has benefits, but some limitations must also be mentioned. This approach does not identify metaphors in which the target domain precedes the source domain, such as "securities that turned toxic" (Cox et al., 2009, p. B2, emphasis added). While a substantial number of metaphors are identified through this method, structurally more complex metaphors in which the target domain precedes the source domain, in which the target domain consists of multiple words, or both are not identified with

this method.

Although the structural-space method employed in this study revealed important findings regarding the use of the *toxic* metaphor family, this method is in its incipient stages, being developed and tested. For further use and testing, an interesting approach could combine the structural roles with the composition of the concepts directly connected to the leading metaphor (i.e., egocentric networks).

While our method of classifying semantic network nodes into one of the four structural roles was used to highlight only eight common metaphors, the classification may easily be broadened to the potential roles of any number of metaphors (e.g., top 10) or variations in semantic networks.

If metaphors are the lenses through which we make sense of our daily lives, then scientists from all domains must acknowledge the importance of studies that elucidate their roles and dynamics in various discourses. In the financial crisis debate, the use of metaphors is important in terms of the images they create, the meanings readers associate to the issue reported, and the potential subsequent behavior changes in the decision-making processes of these readers. Large metaphor families using source domains such as *toxic* may produce overly negative portrayals of the events through their persuasive character (Sopory & Dillard, 2002), which in turn can have consequences for the magnitude of the crisis on the financial markets by creating, for example, panic among consumers of financial products (Kleinnijenhuis et al., 2013). The growing use of *toxic* metaphors can also potentially strengthen this metaphor family across news sources by conventionalizing them – by making them part of our everyday conventional language – and by influencing the ways in which we understand text (Allbritton, 1995).

5 Europe Talks: An Analysis of Discursive Practices, Position Taking, and the Left-Right Ideological Spectrum in the 7TH European Parliament

Abstract

European Parliament speeches are an untapped source of information regarding political positions taken on important issues such as economic policy. The study of parliamentary speeches contributes to an understanding of how policy issues are debated, the discursive practices employed, and the ideological divide in the 7thEuropean Parliament. Even more, these speeches have the potential to exposes critical disparities between two of the largest political groups of the European Parliament, on such critical issues as socio-economic policy in times of crisis. Exploring speeches surrounding the recent Eurozone financial crisis, this study exposes discursive practices employed by members of the two largest political groups in the European Parliament: the European People's Party, a right wing group, and the Progressive Alliance of Socialists and Democrats, a left wing group. Using topic modeling and structural analysis of topic networks, we expose a unified discursive space, with high levels of agreement on most issue of the Eurozone financial crisis. Our results show evidence of a moderate left-right ideological divide only when the unique topics of these two groups are investigated.

Keywords:

5.1 Introduction

Language is not just a means of communicating information but a form of empowerment. In verbal or written form, language can persuade, can alter perceptions of reality, and it can influence the ways in which we give meaning

to different experiences. The power of words, language, and discourse has been acknowledged by a long tradition of research from various disciplines, from rhetorics to linguistics and to critical discourse analysis (van Dijk & Kintsch, 1983). But the power of language does not lay in the words themselves, but rather, language gains power in the hands of the powerful. Words can become tools of power and deception, and the responsibility lies with the users, who try to change the perceptions of their audiences (Wodak, 1989).

In the political realm, the power and importance of language and discourse have been vastly recognized and researched (e.g., Campbell & Jamieson, 1990; Gray & Griffin, 2014; Harris, 1991; Maynard, 1994; Seidel, 1988; Zupnik, 1994). Political activity does not exist without the use of language; political interaction requires language structures, and linguistic behavior necessarily involves structures of domination and legitimation (Giddens, 1984; Gastil, 1992). Language allows politicians to produce world views aligned with their own goals, and to demote, refute, or exclude alternative representations. Through various language strategies, such as the use of metaphors (Bosman, 1987; Lakoff, 1995; Musolff, 2004), framing and reframing (Entman, 1993; Lakoff, 2004), emphasis, or repetition (Wilson, 2015), politicians are able to accentuate certain semantic features of a given utterance, while at the same time obscuring others. In other words, in political arenas, language is used to bring a certain perspective to the fore and compete over the establishment of dominant perspectives.

The present paper focuses on a source of political discourse that has been seldom explored, namely speeches of the plenary sessions in the European Parliament (EP). Our aims are to uncover the dominant discursive practices and the positions taken by the two largest political groups of the 7^{th} EP in the context of the recent financial crisis. While previous research has addressed many aspects of the complex and multidimensional character of the EP (e.g., coalition formation in the EP (Kreppel & Tsebelis, 1999), the party system of the EP (Kreppel, 2002; Hix et al., 2003), voting behavior (Attiná, 1990), the empowerment of the EP (Hix & Høyland, 2013) etc.), research on the manner in which the recent financial crisis has reflected in the discursive space of the EP is still in its incipient stages. In order to address this research gap, we focus

on three main questions: (1) what are the *dominant discursive practices* of the two largest political groups of the EP in the context of the financial crisis? (2) what are the *positions* these groups take on issues of the financial crisis? and (3) is there evidence of a left-right ideological divide in the speeches that specifically address the *financial crisis*? Firstly, uncovering the dominant discursive practices of the two political groups that together hold 61% of the EP seats leads to a better understanding of how the financial crisis was portrayed, the themes and topics that were mostly emphasized, and hence, the positions these groups occupy in relation to the financial crisis. As an equal partner of the EU Council in almost all policy areas, including internal market legislation, the new and 'empowered' 7$^{\text{th}}$ EP played a key role in the EU's efforts to combat the rippling effects of the Eurozone crisis through the large number of socio-economic legislative proposals adopted in a relatively short time. Thus, understanding the EP political space and the policy positions of its members is of paramount importance. Secondly, when investigating a complex and disputed discursive space as that of the EP, we must also recognize the ideological diversity of its members. Whilst the two political groups we focus on are right-wing and left-wing oriented, and their ideological positions on the left-right spectrum have been established by previous research (e.g., Proksch & Slapin, 2010), we reveal the persistence of this divide when issues of the financial crisis are debated. Furthermore, while previous research has revealed the presence of the right-left divide based on party manifestos (e.g., Gabel & Hix, 2002), roll-call voting (e.g., Hix et al., 2005), and expert surveys (e.g., McElroy & Benoit, 2012), we demonstrate the prominence of this divide in those speeches (of their members) that specifically address issues of the financial crisis.

Before addressing our research questions, we will discuss the new and empowered role of the 7$^{\text{th}}$ EP in the financial crisis oriented policy adoption process in the European Union (EU), the structure and roles of the political groups of the EP, and the issue of political ideology surrounding these groups. Lastly, while acknowledging previous research investigating party manifestos, roll-call voting, and expert surveys, we present an automated method which allows for the identification of dominant themes and topics in large text corpora, in our

case a large collection of speeches.

5.2 Background

As the only directly-elected institution of the EU, the European Parliament (EP) is one of the most powerful and influential of the EU bodies (Hix, 2011). Since its establishment as the Assembly of the European Coal and Steel Community, the EP has gone through a process of empowerment. This process began with a series of formal treaties and rulings of the European Court of Justice relating to the legislative powers of the EU institutions and the evolution of the EP's own "Rules of Procedure." Today, the EP is an equal partner of the EU Council in almost all policy areas, including internal market legislation, adopts or amends proposals from the European Commission (EC), supervises the work of the EC, and adopts the EU's budget. The EP's powers in the area of financial regulations were further increased through the Lisbon Treaty, which ensured an even greater role for the 7th EP.

Elected just months before the first signs of the Eurozone crisis emerged, the 7th EP was soon faced with increasing financial instability on the European markets, accelerating debt levels of EU member states, and banking system bailouts. The newly elected EP became an important actor in the policy making process of the European Union during the Eurozone financial crisis, carrying out a rapid adoption of legislative proposals (see Annex 1) that restructured the EU's financial socio-economic system. These proposals resulted in fundamental changes to the financial system and to the manner in which institutions do business (Broin, 2012). From the examples of the 7th EP legislative's actions presented in Annex 1, it becomes clear that the EP's involvement in policy decisions having direct impacts on the European financial markets spans a wide range of areas and issues. The new and 'empowered' 7th EP played a key role in the EU's efforts to combat the rippling effects of the Eurozone crisis through the large number of socio-economic legislative proposals adopted in a relatively short time. Thus, understanding the EP political space and the policy positions of its members is of paramount importance.

In terms of decision-making, the EP is assumed to be a unitary actor because decisions are made either by simple majority of the members present or by absolute majority of all Members of the European Parliament (MEPs). However, once elected[1], MEPs generally organize into political groups along political affiliation lines of their national parties.[2] Currently there are seven transnational groups within the European Parliament[3], with the two largest groups being the European People's Party (EPP) and the Progressive Alliance of Socialists and Democrats (S&D). Theses two groups, the EPP and the S&D, are the groups this paper focuses on.

During the 2009–2014 term, the EPP has 265 MPE members and brings together centre and centre-right pro-European members from EU states. On the other hand, the S&D is a centre-left political group with 184 MEP members from all 27 EU countries. These two political groups have been in existence for over 50 years and have (together) continuously controlled between 50 and 70% of the seats in the EP (Kreppel, 2002).

In the past decade, the balance of power in the EP has remained stable with the EPP's overall share of seats remaining surprisingly constant at around 36 percent and the S&D's overall share of seats between 25 and 28 percent (McElroy & Benoit, 2012). This power balance has not been radically affected by the EP elections in 2009.

As the two largest political groups of the EP, the EPP and the S&D are important actors in the legislative process of the 7th EP by holding 61% of the seats, and hence the votes in the parliament. Moreover, these groups are also major players in other decision-making bodies of the EU, and their influence is a direct function of their size (McElroy & Benoit, 2007). Because these two groups are highly influential and strongly impact the policy-making process of

[1]The seats in the EP are shared out proportionately to the population of each member state of the EU, with the maximum seats per member state being 96 (Germany) and the minimum being six (Malta).

[2]Some MEPs do not belong to any political group and are known as non-attached members.

[3](1) The European People's Party, (2) the Progressive Alliance of Socialists and Democrats in the European Parliament, (3) the European Conservatives and Reformists Group, (4) the Alliance of Liberals and Democrats for Europe, (5) the Confederal Group of the European United Left - Nordic Green Left, (6) the Group of the Greens/European Free Alliance, and (7) the Europe of Freedom and Direct Democracy Group.

the EP, there is a need to understand how they position themselves within the debate of financial crisis, the topics and issues they emphasize when discussing the crisis, and hence, and the discursive practices they employ. However, these two political groups are situated on different sides of the ideological spectrum, with the EPP being a center-right group, while the S&D is a centre-left group.

Ideology of EP Political Groups

Political ideology is a concept that has prompted countless definitions in political research (e.g., Gerring, 1997; Jost, 2006; Jost et al., 2009; Lane, 1967). For the purpose of this study, we employ a text-book definition of political ideology offered by Erikson & Tedin (2003): "a set of beliefs about the proper order of society and how it can be achieved" (p. 64). Furthermore, ideologies structure and communicate the shared beliefs, opinions, and values of an identifiable group, class, constituency, or society (Freeden, 2001; Knight, 2006). Ideology functions as a way of structuring political knowledge and expertise. Although political ideology can be categorized in multiple ways, our focus revolves around the left-wing and the right-wing ideology divide.

Within the right-wing spectrum, views differ on whether hierarchy and inequality stem from traditional social differences or from competition in market economies (Adams, 2001). The right-wing ideology proposes that forms of social stratification or social inequality are either inevitable, natural, normal, or desirable, justifying this position on the basis of natural law or tradition (Bobbio, 1996; Carlisle, 2005; Lukes, 2008). On the other hand, the left-wing political ideology is often characterized as acceptant or supportive of social equality and egalitarianism, and often opposes social hierarchy and social inequality. Typically, it involves concern for those in society perceived as disadvantaged and a belief that there are unjustified inequalities that need to be reduced or abolished (Bobbio, 1996; Lukes, 2008; Smith & Tatalovich, 2003; Thompson, 1996). Nevertheless, the two fundamental facets of the left-right ideological divide are change versus stability and equality versus inequality.

Research has found that EP "party groups not only occupy the entire range of the left-right spectrum, but also are clearly distinguishable from one another

in policy terms." (McElroy & Benoit, 2012, p. 1). Also, the evolution of party positions over time has been found to be remarkably stable (McElroy & Benoit, 2012). This attests to the ideological and political diversity of the EP, a policy space that became increasingly organized and competitive in the past decades (Hix et al., 2005).

The 2010 expert survey conducted by McElroy & Benoit (2012)[4] confirmed the right-of-center and left-of-center positions of the EPP and S&D in the ideological space of the EP. Also, the policy positions of these two groups have been found to be at the centre of the distribution of the positions of their member national parties (McElroy & Benoit, 2012), suggesting a degree of ideological coherence among the members of these political groups. On issues related to economic, social, and environmental regulation of the single European market, the EP political groups are also more likely to be divided along left–right lines (Hix et al., 2003, 2005).

In this article, we do not take a stance on whether ideological orientations reflect the personal belief systems of specific politicians, but rather we assume that ideology ultimately influences political behavior, such as discursive practices. The studies mentioned above investigated the right-left ideological spectrum of the EP based on expert surveys (McElroy & Benoit, 2012) and roll-call votes (Hix et al., 2003, 2005), while others have exposed the traditional left-right divide of the EP by investigating party manifestos (e.g., Gabel & Hix, 2002). In this study, we bring a contribution to the area of political ideology research by revealing the manifestations of the left-right divide in the dominant discursive practices of the EPP and the S&D surrounding the financial crisis. This is to say that when exposing the dominant discursive practices of these two groups, we draw a parallel between the themes and topics they emphasize and their ideological orientation. We expect the left-right divide to become even more salient in times of crisis, when the EP has to vote on economic policy. For instance, while left-wing ideology calls for a welfare state with a nationalized economy, the right-wing ideology supports capitalism, economic freedom, and

[4]The survey places political groups on left-right scale from 1 to 20, where 1 indicates far left ideological positions, and 20 indicates far right positions. On this scale, the EPP scored 13.5, and the S&D scored 7.8.

a decentralized economy. It is then plain to see how this divide would play an important role in the policy space of the EP and in the discursive practices of the EPP and the S&D when addressing policy issues meant to mitigate the effects of the financial crisis.

Political Speeches

As mentioned above, previous research into the European Parliament has mostly focused on party manifestos, expert surveys, and roll-call votes. But one of the of the ways in which EP political groups make their positions known (on various legislative issues) is through speeches given by their members in the plenary sessions of the EP. These sessions occur every month for a week, and the largest proportion of speaking time during these sessions is allocated to the political groups of the EP. Each political group receives speaking time (roughly) in proportion to its seat share. Party groups decide internally how to divide time among their MEPs, with the time for individual speeches being strictly limited, usually not more than three minutes.

Through speeches and the discursive practices employed, MEPs offer arguments that support their positions (and hence the positions of their political groups) on the issues debated. Political actors use speeches to "manipulate, strategize and fight to have their frame accepted as the dominant narrative." (Boin et al., 2009, p. 82) This is to say that through speeches and discursive strategies, such as frames or emphasis through repetition, a political group may effectively control discussion and perceptions of an issue. Repetition, or emphasizing of issues, helps embed specific interpretations, and once particular representations are established they are hard to shift (Wilson, 2015).

Just as in the case of the EP, previous research on policy positions of other political parties or groups has also mainly focused on party manifestos (e.g., Benoit et al., 2009; Franzmann & Kaiser, 2006; Gabel & Huber, 2000), voting behavior (e.g., Hix et al., 2003, 2006; Poole & Rosenthal, 2000; Snyder Jr & Groseclose, 2000), and expert surveys (e.g., McElroy & Benoit, 2010; Ray, 1999). In most cases, the data for such studies has been obtained from the Manifesto Research Group (MRG), group which has developed its own coding

scheme and has analyzed nearly all manifestos of political parties from 50 countries covering all free, democratic elections since 1945. However, most of these studies involve a degree of manual coding and/or human reading of the full texts, the use of pre-developed techniques of transforming text documents into numerical data (Gabel & Huber, 2000), or the involvement of expert coders. These approaches to the analysis of political texts, although valuable, involve time consuming coding techniques, coder bias, and limitations in terms of the size of the corpora analyzed.

In this paper we employ an exclusively automated method to text analysis, capable of extracting the predominant recurring topics in text corpora of any size. Such automated approaches to the analysis of speeches in the EP have not been widely used. One such study is the one of Høyland et al. (2014), study which automatically predicted party group affiliation of participants in European Parliament based on the content of their speeches using a support vector machine multi-class model. Another study comes from Proksch & Slapin (2010), and it estimated the principal latent dimension of spoken conflict using word counts from legislative speeches of the EP, and how national parties and their members position themselves in the EP. Joining this underdeveloped research stream of automated analysis of speeches in the EP, we employ topic modeling on a collection of approximately 4000 speeches given in the plenary sessions of the 7^{th} EP on the issues relevant to the financial crisis. A more detailed description of the analysis methods employed is given below, after we describe the data and the data collection process.

5.3 Data

As mentioned above, this study focuses on the speeches given by members of the two largest political groups of the European Parliament, namely the EPP and the S&D. Hence, we have collected all the speeches given by the members of these two political groups between the 14th of June 2009 and the 30th of June 2014 from the Talk of Europe project database (van Aggelen & Hollink, 2015). The interval selected represents the mandate of the of the 7^{th}

European Parliament, which has been elected in June 2009. The first meeting (constitutive session) of the Seventh Parliament took place on July 14, 2009, when Jerzy Buzek was elected as President of the European Parliament.

To assess the ways in which members of the two political groups position themselves in relation to the financial crisis, and what discursive practices they employed in regards to the crisis, only those speeches containing "financial crisis" and/or "economic crisis" have been collected. After duplicate removal, a total of 2499 speeches given my members of the EPP and 1456 speeches given by members of the S&D have been included in the analysis. In Figure 1 we present an overview of the number of speeches addressing financial crisis issues between 2009 and 2014 for each of the two political groups. We must note that because the 7^{th} European Parliament started its legislative activities in June 2009, the speeches collected for the year 2009 span approximately six months of data. This is also the case for the year 2014, when the 7^{th} European Parliament was replaced by the 8^{th} Parliament at the end of June. This is clearly visible in Figure 5.1, in which the counts for 2009 and 2014 are very low. However, the low number of speeches given by both parties in 2013 (on the topic of the financial crisis) is not an artifact of the data collection process. This is an interesting finding considering the financial troubles in the European Union were far from over in 2013, when the precarious state of the Greek economy was still posing a major challenge. Furthermore, unemployment rates in the EU grew to a record high of 27.9% in June 2013, youth unemployment rate rose to as high as 62% (Eurostat, 2013) , and the ECB lowered its bank rate to only 0.25% to aid recovery in the Eurozone (ECB, 2013).

5.4 Methods

Now that we have described the data collection and the scope and aims of this paper, we continue by presenting the methods used in the analysis of our text corpora. We begin by introducing topic models, followed by the structural space approach as an extension of semantic network analysis.

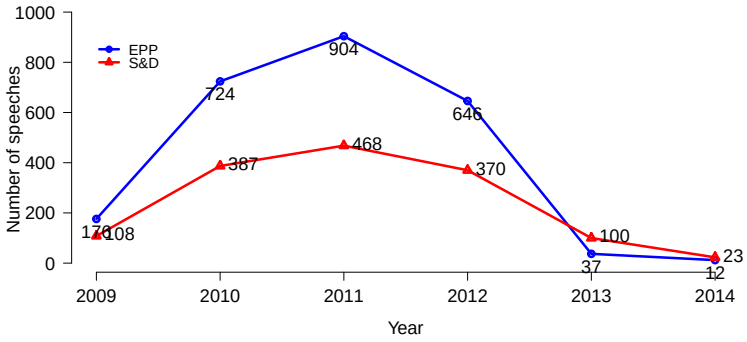


Figure 5.1: Speeches collected for each political group

5.4.1 Topic Modeling

Although much of the EP parliamentary speeches are structured by agenda items, we set to investigate the most predominant topics debated by each political group through topic modeling. Uncovering the most prevalent topics in the speeches of the EPP and S&D helps us better understand the EP discursive space, the discursive practices of these two political groups, and the ways in which they discursively define their positions on the issue of the financial crisis. For this purpose, we use latent Dirichlet allocation (LDA) topic models, a class of automated text analysis tools that seek to identify, extract and characterize the various (latent) topics that are discussed in our collection of speeches.

Topic models provide efficient methods of analysis for large collections of unstructured text. A cluster of words that co-occur frequently across a number of documents constitute a "topic". Using contextual clues, topic models connect words with similar meanings and differentiate between uses of words with multiple meanings. More technically, topic modeling is based on the idea that documents are collections of topics, where a topic represents a probability distribution over words. Each topic is separately meaningful, offering a probability distribution over words which produces a consistent cluster of correlated terms (Blei et al., 2003; Griffiths & Steyvers, 2002, 2003, 2004; Hofmann, 1999, 2001). First described and implemented in the context of natural language processing, topic models use algorithms designed to browse and summarize large archives

of texts. In this paper we employ latent Dirichlet allocation (LDA) (Blei et al., 2003), a three-level hierarchical Bayesian model used to retrieve information contained in large collections of texts. Widely used, topic models have emerged as a powerful new technique for uncovering underlying semantic structure in otherwise unstructured corpora, and they are highly useful exploratory tools in text analysis.

We fit the topic model to the entire collection of speeches of each political group. This procedure will help identify the principal topics in the discourses of the EPP and the S&D surrounding the financial crisis in the 7^{th} European Parliament. After removing a standard list of stop-words from the 2499 speeches of the EPP and 1456 speeches of the S&D and after exploring different levels, we settled on a 15 topic solution and 1000 iterations.

