Ca' Foscari
University
of Venice

Master's Degree Programme
in Language Sciences

Final Thesis

A computational study of metaphor in American presidential speeches

Supervisor
Prof. Gianluca Lebani

Assistant supervisor
Prof.ssa Francesca Santulli

Graduand
Emma Rocco
Matriculation Number 866689

Academic Year
2021/2022

Abstract

Metaphor is a ubiquitous figure of speech generally employed to describe one concept in terms of another. Within the field of study of political discourse, metaphor is additionally described as a persuasive device that is used to achieve rhetoric goals and to organize discourse structure. While the use of metaphor in political discourse has been extensively analyzed in small-scale corpus linguistics studies, there are few large-scale studies that explore metaphor patterns in political discourse due to the complexity of manual metaphor extraction. Automatic metaphor identification is however a growing topic of interest within the field of natural language processing (NLP). Adopting a computational approach to metaphor identification might provide a broader insight into the use of figurative language in political discourse. This thesis adopts a metaphor detection model that was designed by Su et al. (2020) to automatically extract metaphors in a corpus of 1721 American presidential speeches. A quantitative and qualitative analysis is performed on the metaphors extracted by the model. The results of the explorative analysis are compared to the findings of studies that do not rely on NLP for metaphor detection in order to evaluate the efficacy of the use of computational methods for corpus-based studies of this scale.

Table of contents

List of tables

List of figures

Introduction

Metaphor is a ubiquitous figure of speech that can be defined as the result of mapping a concept or domain to another (Lakoff & Johnson, 1980). Metaphorical mappings are typically employed to convey an abstract concept in terms of another that is more concrete. For instance, in sentences such as “She *attacked* my argument” and “My argument had no *strategy*,” the abstract domain of argument is conveyed in terms of the more concrete domain of war.

Within the field of study of political discourse, metaphor is additionally described as a powerful persuasive device that is used to achieve rhetoric goals (Charteris-Black, 2011; Musolff, 2004) and to organize discourse structure (Musolff, 2004; Semino, 2008). While the use of metaphor in political discourse has been extensively analyzed in small-scale corpus linguistics studies (Charteris-Black, 2004; Howe,1988), there are few large-scale studies that explore metaphor patterns in political discourse due to the complexity of manual metaphor extraction. Automatic metaphor identification is however a growing topic of interest within the field of natural language processing (NLP) (Shutova, 2010) and great advances in performance have been achieved in recent years (Dankers et al., 2020; Devlin et al., 2018; Su et al., 2020; Wu et al., 2018).

Adopting a computational approach to metaphor identification, this thesis sets out to provide a broader insight into the use of figurative language in political discourse. Specifically, this thesis will adopt a metaphor detection model, the DeepMet model (Su et al., 2020), to automatically extract metaphors in an ad-hoc corpus of 1721 American presidential speeches, which will be referred to as the APS corpus. A quantitative and qualitative analysis will then be performed on the metaphors extracted by the model. The results of the explorative analysis will be compared to the findings of studies that do not rely on NLP for metaphor detection (Charteris-Black, 2004) to evaluate the efficacy of the use of computational methods for corpus-based studies of this scale.

Thus, the first aim of the study is to overcome the challenge of metaphor extraction in large corpora such as the one constructed for this thesis by employing a computational model for automatic metaphor identification, namely the DeepMet model (Su et al., 2020). The second aim of this thesis is to quantitatively and qualitatively analyze the results of the metaphor identification process to gain broader insights into the use of figurative language in American presidential political discourse, and to further compare these findings to the results of studies that do not rely on computational models for metaphor detection, specifically, Charteris-Black's (2004) research on US presidential inaugural speeches.

The thesis is organized as follows: Chapter 1 examines the relevant literature on metaphor within the fields of corpus linguistics, computational linguistics and political discourse analysis; Chapter 2 covers the material that will be analyzed, the APS corpus, the model selected for metaphor identification, the DeepMet model (Su et al., 2020), and the methods adopted to analyze the metaphors that are identified in the corpus; Chapter 3 reports the quantitative and qualitative analysis of the metaphors identified in the APS corpus and presents the general discussion of the results of the analysis; lastly, Chapter 4 summarizes the main findings and contributions of the study, discusses the limitations and challenges of the analysis, and suggests directions for future research.