Topic Networks

Furthermore, we also assess the level of emphasis the EPP and S&D place on each of the topics identified and the diversity of their discourse by generating topic by topic networks, in which the links represent concepts shared by the topics. Thus, if a word is a member of two topics, a link will be formed between these topics. The values of each link in the topic by topic network represents the number of shared words between topics. The exploration of the topic by topic networks is also done through summed frequencies of the topic members (words) but also through the structural space approach, an extension of semantic network analysis upon which we elaborate below.

Structural Space

The structural space approach draws on semantic network analysis and a combination of centrality measures of semantic network nodes to uncover subtle structural properties of discourse (Nerghes et al., 2014b). This approach to semantic network analysis has been successfully applied to newspaper corpora to dynamically track the evolution of metaphor families (Nerghes et al., 2015a), and to uncover subtle changes of highly structured (formal) organizational discourse (Nerghes et al., 2014b, 2015b). The distinctive feature of the structural

space approach is the identification of four structural roles of semantic network nodes based on the combination of two structural measures: popularity (i.e., total degree centrality) and connectivity (i.e., betweenness centrality). This combination positions the concepts within this structural role space. The structural space builds on the manner in which popular and connecting concepts play different roles in the structure and dynamics of semantic networks.

Because the links in our topic networks represent shared-concepts among topics, we need to briefly consider the meanings of degree centrality and betweenness centrality, and hence the characteristics of each structural role in the context of these topic networks. First, degree centrality in our shared-concepts networks represents the number of concepts a topics shares with other topics, and thus it becomes a similarity measure. A topic with high degree centrality has higher similarity to one or more topics, than a topic with low degree centrality. On the other hand, betweenness centrality in these shared-concepts networks, denotes the connectivity potential of the topic. A topic with high degree betweenness is a gateway to other parts of the topic network, while a topic with low betweenness centrality does not fulfill a connective role.

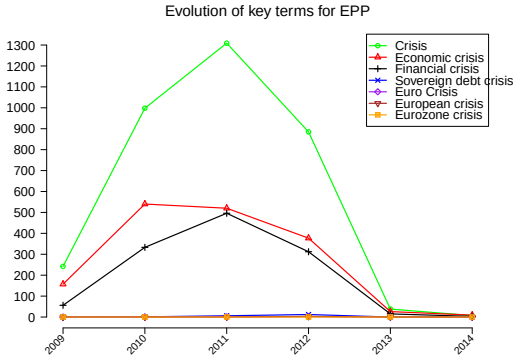
While in semantic networks the four structural roles characterize each node based on popularity and connectivity potential, in the particular case of these topic networks the structural roles will characterize topics based on similarity (rather than popularity), and connectivity potential. Thus, the *Globally Central (GC)* role, which includes topics with both high degree centrality and high betweenness centrality, will identify topics with high similarity to other topics due to a high number of shared concepts (degree centrality), but also topics with connective roles due to their high betweenness centrality. The *Locally Central (LC)* role highlights topics with high degree centrality and low betweenness centrality. The LC role will include topics with high similarity but low connective potential. Topics with low degree centrality and high betweenness centrality become part of the *Gatekeeper (G)* role. The G topics, probably the most informative for this type of networks, are topics with high connective potential and low similarity to other topics. Lastly, *Marginal (M)* topics have low similarity to other topics and low connectivity potential, which

could be indicative of unique or very distinct topics. M topics are topics with both low degree centrality and low betweenness centrality.

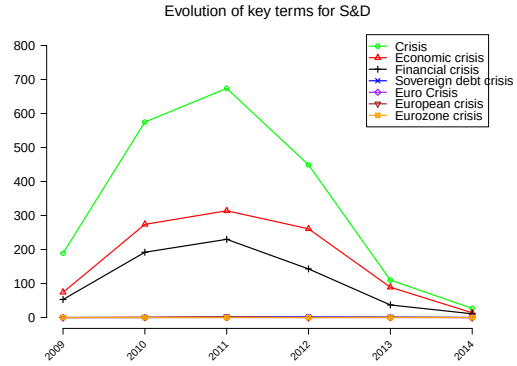
5.5 Results

We begin reporting our results by highlighting some general findings in the discourses of the EPP and S&D, followed by the topic modeling results, the topic by topic networks, and the structural space analysis of the topic networks.

General



(a) EPP



(b) S&D

Figure 5.2: Evolution of key terms

An interesting finding (shown in Figure 5.2) is that neither the members of the EPP, nor those of the S&D use such key concepts as *'Euro Crisis', 'European Crisis', 'Eurozone Crisis'*, or *'Sovereign debt crisis'* in their discourse. While concepts such as *'Crisis', 'Economic crisis'*, and *'Financial Crisis'* are abundant in the discourse of these two political groups, very little association is made between the crisis and european matters. This finding indicates there was very little focus on the crisis as an European event. The members of the two largest political groups of the EP did not frame the crisis as an European event, but rather employed terminology of a more general nature to characterize the eurozone crisis.

Topic Modeling

By fitting the LDA topic model to our corpora of speeches, we identify the 15 principal topics in the discourses of the EPP and the S&D surrounding the financial crisis. We present each of these topics in Tables 5.1 and 5.2, where the 'Topic members' column lists the most representative 10 words for each of the topics. The 'Topic' column lists the topic labels, which are not provided by the procedure but they have been assigned by us as a way of summarizing the character of these topics. To exemplify, we assigned the 'Gender Equality' label to the first identified topic for the EPP corpus to summarize the following set of most probable words: Women, Men, Equality, Gender, Right, Poverty, Equal, Mother, Economic, and Life.

Comparing the topics that emerged from the EPP and the S&D speeches, we note that ten of the topics are common to both discourses and five are unique. While the common topics may in part driven by the structure of the European Parliament sessions, they are also driven by the agenda-setting abilities of the two political groups. Each political group is able to bring the requests of his/her group regarding agenda items. Also, during the meetings of the political groups in the week prior to the plenary session, they scrutinize the draft agenda and decide wether to push for further changes. The drawing of the EP agenda is an elaborate process, which allows a degree of liberty for political groups to decide on the issues they bring forward for debate.

Irrespective of the way in which these topics were included on the agenda of the EP, the high number of common topics emerging from the discourses of these two political groups is a noteworthy and important finding. This high number of common topics reveals a unified discursive space in the EP and it indicates a high level of agreement between these two groups in the debates over the financial crisis. However, the presence of these topics does not divulge the position of the political groups discussing them. In other words, a topic that is present in both discourses can also be discussed in opposing terms. To further explore this particular finding and the level of agreement on these topics, we present examples of speech fragments for each political group and each of the common topics. These fragments were selected randomly from our data sets and we present them in Section 5.C. This closer investigation into the discursive practices employed by the MEPs of these groups on all of the ten common topics confirms our initial finding. In fact, the similarity of the positions taken by these two groups on all of the common topics creates a discursive space in which no evidence of the left-right ideological divide can be found. Hence, we can conclude that (1) these ten topics are of common interest to both parties in the debate of the financial crisis, and (2) that when debating these issues of common interest, the EPP and the S&D form a common front which is not impacted by political competition nor ideology.

While the common topics of the EPP and the S&D reveal unified and non-ideological discursive practices, the same is not true for their unique topics. Most of the five unique topics of each political group are highly representative of their ideological orientation and they are indicative of the individual and distinct positions these groups occupy in relation to the financial crisis.

For the S&D, the unique topics are 'European Crisis', 'Banking System', 'Taxation', 'EU Accession', and 'Labour Rights'. As a center-left group, the focus of the S&D is on a 'regulated economy' via governing structures as shown by topics such as 'Banking System' and 'Taxation'. Also, the 'EU Accession', and 'Labour Rights' topics can be linked to the political orientation of the S&D as a left-wing party, orientation that encourages egalitarianism and social equality. Interestingly, while the speeches collected for the purpose of this

article are those that focus on the financial crisis, the topic of the European Crisis is only present in the discourse of the S&D.

In the case of the EPP, the unique topics are 'Globalization', 'Human Rights', 'Growth', 'Fishing Industry', and 'EU Governance'. The center-right orientation of this group suggests a focus on maintaining the status-quo, economic freedom, and a decentralized economy. That being said, the presence of 'EU Governance' as one of the unique topics of the EPP is at the very least intriguing. To further explore the context around this particular topic, we show part of a speech given by Joseph Daul, the president of the EPP, in the plenary session of the EP. In this speech, the position taken by Daul on the issue of 'EU Governance' becomes clear and it is highly representative of the ways in which members of the EPP addressed this topic in their speeches[5]:

> "In a period of growth it is generally felt that one has the right to conduct one's own budgetary, fiscal, and social policy without really worrying about anyone else. However in a period of crisis those who have spent the most call for solidarity from those who have been let us say more sensible. Can this continue? I do not think so. It is time for the Member States to coordinate their budgetary, fiscal, and social policies better and let us not be afraid to say it: we want more European governance!" [speech given on 24 March 2010 in the plenary session of the EP]

In the above speech, the position taken by the president of the EPP seems at the very least contrary to what one would expect from a right-wing political group. Joseph Daul calls for 'more European governance' and coordination of 'budgetary, fiscal, and social policies' of Member States as a strategy to combat the effects of the crisis. Essentially calling for an increase in EU involvement in combating the effects of the financial crisis, the speech of Joseph Daul contradicts the right-wing support for limited role of governing bodies into the financial markets and a decentralized economy.

The 'Globalization' topic in the EPP discourse can be linked to the pro free-trade inclinations of the right-wing politics, while the topic of 'Human Rights' can be linked to the right-wing support of personal liberties and moral-order.

[5]Due to space limitations, we selected the speech of the EPP president as an example. However, the speeches of other EPP members are similar in the stance taken towards the topic of European Governance.

The topic of 'Growth' as a prominent topic in the EPP discourse indicates the support of capitalism of a right-wing political group.

In sum, the unique topics of the EEP and the S&D revealed those issues that were distinct in their crisis-oriented discourses. Unlike the common topics, these unique topics also provide evidence of the left-right ideological differences between these two groups. However, the presence of 'EU Governance' as a unique topic in the discourse of the EPP raises intriguing questions regarding the demand for increased European governance from a right-wing political group.

Next, we move further to explore the emphasis the two political groups put on the topics revealed, the presence of these topics in their respective discourses, and the structure of shared concepts among topics. In order to reach these goals, a two step procedure was employed in transforming our topic tables into topic by topic networks.

The first step involved a transformation of the topics and the words that belong to each topic into topic by concept networks for each political group (see Section 5.B). While describing these concept by topic networks would be a repetition of the information already presented in Tables 5.1 and 5.2, we must note the concept counts (94 for the EPP and 98 for the S&D) and the densities of these networks (0.106 for the EPP and 0.102 for the S&D), which suggest a slightly more repetitive discourse of the EPP. The links counts of the two networks are identical (150). In the second phase, we further transform (fold) our concept by topic networks into topic by topic networks. In these networks the links represent shared topic members, and the value of each link indicates the number of shared members.

We present our topic by topic networks in Figures 5.3a and 5.3b, where each topic is sized by the sum of the frequencies of its top 10 topic members, and the links are weighted by their values. Based on these networks, we clearly see that the most emphasized topics in the speeches of EPP members are 'Growth' and 'EU Budget'. The S&D members, on the other hand, focused on 'EU Budget' and 'Market Regulations' in almost equal measures. These prominent topics are what we call the dominant discursive practices of these two political groups,

Table 5.1: Topic modeling of the EPP corpora

Topic	**Topic members**
Gender Equality^{T0}	Women Men Equality Gender Right Poverty Equal Mother Economic Life
Globalization^{T1}	Global Fund European Globalization Adjustment Crisis Redundant Vote Financial Economic
Human Rights^{T2}	European Union Country Right Political Violation Human Economic Support Agreement
Global Aid^{T3}	Country Trade Economic Global International World Development Developing Aid Agreement
Agriculture^{T4}	Food Financial Agricultural European Farmer Supervision Sector Authority Price Crisis
Labour Market^{T5}	People Education Young Social Labour Member Poverty Unemployment Employment Crisis
Growth^{T6}	Economic European Crisis Vote Financial Europe Policy Union Growth Strategy
Fishing Industry^{T7}	Sector Ireland Transport Fishing Industry Economic Corruption Country National Fishery
SMEs^{T8}	Market SMEs Crisis Member Competition Enterprises Tax Medium-sized Commission Business
Climate Change^{T9}	Energy Research Innovation Climate Industry Environmental Development Change Emission Efficiency
EU Budget^{T10}	European Budget Member Financial State Crisis Commission Vote Fund Economic
EU Governance^{T11}	Economic Crisis Euro Member State Area European Governance Financial Council
Labour Migration^{T12}	Economic European Crisis Globalisation Company Region Resource Worker Support Mobilisation
Parliament Debates^{T13}	European President Europe Debate Parliament Crisis Commission Council Union Work
Market Regulations^{T14}	Financial Crisis Vote European System Market Tax Regulation Credit Risk

Table 5.2: Topic modeling of the S&D corpora

Topic	**Topic members**
European Crisis^{T0}	Crisis Economic State Financial Member European Euro Public Area Country
Market Regulations^{T1}	Economic European Crisis Vote Market Policy Member Current Union State Financial
Agriculture^{T2}	Food Programme Agricultural Price Aid Deprived Million Crisis Farmer Sector
EU Budget^{T3}	Budget Policy Financial European Cohesion Fund Resource Member Programme Region
SMEs^{T4}	Energy European SMEs Transport Economic Sector Job Financial Efficiency Innovation
Gender Equality^{T5}	Women Men Gender Equality Crisis Pay Care Economic Leave Work
Climate change^{T6}	Trade Climate Agreement Change Global Environmental International World Country Negotiation
Banking System^{T7}	Financial Credit Banking Bank System Company Crisis Rating Supervision Market
Taxation^{T8}	Tax Financial Sector Transaction Market Taxation Crisis VAT System Transaction
Labour Market^{T9}	Social People Education Unemployment Crisis Economic Employment Young European Poverty
EU Accession^{T10}	Enlargement Europe Country European Democracy People Hungary Cyprus Accession Treaty
Labour Rights^{T11}	Right State Member Health Human Worker Economic Protection Labour Service
Parliament Debates^{T12}	European Debate President Parliament Crisis Europe Commission Union Council Work
Global Aid^{T13}	Countries Developing Cooperation World Security Defence Global Aid Crisis International
Labour Migration^{T14}	Worker Crisis Financial Fund European Globalization Support Economic Migration Redundant

which reveal the positions these groups occupied in relation to the financial crisis debate in the 7th EP. The emphasis the EPP placed on 'Growth' is not surprising, and hints towards the right-wing orientation of this group. The political right-wing considers growth as a key factor for economic stability, and more specifically, growth through investment. To exemplify how this emerges in the EPP discourse, we show a fragment of a speech given by Joseph Daul in the State of the Union debate (also shown in Section 5.C):

> "The European budget is not a budget of spending but a budget of investing. It is a forward-looking budget, a budget for growth." [speech given on 12 September 2012]

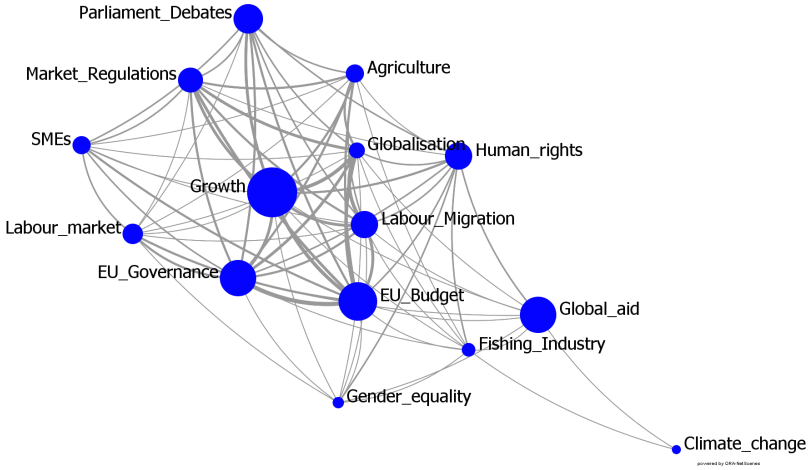
In this particular speech, Daul also touches upon the second most emphasized topic of the EPP, namely 'EU Budget'. This is a topic that is also very prominent in the discourse of the S&D, and as shown below in the speech given by Martin Schulz (also shown in Section 5.C), the president of the S&D between 2004 and 2012 (and again in 2014), the positions taken by these groups on the issue of the EU budget are very similar. Both groups view the EU budget as an investment budget meant to foster economic growth, a finding which does not seem to bear any evidence of a left-right cleavage.

> "The EU budget is an investment budget that is used to leverage economic growth and create jobs. Anyone wielding the shears is robbing us of our common future." [speech given on 9 May 2012]

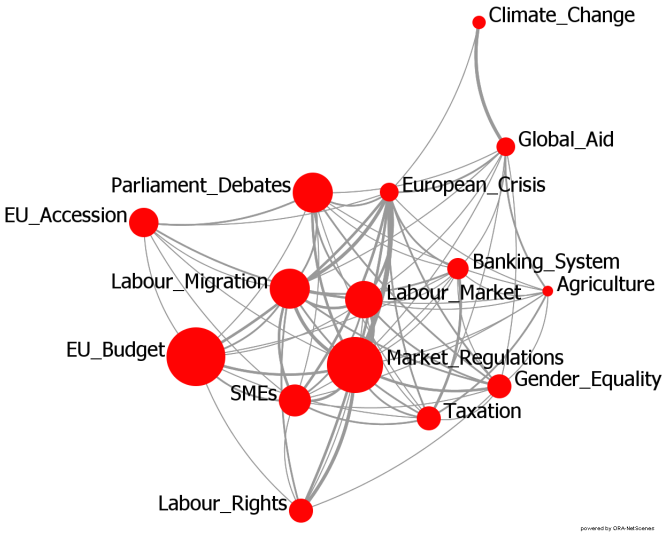
While the approach to the EU budget is a unified one leaving little room for interpretation in terms of political ideology, the focus of the S&D on 'Market Regulations' clearly points towards the left-wing orientation of this group, which emphasizes central planning, nationalized economies, and an anti-free trade movement.

5.5.1 Topic Structural Space

In figures Figures 5.4a and 5.4b, we plot the structural space of the two topic networks. The nodes (topics) are colored based on their frequency (the sum of the frequencies of their top 10 topic members), with red representing the



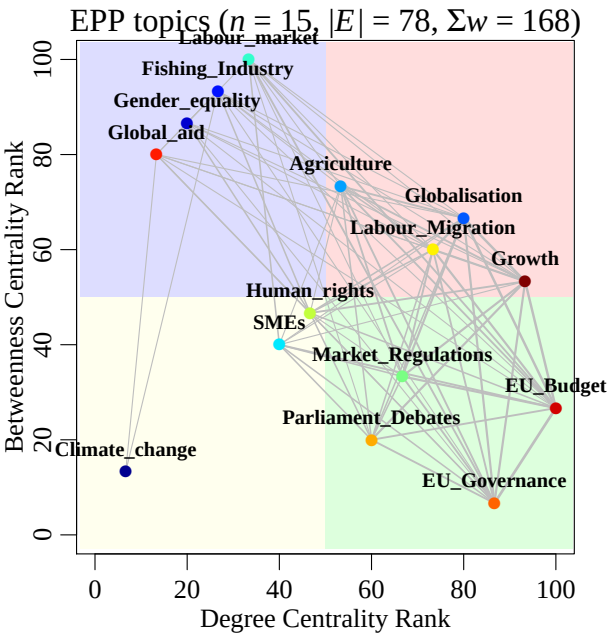
(a) EPP topics



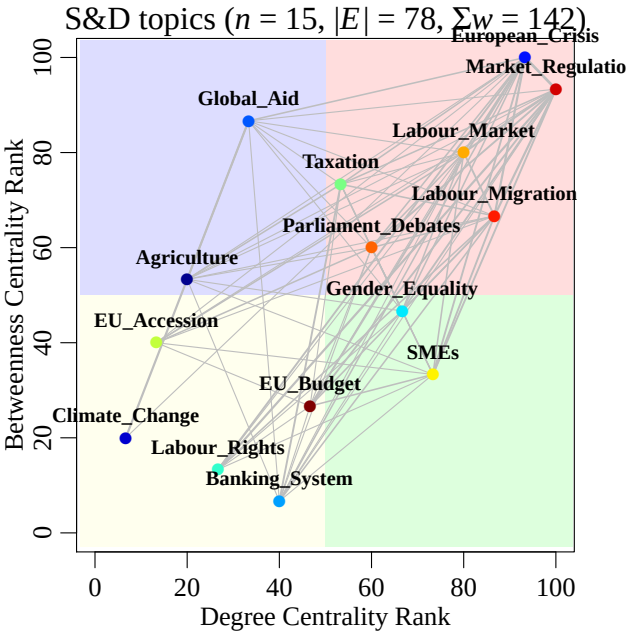
(b) S&D topics

Figure 5.3: Topic presence in the EPP and S&D discourses

highest frequency and blue the lowest frequency. These structural space plots reveal differences between the topic network of the EPP (Figure 5.4a) and that of the S&D (Figure 5.4b). While most of the topics in the EPP discourse are clustered in the GC and LC quadrants, the S&D's topics are spread across the GC and M quadrants. This finding suggests a notable variance in the emergence



(a) EPP topics



(b) S&D topics

Figure 5.4: EPP and S&D topics structural space

of overarching themes in the two discourses. Based on the clustering of topics in these areas, we can argue that the roles of the topics uncover the overarching themes in the discourses of the EPP and S&D, as well as the very distinct themes in these discourses. Because the topics in the Locally Central and the Globally Central quadrants have high degree centrality (a high number of shared words), the topics falling into these roles can be considered to be part of an overarching theme. Conversely, the topics in the Gatekeeper and Marginal roles are distinct ones due to their low number of shared words. Based on this categorization of topics, the structural space approach uncovers interesting characteristics of the two discourses. On the one hand, the EPP's topics are highly clustered in the high similarity area, indicating an overarching theme that spans ten of their topics. Interestingly, four of the five unique topics of the EPP are part of this overarching theme. On the other hand, only eight of the S&D topics are part of the high similarity cluster. Also, only two of their unique topics are highly similar to other topics, while the other three are part of the Marginal quadrant, and hence bear little similarity to other topics.

At the top of each graph, besides the number of topics n, we also report the number of links in the entire topic network ($|E|$) and the sum of the link weights (Σw). These measures are indicative of the overlap between the topics in the network. $|E|$ is the extent of concept sharing or the breadth of topics that share concepts, while Σw represents the extent of total sharing of concepts. We also calculate the $\overline{w} = \frac{\Sigma w}{|E|}$ (average link weight), which indicates the average extent of concept sharing. Interestingly, the number of links $|E|$ or breadth of sharing are equal for the EPP and the S&D topic networks. This is most likely coincidental given that nothing, as far as we can tell, forces the link counts (breadth of sharing) to be equal. While the predefined number of topics and concepts in the topics will define the lower and upper bounds of $|E|$ (in this case 0 and 105 respectively),[6] they do not isolate $|E|$ to a specific quantity. So then, the breadth of sharing is contingent upon the structure of the text documents, constrained by the number of topics and concepts in topics. However, it

[6]The upper bound is simply the count of links in a maximally connected graph or clique. For our undirected topic networks, this upper bound is $n(n-1)/2$ where n is the number of topics (15).

is reasonable to believe that the breadth of sharing will bear some similarity in political discourse because this type of discourse is structured and it includes overlapping topics within and across documents. This assumption bears further investigation which is beyond the scope of this paper. Furthermore, the average link weights of the two networks, $\overline{w}_{\text{EPP}} = 2.15$ and $\overline{w}_{\text{S\&D}} = 1.82$ respectively, reveal a higher average extent of concept sharing in the topics of the EPP compared to those of the S&D, which suggests a slightly more repetitive discourse of the EPP.

To summarize, topic modeling in combination with the structural space approach have exposed important characteristics of the EPP and the S&D discourses on the issues of the financial crisis. We have revealed a unified discursive space with high levels of agreement through the high number of common topics addressed by these groups in the financial crisis debate. Through the unique topics of their discourses, we have also shown that in the debate over the financial crisis the focus of these groups was in different areas in accordance with their position on the left-right ideological spectrum. Furthermore, we find the discourse of the EPP more repetitive and structured than that of the S&D, and we further confirm this through the higher similarity of their topics in the structural space.

5.6 Conclusion

The goal of this paper was to identify some of the basic elements and patterns of rhetorical form that make up the grammar of motives employed when speaking in the sessions of the EP. With a focus on the discourses of two largest political groups of the 7^{th} European Parliament on issues pertaining to the Eurozone financial crisis, we uncovered their dominant discursive practices, the positions these groups occupied in relation to issues of the financial crisis, and the evidence of a left-right ideological divide in the speeches that specifically addressed the financial crisis.

Our analysis showed that when debating the issues of the Eurozone financial crisis, both political groups employed frames that characterized the financial

crisis in general terms and not as a European crisis. This specific framing of the crisis shifts the attention from a local (European) crisis to a general crisis, or perhaps even a global crisis. Hence, the way in which MEPs of these two groups frame the financial crisis could potentially indicate that they do not perceive it as a European issue, but rather a global one that is also affecting Europe. Such a framing process, in which the crisis is portrayed as an external event, may have serious implications for the policy-making process of the EP.

Moreover, we find a very low number of speeches on the issues of the Eurozone crisis in 2013, the year before the 7th EP ended its activities. While the Eurozone crisis was far from over in 2013, little attention has been given to this topic in the sessions of the EP by members of the EPP and S&D. Whether this was because the attention of the EP was directed towards other emerging issues, or whether members of these two groups debated issues relevant to the Eurozone financial crisis without employing the terms used in our data collection process ('financial crisis' and/or 'economic crisis') remains a question worth exploring in future work. Further investigation is also warranted by the fact that the topic 'European crisis' was only present in the S&D discourse, even though all the speeches collected for this particular study contained the phrases 'financial crisis' and/or 'economic crisis'.

The topic analysis employed, showed that most of the prominent topics in relation to the financial crisis were addressed by both groups through the speeches of their MEPs, and only five unique topics have been identified for each group. In spite of their different ideological orientations, exploring the ways in which the common topics were addressed by the members of the EPP and the S&D we show a high level of agreement and joint interests of these groups. This is an important finding because the EP arena is considered to be a highly competitive policy space (Hix et al., 2005), where political groups dispute views and positions in their attempts to impose the dominant narrative (Boin et al., 2009). Contrary to this argument, our findings show that on the topic of the financial crisis, the EPP and the S&D were united in the positions taken. This particular finding is even more surprising when taking into consideration the diversity of the members of these groups. While our data sets are not suited for

such purposes, we encourage further research into the relationship between this discursive coherence on the issues of the financial crisis and the rapid adoption of legislative proposals by the 7^{th} EP.

Furthermore, our topic analysis also revealed that evidence of the ideological orientation of these two groups can only be found in their unique topics. For each political group, one of the most emphasized topics is a unique one while the other is a shared one. The unique and most prominent topic of the EPP is 'Growth', topic which can be linked to one of the core values of a right-wing political group, namely a support of capitalism. More precisely, the right-wing considers growth through investment as key to economic prosperity. In the case of the S&D, the unique and most emphasized topic is 'Market Regulations'. As a left-wing political groups, the S&D supports central planning and governing bodies involvement into financial markets. The common topic that is most prominent in both discourses is 'EU Budget' and according to our findings the positions of the EPP and the S&D on this topic are similar. Both groups see the EU budget as an investment tool and a catalyst for economic growth.

The presence of 'EU Governance' as an imbedded and prominent topic in the EPP discourse is an important and unexpected finding that prompted a more careful look into the context in which it has been employed. As a right-wing political group, the EPP is expected to support deregulation and a limited role of governing bodies, values that contrast with the attention given to the topic of 'EU Governance' in their discourse. Calling for 'more European governance', as seen in the speech of Joseph Daul, the position of the EPP on this topic conflicts with the traditional right-wing ideology.

Thus, when discussing topics of common interest these groups exhibit high levels of agreement which do not seem to be driven by their political ideology. On the other hand, when debating distinct topics, more evidence of a left-right ideological divide emerges.

Lastly, the structural space analysis of the topic networks revealed notable differences between the two discourses. This approach reveled a larger overarching theme spanning a large proportion of the topics in the discourse of the EPP, based on the high number of shared concepts across topics. We also re-

vealed that the discourse of the EPP was slightly more structured and repetitive than that of the S&D, based on the repetitive terminology used across topics. This finding could be in part explained by the larger number of members of the EPP (compared to those of the S&D) but also by the general nature of political discourse, which is generally highly formalized and structured. Such discourses are commonly more repetitive and structured, employing less diverse language. Conversely, the more diverse nature of the S&D discourse can be attributed to the left-wing political orientation of this political group, ideology which is arguably a more populist one and hence employs richer and more accessible language.

Whilst our study sets a stepping stone for automated analysis of the discursive space of the European Parliament, revealing some of the fundamental characteristics of such speech acts, it is essential that we postulate some of the challenges and limitations the analysis of large collections of political speech transcripts entails. Firstly, in order to expose the discursive speeches of the two political groups (the EPP and the S&D) we used topic modeling, and more specifically the Latent Dirichlet allocation (LDA) model. Although these generative models are very efficient in inferring co-occurring clusters of words (i.e., topics) in large corpora, they also have a number of underlaying assumptions which should be carefully considered. The LDA model assumes that each document in the corpora is a mix of different topics, and thus fitting this topic model to a collection will yield patterns within the corpus whether or not they are "naturally" there. The extent to which this particular limitation of topic modeling impacts the validity of the results is highly dependent on the type of text documents used. This being said, the researchers' familiarity with the data to be modeled becomes crucial in mitigating such constraints. Furthermore, when using LDA models, the analyst has to pre-select the number of topics to be identified. The pre-selection of topics to be identified is often done through the exploration of various solutions, which increases the bias of such analysis. An alternative to LDA models is the hierarchical Dirichlet process (HDP), which allows the number of topics to be unbounded and learnt from data.

Secondly, in order to expose the underlying, subtle characteristics of the discursive practices employed by the selected political groups of the EP, we used the structural space method. This approach had proven effective in previous studies on the dynamics of metaphor families (Nerghes et al., 2015a) and organizational discourses of central banks (Nerghes et al., 2014b, 2015b). However, in the earlier studies mentioned, the approach has been applied to semantic networks, in which the degree and betweenness centrality measures could easily be interpreted as popularity and connectivity potential. In this particular study, we have applied this approach, that ranks nodes based on both measures (degree and betweenness centrality), to topic networks. Because the links in our topic networks represent the number of words these topics share, the meaning behind the two measures, and consequently the meanings of the four structural roles exposed by the structural space approach, have to be carefully reassessed. As detailed in our method sections, we posit that in this particular type of networks, total degree centrality becomes a topic similarity measure, while betweenness centrality remains a connectivity measure. Nevertheless, as it is often the case in social and semantic network research, the treatment of degree centrality in weighted networks deserves further investigation. Specifically, weights and the number of distinct ties should be considered separately, as the same total degree centrality score of a node can arise from different ego-centric structures. For instance, a topic sharing one word with three other topics will have the same degree centrality as a topic sharing three words with one other topic. This particular situation leads to very different conclusions regarding topic similarity. This being said, we encourage future research to explore the use of weighted degree centrality in semantic and social network analysis and the implications behind its interpretations for different types of networks.

Finally, we must address the multilingual character of the European Union, and implicitly of the European Parliament. EP documents are published in all the official languages of the EU and every MEP has the right to speak in the official language of their choice. Thus, one of the challenges this study faces is the loss of meaning through translations. MEPs are able to address the EP in their (official EU) language of choice, and these speeches are later translated.

Due to this multilingual character of the EP, all legislative speech occurs in translation. Hence, all of the EP's business occurs in multiple languages and therefore in translation. Even though so much of international politics occurs in translation, scholars have not paid significant attention to the effects of translation when using computer-based content analysis. Recognizing the multilingual character of the EP's, and some of the implicit limitations this imposes on our study, we argue that analyzing the speeches of the EP political groups (in translation) remains an important source of information for discourse analysts and political scientists alike. Hence, all the speeches collected for the analysis presented in this articles are in English, either as translations or speeches given in English.

To conclude, we anticipate that the findings of this paper will open new avenues for (semi-)automated research of political discourse and we encourage a more active debate on the relationship between political speeches and the right-left ideological divide of European parties. While previous research has focused on voting behavior and expert surveys to assess the dynamics of the European Parliament, the speech acts of the plenary sessions have seldom been studied. As the EP becomes a more powerful and strategic actor of the European Union, the ways in which its discursive space is being shaped by the speeches of the MEPs becomes a valuable source of information.

5.A Appendix 1

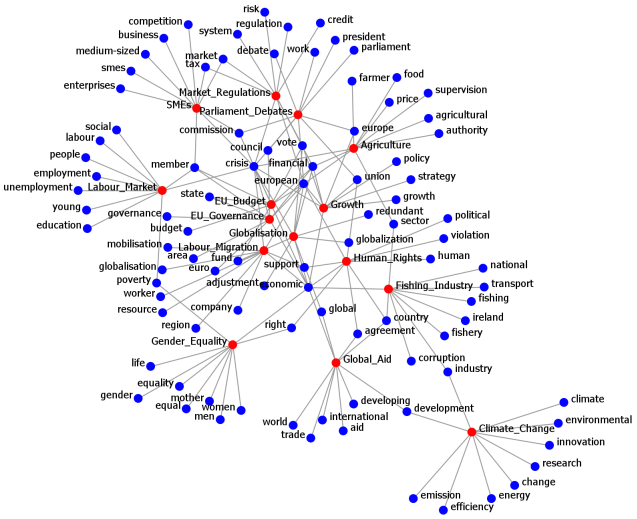
Socio-Economic Legislative Proposals Adopted by the 7th EP

- The establishment of the *European System of Financial Supervision* (September 2010), which assesses financial trends or shocks posing severe threats to the European financial system;
- The adoption of the *Alternative Investment Funds Directive* (November 2010) which demands compulsory registration, reporting, and initial capital requirements from all alternative investment funds;
- The ban on *Naked Sovereign Credit Default Swaps (CDS)* in October 2011;
- The *Credit Rating Agency (CRAs) Regulation* (2009) which obliges each CRA to apply for a licence from the European Securities and Markets Authority and to comply with a number of procedural requirements;
- The adoption of the *Capital Requirements Directives (CRD III and IV)* (July 2010), which includes stipulations regarding bonuses for bankers and proposals on the increase of the amount of capital banks should hold in reserve;
- The adoption of the *European Market Infrastructure Regulation (EMIR)* which reforms the market for over-the-counter (OTC) derivatives;
- The *Solvency II*, a framework for insurance firms and their corporate groups;
- The *Markets in Financial Instruments Directive (MiFID) II*, which proposes reforms in investor protection, transparency, and supervisory requirements for financial products and services provided by banks;
- The *UCITS Directive*, that led to deeper integration in the European investment fund industry;

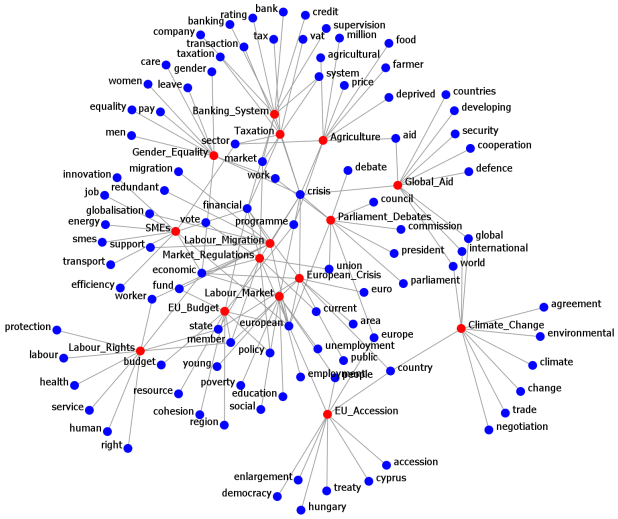
- *Deposit guarantee schemes*, which entitle depositors to insurance that enables them to recover part of their savings in the event of a bank collapse;
- The *Single European Payments Area (SEPA)*, which creates an integrated electronic payments system across the EU and the European Economic Area states;
- *Consumer Rights Directive*, designed to boost consumer confidence in digital businesses.

5.B Appendix 2

Concept by Topic Networks of the EPP and the S&D



(a) EPP



(b) S&D

Figure 5.5: Concept by topic networks

5.C Appendix 3

Examples of Speeches for the Shared Topics

	EPP	**SD**
Agriculture	"Above all we must not forget that hundreds of thousands of jobs in Europe depend on agriculture functioning." (Elisabeth Kostinger)	"Investments in agriculture will also ensure that the European Union and its Member States will be able to retain existing jobs and create other new jobs." (Silvia-Adriana Ticau)
Climate Change	"It is so very important to continue to build towards a sustainable economy and towards preventing climate change." (Corien Wortmann-Kool)	"Ladies and gentlemen the world needs a global agreement to fight climate change; both developed and developing countries need to stop the planet from overheating and so they need to pool their efforts and make a courageous decision." (Edite Estrela)
EU Budget	"The European budget is not a budget of spending but a budget of investing. It is a forward-looking budget, a budget for growth." (Joseph Daul)	"The EU budget is an investment budget that is used to leverage economic growth and create jobs. Anyone wielding the shears is robbing us of our common future." (Martin Schulz)
Gender Equality	"We need to look at achieving a work-life balance and promoting gender equality in the labour market from a broader perspective." (Lena Kolarska Bobinska)	"When preparing the development policy particular attention should be given to the development of gender equality." (Joanna Senyszyn)
Global Aid	"Providing aid to developing countries is one duty that the European Union must not abandon." (Catherine Soullie)	"I believe it to be vital that the European voice should be united and strong with regards to providing aid to developing countries." (Enrique Guerrero Salom)
Labour Market	"In order to reach these objectives labour market reforms must be stepped up and suitable incentives must be provided for individuals and companies to help them invest in training." (Clemente Mastella)	"We need targeted policies and strategies to reduce the percentage of early school leavers proper programming of people's needs for training and closer ties between the education provided and the needs of the labour market." (Antigoni Papadopoulou)

Labour Migration	"Recent studies undertaken by the European Commission have proved that opening our labour market will be beneficial and that worries concerning job losses due to labour migration are totally unfounded." (Traian Ungureanu)	"Labour migration within the EU has long been perceived as a potential problem in terms of shortages of labour supply and as having a negative effect on collective bargaining and wages within host countries. However concerns from the past related mainly to the fear of a large influx of immigrants job losses in respect of host country nationals or pay erosion have not come to fruition. On the contrary there is an evident and clear added value to the benefit of the entire EU." (Monika Flasikova Benova)
Market Regulations	"With this in mind I believe it is important for Parliament to apply pressure regarding the application of internal market regulations." (Carlos Coelho)	"Mr President I fully agree with my fellow Members Mr. Canfin, Mr. Lehne, and Mr. Ferber because I believe that in order to prevent a new financial crisis we need to strengthen market regulation." (Robert Goebbels)
Parliament Debates	"We must remember that in striving for a better Europe the European Parliament plays a special role a role that is not only institutional but also social a deeply symbolic role. The European Parliament is the essence of the European democratic system." (Jerzy Buzek)	"The European Parliament is the representation of European citizens and the guarantor of democratic legitimacy." (Nessa Childers)
SMEs	"The social economic and financial role of SMEs in the European Union is unquestionable. (Jose Manuel Fernandes)	"SMEs are the backbone of the EU economy." (Mitro Repo)

6 Conclusion and Discussion

In the preamble of this thesis, I posed two important questions for the field of text analysis, namely: Can we infer rich information from 'big text data'? And how can we use text-analytical methods to infer such rich information from large text collections with different characteristics? Although these questions are of great importance in this digitized age, when roughly 80% of all existing data is stored in the form of (unstructured) text (Zicari, 2012), they are also very broad questions. These questions have not been (and cannot be) fully answered through the research presented in this dissertation. Rather, in this dissertation I proposed an approach to the analysis of textual data that can aid researchers in their efforts to bring contributions to answering such broad and important questions. This approach, based on the relational perspective to meaning, allows analysts to uncover patterns of language use, subtle discursive shifts, and discursive dynamics in large text corpora with distinct characteristics. Furthermore, through the development of the analytic approach proposed, the empirical chapters of this dissertation also bring important contributions towards answering the above stated questions. The important empirical findings I present, challenge, extend, and refine what we know about central banks, media, and political discourses in the context of the financial crisis. Thus, before reflecting on the methods and approaches employed and expanded here, I summarize the empirical findings of each empirical study.

6.1 Summary of Findings

The empirical chapters of this research focused on the context of the global financial crisis and investigated three distinct discourses of social actors, namely central bank discourses, media discourses, and political discourses. These three discourses present distinctive characteristics, address different audiences, and fulfill different communicative purposes. Central bank discourses are highly formalized, leading to redundant, structured, and even predictable content (Irvine, 1979). They also contain highly specialized information addressed mostly to those financial market actors interested in the actions and decisions of these banks. The media discourses on the other hand, present stories and employ various rhetorical devices to give these stories a newsworthy angle (Pan & Kosicki, 1993). News items are presented to large and relatively diverse audiences. Lastly, political discourses are characterized by normative as well as conventional or strategic practices that make the identification of meaningful information difficult. While in a general sense political discourses are addressed to the constituents or voters, here I focused on parliamentary speeches. These speeches are mostly addressed to other members of the parliament. The analysis of these distinct discourses produced important results, which I summarized below.

Chapter 3 investigates the discourses of the Federal Reserve System (Fed) and the European Central Bank (ECB) in the context of the financial crisis. These financial-regulatory organizations determine the monetary policy for two of the largest currency areas and two of the largest economies in the world. The two central banks have been at the epicenter of the financial crisis, and have been involved in the efforts to combat the impact of the crisis and aid financial recovery.

While previous research established a link between communications of the ECB and the Fed and their respective impacts on the financial markets (Jansen & de Haan, 2005a,b; de Haan, 2008; Hayo & Neuenkirch, 2010; Hayo et al., 2014), the focus of my analysis was directed at uncovering the shifts and adaptations of their discourse in a time of crisis and increased market volatility.

Pursuing these aims, I explored the press releases of the ECB and Fed over a period of eight years, at the different stages of the financial crisis (i.e., pre-crisis, crisis, post-crisis, and the recovery period). Methodologically, the analysis of these press release was done through semantic network analysis and the structural space approach, a combination of popularity and connectivity of semantic network nodes. In the particular case of the ECB and Fed, the structural space approach I have developed was a particularly valuable instrument for change detection, due to the highly formalized nature of their discourses. Such formal organizational discourses, as the ones of the ECB and Fed, contain repetitive top key concepts, indicative of the obvious and perhaps uninformative central topics of organizations. By looking beyond the core of the network structure, the structural space approach revealed the discourse of the Fed exhibiting greater attentiveness to the financial crisis, while the ECB's attention was delayed and increasing steadily. Furthermore, in this study both the Fed's and the ECB's discourses are shown to be transitioning into a new "hybrid" state in economic recovery period (2012-2013), rather than returning to the pre-crisis status quo. In sum, with this study I bring an important contribution to the understanding of financial-regulatory discourses and their dynamics during societal crises and high levels of financial market uncertainty.