1. Literature review

1.1 Metaphor background

The term *metaphor* (from Ancient Greek *μεταφορά* (metaphorá) for "transfer", "carrying over") typically refers to a rhetorical device whose purpose is to describe a transfer of meaning between two terms that possess similar attributes. This transfer is used to convey a concept in terms of the properties of another concept. Below are some examples of metaphor.

(1) How can I *kill* a process? (Martin, 1988, p. 396)
(2) I *demolished* his argument. (Lakoff & Johnson, 1980, p. 4)
(3) All the world's a *stage*, and all the men and women merely *players*[1].
(4) And then my heart with pleasure *fills*,
And *dances* with the daffodils[2].

Metaphors are ubiquitous in language, we produce and comprehend them in various forms both in everyday speech, in the form of conventional metaphors such as those in (1) and (2), and in poetic language, with novel metaphors such as those in example (3) and (4).

In metaphorical expressions, specific features of one concept are transferred to another concept. For instance, in example (1) a computational process is conveyed as something that can be killed and therefore as something that is alive, here the properties of a living thing are associated with an artificial computer process to convey its forced termination as an act of killing.

Metaphors are often employed to describe abstract concepts in terms of more concrete or physical experiences. In (2) for instance the act of proving someone's argument wrong is associated with the physical verb *demolish*, the abstract concept of an argument is thus conveyed as something that can be physically destroyed.

[1] taken from the play "As You Like It" by William Shakespeare (1623)

[2] taken from the verse "I wandered lonely as a cloud" by William Wordsworth (1804)

1.1.1 An overview of the major views on metaphor

Various views on the nature and role of metaphor have been developed throughout history in different areas of study such as rhetoric, philosophy, and linguistics. Discussion on the subject of metaphor dates back to the classical Greek and Roman traditions, where metaphors were regarded as rhetorical devices with a dual purpose. Metaphorical expressions could either be used to embellish prose and speeches, or they were employed to help construct more persuasive and incisive arguments, conveying a concrete objective correlative in order to give substance to more abstract concepts (Dalla Libera, 2017).

Aristotle was the first to assign more importance to metaphor's rhetorical and poetic values than to the aforementioned aesthetic ones (Kirby, 1997). The Aristotelian definition refers to metaphor as a language tool where one word is used instead of another to mean the same. This definition was favored in classical rhetorics, overshadowing explorations into possible cognitive aspects of metaphor (Guastini, 2004).

In the Middle Ages the notion of metaphor as a means for stylistic embellishment and argument accessory persisted. During the Italian Enlightenment however philosopher and rhetorician Vico presented a theory that described metaphor as a cognitive basis for primitive people - metaphors being fundamental for language development based on the assumption that figurative speech is antecedent to analytic and rational speech (Danesi, 2011). This innovative approach that viewed metaphor as a mental faculty was not widely recognized and thus the idea of metaphors as exclusively stylistic and rhetorical accessories was not challenged until decades later.

In the second half of the 1900s scientific enquiry on the subject of metaphor gained popularity in the fields of linguistics and philosophy. During this time, different but often overlapping approaches to metaphor were established as what can be considered the prominent approaches to metaphor to this day. A selection of the major views on metaphor developed in the second half of the 1900s include: the *substitution view* (Winner, 1988); the *comparison view* (Gentner, 1983); the *anomaly view* (Wilks, 1975, 1978); the *class inclusion view* (Davidson, 1978; Glucksberg et al., 1997); the *interaction view* (Black, 1962); and *conceptual metaphor theory* (Lakoff & Johnson, 1980; Lakoff, 1993).

Winner's (1988) *substitution view* presents metaphor as an imprecise lexical substitute that is used in the absence of a clearer literal expression and as a verbal embellishment. In the context of this approach, metaphorical expressions always need to be first converted to literal ones in order to be interpreted correctly. Winner further states that metaphors' main purpose has an aesthetic basis to amuse the reader and that metaphors' imprecision and lack of literality makes them a less desirable form of communication than a literal one. This means that metaphorical expressions such as (5) below, are considered rhetorical embellishments with no truth value unless they are paraphrased into literal expressions that make sense. However, finding a precise literal substitution is not always possible, in (5), for example, *shark* could be substituted with different properties associated with the animal such as *aggressive* or *intelligent*, and choosing one could lead to an imprecise interpretation and unclear communication.