The second empirical study presented in this dissertation (Chapter 4) focused on a different discursive space, namely that of the media. Investigating news items published by three newspapers (i.e., the Financial Times, the New York Times, and the Sun), I expose the ways in which metaphor families (i.e., sets of related metaphors) fulfill a translator role for emerging financial terminology in the media. Employing a similar methodological approach as the on in Chapter 3, this study brings a contribution to metaphor theory by widening the focus from conceptual metaphors to metaphor families, and exposes the expansion and evolution of the 'toxic' metaphor family, revealing subtle changes of metaphor use in three newspapers (i.e., the Financial Times, the New York Times, and the Sun) over time. I show a transition from generic image-creating metaphors toward financial-instrument targeted metaphors across three stages of the financial crisis, spanning a period of five years. Additionally, the results

reveal that most of the 'toxic' metaphor family variations were created in the actual crisis period, not during the pre-crisis period. Overall, the findings of this study suggest that metaphor families are used as translating devices for unfamiliar terminology, practice which has the potential of influencing perceptions of the financial crisis.

The last empirical study included in this dissertation, presented in Chapter 5, provided an account of the discursive space of the European Parliament (EP) in the context of the Eurozone financial crisis. The analysis in this study focused on the 7th European Parliament (2009-2014), whose rapid adoption of socio-economic legislative proposals played a key role in the European efforts to combat the rippling effects of the Eurozone financial crisis. Particularly, I investigated the discursive practices employed by Members of the European Parliament (MEPs) in their speeches on the issues of the Eurozone financial crisis. The MEPs included in this study are those affiliated with the two of the largest political groups of the European Parliament (the European People's Party and the Progressive Alliance of Socialists and Democrats), groups holding 61% of the seats, and hence votes in the EP.

With this study of parliamentary speeches I contribute to a better understanding of how policy issues are debated, the discursive practices employed, and the ideological divide in the European Parliament. Situated on different sides of the left-right ideological spectrum, the speeches of these two EP political groups have the potential to exposes critical disparities on such critical issues as socio-economic policy in times of crisis. Thus, by identifying some of the basic elements and patterns of rhetorical form that make up the grammar of motives employed when speaking in the sessions of the EP, I set a stepping stone towards a better understanding of the complex and multidimensional discursive space of the EP. To expose the dominant discursive practices and topics in the speeches of these political groups, I employed a combination of topic models, network representations of topic networks, and the structural space approach. This combinations of methods, uncovered a unified discursive space, with high levels of agreement on most issues of the Eurozone financial crisis between the two political groups, and a moderate left-right ideological

divide only when the unique topics of each political group are investigated.

The findings of these three empirical studies, summarized in Section 6.1, show how different sets of social actors capture different aspects of the financial crisis through their discourses. For instance, while central banks focus on the immediate and localized events of the crisis within the financial markets, the medias' focus is on reporting emerging events and informing the public by covering a wider range of topics and events. However, in their role of conveying information about the crisis to their readers, the media dramatizes stories and provides evaluations of events. As shown in Chapter 4, media often employs rhetorical devices meant to give stories a newsworthy angle, but also to translate unfamiliar topics and terminology in their reporting. Employing such devices, as the 'toxic' metaphor family, creates highly negative evaluations of the financial crisis that may impact the perceptions of their readers. In contrast, the discourses of central banks are focused on disseminating highly specialized information addressed mostly to those financial market actors interested in the actions and decisions of these banks. Hence, their discourses are oriented towards the immediate and relatively isolated effects of the crisis. The discourses of central banks, then, will have a narrow focus on the financial markets they regulate. Contrariwise, political discourse exhibits a broader perspective on the financial crisis. The discourses of the two European Parliament political groups addressed the financial crisis as an event affecting many areas of society. The topics emerging from the financial crisis discourse of these two EP political groups span a wide range of societal areas, such as gender equality, globalization, human rights, labour migration etc. As such, it can be argued that as a policy-making body, the perspective of the EP on the financial crisis extends beyond the financial markets, to most areas of society affected by the crisis.

In sum, by investigating these discourses it is clear to see that, in a sense, a different 'version' of the crisis emerges from each of these discourses. Can we then talk about 'the financial crisis' as a unitary and single global event? If the argument is that a global 'definition' of the crisis emerges from combining the perspectives offered by various social discourses, what is the contribution

of each perspective? While attaining such a global definition of the crisis might be intractable, these different perspectives have the potential to steer manifestations of the crisis by altering perceptions and prompting action. Hence, the analysis of multiple discursive sites can provide a more inclusive depiction of some of the mechanisms that contribute to the containment and/or amplification of a crisis.

If we look at the three sets of actors investigated in this dissertation, the results suggest that the discourses of central banks, as well as that of the European Parliament, were focused on system stabilizing elements. To this end, the results presented in Chapter 3 suggest that the discourses of central banks (ECB and Fed) are neutral and, to some extent, they present very stable explanations and reactions to the global financial crisis. Similarly, the discursive space of the European Parliament, investigated in Chapter 5, exhibits a striking focus on system stabilizing elements. However, while the discourses of the central banks focused on those system stabilizing elements directly relevant to the financial markets, the discourse of the European Parliament focused on elements relevant to a wider range of social areas (e.g., agriculture, labour rights, gender rights etc.). In contrast, the discourses of the media actors investigated (i.e., The Financial Times, the New York Times, and the Sun) exhibit different dynamics, moving away from neutral, stabilizing elements. The three newspapers analyzed in Chapter 4 focused on dramatizing and escalating the elements of the crisis through highly negative portrayals of various events relevant to the global financial crisis.

Considering the different characteristics of these three discourses, which do not overlap and which are only connected through the topic they address, it is natural to assume that their contributions towards the escalation or stabilization of crisis events will weigh differently. In the particular case of the three discourses explored in this dissertation, I can conclude that while central banks and the European Parliament discourses presented as neutral and relatively stable across the different stages of the crisis, the newspapers investigated exposed rapidly changing discourses containing highly negative evaluative elements. In this sense, then, the discourses of the central banks and that of the

European Parliament are more likely to have contributed to the containment of the crisis, while the media discourses have contributed to the escalation of the events by creating highly negative images that drove the increasing levels of market uncertainty. Hence, exploring an ensemble of societal discourses, relevant to such events as the global financial crisis, has the potential to uncover those discourses that escalate certain developments and those that contain or stabilize them.

Table 6.1: View of the financial crisis in central banks, political, and media discourses

Discourse	**Findings**
Central banks	Focus on the overwhelming market defaults and not towards the crisis as a whole; Focus on the immediate (relatively isolated) effects of the crisis; Narrow and localized perspective disregarding the global nature of the crisis; Shift of focus away from their main objectives; Transition to a new 'hybrid' state post-crisis;
Media	Dramatizing the events of the financial crisis; Highly negative evaluations of the crisis; Metaphor families used as translating devices for unfamiliar terminology; Most 'toxic' metaphor family variations created at the hight of the crisis; Transition from generic image-creating metaphors toward financial-instrument targeted metaphors;

Political	Unified discursive space in the EP; The financial crisis framed as a general crisis, not a European one; Broader perspective on the financial crisis as an event affecting many areas of society; High levels of agreement on issues of the Eurozone financial crisis on topics of common interest; Moderate left-right ideological divide.

However, as argued throughout this dissertation, exploring such discourses in a comprehensive and inclusive manner entails analysis of large text collections, containing complex manifestations of discourses with distinct characteristics. For example, central bank, media, and political discourses present different levels of codification (formalization), different levels of specialized information contained, and different levels of normative and strategic discursive practices. As such, accomplishing an inclusive analysis of an ensemble of distinct discourses, produced in different social settings, requires flexible methods that can be attuned to accommodate the specific elements and characteristics of these discourses. This is to say that text analytical methods, computer-aided or otherwise, cannot be indiscriminately employed to analyze multiple discourses of different types. For instance, while the methods of relational meaning analysis applied and expanded throughout this dissertation demonstrated their value as instruments for subtle change detection across time and across distinct discourses, the application of these methods has been adapted to the characteristics of each discourse analyzed and to specific analytical goals. The adaptation of these methods can be done through a selection of coding choices, as well as through a selection of specific combinations of analytical procedures to be performed. Whereas the combinations of analytical procedures performed in the empirical studies of this dissertation have been detailed in Table 2.1 on page 26, the selection of coding choices will be further dis-

cussed in the following section (Section 6.2). However, before discussing such selections, their implications, and limitations, I will briefly discuss the main methodological contribution of my research.

The main methodological contribution of this dissertation is an analytic approach to large collections of texts that allows for a comprehensive investigation of subtle discursive manifestations and their temporal dynamics. Particularly, this type of analysis (i.e., the structural space approach) proved to be more explanatory regarding the overtime subtle shifts and changes in discourse, by combining structural measures and looking beyond the core of semantic and topic network structures. As it has been demonstrated, this approach has also been effective in identifying emerging topics. Furthermore, this approach has revealed important results at the level of three distinct discourses analyzed throughout this dissertation, demonstrating its adaptability and effectiveness in exposing substantial and imperative shifts across distinct discourses. Outperforming more direct text-analytic approaches, the structural space approach opens new avenues for text analysis research in a variety of contexts and topics. However, it is important to note that the effectiveness of the structural space approach across distinct discourses is dependent on the (semantic or topic) networks to which it is applied. Thus, as elaborated earlier, even a text analysis method that proved to be effective across distinctive discourses, cannot be applied broadly without careful consideration of the type of discourses analyzed and the coding and analytical choices to be made when generating and analyzing semantic or topic networks. As it will become evident in the following section, these choices have important consequences for the subsequent results and the ways in which results can be interpreted and generalized. The following section starts from the more general level of text-analysis and moves to the particular methods employed and developed here.

6.2 Considerations and Directions for Further Research

The field of text analysis has made sizeable strides in recent years, producing more and more sophisticated and time-efficient procedures of analysis, especially in the case of large text corpora analysis. Part of the appeal of such methods in the social sciences rests in their flexibility and adaptability for a wide-range of research goals and ease with which these methods can be applied to large collections of texts. However, while methods and tools implementing them evolved rather rapidly, the same cannot be said about theories supporting this field of research. The field of text-analysis has mostly been driven by empirical and methodological goals. Take for instance the case of the relational approaches to text analysis employed in this dissertation. The theoretical foundations of these approaches come from the 1940s, 1950s, and 1970s (e.g., Cassirer, 1944; Eco, 1979; Geertz, 1973; Minsky, 1975; de Saussure, 1959). Although the persistence of such theories speaks to their value, the fast pace at which larger and larger text corpora are analyzed and the rate at which new methods are proposed, calls for a reevaluation of these theories and the challenges big text data poses for the social sciences. Additionally, the fragmented development of work in the area of text and discourse analysis makes it increasingly difficult to gain an overview of the various contributions brought by scientists from multiple disciplines. Linguists, humanists, computer scientists, and social scientists develop their respective text-analysis methods in parallel, rarely referring to each other's work (Pollach, 2012; Popping, 2000). Furthermore, the wide array of methods and approaches to text analysis led to a large assortment of terminology. As such, across disciplines, different terms are used to refer to similar concepts and processes. Agreement on and development of a 'standardized' glossary of terminology for basic concepts in text-analysis would contribute to creating a more coherent research field.

The development of the text analysis field, then, requires multidisciplinarity. In such cross-disciplinary and unified approaches, linguists would contribute by

elucidating the structure and grammar of texts, computer scientists can provide appropriate tools and analysis algorithms, while humanists and social scientists would relate texts to the social context in which they were produced and the mechanisms according to which words influence and are influenced by human behavior. The different perspectives from which these disciplines approach the analysis of texts are not mutually exclusive and hold great potential of informing one another.

6.2.1 Relational Approaches to Text Analysis

Semantic network analysis and topic modeling, upon which I build and expand, are based on the presumption that meaning in language and texts is dependent on the structure of the words and sentences they contain. These relational methods of text analysis bridge the gap between qualitative and quantitative text analysis by taking advantage of recent technological and methodological advances, while at the same time allowing the analysts to stay close to the text throughout the processing steps. Operationalized through computer-aided automated or semi-automated processing tools, these methods are relatively easy to use and provide fast procedures for the analysis of large collections of text. Offering a degree of convenience for text-analysts, these methods remain flexible enough to allow researchers ample opportunities to engage into coding decisions based on knowledge of the corpora, and to elucidate the results in an interpretative, qualitative manner. These coding decisions, as discussed below, are necessary in order to attune the analysis to the specific discourses to be investigated and they have important consequences for the subsequent results and the ways in which these results can be interpreted and generalized. Below, I discuss in more detail the coding choices and selections relevant to semantic network analysis, topic modeling, and the structural space approach. For each of these analysis methods, I also review their strengths and limitations. But first, I note the importance of carefully selecting text corpora relevant to the phenomenon under study.

Corpora Selection

The selection of corpora to be analyzed with these methods should be carefully considered. On the one hand, being able to access, collect, and process extensive collections of texts opens ample opportunities for social scientists to test broad hypotheses and pose wide-ranging research questions. On the other hand, the increased availability of textual data and computer-aided processing tools warrant an informed selection process guided by well defined research goals. While the issue of data selection is not unique to text analysis, the rising volumes of data available and the efforts towards overly inclusive data analyses, may result in greater levels of error and imprecision. Whether employing semantic network analysis, topic modeling or other text-analytic methods, careful consideration, guided by firm analytical frameworks, is desirable. In this way, researchers can limit the materials selected to those relevant to their research goals, thereby reducing the levels of error and imprecision.

Semantic Networks, Processing, and Coding Decisions

While semantic networks are a valuable approach to knowledge and meaning extraction and representation, the ambiguity and complexity of texts from which they are generated, require a number of choices and preprocessing steps. The process of transforming textual data into networks of concepts (or words) implies a series of coding choices which can greatly impact the results of the analysis. That is, the techniques used when preprocessing the raw text (e.g., removing noise words, removing numbers, etc.), the identification of nodes to be included in the network, and/or the parameters used for the creation of links (e.g., window size and/or stop unit) can strongly impact the structure of the resulting network. As such, these coding choices should be closely aligned to the objectives of the researcher and should be chosen with care. Below I summarize a number of choices the analysis should consider. While there is no one right classification scheme or preprocessing sequences, an understanding of the benefits and effects of these various procedures can guide the analyst towards the appropriate procedure for specific research goals.

Concept generalization: The question to be answered here is: What are the nodes in the semantic network? In answering such question, the analyst has to decide wether a node in the semantic node is a word or multiple words recoded as a concept. To exemplify, the word 'crisis' can be a node in a semantic network and so can the concept 'global crisis'. This form of concept generalization is generally done through the identification of n-grams. N-grams are commonly used multi-word expressions which are meaningful together. Identification of n-grams can be done for common words (e.g., 'civil war', 'interest rate', etc.) but also for named-entity identification (e.g., named actors, organizations, locations, etc.). This allows the analyst to distinguish, for example, the "White House" from the color and the domicile. While such conversions will arguably be beneficial in interpreting the semantic networks generated and the meaning structures that emerge, n-gram coding can become time consuming for larger text collections (Grimmer & Stewart, 2013; Hopkins & King, 2010). Although automated tools such as Automap (Carley et al., 2013a) provide n-gram detection functions based on preexisting lists of common multi-word expressions, the selection of relevant or appropriate n-grams has to be done manually by the analyst, based on knowledge of the corpora.

Furthermore, generalization of concepts can also be done by 'translating' text-level concepts into higher-level concepts. An example of such procedure is the conversion of 'President Barak Obama', 'Barak Obama', and 'President of the United States' into a single concept (e.g., Barak_Obama). Through this type of generalization, the researcher can reduce the complexity of the resulting semantic network, while at the same time identifying synonymous multi-word expressions. However, such generalization may affect the results of the subsequent analysis by masking the subtle linguistic variations in the corpora explored.

Identification of relevant concepts: It is common procedure in semantic network extraction to remove those parts of speech that are considered superfluous, such as prepositions (e.g., to, at, after, on, but), conjunctions (e.g., and, but, when), adverbs (e.g., quickly, silently, well, badly, very, really), or pronouns (e.g., I, you, he, she, some) (Diesner & Carley, 2005; Martin et al., 2013). In

general, such text pre-processing techniques are employed to reduce the size of resulting semantic networks to a more manageable size, that allows for meaningful interpretation (Carley, 1993). Thus, the removal of these so called noise words subsequently impacts the structure of the semantic networks generated, which in turn will impact analysis results. Moreover, removing contractions as 'don't', 'doesn't' or 'aren't' (or words such as 'not' and 'no') may impair meaningful inference if the goals of the analysis involve the identification of preferences, disagreements or negations. Such selections, then, should be made under the understanding of their full effects on the structure of the networks, and their impact on the analysis results.

Range of relations or window size: Once the pre-processing (i.e., the concept generalizations and the identification of relevant concepts) of the corpora has been completed and the semantic networks can be generated, a decision has to be made regarding the range in which connections between text concepts are created. Also known as a window size, this procedure delimits a sequentially moving window in which connections are created and it can be two words, a clause, a sentence, a paragraph or an entire document. The range of links in semantic networks affects the structure of these networks and influences the interpretation of results. Thus, window sizing decisions should be made based on knowledge of the texts analyzed and the questions to be answered. For instance, the types of texts analyzed can give an indication on whether meaningful ideas are conveyed at the level of a sentence, a clause or a paragraph. Diesner (2012b) provides a detailed account of the impact window sizing has at the level of different types of textual data.

Directionality of relations: The relations (or links) in semantic networks can be extracted as undirectional or directional. Directional links account for the direction of word associations, thus providing an indication of the positions of concepts relative to one another. While some authors have argued in favor of this directional approach (e.g. Carley, 1993; Sowa, 1992), others maintain that the inherent meaning in texts is undirected (e.g., Danowski, 1993; Popping, 2000). Whereas the undirectional approach to semantic network links is supported by work on semantic memory and the hierarchical association of

words at a cognitive level (Collins & Quillian, 1969; Chang, 1986), the directional approach can be valuable when the analyst is specifically interested in the grammatical structure of sentences.

Value of relations: Because the links in most semantic networks are based on co-occurrences, they can be unweighted as well as weighted. An unweighted link in a semantic network represents the existence of a relation (e.g., two words co-occurred in the specified window), while a weighted link also shows the intensity of that relation (e.g., how often two words co-occurred within the specified window across the corpora). The selection of unweighted or weighted links can impact further analysis. For example, if the analysts intends to calculate various network metrics (e.g., centrality) the value of the links should be carefully considered. While semantic networks with unweighted links are easier to compare, networks with weighted links allow the researcher to retain more information regarding the underlying text (Carley, 1993). Furthermore, extracting weighted relations among text concepts exposes the emphasis that is placed on the relationships between the concepts. (Johnson-Laird et al., 1984).

Network complexity and thresholds: Once semantic networks are generated from large collections of texts, they are often large and complex and exhibit highly intricate network structures (Bales & Johnson, 2006; Steyvers & Tenenbaum, 2005; Postma et al., 2000). In order to analyze such complex networks and their array of global and local features, researchers often employ thresholds. More precisely, it is often the case that nodes (i.e., concepts) with low frequency counts are removed, in an attempt to reduce the size and complexity of such networks. Such selections have great impacts on network metrics and should be taken into consideration only when their effect is fully understood. Moreover, unlike content analysis where the importance of concepts is determined based on frequency, semantic network analysis is able to identify important concepts based on their relations to other concepts. For instance, in the network approach, those connecting concepts bridging different topics in texts can be identified. However, such theme bridging concepts are often removed by imposed thresholds due to their low frequency counts.

All these different coding choices are useful in reducing the complexity of

the resulting semantic networks or in ensuring a proper representation of language relationships within and across text documents (Carley, 1993). However, careful consideration of their effects and impacts on the resulting semantic networks is imperative. As elaborated in the introductory chapters of this dissertation, semantic network analysis, or any other (semi-)automated methods of text analysis for that matter, do not remove the need of in-depth knowledge of the corpora analyzed and the social context from which the corpora emerged (Grimmer & Stewart, 2013). Coding and selection decisions should be guided by a well-founded analytical framework, specifically designed to incorporate elements of social context awareness and an understanding of the specific characteristics of various discourses. While text-analytic methodologies evolved considerably in the past decades, the complexities of language make it impossible to develop a universal method of text or discourse analysis. The tractability and social character of language have constantly posed a challenge for social scientists, humanists, computer scientists, and linguists alike.

Moreover, depending on perspective, the necessity for in-depth knowledge of the corpora analyzed and the analysts' opportunities to interact with the texts, for coding or interpretation, can be seen as an advantage as well as a limitation. In this sense, empirically investigating textual data is always prone to a certain degree of bias introduced through human selection of corpora, coding choices or interpretation of results. Even methods of translating texts in numerical data involve methods of coding developed by the analyst.

Lastly, while semantic networks are powerful meaning extraction and representation tools for corpora of any size, it is easy to attribute them too much power (Jonassen & Marra, 1994). Semantic networks extracted from text represent static snapshots of text and discourse and thus, they do not capture the evolving character of language and meaning. Indeed, the analyst can surely discover discursive changes by comparing semantic networks extracted from temporal text data, but even then, the networks will represent snapshots that will not capture the more fine grained evolution of meaning and language. Meaning is constructed in social interaction (Fairclough, 2001) and thus, its evolution is influenced by a multitude of factors such as context and purpose.

Thus, semantic networks should not be understood or interpreted as literal, static, and immutable representations of meaning. Capturing evolving meanings would require extraction of minute-by-minute, context-by-context changes in the concepts, relationships and structures that are represented in order to more accurately map the evolving nature of meaning. To conclude, while the strength of semantic networks lies in making explicit what is normally implicit, they should be interpreted cautiously and understood as a representation tool for inferring socially negotiated meanings and knowledge, which are constantly evolving.

Topic Modeling

Topic models are powerful exploratory tools for understanding large archives of text documents (DiMaggio et al., 2013). One of the core strengths of topic models is their ability to expose *what* is being talked about in the corpora analyzed, but also *how* issues are being talked about. Just as in the case of semantic networks, the relational aspect of topic models renders them effective in extracting more than just list of words ordered by their prominence in text collections. Topic models "can identify discourses as well as subject categories and embedded languages." (see Goldstone & Underwood, 2012, section *Using LDA as evidence*, para. 1)

However, whether using topic models to investigate Flickr groups (Wang et al., 2012), Twitter feeds (Ramage et al., 2010), or scientific topics (Griffiths & Steyvers, 2004), the analyst has to consider a number of underlaying implementation assumptions of topic models. As Goldstone and Underwood (2012) show, these assumptions in combination with different methodological choices may lead to contrasting results. In the following paragraphs I discuss the limitations of topic models, specifically the latent Dirichlet allocation (LDA) model, as well as the implications of various methodological choices.