(5) My lawyer is an *old shark*. (Rai & Chakraverty, 2020, p. 4)

Critics of the substitution view have pointed out that metaphors are too complex and nuanced for them to be rephrased literally only based on perceived similarity (Rai & Chakraverty, 2020). Furthermore, claims of metaphor as a form of unclear communication have been contested by a number of studies that argue for metaphor's ability to in fact facilitate communication, rather than hinder it (Nguyen et al., 2015; Thibodeau and Boroditsky, 2011).

Gentner's (1983) *comparison view* of metaphors reiterates the aforementioned principle of analogy proposed by Aristotle and introduces the concept of structure mapping to convey analogy relations. According to Gentner, metaphor is used to compare two or more semantic domains on the basis of attributes that both domains share and based on the similarities shared by the attributes themselves. In the comparison view, similarly to the substitution view, metaphorical expressions are considered as mostly decorative and can always be replaced by a literal counterpart that resembles a simile. For example, for a metaphor such as the one in (5), the literal simile counterpart would be "My lawyer is like an old shark" and the interpretation of this comparison is based on the identification of the perceptual properties that are applicable to both *old shark* and *lawyer* such as *aggressive*, *intelligent* and *cunning*.

Critics of the comparison view point out that metaphors and similes are not always interpreted in the same way, metaphors usually conveying a more energetic picture than similes (Nguyen et al. 2015; Thibodeau and Boroditsky 2011) and also note that similarities between domains are not consistently easy to identify.

The *anomaly view* of metaphors (Wilks, 1975) highlights the dissimilarities between the connected semantic domains rather than their similarities. As such, metaphorical expressions are defined as semantic incongruities that generate from a violation of selectional preference (the preference patterns for linguistic classes towards certain semantic classes). For example, sentence (6) is semantically incongruous: the verb *drink* usually selects an animate agent and an edible liquid theme, since *car* is inanimate, the use of the verb *drink* is in this case considered metaphorical.

(6) My car *drinks* gasoline. (Wilks, 1978, p. 199)

Critics of the anomaly view have however pointed out cases in which an apparent violation of selectional preference could be mistaken for an indicator of metaphor, while the correct interpretation is actually literal and heavily relies on contextual information. An example of this is evident in the two utterances shown in (7) below: while the verb *brand* that selects the theme *heart* is metaphorical in (7a), the contextual information underlined in (7b) leads to a literal interpretation of the same verb selecting the same theme.

(7) a. She *branded* his heart *with her hateful jibes*. (Rai & Chakraverty, 2020, p.6)
b. She branded his heart *with a satanic seal*.

Glucksberg et al. (1997), following Davidson's (1978) *class inclusion view*, propose that metaphors are categorical assertions rather than comparison based structure mappings. According to this view of metaphor, in metaphorical statements such as "x is y", the former element is usually established as part of a superordinate category represented by the latter. For example, in an utterance such as "my job is a jail" (Rai & Chakraverty, 2020, p. 24), the metaphorical noun *jail* represents a superordinate class of unpleasant, confining places or situations, and the noun *job* is asserted as being included in that specific class.

Black's (1962) *interaction view* proposes that metaphor meaning is derived from a bidirectional interaction established between the two constituent parts of a metaphor rather than from a literal substitution or a comparison of shared properties. According to the interaction view, a metaphor is composed of a *metaphorical focus*, shown in italics in example (8), and a *literal frame*, underlined in (8).

(8) The chairman *plowed* through the discussion. (Black, 1962, p. 26)

In order to process the metaphorical meaning of the verb *plow* in (8) one first has to know its literal meaning (*plow* as in *plowing soil*). Black (1962) proposes that metaphor acts as a filter that controls the interaction between the literal and metaphorical meanings of the focus, and it establishes the culturally common attributes between the two mapped concepts also in view of the literal frame. In (8), for example, the act of *participating in discussion* is filtered through the screen of *plowing the earth* and at the same time the concept of *plowing* is restructured in terms of the act of discussion, resulting in a bidirectional interaction between the mapped concepts.

Because of the centrality given to the idea of metaphor interconceptual mapping, the interactionist theory is considered to be the forerunner of the most prominent approach to metaphor: the *cognitive* or *conceptual mapping view* of metaphor introduced by Lakoff and Johnson (1980).

The remainder of this section will discuss Lakoff and Johnson's (1980) *conceptual metaphor theory* in more detail, since this approach was employed as a theoretical basis for the vast majority of corpus linguistics research carried out on the subject of metaphor, for various computational models designed for metaphor identification, as well as in the analysis presented in this thesis.

1.1.2 Conceptual metaphor theory

Conceptual metaphor theory (CMT) is the most prominent view of metaphor to this day. It was proposed by Lakoff and Johnson (1980) as a way to define metaphor not only from a linguistic point of view but also from a cognitive standpoint. CMT rejects any similarity or dissimilarity based views of metaphor by defining a metaphorical expression as a systematic conceptual mapping between a *target domain* and a *source domain*, or alternatively, as a cognitive reconceptualization of a domain.

Moreover, differently from Black's (1962) interaction view, the mapping in CMT is not based on analogy relations, rather, it is a mapping between two concepts, one of which is from a more abstract domain (the target) and the other from a more concrete domain (the source), so that the first less familiar concept can be understood in terms of the second well understood concept.

A well known example of conceptual metaphor in the form of "target domain is source domain" is ARGUMENT IS WAR, where the concept of *argument* is reconceptualized to that of *war*. This kind of conceptual metaphor is realized in everyday language in the form of different linguistic expressions such as the ones presented in (9):

(9) a. Your claims are *indefensible*. (Lakoff & Johnson, 1980, p. 4)
b. He *attacked every weak* point in my argument.
c. His criticisms were *right on target*.
d. I *demolished* his argument.
e. I've never *won* an argument with him.
f. You disagree? Okay, *shoot*!
g. If you use that *strategy*, he'll *wipe you out*.
h. He *shot down* all of *my* arguments.

It is evident from the linguistic expressions in (9) how the target domain ARGUMENT is systematically mapped to the source domain WAR: what is usually experienced in an argument is reconceptualized with reference to war actions such as *shooting* and *attacking* and war entities such as *opponents* and *targets*.

Lakoff and Johnson (1980) and other cognitive theorists such as Kövecses (2005), following Black (1962), consider metaphor as a cognitive process rather than just a rhetorical figure. They claim that the metaphorical thought process can facilitate our understanding of abstract concepts since these are usually reconceptualized with the use of concepts that are more familiar and related to our physical senses. Below are a few other examples of common metaphorical mappings that associate abstract concepts with more concrete concepts.

- TIME IS MONEY, e.g. "You're *wasting* my time." (Lakoff & Johnson, 1980, p. 7)
- THE MIND IS A MACHINE, e.g. "We've been working on this problem all day and now we're *running out of steam*." (Lakoff & Johnson, 1980, p. 27)
- LOVE IS A JOURNEY, e.g. "I don't think this relationship is *going anywhere*." (Lakoff & Johnson, 1980, p. 44)
- THEORIES ARE BUILDINGS, e.g. "The theory needs more *support*." (Lakoff & Johnson, 1980, p. 46)
- UNDERSTANDING IS SEEING, e.g. "It *looks* different from my *point of view*." (Lakoff & Johnson, 1980, p. 48)
- LIFE IS A CONTAINER, e.g. "I've had a *full* life." (Lakoff & Johnson, 1980, p. 51)

According to Lakoff and Johnson's (1980) conceptual metaphor theory there are three main types of metaphor: *structural*, *orientational*, and *ontological metaphors*.

Structural metaphors are the ones in which "one concept is metaphorically structured in terms of another" (Lakoff & Johnson, 1980, p. 14). An example of structural metaphor is the already mentioned ARGUMENT IS WAR.

Orientational metaphors are, on the other hand, those metaphorical expressions that originate from our awareness of spatial relationships such as "up-down, in-out, front-back, on-off, deep-shallow, central-peripheral" (Lakoff & Johnson, 1980, p. 14). For instance, the HAPPY IS UP and SAD IS DOWN as in the utterances "My spirits *rose*" and "I'm feeling *down*" (Lakoff & Johnson, 1980, p. 15), are common examples of orientational metaphors.