Firstly, LDA topic models rely upon the "bag-of-words" assumption, disregarding the order of words within a text (Meeks & Weingart, 2012) and the order of documents within a text corpora (Blei & Lafferty, 2009). Thus, LDA models assume that documents and words are interchangeable and they will

yield different results when the starting point of the procedure is different. This is perhaps the most common limitation of LDA topic models mentioned in the literature (Chang et al., 2009). But, as Mohr and Bogdanov (2013) explain "the real genius of topic models is precisely that for this specific type and level of meaningful content, it appears as though relationality trumps syntax." (p. 23) Thus, removing the semantic and syntactic information provided by sentence order does not influence the robustness of LDA topic model results (Mohr & Bogdanov, 2013). However, caution is advisable when analyzing text corpora with evolving content (e.g., email messages).

Secondly, the LDA model assumes that each text in the corpora is a mix of different topics, and thus fitting this topic model to a collection will yield patterns within the corpus whether or not they are "naturally" there (Blei & Lafferty, 2009). The extent to which this particular limitation of topic modeling impacts the validity of the results is highly dependent on the type of text documents used. This being said, the researchers' familiarity with the data to be modeled becomes crucial in mitigating such constraints (Mohr & Bogdanov, 2013). Any effort to apply topic modeling to a corpus to answer interpretive questions must include a subject-area specialist able to asses the interpretability of the results (DiMaggio et al., 2013). Subject-area knowledge will allow researchers employing topic models to extensively and critically asses the extent to which individual words fit within the topics identified and to asses how grounded their inferences are (Schmidt, 2012).

Furthermore, when using LDA models, the analyst has to pre-select the number of topics to be identified. The pre-selection of topics to be identified is often done through the exploration of various solutions. Often, choosing the number of topics when applying LDA models is seen as a limitation (Blei & Lafferty, 2009; Tang et al., 2014). However, as the analyst changes the number of topics to be identified (and other parameters), the LDA model provides different depictions of the same underlying collection of texts. Hence, this should not be understood as evidence of unreliability for topic models. Rather, this multiplicity of perspectives should be understood as a strength of topic models because it supports different interpretations of the same data. These

interpretations will be compatible by definition because they represent different distances and different levels of granularity of the same textual data (Goldstone & Underwood, 2012). Furthermore, selecting the appropriate number of topics for a given text corpora can be done iteratively, comparing different solutions and the overall level of interpretability of the resulting sets of topics (Blei & Lafferty, 2009).

Lastly, as with most other easily available text analysis tools, the sheer simplicity of topic models and the ease of use as a textual analysis method raise some important concerns (Grimmer & Stewart, 2013; Schmidt, 2012). Implemented in readily-available, user friendly software packages like ORA (Carley et al., 2013b), ConText (Diesner et al., 2015), and MALLET[1], topic modeling may easily be misinterpreted as a 'one-size-fits-all' tool for text analysis, so researchers should be judicious in its use. As with any scholarly pursuit, analysis methods should be selected based on clear analytical frameworks and research goals and not based on their simplicity or ease of use. Once these considerations have been addressed, the ability of topic models to quickly and efficiently parse through large text corpora may encourage social scientists to address big-picture questions regarding huge textual libraries (Schmidt, 2012).

On a final note, topic models may lead to interpretations having more of a focus on the topics themselves rather than the language, which as Meeks and Weingart (2012) explain "might provide the false security of having resolved the distinction between a word and the thing that it represents." (see Section 2, para. 4)

The Structural Space Approach

Expanding semantic network analysis, the structural space approach revealed important findings regarding subtle discursive manifestations and their dynamics across time in distinct discourses. The approach can reveal the overall meaning and the latent agenda of the analyzed corpora and also the elements of text (topics or concepts) that are potentially impactful and deserving of further attention.

[1]Available at: `https://code.google.com/p/topic-modeling-tool/`

Based on the relevance of the results uncovered by this method, I advocate for further development and testing of the structural space as an analysis method of semantic and topic networks. One avenue to further test and demonstrate the value of this approach is to apply it to temporal text data representing smaller time slices. In the studies presented, the temporal data explored has been aggregated into two year periods. Arguably, smaller data time slices could potentially reveal even subtler aspects in the dynamics of discourse and fluctuations in terminology roles.

One of the most important limitations of the structural space approach stems from the specific types of networks it has been applied to. More specifically, when generating these (semantic and topic) networks, links were created to represent the existence of a relation between nodes (words or topics) but the intensity of the relations has also been taken into account by adding co-occurrence values to each link connecting a pair of nodes. Consequently, the resulting networks were weighted and not binary. Hence, when applying the structural space approach, which computes ranks of network nodes based on the combination of betweenness centrality and degree centrality, the treatment of degree centrality raises some concerns.

Initially designed only for binary networks (Freeman, 1979), degree centrality represents the number of nodes that a focal node is connected to. Degree centrality was extended to weighted networks by Barrat et al. (2004) and defined as the sum of the weights attached to the ties connected to a node. Thus, a network node connected to three other nodes will have the same degree centrality as a node connected by a link with a value of three to one other node. While typical in network research, this naïve treatment of weighted degree centrality ignores the relative importance of tie weights to the number of ties. Thus, weighted degree centrality puts equal importance on the distinctiveness of a network link and the intensity or frequency of the link. If nodes having a higher number of links are of greater consequence to the object of study, compared to nodes having fewer but highly weighted links, weighted degree centrality will not appropriately distinguish these nodes.

This particular limitation of weighted degree centrality has been addressed

by recent work attempting to allow for separate assignments of importance on the intensity and distinctiveness of a link in constructing centrality measures (Opsahl et al., 2010). However, the approach developed by Opsahl et al. (2010) requires a subjective or a priori assignment of this relevance, and thus does not solve the issue entirely. While the exploration of this issue is beyond the scope of this dissertation, I hope (and expect) that future research will further address the use of weighted degree centrality in semantic and social network analysis, and the implications behind its interpretations for different types of networks.

Furthermore, because the structural space approach ranks semantic network nodes based on both degree centrality and betweenness centrality, within-network correlations for the two centrality measures will affect the structural role assignment of nodes. Highly correlated centrality metrics will result in a structural space in which the majority of nodes will be placed along the diagonal between the GC and M roles. In the particular case of the semantic networks explored through this approach in this dissertation, correlations for the two centrality measures were modestly high, echoing other findings, (e.g., Valente et al., 2008; Lee & Pfeffer, 2015a) and also highlighting the usefulness of the structural role approach in identifying outliers in the G and LC roles.

The demonstrated value of the structural space approach warrants efforts to further develop it and expand it. Future research employing this method should explore the inclusion of other structural measures, such as closeness centrality, clique counts or clustering coefficient. During the preliminary testing of this approach, the inclusion of eigenvector centrality in the structural space was tested, which proved to be highly correlative to total degree centrality, and thus did not add to the informative value of the structural roles. Additionally, a further development of this method would involve the establishment of thresholds that clearly define the boundaries of each role. Such thresholds can, for example, be based on the actual values of the centrality measure of the empirical networks, and may help better categorize those nodes spanning multiple structural roles.

Finally, while it is common to apply social network analysis metrics to semantic networks (e.g., Hoser et al., 2006), very little effort has been dedicated

to theorizing on how these metrics apply to networks in which the nodes are concepts or words. Consequently, researchers have only attempted to conceptualize the significance of degree centrality and betweenness centrality in semantic networks by focusing on their specific studies of interest (e.g., Grebitus & Bruhn, 2008; Henderson et al., 1998; Hill & Carley, 1999; Hooper et al., 2012). Arguably, a more wide-ranging conceptualization of centrality metrics for semantic networks would guide researchers in selecting those centrality metrics appropriate for their research goals and would support the inference of more robust interpretations of results. To this end, through the structural space approach, I bring a contribution to the reconceptualization of total degree and betweenness centrality to the specific case of semantic networks. To briefly reiterate this contribution, semantic network nodes with high degree centrality values function as a hot topic's central key concepts, while high betweenness centrality values in semantic networks indicate the influence of nodes as gatekeepers between different discursive themes (see Chapter 3 on page 47 for a more detailed discussion).

6.3 Final Remarks

Language and text are consequential to all aspects of social life and essential to our understanding of reality. Rather than being peripheral emanations of social activity, texts are collective products of social interaction and important instruments through which we constitute and articulate our world. In written or spoken form, texts are fundamentally interactive (Halliday, 1978) and they have social effects through the discursive strategies and devices contained, which in turn shape and construct our social reality (Fairclough, 2001). Therefore, the question is no longer whether or not texts should be investigated, but rather which approach is the most insightful for a given set of texts and a specific research goal. To this end, in this dissertation I argued and presented evidence in favor of relational text-analysis methods. By applying and expanding these methods, the contributions of my research have been both methodological as well as empirical. Methodologically, through the combina-

tion of text-analytical methods employed, and more specifically through the structural space approach, I provided compelling evidence for the great potential of relational text-analytic perspectives. While this dissertation is perhaps a (modest) first step, it indubitably confirms the remarkable value of these methods and the need for further methodological and theoretical development of relational text-analytical methods. Empirically, my research exposed key findings regarding the subtle shifts in the discursive practices of distinct social actors (e.g., central banks, media, parliament) in the context of the global financial crisis. By investigating multiple discourses with different characteristics, I have shown that discourses of distinct social actors capture different aspects of a societal crisis. Understanding how these different aspects are manifested in these discourses can reveal the extent to which the discourses of certain social actors (organizations or individuals) contribute towards the containment or escalation of a societal crisis. I anticipate these noteworthy empirical results, combined with my methodological contributions, to open new avenues for large scale textual data research dealing with various discourses, even beyond the context of the financial crisis.

And since the very first sentences of this dissertation posed two important but broad questions, it is only proper to conclude on a similar note, by encouraging further research to participate in answering a fundamental question relevant to the recent textual focus in the social sciences: What is the role of texts, and the discursive manifestations they contain, in creating a global definition of societal problems and crises?

Summary

"Language is no longer regarded as peripheral to our grasp of the world in which we live, but as central to it. Words are not mere vocal labels or communicational adjuncts superimposed upon an already given order of things. They are collective products of social interaction, essential instruments through which human beings constitute and articulate their world." (Harris, 1988, p. ix)

Language and text are consequential to all aspects of social life and essential to our understanding of reality. Rather than being peripheral emanations of social activity, texts are collective products of social interaction and important instruments through which we constitute and articulate our world. Most of our knowledge of current and past world events, advances in science, or even elements of culture are gained, formed, and passed on through written and verbal text. But texts are not mere vehicles of factual information. They contain significant information about the orientations and beliefs of the actors generating them and the ways in which meanings are attributed. Furthermore, texts and the discursive strategies they contain have the potential to shape a reader's perceptions. In this sense, texts may become tools of influence, deception, and manipulation in the 'hands' of the writer or speaker, who may try to change the perceptions of their audiences or to impose a dominant narrative. Additionally, in the past decades, the increasing availability of such valuable textual information opened new venues for large-scale research in the social sciences, fostering an increase in attention given to text. The ease with which large volumes of data capturing social communication can be stored, accessed, and collected has risen to match the ambitions of social scientists in understanding behavior, structures, values or norms.

Given the ample availability of textual data and the valuable social information it contains, the question is no longer whether or not their content is worth investigating, but rather which approach is the most insightful for a given research goal. Hence, what motivated the aims and outcomes of the research

presented in this dissertation was the ambition to provide a model of inferring rich information from 'big text data' with different characteristics through relational text-analytical methods.

Exploiting automated and semi-automated text analysis tools, in this dissertation, I focused on expanding two relational analysis methods to text and meanings, namely semantic network analysis and topic modeling. The relational perspective on language and meaning posits that meaning is a relational phenomenon. Thus, it is not the individual words that generate meaning, but rather meaning is created through interrelated sets of words and concepts. Hence, words are mere symbols whose meaning is dependent on their use; that is, their relations to other words and the similarities and differences between these interconnected words. Analyzing text documents guided by this relational perspective on meaning entails moving beyond content analytic methods, towards the extraction of not only words and concepts, but also the complex relations that connect them. While semantic network analysis translates pre-selected text into networks of concepts and the links between them, topic modeling seek to identify, extract, and characterize the various (latent) topics contained by collections of texts. Topic models define a coherent topic as a set of word clusters, based on the same assumption regarding the relational aspect of meaning as semantic networks. These two text-analytical methods have become increasingly popular in many areas of social science research.

The main methodological contribution of this dissertation is an analytical approach to large collections of texts that allows for a comprehensive investigation of subtle discursive manifestations and their temporal dynamics. The structural space approach, based on the relational perspective to text and meaning, combines two classic social network analysis structural measures, degree centrality (i.e., popularity) and betweenness centrality (i.e., connectivity) of concepts, to create four structural roles for network nodes. This combination of structural measures, builds on the manner in which popular and connecting concepts play different roles in the structure and dynamics of texts, and allows for the identification of four structural roles. Ranking text concepts into four structural roles, then, exposes not only popular concepts, but also purely con-

nective concepts, marginal concepts, and those concepts that are both popular and connective. This approach allows analysts to uncover patterns of language use, subtle discursive shifts, and discursive dynamics in large text corpora with distinct characteristics.

Through the development of the analytic approach proposed (i.e., the structural space approach), the empirical chapters of this dissertation exposed key findings regarding the subtle shifts in the discursive practices of distinct social actors (e.g., central banks, media, parliament) in the context of the global financial crisis.

The three distinct discourses of social actors investigated in this dissertation (central bank discourses, media discourses, and political discourses), present distinctive characteristics, address different audiences, and fulfill different communicative purposes. Central bank discourses are highly formalized, leading to redundant, structured, and even predictable content. They also contain highly specialized information addressed mostly to those financial market actors interested in the actions and decisions of these banks. The media discourses on the other hand, present stories and employ various rhetorical devices to give these stories a newsworthy angle. News items are presented to large and relatively diverse audiences. Lastly, political discourses are characterized by normative as well as conventional or strategic practices that make the identification of meaningful information difficult. While in a general sense political discourses are addressed to the constituents or voters, here I focused on parliamentary speeches. These speeches are mostly addressed to other members of the parliament. The analysis of these distinct discourses produced important results, which I summarized below.

Chapter 3 investigates the discourses of the Federal Reserve System (Fed) and the European Central Bank (ECB) in the context of the financial crisis. The focus of the analysis in this chapter was directed at uncovering the shifts and adaptations of their discourse in a time of crisis and increased market volatility. Pursuing these aims, I explored the press releases of the ECB and Fed over a period of eight years, at the different stages of the financial crisis (i.e., pre-crisis, crisis, post-crisis, and the recovery period). The results revealed the

discourse of the Fed exhibiting greater attentiveness to the financial crisis, while the ECB's attention was delayed and increasing steadily. Furthermore, in this study both the Fed's and the ECB's discourses are shown to be transitioning into a new "hybrid" state in the economic recovery period (2012-2013), rather than returning to the pre-crisis status quo. Overall, this empirical chapter brings an important contribution to the understanding of financial-regulatory discourses and their dynamics during societal crises and high levels of financial market uncertainty.

The second empirical study presented in this dissertation (Chapter 4) focused on a different discursive space, namely that of the media. Investigating news items published by three newspapers (i.e., the Financial Times, the New York Times, and the Sun), I expose the ways in which metaphor families (i.e., sets of related metaphors) fulfill a translator role for emerging financial terminology in the media. This study brings a contribution to metaphor theory by widening the focus from conceptual metaphors to metaphor families, and exposes the expansion and evolution of the 'toxic' metaphor family, revealing subtle changes of metaphor use in three newspapers (i.e., the Financial Times, the New York Times, and the Sun) over time. I show a transition from generic image-creating metaphors toward financial-instrument targeted metaphors across three stages of the financial crisis, spanning a period of five years. Additionally, the results reveal that most of the 'toxic' metaphor family variations were created in the actual crisis period, not during the pre-crisis period. Overall, the findings of this study suggest that metaphor families are used as translating devices for unfamiliar terminology, practice which has the potential to influence perceptions of the financial crisis.

The last empirical study included in this dissertation, presented in Chapter 5, provided an account of the discursive space of the European Parliament (EP) in the context of the Eurozone financial crisis. The analysis in this study focused on the 7$^{\text{th}}$ European Parliament (2009-2014), whose rapid adoption of socio-economic legislative proposals played a key role in the European efforts to combat the rippling effects of the Eurozone financial crisis. With this study of parliamentary speeches I contribute to a better understanding of how pol-

icy issues are debated, the discursive practices employed, and the ideological divide in the European Parliament. The results presented in this chapter, uncovered a unified discursive space, with high levels of agreement on most issues of the Eurozone financial crisis between the two political groups investigated, and a moderate left-right ideological divide only when the unique topics of each political group are investigated.

The findings of these three empirical studies, show how different sets of social actors capture different aspects of the financial crisis through their discourses. For instance, while central banks focus on the immediate and localized events of the crisis within the financial markets, the medias' focus is on reporting emerging events and informing the public by covering a wider range of topics and events.

By investigating these discourses it is clear to see that, in a sense, a different 'version' of the crisis emerges from each of these discourses. Can we then talk about 'the financial crisis' as a unitary and single global event? If the argument is that a global 'definition' of the crisis emerges from combining the perspectives offered by various social discourses, what is the contribution of each perspective? While attaining such a global definition of the crisis might be intractable, these different perspectives have the potential to steer manifestations of the crisis by altering perceptions and prompting action. Hence, the analysis of multiple discursive sites can provide a more inclusive depiction of some of the mechanisms that contribute to the containment and/or amplification of a crisis.

Throughout this dissertation I argued and presented evidence in favor of relational text-analysis methods. By applying and expanding these methods, the contributions of my research have been both methodological as well as empirical. Methodologically, through the combination of text-analytical methods employed, and more specifically through the structural space approach, I provided compelling evidence for the great potential of relational text-analytic perspectives. While this dissertation is perhaps a (modest) first step, it indubitably confirms the remarkable value of these methods and the need for further methodological and theoretical development of relational text-analytical meth-

ods. Empirically, my research exposed key findings regarding the subtle shifts in the discursive practices of distinct social actors (e.g., central banks, media, parliament) in the context of the global financial crisis. By investigating multiple discourses with different characteristics, I have shown that discourses of distinct social actors capture different aspects of a societal crisis. Understanding how these different aspects are manifested in these discourses can reveal the extent to which the discourses of certain social actors (organizations or individuals) contribute towards the containment or escalation of a societal crisis. I anticipate these noteworthy empirical results, combined with my methodological contributions, to open new avenues for large scale textual data research dealing with various discourses, even beyond the context of the financial crisis.

List of Figures

List of Tables

Bibliography

Adams, I. (2001). *Political Ideology Today*. Manchester, UK: Manchester University Press.

Alejo, R. (2010). Where does the money go? An analysis of the container metaphor in economics: The market and the economy. *Journal of Pragmatics*, 42(4), 1137–1150. doi:10.1016/j.pragma.2009.08.013.

Allbritton, D. W. (1995). When metaphors function as schemas: Some cognitive effects of conceptual metaphors. *Metaphor and Symbolic Activity*, 10(1), 33–46. doi:10.1207/s15327868ms10014.

Alvesson, M. & Karreman, D. (2000). Varieties of discourse: On the study of organizations through discourse analysis. *Human Relations*, 53(9), 1125–1149. doi: 10.1177/0018726700539002.

Amato, J. D., Morris, S., & Shin, H. S. (2002). Communication and Monetary Policy. *Oxford Review of Economic Policy*, 18(4), 495–503. doi: 10.1093/oxrep/18.4.495.

An, S. K. & Gower, K. K. (2009). How do the news media frame crises? A content analysis of crisis news coverage. *Public Relations Review*, 35(2), 107–112. doi:10.1016/j.pubrev.2009.01.010.

Anderson, C. J., Wasserman, S., & Faust, K. (1992). Building stochastic blockmodels. *Social Networks*, 14(1), 137–161. doi:10.1016/0378–8733(92)90017–2.

Attiná, F. (1990). The voting behaviour of the European Parliament members and the problem of the Europarties. *European Journal of Political Research*, 18, 557–579. doi: 10.1111/j.1475–6765.1990.tb00248.x.

Bales, M. E. & Johnson, S. B. (2006). Graph theoretic modeling of large-scale semantic networks. *Journal of Biomedical Informatics*, 39(4), 451–464. doi:10.1016/j.jbi.2005.10.007.

Ball, S. J. (2012). *Politics and Policy Making in Education: explorations in sociology*. London and New York: Routledge.

Ball-Rokeach, S. J. (1985). The Origins of Individual Media-System Dependency: A Sociological Framework. *Communication Research*, 12(4), 485–510. doi: 10.1177/009365085012004003.

Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512. doi: 10.1126/science.286.5439.509.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. In *PNAS* (pp. 3747–52).

Belke, A. & Klose, J. (2010). (How) do the ECB and the fed react to financial market uncertainty? The Taylor rule in times of crisis. Discussion Paper No. 972. Retrieved from `http://www.econstor.eu/handle/10419/88238`.

Benoit, K. (2011). Data, Textual. In B. Badie, D. Berg-Schlosser, & L. Morlino (Eds.), *International Encyclopedia of Political Science* (pp. 526–531). Thousdand Oaks,CA: Sage.

Benoit, K., Laver, M., & Mikhaylov, S. (2009). Treating words as data with error: Uncertainty in text statements of policy positions. *American Journal of Political Science*, 53(2), 495–513. doi: 10.1111/j.1540–5907.2009.00383.x.

Berdayes, L. C. & Berdayes, V. (1998). The information highway in contemporary magazine narrative. *Journal of Communication*, 48(2), 109–124. doi: 10.1111/j.1460–2466.1998.tb02750.x.

Berg, B. L. (2007). *Qualitative Research Methods for the Social Sciences*. Boston, MA: Allyn and Bacon.

Berge, C. (1958). *Théorie des Graphes et ses Applications*. Paris, France: Dunod.

Bernanke, B., Reinhart, V., & Sack, B. (2004). Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment. Working paper No. 2004-48. Retrieved from `http://core.ac.uk/download/pdf/6893868.pdf`.

Bernard, H. R. & Ryan, B. (1998). Text Analysis. In H. R. Bernard (Ed.), *Handbook of Methods in Cultural Anthropology* (pp. 595–645). London and New Delhi: Altamira Press.

Bhatia, V. K., Flowerdew, J., & Jones, R. H., Eds. (2008). *Advances in Discourse Studies*. London and New York: Routledge.

Birke, J. & Sarkar, A. (2006). A Clustering Approach for the Nearly Unsupervised Recognition of Nonliteral Language. In *EACL* (pp. 329–336).

Black, M. (1962). *Models and Metaphors*. Ithaca and London: Cornell University Press.

Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.

Blei, D. M. & Lafferty, J. D. (2009). Topic models. In A. N. Srivastava & M. Sahami (Eds.), *Text Mining: Classification, Clustering, and Applications* chapter 4, (pp. 71–94). London: Taylor Taylor and Francis.

Blinder, A. S., Ehrmann, M., Fratzscher, M., Haan, J. D., & Jansen, D.-J. (2008). Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence. NBER Working Paper No. 13932. Retrieved from `http://www.nber.org/papers/w13932`.

Bobbio, N. (1996). *Left and Right: The Significance of a Political Distinction*. Chicago, IL: University of Chicago Press.

Bobrow, D. G. & Collins, A., Eds. (1975). *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press.

Boin, A., Hart, P., & McConnell, A. (2009). Crisis exploitation: political and policy impacts of framing contests. *Journal of European Public Policy*, 16(1), 81–106. doi:10.1080/13501760802453221.

Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5), 1170–1182.

Boorman, S. A. & White, H. C. (1976). Social structure from multiple networks. II. Role structures. *American Journal of Sociology*, (pp. 1384–1446).

Bordo, M. D. (2008). An Historical Perspective on the Crisis of 2007-2008. NBER Working Paper No. 14569. Retrieved from `http://www.nber.org/papers/w14569`.

Borgatti, S. P. & Everett, M. G. (1992). Notions of position in social network analysis. *Sociological Methodology*, 22(1), 1–35. doi:10.2307/270991.

Borge-Holthoefer, J. & Arenas, A. (2010). Semantic Networks: Structure and Dynamics. *Entropy*, 12(5), 1264–1302. doi:10.3390/e12051264.

Bosman, J. (1987). Persuasive effects of political metaphors. *Metaphor and Symbolic Activity*, 2(2), 97–113. doi:10.1207/s15327868ms02022.

Bounegru, L. & Forceville, C. (2011). Metaphors in editorial cartoons representing the global financial crisis. *Visual Communication*, 10(2), 209–229. doi:10.1177/1470357211398446.

Boyd, J. P. & Everett, M. G. (1999). Relations, residuals, regular interiors, and relative regular equivalence. *Social Networks*, 21(2), 147–165. doi:10.1016/S0378–8733(99)00006–4.

Brandes, U. & Corman, S. R. (2003). Visual unrolling of network evolution and the analysis of dynamic discourse. *Information Visualization*, 2(1), 40–50. doi:10.1057/palgrave.ivs.9500037.

Broekel, T. & Boschma, R. (2012). Knowledge networks in the dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, 12(2), 409–433. doi:10.1093/jeg/lbr010.

Broin, P. O. (2012). The Euro Crisis: EU Financial Reform and the European Parliament. IIEA Working paper No. 7. Retrieved from `http://www.iiea.com/publications/the-euro-crisis-eu-financial-reform-and-the-european-parliament`.

Bucchi, M. (1998). *Science and the Media: Alternative Routes in Scientific Communication*. London and New York: Routledge.

Budd, R. W., Thorp, R. K., & Donohew, L. (1967). *Content Analysis of Communications*. New York: Macmillan.

Burt, R. (1990). Detecting Role Equivalence. *Social Networks*, 12(1), 83–97. doi:10.1016/0378–8733(90)90023–3.

Burt, R. S. (1978). Cohesion Versus Structural Equivalence as a Basis for Network Subgroups. *Sociological Methods & Research*, 7(2), 189–212. doi: 10.1177/004912417800700205.

Burt, R. S. (2009). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Caballero, R., Farhi, E., & Gourinchas, P.-O. (2008). Financial crash, commodity prices and global imbalances. NBER Working Paper No. 14521. Retrieved from `http://www.nber.org/papers/w14521.ack`.

Caballero, R. J. & Simsek, A. (2009). Complexity and Financial Panics. NBER Working Paper No. 14997. Retrieved from `http://www.nber.org/papers/w14997`.

Campbell, K. K. & Jamieson, K. H. (1990). *Deeds Done in Words: Presidential Rhetoric and the Genres of Governance*. Chicago, IL: University of Chicago Press.

Carley, K. M. (1984). *Constructing Consensus*. PhD thesis, Harvard University.

Carley, K. M. (1986a). An approach for relating social structure to cognitive structure. *The Journal of Mathematical Sociology*, 12(2), 137–189. doi:10.1080/0022250X.1986.9990010.

Carley, K. M. (1986b). Knowledge Acquisition as a Social Phenomenon. *Instructional Science*, 14(3), 381–438.

Carley, K. M. (1988). The social construction of knowledge. In K. M. Carley, R. Lawler, & G. Drescher (Eds.), *Experiements in Epistemology: Learning in People and Machines*. New York: Ablex.

Carley, K. M. (1990). Content Analysis. In R. E. Asher & J. M. Y. Simpson (Eds.), *The Encyclopidia of Language and Linguistics* (pp. 725–730). Edinburgh, UK: Pergamon Press.

Carley, K. M. (1993). Coding Choices for Textual Analysis: A Comparison of Content Anlaysis and Map Analysis. *Sociological Methodology*, 23, 75–126.

Carley, K. M. (1997a). Extracting team mental models through textual analysis: Summary. *Journal of Organizational Behavior*, 18, 533–558.

Carley, K. M. (1997b). Network text analysis: The network position of concepts. In C. W. Roberts (Ed.), *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts* chapter 4, (pp. 79–102). Mahwah, NJ: Lawrence Erlbaum.

Carley, K. M., Columbus, D., & Landwehr, P. (2013a). *AutoMap User's Guide 2013*. Technical Report CMU-ISR-13-105, Carnegie Mellon University, Pittsburgh, PA.

Carley, K. M. & Kaufer, D. S. (1993). Semantic connectivity: An approach for analyzing symbols in semantic networks. *Communication Theory*, 3(3), 183–213. doi:10.1111/j.1468–2885.1993.tb00070.x.

Carley, K. M. & Palmquist, M. (1992). Extracting, Representing, and Analyzing Mental Models*. *Social Forces*, 70(3), 601–636. doi:10.1093/sf/70.3.601.

Carley, K. M., Pfeffer, J., Reminga, J., Storrick, J., & Columbus, D. (2013b). *ORA User's Guide 2013*. Technical report, CMU-ISR-13-108, Pittsburgh, PA.

Carlisle, R. P. (2005). *Encyclopedia of Politics: The Left and the Right*, volume 1. Thousand Oaks, CA: Sage Publications.

Cassirer, E. (1944). *An Essay on Man. An Introduction to a Philosophy of Human Culture.* Cambridge: Yale University Press.

Castells, M. (2000). The Rise Of The Network Society. In *The Information Age: Economy, Society and Culture* (pp. 594). Oxford: Blackwell Publishing, 2nd edition.

Castells, M. (2007). The Power of Identity. In *The Information Age: Economy, Society and Culture* (pp. 538). Oxford: Blackwell Publishing, 2nd edition.

Cavanagh, J. & Mader, J. (2004). *Alternatives to Economic Globalization. A Better World is Possible.* San Francisco: Berrett-Koehler.

Chalaby, J. K. (1996). Beyond the prison-house of language : discourse as a sociological concept. *The British Journal of Sociology*, 47(4), 684–698.

Chambers, E. (2001). *Contemporary Themes in Humanities Higher Education.* Dordrecht: Kluwer Academic Publishers.

Chang, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading Tea Leaves: How Humans Interpret Topic Models. In *Advances in Neural Information Processing Systems* (pp. 288–296).

Chang, T. M. (1986). Semantic memory: Facts and models. *Psychological Bulletin*, 99(2), 199–220.

Charteris-Black, J. & Musolff, A. (2003). 'Battered hero' or 'innocent victim'? A comparative study of metaphors for euro trading in British and German financial reporting. *English for Specific Purposes*, 22(2), 153–176. doi:10.1016/S0889–4906(02)00012–1.

Chilton, P. & Ilyin, M. (1993). Metaphor in Political Discourse: The Case of the 'Common European House'. *Discourse & Society*, 4(1), 7–31. doi: 10.1177/0957926593004001002.

Chirinko, R. S. & Curran, C. (2006). Greenspan Shrugs: Formal Pronouncements , Bond Market Volatility, And Central Bank Communication. In *Proceedings of the American Economic Association Meetings* (pp. 1–33).

Clayman, S. E. & Gill, V. T. (2012). Conversation analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 120–134). London and New York: Routledge.

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240–247. doi:10.1016/S0022–5371(69)80069–1.

Condor, S. & Antaki, C. (1997). Social cognition and discourse. In T. A. van Dijk (Ed.), *Discourse as Structure and Process* (pp. 1–34). London: Sage.

Congleton, R. D. (2009). On the political economy of the financial crisis and bailout of 2008–2009. *Public Choice*, 140(3-4), 287–317. doi:10.1007/s11127–009–9478–z.

Connolly, E. & Kohler, M. (2004). News and Interest Rate Expectations: A Study of Six Central Banks. Research Discussion Paper No. 2004-10. Economic Group Reserve Bank of Australia. Retrieved from `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.391.7106&rep=rep1&type=pdf`.

Coombs, W. T. & Holladay, S. J. (1996). Communication and Attributions in a Crisis: An Experimental Study in Crisis Communication. *Journal of Public Relations Research*, 8(4), 279–295. doi:10.1207/s1532754xjprr080404.

Corman, S., Kuhn, T., McPhee, R. D., & Dooley, K. J. (2002). Studying Complex Discursive Systems. *Human Communication Research*, 28(2), 157–206. doi:10.1111/j.1468–2958.2002.tb00802.x.

Cox, R., Beales, R., & Dixon, H. (2009). Moody's Mirrors Wall St. Salaries. *The New York Times*, (pp.B̃2).

Crotty, J. (2009). Structural causes of the global financial crisis: a critical assessment of the 'new financial architecture'. *Cambridge Journal of Economics*, 33(4), 563–580. doi:10.1093/cje/bep023.

Cucchiarelli, A., D'Antonio, F., & Velardi, P. (2012). Semantically interconnected social networks. *Social Network Analysis and Mining*, 2(1), 69–95. doi:10.1007/s13278–011–0030–z.

Danowski, J. (1993). Network analysis of message content. In W. D. Richards & G. A. Barnett (Eds.), *Progress in Communication Science* (pp. 197–222). Norwood, NJ: Ablex.

Danowski, J. A. (2012). Analyzing Change over Time in Organizations' Publics with a Semantic Network Include List: An Illustration with Facebook. In *ASONAM* (pp. 954–959).: IEEE.

D'Arcais, G. B. F. & Schreuder, R. (1987). Semantic activation during object naming. *Psychological Research*, 49(2-3), 153–159. doi:10.1007/BF00308681.

Davies, P. J. (2006). Axa IM: Debt markets are in the blood of derivatives specialist. *The Financial Times*, (pp.4̃).

de Haan, J. (2008). The effect of ECB communication on interest rates: An assessment. *The Review of International Organizations*, 3(4), 375–398. doi:10.1007/s11558–008–9048–z.

de Saussure, F. (1959). *Course in General Linguistics*. New York: Philosophical Library.

Dekker, D., Krackhardt, D., & Snijders, T. A. (2003). Multicollinearity robust qap for multiple regression. In *1st Annual Conference of the North American Association for Computational Social and Organizational Science* (pp. 22–25).

Dekker, D., Krackhardt, D., & Snijders, T. A. (2007). Sensitivity of mrqap tests to collinearity and autocorrelation conditions. *Psychometrika*, 72(4), 563–581.

DellaVigna, S. & Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics*, 122(3), 1187–1234. doi:10.1162/qjec.122.3.1187.

Diamond, D. W. & Rajan, R. (2009). The Credit Crisis: Conjectures About Causes And Remedies. NBER Working Paper No. 14739. Retrieved from `http://www.nber.org/papers/w14739`.

Diesing, P. (1991). *How Does Science Work? Reflections on Process.* Pittsburgh, PA: University of Pittsburgh Press.

Diesner, J. (2012a). *Uncovering and Managing the Impact of Methodological Choices for the Computational Construction of Socio-Technical Networks from Texts.* Technical Report CMU-ISR-12-101, Carnegie Mellon University, Pittsburgh, PA.

Diesner, J. (2012b). *Uncovering and Managing the Impact of Methodological Choices for the Computational Construction of Socio-Technical Networks from Texts.* PhD thesis, Carnegie Mellon University.

Diesner, J. (2013). From Texts to Networks: Detecting and Managing the Impact of Methodological Choices for Extracting Network Data from Text Data. *Künstliche Intelligenz*, 27(1), 75–78. doi:10.1007/s13218–012–0225–0.

Diesner, J., Aleyasen, A., Chin, C., Mishra, S., Soltani, K., & Tao, L. (2015). Context: Network construction from texts *[Software]*. Available from `http://context.lis.illinois.edu/`.

Diesner, J., Aleyasen, A., Kim, J., Mishra, S., & Soltani, S. (2013). Using socio-semantic network analysis for assessing the impact of documentaries. In *WIN (Workshop on Information in Networks)* New York, NY.

Diesner, J. & Carley, K. M. (2005). Revealing social structure from texts: meta-matrix text analysis as a novel method for network text analysis. In V. K. Narayanan & D. J. Amstrong (Eds.), *Causal Mapping for Information Systems* (pp. 81–108). London: Idea Group Publishing.

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, 41(6), 570–606. doi:10.1016/j.poetic.2013.08.004.

Doerfel, M. L. & Barnett, G. a. (1999). A Semantic Network Analysis of the International Communication Association. *Human Communication Research*, 25(4), 589–603. doi:10.1111/j.1468–2958.1999.tb00463.x.

Downe-Wamboldt, B. (1992). Content analysis: Method, applications, and issues. *Health Care for Women International*, 13(3), 313–321. doi:10.1080/07399339209516006.

ECB (2013). Retrieved from `https://www.ecb.europa.eu/ecb/html/index.en.html`. 17.06.2015.

Eco, U. (1979). *The Role of the Reader: Explorations in the Semiotics of Texts*. Bloomington, IN: Indiana University Press.

Ehrmann, M. & Fratzscher, M. (2007). Communication by Central Bank Committee Members: Different Strategies, Same Effectiveness? *Journal of Money, Credit and Banking*, 39(2), 509–541. doi:10.1111/j.0022–2879.2007.00034.x.

Entman, R. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 50–58. doi:10.1111/j.1460–2466.1993.tb01304.x.

Erdős, P. & Rényi, A. (1959). On random graphs. I. *Publicationes Mathematicae (Debrecen)*, 6, 290–297.

Erikson, R. S. & Tedin, K. L. (2003). *American Public Opinion: Its Orgin, Contents, and Impact*. New York: Longman Publishing Group.

European Union (1997). Treaty of Amsterdam amending the Treaty on European Union, the Treaties establishing the European Communities and related acts. *Official journal of the european communities*, (pp. 144).

Eurostat (2013). Retrieved from `http://ec.europa.eu/eurostat/`. Last accessed: 17-06-2015.

Fairclough, N. (1992). *Discourse and Social Change*. Cambridge, UK: Polity press.

Fairclough, N. (2001). *Language and Power*. Essex, UK: Pearson Education, 2nd edition.

Fairclough, N. (2012). Critical Discourse Analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 9–34). London and New York: Routledge.

Fass, D. (1991). met*: A method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, 17(1), 49–90.

Fauconnier, G. (1994). *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge: Cambridge University Press.

Federal Reserve System (2005). *The Federal Reserve System Purposes and Functions*. Washington DC: Board of Governors of the Federal Reserve System, 9th edition.

Ferrer I Cancho, R. & Solé, R. V. (2001). The small world of human language. In *Proceedings of Biological Sciences* (pp. 2261–5).

Financial Crisis Inquiry Commission (2011). *The Financial Crisis Inquiry Report: Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*. Technical report, The Financial Crisis Inquiry Commission, Washington, D.C.

Franzmann, S. & Kaiser, A. (2006). Locating Political Parties in Policy Space: A Reanalysis of Party Manifesto Data. *Party Politics*, 12(2), 163–188. doi:10.1177/1354068806061336.

Franzosi, R. (1998). Narrative Analysis—Or Why (and How) Sociologists Should Be Interested In Narrative. *Annual Review of Sociology*, 24(1), 517–554. doi:10.1146/annurev.soc.24.1.517.

Freeden, M., Ed. (2001). *Reassessing Political Ideologies: The Durability of Dissent*. London: Routledge.

Freeman, L. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1), 35–41. doi:10.2307/3033543.

Freeman, L. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239. doi:10.1016/0378–8733(78)90021–7.

Friedman, B. M. (2002). The use and meaning of words in central banking: Inflation targeting, credibility, and transparency. NBER Working Paper No. 8972. Retrieved from `http://www.nber.org/papers/w8972`.

Gabel, M. & Hix, S. (2002). Defining the eu political space: An empirical study of the european elections manifestos, 1979-1999. *Comparative Political Studies*, 35(8), 934–964. doi:10.1177/001041402236309.

Gabel, M. & Huber, J. D. (2000). Putting Parties in Their Place: Inferring Party Left-Right Ideological Positions from Party Manifestos Data. *American Journal of Political Science*, 44(1), 94–103. doi:10.2307/2669295.

Gamson, W. A. & Modigliani, A. (1989). Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach. *American Journal of Sociology*, 95(1), 1–37. doi:10.1086/229213.

Gastil, J. (1992). Undemocratic discourse: a review of theory and research on political discourse. *Discourse & Society*, 3(4), 469–500. doi:10.1177/0957926592003004003.

Gedigian, M., Bryant, J., Narayanan, S., & Ciric, B. (2006). Catching metaphors. In *Proceedings of the 3rd Workshop on Scalable Natural Language Understanding* (pp. 41–48). New York: ACL.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz (Ed.), *The Interpretation of Culture* (pp. 3–30). New York: Basic Basic Books.

George, A. (1959). Quantitative and qualitative approaches to content analysis. In I. de S. Pool (Ed.), *Trends in Content Analysis.* (pp. 7–32). Chicago, IL: University of Illinois Press.

Gerber, A. S., Karlan, D., & Bergan, D. (2009). Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and

political opinions. *American Economic Journal: Applied Economics*, 1(2), 35–52.

Gerring, J. (1997). Ideology: A Definitional Analysis. *Political Research Quarterly*, 50(4), 957–994. doi:10.2307/448995.

Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley, CA: University of California Press.

Gloor, P. & Krauss, J. (2009). Web science 2.0: Identifying trends through semantic social network analysis. In *Proceedings of the International Conference on Computational Science and Engineering* (pp. 215–222).: IEEE.

Goatly, A. (1997). *The Language of Metaphors*. London: Routledge.

Goldman, a. (2008). Consultant and Critics on the Couch. *Journal of Management Inquiry*, 17(3), 243–249. doi:10.1177/1056492608318160.

Goldstone, A. & Underwood, T. (2012). What can topic models of pmla teach us about the history of literary scholarship? *Journal of Digital Humanities*, 2(1).

Gollob, H. F. (1968). Impression formation and word combination in sentences. *Journal of Personality and Social Psychology*, 10(4), 341–353. doi:10.1037/h0026822.

Gower, J. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3), 325–338. doi:10.2307/2333639.

Grant, D. & Hardy, C. (2004). Introduction: Struggles with Organizational Discourse. *Organization Studies*, 25(1), 5–13. doi:10.1177/0170840604038173.

Gray, D. & Griffin, C. (2014). A journey to citizenship: Constructions of citizenship and identity in the British Citizenship Test. *British Journal of Social Psychology*, 53(2), 299–314.

Grebitus, C. & Bruhn, M. (2008). Analyzing semantic networks of pork quality by means of concept mapping. *Food Quality and Preference*, 19(1), 86–96. doi:10.1016/j.foodqual.2007.07.007.

Griffiths, T. & Steyvers, M. (2002). A probabilistic approach to semantic representation. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 381–386).

Griffiths, T. L. & Steyvers, M. (2003). Prediction and semantic association. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Neural Information Processing Systems 15* (pp. 11–18). Cambridge, MA: MIT Press.

Griffiths, T. L. & Steyvers, M. (2004). Finding scientific topics. In *Proceedings of the National Academy of Sciences* (pp. 5228–5235).: National Academy of Sciences.

Grimmer, J. & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297. doi:10.1093/pan/mps028.

Grishman, R. (1986). *Computational Linguistics: An Introduction*. Cambridge, UK: Cambridge University Press.

Gürkaynak, R. S., Sack, B., & Swanson, E. (2005a). Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. *International Journal of Central Banking*, 1(1), 55–93.

Gürkaynak, R. S., Sack, B., & Swanson, E. (2005b). The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *The American Economic Review*, 95(1), 425–436. doi:10.2307/4132689.

Haass, R. & Litan, R. E. (1998). Globalization and Its Discontents-Navigating the Dangers of a Tangled World. *Foreign Affairs*, 2(June), 2–5.

Hallahan, K. (1999). Seven models of framing: Implications for public relations. *Journal of Public Relations Research*, 11(3), 205–242. doi:10.1207/s1532754xjprr1103_02.

Halliday, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. Baltimore, MD: University Park Press.

Hansen, D., Shneiderman, B., & Smith, M. A. (2011). *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*. Boston, MA: Elsevier.

Hardy, C., Palmer, I., & Phillips, N. (2000). Discourse as a Strategic Resource. *Human Relations*, 53(9), 1227–1248. doi:10.1177/0018726700539006.

Harris, R. (1988). *Language, Saussure and Wittgenstein. How to play games with words*. London and New York: Routledge.

Harris, S. (1991). Evasive action: How politicians respond to questions in political interviews. In *Broadcast Talk* (pp. 76–99). London: Sage Publications.

Hayes, M. (1997). The use of metaphor in reporting financial market transactions. *Cuadernos de Filología Inglesa*, 612, 233–245.

Hayo, B., Kutan, A. M., & Neuenkirch, M. (2014). Financial Market Reaction to Federal Reserve Communications: Does the Crisis Make a Difference? *Empirica*, 42(1), 185–203. doi:10.1007/s10663–014–9258–y.

Hayo, B. & Neuenkirch, M. (2010). Do Federal Reserve Communications Help Predict Federal Funds Target Rate Decisions? *Journal of Macroeconomics*, 32(4), 1014–1024. doi: 10.1016/j.jmacro.2010.06.003.

Heath, R. L. (2010). Introduction Crisis Communication: Defining the Beast and De-marginalizing Key Publics. In W. T. Coombs & S. J. Holladay (Eds.), *The Handbook of Crisis Communication* (pp. 1–13). Oxford: Wiley-Blackwell.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.

Heise, D. R. (1969). Affectual dynamics in simple sentences. *Journal of Personality and Social Psychology*, 11(3), 204–313. doi:10.1037/h0027050.

Heise, D. R. (1970). Potency dynamics in simple sentences. *Journal of Personality and Social Psychology*, 16(1), 48–54. doi:10.1037/h0030084.

Hellsten, I. (2000). Dolly: Scientific breakthrough or Frankenstein's monster? Journalistic and scientific metaphors of cloning. *Metaphor & Symbol*, 15(4), 213–221. doi:10.1207/S15327868MS15043.

Hellsten, I. (2002). *The Politics of Metaphor: Biotechnology and Biodiversity in the Media*. Phd dissertation, University of Tampere.

Henderson, G., Iacobucci, D., & Calder, B. (1998). Brand diagnostics: mapping branding effects using consumer associative networks. *European Journal of Operational Research*, 111(2), 306–327. doi:10.1016/S0377–2217(98)00151–9.

Hernández-Guerra, C. (2013). Textual, intertextual and rhetorical features in political discourse: the case of President Obama in Europe. *Revista de Lingüística y Lenguas Aplicadas*, 8, 59–65. doi:10.4995/rlyla.2013.1175.

Hill, V. & Carley, K. M. (1999). An approach to identifying consensus in a subfield: The case of organizational culture. *Poetics*, 27(1), 1–30. doi:10.1016/S0304–422X(99)00004–2.

Hix, S. (2011). Parliamentary Behavior with Two Principals: Preferences, parties, and voting in the European Parliament. *American Journal of Political Science*, 46(3), 688–698. doi:10.2307/3088408.

Hix, S. & Høyland, B. (2013). Empowerment of the European Parliament. *Annual Review of Political Science*, 16, 171–189. doi:10.1146/annurev–polisci–032311–110735.

Hix, S., Kreppel, A., & Noury, A. (2003). The Party System in the European Parliament: Collusive or competitive? *JCMS: Journal of Common Market Studies*, 41(2), 309–331. doi:10.1111/1468–5965.00424.

Hix, S., Noury, A., & Roland, G. (2005). Power to the Parties: Cohesion and Competition in the European Parliament, 1979-2001. *British Journal of Political Science*, 35(2), 209–234. doi:10.1017/S0007123405000128.

Hix, S., Noury, A., & Roland, G. (2006). Dimensions of politics in the european parliament. *American Journal of Political Science*, 50(2), 494–520.

Hofmann, T. (1999). Probabilistic latent semantic indexing. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 50–57).: ACM.

Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42(1), 177–196.

Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley.

Hooper, C. J., Marie, N., & Kalampokis, E. (2012). Dissecting the Butterfly: Representation of Disciplines Publishing at the Web Science Conference Series. In *WebSci* (pp. 197–200).

Hopkins, D. J. & King, G. (2010). Extracting Systematic Social Science Meaning from Text. *American Journal of Political Science*, 54(1), 229–247.

Hoser, B., Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006). Semantic network analysis of ontologies. In Y. Sure & J. Domingue (Eds.), *The Semantic Web: Research and Applications*, Lecture Notes in Computer Science (pp. 514–529). Springer.

Høyland, B., Godbout, J.-F., Lapponi, E., & Velldal, E. (2014). Predicting Party Affiliations from European Parliament Debates. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science* (pp. 56–60). Baltimore, Maryland, USA: ACL.

Huang, S., Lv, T., Zhang, X., Yang, Y., Zheng, W., & Wen, C. (2014). Identifying node role in social network based on multiple indicators. *PloS one*, 9(8), e103733. doi:10.1371/journal.pone.0103733.

Hulst, R. C. (2008). Introduction to Social Network Analysis (SNA) as an investigative tool. *Trends in Organized Crime*, 12(2), 101–121. doi:10.1007/s12117–008–9057–6.

Iacobucci, D., Henderson, G., Marcati, A., & Chang, J. (1996). Network analyses of brand switching behavior. *International Journal of Research in Marketing*, 13, 415–429. doi:10.1016/S0167–8116(96)00022–5.

Irvine, J. T. (1979). Formality and Informality in Communicative Events. *American Anthropologist*, 81(4), 773–790. doi:10.1525/aa.1979.81.4.02a00020.

Iyengar, S. (1987). Television news and citizens' explanations of national affairs. *American Political Science Review*, 81, 815–831. doi:10.2307/1962678.

Jansen, D. & de Haan, J. (2005a). Is a word to the wise indeed enough? ECB statements and the predictability of interest rate decisions. DNB Working Paper No. 75. Retrieved from `http://www.dnb.nl/binaries/Working%20Paper%2075_tcm46-146732.pdf`.

Jansen, D. J. & de Haan, J. (2005b). Were verbal efforts to support the euro effective? A high-frequency analysis of ECB statements. *European Journal of Political Economy*, 23(1), 245–259. doi:10.1016/j.ejpoleco.2006.09.017.

Jin, Y., Lin, C.-Y., Matsuo, Y., & Ishizuka, M. (2012). Mining dynamic social networks from public news articles for company value prediction. *Social Network Analysis and Mining*, 2(3), 217–228. doi:10.1007/s13278–011–0045–5.

Johnson, M. (1981). Introduction: Metaphor in the Philosophical Tradition. In *Philosophical Perspectives on Metaphor* (pp. 3–47). Minneapolis: University of Minnesota Press.

Johnson-Laird, P. N., Herrmann, D. J., & Chaffin, R. (1984). Only connections: A critique of semantic networks. *Psychological Bulletin*, 96(2), 292–315. doi:10.1037/0033–2909.96.2.292.

Johnston, H. (2002). Verification and Proof in Frame and Discourse Analysis. In B. Klandermans & S. Staggenborg (Eds.), *Methods of Social Movement Research* (pp. 62–91). Minneapolis: University of Minnesota Press.

Jonassen, D. & Marra, R. (1994). Concept mapping and other formalisms as mindtools for representing knowledge. *Research in Learning Technology*, 2(1), 50–56. doi:10.3402/rlt.v2i1.9573.

Jost, J. T. (2006). The end of the end of ideology. *The American psychologist*, 62(7), 1077–1080. doi:10.1037/0003–066X.62.9.1077.

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political Ideology: Its structure, functions, and elective affinities. *Annual review of psychology*, 60, 307–337. doi:10.1146/annurev.psych.60.110707.163600.

Kennedy, V. (2000a). Intended Tropes and Unintended Metatropes in Reporting on the War in Kosovo. *Metaphor & Symbol*, 15(4), 253–265. doi:10.1207/S15327868MS15046.

Kennedy, V. (2000b). Metaphors in the news-introduction. *Metaphor & Symbol*, 15(4), 209–211. doi:10.1207/S15327868MS15042.

Kilduff, M. & Krackhardt, D. (1994). Bringing the individual back in: A structural analysis of the internal market for reputation in organizations. *Academy of Management Journal*, 37(1), 87–108. doi:10.2307/256771.

Kim, L. (2011). Denotation and connotation in public representation: semantic network analysis of Hwang supporters' internet dialogues. *Public Understanding of Science*, 22(3), 335–350. doi:10.1177/0963662511401784.

Kitis, E. & Milapides, M. (1997). Read it and believe it: How metaphor constructs ideology in news discourse. A case study. *Journal of Pragmatics*, 28, 557–590. doi:10.1016/S0378–2166(97)00075–1.

Kleinnijenhuis, J., Schultz, F., Oegema, D., & van Atteveldt, W. (2013). Financial news and market panics in the age of high-frequency sentiment trading algorithms. *Journalism*, 14(2), 271–291. doi:10.1177/1464884912468375.

Kleinnijenhuis, J., van Hoof, A. M. J., Oegema, D., & de Ridder, J. A. (2007). A test of rivaling approaches to explain news effects: News on issue positions of parties, real-world developments, support and criticism, and success and failure. *Journal of Communication*, 57(2), 366–384. doi:10.1111/j.1460–2466.2007.00347.x.

Knight, K. (2006). Transformations of the Concept of Ideology in the Twentieth Century. *American Political Science Review*, 100(4), 619–626. doi:10.1017/S0003055406062502.

Kohn, D. L. & Sack, B. P. (2003). Central Bank Talk: Does It Matter and Why? FEDS Working Paper No. 2003-55. Retrieved from `http://www.federalreserve.gov/pubs/feds/2003/200355/200355pap.pdf`.

Krackhardt, D. (1987). QAP Partialling as a Test of Spuriousness. *Social Networks*, 9, 171–186. doi:10.1016/0378–8733(87)90012–8.

Kreppel, A. (2002). *The European Parliament and Supranational Party System: A Study in Institutional Development*. Cambridge: Cambridge University Press.

Kreppel, A. & Tsebelis, G. (1999). Coalition Formation in the European Parliament. *Comparative Political Studies*, 32(8), 933–966. doi: 10.1177/0010414099032008002.

Krippendorff, K. (1980). *Content Analysis. An Introduction to its Methodology*. London: Sage.

Krippendorff, K. (2004). Measuring the Reliability of Qualitative Text Analysis Data. *Quality & Quantity*, 38(6), 787–800. doi:10.1007/s11135–004–8107–7.

Krishnakumaran, S. & Zhu, X. (2007). Hunting elusive metaphors using lexical resources. In *Proceedings of Workshop Computational Approaches to Figurative Language* (pp. 13–20). Rochester, New York: ACL.

Lakoff, G. (1995). Metaphor, morality, and politics, or, why conservatives have left liberals in the dust. *Social Research*, 62(2), pp. 177–213.

Lakoff, G. (2004). *Don't Think of an Elephant! Know Your Values and Frame the Debate*. White River Junction, VT: Chelsea Green Publishing.

Lakoff, G. & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, 4(2), 195–208. doi:10.1016/S0364–0213(80)80017–6.

Lakoff, G. & Johnson, M. (2003). *Metaphors We Live By.* Chicago: Chicago London, 2nd edition.

Lane, R. E. (1967). *Political Ideology.* Oxford, UK: Free Press.

Lasswell, H. D. (1948). The structure and function of communication in society. In L. Bryson (Ed.), *The Communication of Ideas* (pp. 37–51). New York: Harper & Row.

Laver, M., Benoit, K., & Garry, J. (2003). Extracting Policy Positions from Political Texts Using Words as Data. *American Political Science Review*, 97(2), 311–331. doi:10.1017/S0003055403000698.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Computational social science. *Science*, 323, 721–723. doi:10.1126/science.1167742.

Leach, J. (2003). Rhetorical Analysis. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative Researching With Text, Image and Sound. A Practical Handbook* (pp. 207–226). London: Sage Publications.

Lee, J.-S. & Pfeffer, J. (2015a). Estimating centrality statistics for complete and sampled networks: Some approaches and complications. In *Proceedings of the 48th HICSS* (pp. 1686–1695).

Lee, J.-S. & Pfeffer, J. (2015b). Robustness of network centrality metrics in the context of digital communication data. In *Proceedings of the 48th HICSS* (pp. 1789–1870).

Lehmann, F. (1992). Semantic networks. *Computers and Mathematics with Applications*, 23(2), 1–50.

Leydesdorff, L. (1995). *The Challenge of Scientometrics: The development, measurement, and self-organization of scientific communications.* Leiden: DSWO Press.

Leydesdorff, L. (2013). FullText.exe. Available at `http://www.leydesdorff.net/software/fulltext/index.htm`.

Leydesdorff, L. & Hellsten, I. (2005). Metaphors and Diaphors in Science Communication: Mapping the Case of Stem Cell Research. *Science Communication*, 27(1), 64–99. doi:10.1177/1075547005278346.

Leydesdorff, L. & Hellsten, I. (2006). Measuring the meaning of words in contexts: An automated analysis of controversies about 'Monarch butterflies,' 'Frankenfoods,' and 'stem cells'. *Scientometrics*, 67(2), 231–258. doi:10.1007/s11192–006–0096–y.

Leydesdorff, L. & Welbers, K. (2011). The semantic mapping of words and co-words in contexts. *Journal of Informetrics*, 5(3), 469–475. doi:10.1016/j.joi.2011.01.008.

Leydesdroff, L. (1989). Words and co-words as indicators of intellectual organization. *Research Policy*, 18(4), 209–223. doi:10.1016/0048–7333(89)90016–4.

Li, C., Li, Q., Van Mieghem, P., Stanley, H. E., & Wang, H. (2015). Correlation between centrality metrics and their application to the opinion model. *The European Physical Journal B*, 88(65), 1–13. doi:10.1140/epjb/e2015–50671–y.

Lindkvist, K. (1981). Approaches to textual analysis. In *Advances in Content Analysis* (pp. 23–41). Beverly Hills, CA: Sage Publications.

Loewenstein, J., Ocasio, W., & Jones, C. (2012). Vocabularies and vocabulary structure: A new approach linking categories, practices, and institutions. *The Academy of Management Annals*, 6(1), 41–86. doi:10.1080/19416520.2012.660763.

López, A. M. & Llopis, M. A. (2010). Metaphorical pattern analysis in financial texts: Framing the crisis in positive or negative metaphorical terms. *Journal of Pragmatics*, 42(12), 3300–3313. doi:10.1016/j.pragma.2010.06.001.

Lukes, S. (2008). Epilogue: The grand dichotomy of the twentieth century. In T. Ball & R. Bellamy (Eds.), *The Cambridge History of Twentieth-Century Political Thought* (pp. 602–626). Cambridge, UK: Cambridge University Press.

Lynch, M. & Woolgar, S. (1988). Introduction: Sociological orientations to representational practice in science. *Human Studies*, 11(2-3), 99–116. doi:10.1007/BF00177300.

Mandler, J. M. (1982). Some uses and abuses of a story grammar. *Discourse Processes*, 5(3-4), 305–318. doi:10.1080/01638538209544547.

Mandler, J. M. (1984). *Stories, Scripts, and Scenes: Aspects of Schema Theory.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Martin, M. K., Pfeffer, J., & Carley, K. M. (2013). Network text analysis of conceptual overlap in interviews, newspaper articles and keywords. *Social Network Analysis and Mining*, 3(4), 1165–1177. doi:10.1007/s13278–013–0129–5.

Mason, Z. J. (2004). CorMet: A Computational, Corpus-Based Conventional Metaphor Extraction System. *Computational Linguistics*, 30(1), 23–44. doi:10.1162/089120104773633376.

Maynard, S. K. (1994). Images of involvement and integrity: Rhetorical style of a japanese politician. *Discourse & Society*, 5(2), 233–261. doi:10.1177/0957926594005002005.

Mazzoleni, G. & Schulz, W. (1999). "Mediatization" of Politics: A Challenge for Democracy? *Political Communication*, 16(3), 37–41. doi:10.1080/105846099198613.

McElroy, G. & Benoit, K. (2007). Party groups and policy positions in the european parliament. *Party Politics*, 13(1), 5–28. doi:10.1177/1354068806068593.

McElroy, G. & Benoit, K. (2010). Party policy and group affiliation in the european parliament. *British Journal of Political Science*, 40(02), 377–398. doi:10.1017/S0007123409990469.

McElroy, G. & Benoit, K. (2012). Policy positioning in the European Parliament. *European Union Politics*, 13(1), 150–167. doi:10.1177/1465116511416680.

McEnery, A. M. (1992). *Computational Linguistics: A Handbook and Toolbox for Natural Language Processing*. Wilmslow, UK: Sigma.

McTavish, D. G. & Pirro, E. B. (1990). Contextual Content Analysis. *Quality & Quantity*, 24, 245–265. doi:10.1007/BF00139259.

Meeks, E. & Weingart, S. B. (2012). The digital humanities contribution to topic modeling. *Journal of Digital Humanities*, 2(1).

Merton, R. K. (1957). *Social Theory and Social Structure*. Glencoe, IL: Free Press.

Mikhaylov, S., Laver, M., & Benoit, K. R. (2012). Coder Reliability and Misclassification in the Human Coding of Party Manifestos. *Political Analysis*, 20(1), 78–91. doi:10.1093/pan/mpr047.

Miller, G., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. (1990). Introduction to wordnet: An on-line lexical database*. *International Journal of Lexicography*, 3(4), 235–244.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The Psychology of Computer Vision* (pp. 211–277). New York: McGraw-Hill.

Mohr, J. W. & Bogdanov, P. (2013). Introduction-Topic models: What they are and why they matter. *Poetics*, 41(6), 545–569. doi:10.1016/j.poetic.2013.10.001.

Monge, P. R. & Contractor, N. S. (2001). Emergence of communication networks. In F. M. Jablin & L. Putnam (Eds.), *The New Handbook of Organizational Communication: Advances in theory, research and methods* chapter 12, (pp. 440–502). Thousand Oaks, CA: Sage Publications.

Murphy, P. (1996). Chaos theory as a model for managing issues and crises. *Public Relations Review*, 22(2), 95–113. doi:10.1016/S0363–8111(96)90001–6.

Musolff, A. (2004). *Metaphor and Political Discourse. Analogical Reasoning in Debates About Europe*. New York: Palgrave Macmillan, 1st edition.

Nelson, S. C. & Katzenstein, P. J. (2014). Uncertainty, Risk, and the Financial Crisis of 2008. *International Organization*, 68(2), 361–392. doi:10.1017/S0020818313000416.

Nerghes, A., Hellsten, I., & Groenewegen, P. (2015a). A Toxic Crisis: Metaphorizing the Financial Crisis. *International Journal of Communication*, 9, 106–132. doi:1932–8036/20150005.

Nerghes, A., Lee, J.-S., & Groenewegen, P. (2013). EU Regulators: A Structure and Content Analysis. XXXIII Sunbelt Social Networks Conference of INSNA, Hamburg, Germany.

Nerghes, A., Lee, J.-S., & Groenewegen, P. (2014a). Market Player or Regulator? A Semantic Network Analysis of the Shifting Roles of the ECB During the Financial Crisis. XXXIV Sunbelt Social Networks Conference of INSNA, St. Pete Beach, FL, USA.

Nerghes, A., Lee, J.-S., Groenewegen, P., & Hellsten, I. (2014b). The shifting discourse of the European Central Bank: Exploring structural space in semantic networks. In K. Yetongnon & A. Dipanda (Eds.), *Proceedings of the International Conference on Signal-Image Technology & Internet-Based Systems* (pp. 447–455). Marrakesh, Morocco: IEEE.

Nerghes, A., Lee, J.-S., Groenewegen, P., & Hellsten, I. (2015b). Mapping discursive dynamics of the financial crisis: a structural perspective of concept roles in semantic networks. *Computational Social Networks*, 2(16), 1–29. doi:10.1186/s40649–015–0021–8.

Nerlich, B., Clarke, D., & Dingwall, R. (2000). Clones and Crops: The Use of Stock Characters and Word Play in Two Debates About Bioengineering. *Metaphor & Symbol*, 15(4), 223–239. doi:10.1207/S15327868MS15044.

Nerlich, B. & Jaspal, R. (2012). Metaphors We Die By? Geoengineering, Metaphors, and the Argument From Catastrophe. *Metaphor & Symbol*, 27(2), 131–147. doi:10.1080/10926488.2012.665795.

Norris, F. (2004). Speculating in Worthless Stock While the Trains Speed to London. *The New York Times*, (pp.1̃).

Nystrand, M. & Himley, M. (1984). Written text as social interaction. *Theory Into Practice*, 23(3), 198–207. doi:10.1080/00405848409543114.

Obstfeld, M. & Rogoff, K. (2009). Global Imbalances and the Financial Crisis: Products of Common Causes. In *Federal Reserve Bank of San Francisco Global recovery: Asia and the new financial landscape Conference Proceedings* (pp. 173–177).: Federal Reserve Bank of San Francisco.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3), 245–251. doi:10.1016/j.socnet.2010.03.006.

Ortony, A. (1993). Metaphor, Language, and Thought. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 696). Cambridge: Cambridge University Press, 2nd editio edition.

Osgood, C. E. (1959). The representational model and relevant research methods. In I. de S. Pool (Ed.), *Trends in Content Analysis* (pp. 33–88). Champaign, IL: University of Illinois Press.

Pan, Z. & Kosicki, G. M. (1993). Framing analysis: An approach to news discourse. *Political Communication*, 10(1), 55–75. doi:10.1080/10584609.1993.9962963.

Paris, R. (2002). Kosovo and the Metaphor War. *Political Science Quarterly*, 117(3), 423–450. doi:10.2307/798263.

Peckham, R. (2013). Economies of Contagion: Financial Crisis and Pandemic. *Economy and Society*, 42(2), 226–248. doi:10.1080/03085147.2012.718626.

Pendery, D. (2009). Three Top Economists Agree 2009 Worst Financial Crisis Since Great Depression; Risks Increase if Right Steps are Not Taken. *Reuters*, (pp.1̃).

Peters, W. & Peters, I. (2000). Lexicalised Systematic Polysemy in WordNet. In *Proceedings of LREC* Athens, Greece.

Petersen, A. (2005). The Metaphors of Risk: Biotechnology in the news. *Health, Risk & Society*, 7(3), 203–208. doi:10.1080/13698570500229572.

Phillips, N., Sewell, G., & Jaynes, S. (2008). Applying critical discourse analysis in strategic management research. *Organizational Research Methods*, 11(4), 770–789. doi:10.1177/1094428107310837.

Pinkleton, B. E., Austin, E. W., & Fortman, K. K. J. (1998). Relationships of media use and political disaffection to political efficacy and voting behavior. *Journal of Broadcasting & Electronic Media*, 42(1), 34–49. doi:10.1080/08838159809364433.

Podpiera, A. M. & Otker-Robe, I. (2013). The social impact of financial crises: evidence from the global financial crisis. Working Paper No. WPS6703, The World Bank. Retrieved from: `http://documents.worldbank.org/curated/en/2013/11/18512644/social-impact-financial-crises-evidence-global-financial-crisis`.

Polanyi, M. (1962). *Personal Knowledge: Towards a Post-Critical Philosophy.* Chicago, IL: University of Chicago Press.

Pollach, I. (2012). Taming textual data: The contribution of corpus linguistics to computer-aided text analysis. *Organizational Research Methods*, 15(2), 263–287. doi:10.1177/1094428111417451.

Pool, I. D. S. (1959). *Trends in Content Analysis.* Champaign, IL: University of Illinois Press.

Poole, K. T. & Rosenthal, H. (2000). *Congress: A Political-Economic History of Roll Call Voting.* New York: Oxford University Press.

Popping, R. (2000). *Computer-Assisted Text Analysis*. Thousand Oaks, CA: Sage Publications.

Popping, R. (2003). Knowledge Graphs and Network Text Analysis. *Social Science Information*, 42(1), 91–106. doi:10.1177/0539018403042001798.

Postma, E., Wiesman, F., & van den Herik, H. (2000). Small-world semantic networks. In *Proceedings of the Twelfth Belgium-Netherlands Artificial Intelligence Conference* (pp. 191–198). Tilburg: Tilburg University.

Potter, J. & Wetherell, M. (1987). *Discourse and Social Psychology: Beyond Attitudes and Behaviour*. London: Sage.

Potter, J., Wetherell, M., Gill, R., & Edwards, D. (1990). Discourse: noun, verb or social practice? *Philosophical Psychology*, 3(2), 205–217. doi:10.1080/09515089008572999.

Proksch, S.-O. & Slapin, J. B. (2010). Position Taking in European Parliament Speeches. *British Journal of Political Science*, 40(December), 587–611. doi:10.1017/S0007123409990299.

Quarles, R. C. (1979). Mass Media Use and Voting Behavior The Accuracy of Political Perceptions among First-Time and Experienced Voters. *Communication Research*, 6(4), 407–436. doi:10.1177/009365027900600402.

Ramage, D., Dumais, S., & Liebling, D. (2010). Characterizing Microblogs with Topic Models. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media* (pp. 130–137).: AAAI.

Ray, L. (1999). Measuring party orientations towards European integration: results from an expert survey. *European Journal of Political Research*, 36, 283–306. doi:10.1111/1475–6765.00471.

Reinhart, C. & Rogoff, K. S. (2008). Is the 2007 US sub-prime financial crisis so different? An international historical comparison. *American Economic Review*, 98(2), 339–344. doi:10.1257/aer.98.2.339.

Reinhart, V. & Sack, B. (2006). Grading the federal open market committee's communications. In *AEA Annual Meetings in Boston, January* (pp. 6–8).

Rhodes, C. & Garrick, J. (2002). Economic metaphors and working knowledge: enter the 'cogito-economic' subject. *Human Resource Development International*, 5(1), 87–97. doi:10.1080/13678860110018803.

Riaz, S. (2009). The global financial crisis: an institutional theory analysis. *Critical Perspectives on International Business*, 5(1/2), 26–35. doi:10.1108/17422040910938668.

Rice, R. E. & Danowski, J. A. (1993). Is it really just like a fancy answering machine? comparing semantic networks of different types of voice mail users. *International Journal of Business Communication*, 30(4), 369–397.

Ricoeur, P. (1991). *From Text to Action. Essays in Hermeneutics, II*. Evanson, IL: Northwestern University Press.

Roberts, A. (2004). Relieving the debt hangover by giving a fine tune to the accounts. *The Financial Times*, (pp.$\tilde{2}$4).

Roberts, C. W. (1989). Other Than Counting Words: Approach to Content Analysis. *Social Forces*, 68(1), 147–177. doi: 10.1177/002194369303000401.

Roberts, C. W. (2000). A Conceptual Framework for Quantitative Text Analysis - On Joining Probabilities and Substantive Inferences About Texts. *Quality & Quantity*, 34(3), 259–274. doi:10.1023/A:1004780007748.

Rosa, C. & Verga, G. (2005). Is ECB communication effective? CEP Discussion Paper No. 682. Retrieved from `https://ideas.repec.org/p/cep/cepdps/dp0682.html`.

Rouse, J. (2007). Practice Theory. In *Handbook of the Philosophy of Science* (pp. 499–540). Division I Faculty Publications, Wesleyan University. Retrieved from `http://wesscholar.wesleyan.edu/div1facpubs/43`.

Sahuc, J. G. & Smets, F. (2008). Differences in Interest Rate Policy at the ECB and the Fed: An Investigation with a Medium-Scale DSGEModel. *Journal of Money, Credit and Banking*, 40(2-3), 505–521. doi:10.1111/j.1538–4616.2008.00124.x.

Schleppegrell, M. J. (2012). Systemic functional linguistic. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 21–34). London and New York: Routledge.

Schmidt, B. M. (2012). Words alone: Dismantling topic models in the humanities. *Journal of Digital Humanities*, 2(1).

Schwartz, N. D. (2010). Bank of America Reports $7.3 Billion Loss, Citing Charges. *The New York Times*.

Scollon, S. W. & de Saint-Georges, I. (2012). Mediated discourse analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 66–78). London and New York: Routledge.

Sculley, D. & Pasanek, B. M. (2008). Meaning and mining: The impact of implicit assumptions in data mining for the humanities. *Literary and Linguistic Computing*, 23(4), 409–424. doi: 10.1093/llc/fqn019.

Seidel, G. (1988). Verbal strategies of the collaborators: A discursive analysis of the july 1986 european parliamentary debate on south african sanctions. *Text-Interdisciplinary Journal for the Study of Discourse*, 8(1-2), 111–128. doi:0165–4888/88/0008–0111.

Sheng, A. (2010). Financial Crisis and Global Governance : A Network Analysis. In M. Spence & D. Leipziger (Eds.), *Globalization and Growth Implications for a Post-crisis World* (pp. 69–93). Washington, DC: Commission on Growth and Development.

Shim, J., Park, C., & Wilding, M. (2015). Identifying policy frames through semantic network analysis: an examination of nuclear energy policy across six countries. *Policy Sciences*, 48(1), 51—83. doi:10.1007/s11077–015–9211–3.

Shutova, E. (2010). Models of metaphor in NLP. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 688–697). Uppsala, Sweden: ACL.

Smart, G. (2012). Discourse-oriented ethnography. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 147–159). London and New York: Routledge.

Smith, N. (2009). Toxic Capitalism. *New Political Economy*, 14(3), 407–412. doi:10.1080/13563460903087623.

Smith, T. A. & Tatalovich, R. (2003). *Cultures at War: Moral Conflicts in Western Democracies*. Peterborough, ON: Broadview Press.

Snow, D. A. & Benford, R. D. (1992). Master Frames and Cycles of Protest. In A. D. Morris & C. M. C. Mueller (Eds.), *Frontiers in social movement theory* (pp. 133–155). New Haven, CT: Yale University Press.

Snyder Jr, J. M. & Groseclose, T. (2000). Estimating party influence in congressional roll-call voting. *American Journal of Political Science*, 44(2), 193–211. doi:10.2307/2669305.

Solé, R. V., Corominas-Murtra, B., Valverde, S., & Steels, L. (2010). Language networks: Their structure, function, and evolution. *Complexity*, 15(6), 20–26. doi:10.1002/cplx.20305.

Sopory, P. & Dillard, J. P. (2002). The Persuasive Effects of Metaphor. A Meta-Analysis. *Human Communication Research*, 28(3), 382–419. doi:10.1111/j.1468–2958.2002.tb00813.x.

Sowa, J. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. Boston, MA: Addison-Wesley Longman.

Sowa, J. (1992). Semantic networks. In S. C. Shapiro (Ed.), *Encyclopedia of Artificial Intelligence* (pp. 291–9). New York: Wiley, 2nd edition.

Spooren, W. & Degand, L. (2010). Coding coherence relations: Reliability and validity. *Corpus Linguistics and Linguistic Theory*, 6(2), 241–266. doi:10.1515/CLLT.2010.009.

Star, S. L. & Griesemer, J. R. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3), 387–420. doi:10.1177/030631289019003001.

Steyvers, M. & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1), 41–78. doi:10.1207/s15516709cog29013.

Stone, P. J., Dunphy, D. C., Smith, M. S., & Ogilvie, D. M., Eds. (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press.

Sturm, J. E. & de Haan, J. (2011). Does central bank communication really lead to better forecasts of policy decisions? New evidence based on a Taylor rule model for the ECB. *Review of World Economics*, 147(1), 41–58. doi:10.1007/s10290–010–0076–4.

Tang, J., Meng, Z., Nguyen, X., Mei, Q., & Zhang, M. (2014). Understanding the limiting factors of topic modeling via posterior contraction analysis. In *Proceedings of the International Conference on Machine Learning* (pp. 190–198). Beijing, China: JMLR.

Tesch, R. (1990). *Qualitative Research: Analysis Types and Software*. New York: Routledge.

The Federal Reserve Board (2014). Mission. Retrieved from `http://www.federalreserve.gov/aboutthefed/mission.htm`.

Thelwall, M. & Price, L. (2006). Language evolution and the spread of ideas on the Web: A procedure for identifying emergent hybrid word family members. *Journal of the American Society for Information Science and Technology*, 57(10), 1–20. doi:10.1002/asi.20437.

Thompson, W. (1996). *The Left In History: Revolution and Reform in Twentieth-Century Politics*. London: Pluto Press.

Thornborrow, J. (2012). Narrative analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 51–65). London and New York: Routledge.

Tourish, D. & Hargie, O. (2012). Metaphors of Failure and the Failures of Metaphor: A Critical Study of Root Metaphors used by Bankers in Explaining the Banking Crisis. *Organization Studies*, 33(8), 1045–1069. doi:10.1177/0170840612453528.

Toxic (2012). Oxford Dictionaries. Retrieved from `http://www.oxforddictionaries.com/definition/english/toxic`. Last accessed: 31-10-2014.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381–403). New York: Academic Press.

Ureña Gómez-Moreno, J. M. & Faber, P. (2010). Strategies for the Semi-Automatic Retrieval of Metaphorical Terms. *Metaphor & Symbol*, 26(1), 23–52. doi:10.1080/10926488.2011.535415.

Valente, T. W., Coronges, K., Lakon, C., & Costenbader, E. (2008). How Correlated Are Network Centrality Measure. *Connections*, 28(1), 16–26.

van Aggelen, A. & Hollink, L. (2015). Plenary debates of the European Parliament as Linked Open Data. Retrieved from `http://www.talkofeurope.eu/data/`.

van Dijk, T. A. (1997a). Discourse as interaction in society. In T. A. van Dijk (Ed.), *Discourses as Social Interaction* (pp. 1–37). London: Sage.

van Dijk, T. A. (1997b). The study of discourse. In *Discourse as Structure and Process* (pp. 1–34). London: Sage.

van Dijk, T. A. (1998). What is Political Discourse Analysis? In *Political Linguistics* (pp. 11–52). Amsterdam: Benjamins.

van Dijk, T. A., Ed. (2011). *Discourse Studies: A Multidisciplinary Introduction*. London: Sage, 2nd edition.

van Dijk, T. A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

Wang, J., Zhao, Z., Zhou, J., Wang, H., Cui, B., & Qi, G. (2012). Recommending Flickr groups with social topic model. *Information Retrieval*, 15(3-4), 278–295. doi:10.1007/s10791–012–9193–0.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393, 440–442. doi:10.1038/30918.

Weber, R. P. (1985). *Basic Content Analysis*. Beverly Hills, CA: Sage.

Wei, T.-H. (1952). *The Algebraic Foundations of Ranking Theory*. PhD thesis, University of Cambridge.

Wetherell, M., Taylor, S., & Yates, S. J. (2001). *Discourse as Data: A Guide for Analysis*. Thousand Oaks, CA: Sage.

White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social structure from multiple networks. I. Blockmodels of roles and positions. *American Journal of Sociology*, 81(4), 730–780.

Wilks, Y. (1978). Making preferences more active. *Artificial Intelligence*, 11, 197–223. doi:10.1007/1–4020–5285–57.

Williams, A. E. (2013). Metaphor, media, and the market. *International Journal of Communication*, 7, 1404–1417. doi:1932–8036/20130005.

Williams, A. E., Davidson, R., & Yochim, E. C. (2011). Who's to blame when a business fails? How journalistic death metaphors influence responsibility

attributions. *Journalism & Mass Communication Quarterly*, 88(3), 541–561. doi:10.1177/107769901108800305.

Wilson, J. (2015). Political Discourse. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The Handbook of Discourse Analysis* (pp. 775–794). Hoboken, NJ, USA: JohnWiley & Sons, Inc., 2nd edition.

Wodak, R. (1989). *Language, Power and Ideology: Studies in Political Discourse*. Amsterdam: John Benjamins Publishing.

Wyatt, S. (2004). Danger! Metaphors at Work in Economics, Geophysiology, and the Internet. *Science, Technology, & Human Values*, 29(2), 242–261. doi:10.1177/0162243903261947.

Xiang, Z., Gretzel, U., & Fesenmaier, D. R. (2009). Semantic representation of tourism on the internet. *Journal of Travel Research*, 47(4), 440–453. doi: 10.1177/0047287508326650.

Zeiss, R. & Groenewegen, P. (2009). Engaging Boundary Objects in OMS and STS? Exploring the Subtleties of Layered Engagement. *Organization*, 16(1), 81–100. doi:10.1177/1350508408098923.

Zicari, R. V. (2012). *On Analyzing Unstructured Data. – Interview with Michael Brands.* Retrieved from `http://www.odbms.org/blog/2012/07/on-analyzing-unstructured-data-interview-with-michael-brands/`. Last accessed: 24-09-2015.

Zupnik, Y.-J. (1994). A pragmatic analysis of the use of person deixis in political discourse. *Journal of Pragmatics*, 21(4), 339–383. doi:10.1016/0378–2166(94)90010–8.

Appendices

A Contributions of Co-Authors for Empirical Chapters

The three empirical studies presented in Chapters 3 and 4 are based on published articles, while the study presented in Chapter 5 is a working paper in preparation for journal submission. Thus, in this Appendix, I acknowledge the valuable contributions of my co-authors to the development of these studies.

The study presented in Chapter 3 was published in the Computational Social Networks journal with the following co-authors: Adina Nerghes, Ju-Sung Lee, Peter Groenewegen, and Iina Hellsten. Importantly, this particular article is an extended version of the article "The shifting discourse of the European Central Bank: Exploring structural space in semantic networks", published in the Proceedings of the Tenth International Conference on Signal Image Technology & Internet Based Systems in 2014 (Nerghes et al., 2014b). However, the contributions of all co-authors has been similar for both the short and the extended version of this study as follows:

- Literature review: Adina Nerghes
- Conceptual development: Adina Nerghes, Ju-Sung Lee, Peter Groenewegen, and Iina Hellsten
- Study design: Adina Nerghes
- Data collection and analysis: Adina Nerghes and Ju-Sung Lee
- Article writing: Adina Nerghes
- Edits and additions: Adina Nerghes, Ju-Sung Lee, Peter Groenewegen, and Iina Hellsten
- Review post journal submission: Adina Nerghes, Ju-Sung Lee

Published in the International Journal of Communication, the article titled "A toxic crisis: Metaphorizing the financial crisis" (Nerghes et al., 2015a) is a

result of a collaboration between myself, Iina Hellsten and Peter Groenewegen. The individual contributions are as follows:

- Literature review: Adina Nerghes
- Conceptual development: Adina Nerghes with guidance from both Iina Hellsten and Peter Groenewegen
- Study design: Adina Nerghes
- Data collection and analysis: Adina Nerghes
- Article writing: Adina Nerghes
- Edits and additions: Adina Nerghes, Iina Hellsten and Peter Groenewegen
- Review post journal submission: Adina Nerghes, Iina Hellsten and Peter Groenewegen

Lastly, the study presented in Chapter 5 is a working paper, currently being prepared for journal submission. The co-authors involved in the development of this study are Adina Nerghes, Peter Groenewegen, Iina Hellsten and Yvette Taminiau, and the individual contributions are as follows:

- Literature review: Adina Nerghes
- Conceptual development: Adina Nerghes with guidance from Iina Hellsten, Peter Groenewegen, and Yvette Taminiau
- Study design: Adina Nerghes
- Data collection and analysis: Adina Nerghes
- Article writing: Adina Nerghes
- Edits and additions: Adina Nerghes
- Review post journal submission: N/A

While this brief summary is necessary when writing a PhD dissertation, the contributions of my collaborators go far beyond these articles by contributing to my development as a scholar and researcher. Brainstorming sessions, comments on various drafts, and prompt emails late at night, although not mentioned in the lists above, have guided me at every stage of this research project.

B Examples Queries for Data Collection

Example 1: All speeches by EPP members between 14th of June 2009 and the 30th of June 2014 that contain the strings 'financial crisis' AND/OR 'economic crisis'

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX lp: <http://purl.org/linkedpolitics/>
PREFIX lpv: <http://purl.org/linkedpolitics/vocabulary/>
PREFIX xml: <http://www.w3.org/XML/1998/namespace>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SELECT DISTINCT ?date ?name ?text
WHERE {
   ?sessionday dcterms:hasPart ?agendaitem.
   ?sessionday dc:date ?date.
   ?agendaitem dcterms:hasPart ?speech.
   ?agendaitem lpv:number ?agendaitemnr.
   ?speech lpv:number ?speechnr.
   ?speech lpv:text ?text.
   ?speech lpv:speaker ?speaker.
   ?speaker foaf:familyName ?name.
   ?speaker lpv:politicalFunction ?function.
   ?function lpv:institution ?party.
   ?party lpv:acronym ?partyname.
```

```
   FILTER (regex(str(?text), "financial crisis")
   || regex(str(?text), "economic crisis"))
   FILTER regex(str(?partyname), "PPE").
   FILTER ( ?date >= "2009-06-14"^^xsd:date
   && ?date <= "2014-06-30"^^xsd:date )
   FILTER(langMatches(lang(?text), "en"))
} ORDER BY ?date ?agendaitemnr ?speechnr
```

Example 2: All speeches by S&D members between 14th of June 2009 and the 30th of June 2014 that contain the strings 'financial crisis' AND/OR 'economic crisis'

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX lp: <http://purl.org/linkedpolitics/>
PREFIX lpv: <http://purl.org/linkedpolitics/vocabulary/>
PREFIX xml: <http://www.w3.org/XML/1998/namespace>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SELECT DISTINCT ?date ?name ?text
WHERE {
   ?sessionday dcterms:hasPart ?agendaitem.
   ?sessionday dc:date ?date.
   ?agendaitem dcterms:hasPart ?speech.
   ?agendaitem lpv:number ?agendaitemnr.
   ?speech lpv:number ?speechnr.
   ?speech lpv:text ?text.
   ?speech lpv:speaker ?speaker.
   ?speaker foaf:familyName ?name.
   ?speaker lpv:politicalFunction ?function.
```

```
   ?function lpv:institution ?party.
   ?party lpv:acronym ?partyname.
   FILTER (regex(str(?text), "financial crisis")
   || regex(str(?text), "economic crisis"))
   FILTER regex(str(?partyname), "S&D").
   FILTER ( ?date >= "2009-06-14"^^xsd:date
   && ?date <= "2014-06-30"^^xsd:date )
   FILTER(langMatches(lang(?text), "en"))
} ORDER BY ?date ?agendaitemnr ?speechnr
```

Words are the windows to society's soul.

Most of our knowledge of current and past world events, advances in science, or even elements of culture are gained, formed, and passed on through written and verbal text. Hence, the question is no longer whether or not text content is worth investigating, but rather which approach is the most insightful for a given research goal.

Expanding methods of relational meaning analysis, this dissertation provides an approach suited for the detection of subtle discursive dynamic shifts in large collections of (temporal) textual data. Alongside the methodological contributions, this dissertation is discourse-centered and focuses on the global financial crisis revealing patterns of discourse, subtle dynamic shifts of discursive practices at different stages of the crisis, and dominant topics at the level of media, financial, and political discourses.

The important empirical findings presented in this dissertation, challenge, extend, and refine what we know about financial, media, and political discourses in the context of the financial crisis and how different aspects of the financial crisis are manifested in discourses of different social actors.

There is no self-understanding that is not mediated by signs, symbols, and texts. – P. Ricoeur

ISBN 978-94-6203-660-4

VU VRIJE UNIVERSITEIT AMSTERDAM | Faculty of Social Sciences