Finally, ontological metaphors refer to either *entity and substance metaphors* or *container metaphors*. Entity and substance metaphors are "ways of viewing events, activities, emotions, ideas, etc., as entities and substances" (Lakoff & Johnson, 1980, p. 25). For instance, THE MIND IS A MACHINE, as in "My mind just isn't operating today" (Lakoff & Johnson, 1980, p. 27). Container metaphors are related to our experiences of territory and visual perception as shown in examples such as "There's a lot of land in Kansas" and "That's in the center of my field of vision" (Lakoff & Johnson, 1980, p. 30).

Lakoff and Johnson's view of metaphor as a mapping has been adopted in numerous studies in the fields of linguistics, philosophy, cognitive science and computational linguistics. However, not every aspect of CMT has been accepted by all researchers. Some critics of the conceptual metaphor theory, for example, challenged the existence of underlying cognitive processes resembling conceptual mappings that, according to cognitive theorists, would be needed to understand conventional metaphors (Keysar et al., 2000). Murphy (1996) has also pointed out the difficulties in reconstructing a target concept when a linguistic expression presents a multiple number of conceptual mappings and further criticized CMT for its lack of consistency and empirical evidence. While the conceptual mappings presented by Lakoff and Johnson (1980) are representative of the linguistic metaphors brought as examples for the theory, they cannot explain the vast majority of linguistic metaphors found in text.

1.2 Types of metaphor

Metaphors can be classified differently depending on what is more relevant in one's research. A possible metaphor categorization is for example Lakoff and Johnson's (1980) classification of metaphors according to their structural mapping relations illustrated in subsection 1.1.2. There are multiple other generally accepted metaphor categorizations either based on the way metaphors are interpreted semantically or on how they manifest linguistically.

This section will outline two metaphor classifications that have been relevant for both corpus linguistics and computational linguistics research on metaphor. The first classification is based on Bowdle and Gentner's (1995, 2005) *career of metaphor* theory, which classifies metaphors on the basis of their semantic relations and degree of metaphoricity. The second classification consists in the categorization of linguistic metaphors according to linguistic type as presented by Rai and Chakraverty (2020).

1.2.1 Metaphor career

According to Bowdle and Gentner (1995, 2005) metaphors can be categorized into three types: *novel metaphors*, also known as *creative*, *conventional metaphors*, and *dead metaphors*. These categories refer to a matephor's level of conventionality or "metaphoricity", which can change over time. Metaphorical expressions always start out as novel and, as a consequence of their repeated use, usually transition to conventional metaphors, and can subsequently become dead metaphors. Bowdle and Gentner (1995, 2005) define this progression of metaphoricity as a *career of metaphor*.

Novel metaphors are described as creative associations of concepts, the type of mappings which are not yet part of standard language lexicon. This category of metaphors does not have a derived metaphorical sense yet, and, because of this, understanding novel metaphors is not automatic, it requires prior knowledge and "a special imaginative leap" (Numberg, 1987). An example of novel metaphor proposed by Bowdle and Gentner (2005) is *glacier* in example (10) below.

(10) Science is a *glacier*. (Bowdle & Gentner, 2005, p. 199)

Here it is evident how the term *glacier* has a literal sense but does not have a generally recognized metaphorical sense yet. In order to interpret (10) in its metaphorical sense, we need to structurally align the familiar literal concept (in this case a glacier is “a large body of ice spreading over a land surface”) with the target concept (“anything that progresses slowly but steadily”) by way of comparison between the two concepts.

Conventional metaphors, on the other hand, are defined by the fact that they have adapted to a particular metaphorical sense over time, to the point where the metaphorical sense in question becomes one of the term’s possible meanings, effectively making the term polysemous. Bowdle and Gentner (1995, 2005) propose that conventional metaphors are interpreted through a process of categorization rather than comparison, the former process being less cognitively demanding than the latter. Bowdle and Gentner use as an example of conventional metaphor the term *blueprint* in the utterance found in (11).

(11) A gene is a *blueprint*. (Bowdle & Gentner, 2005, p. 199)

Here the base term *blueprint* has two closely related senses, one is literal (“a blue and white photographic print showing an architect’s plan”) and the other refers to an associated metaphorical category (“anything that provides a plan”).

In the possible final step of a metaphor’s career, conventional metaphors can become dead metaphors. A dead metaphor behaves like a homonym, meaning that the metaphorical and literal senses evoked by the same term are no longer semantically related in any way, essentially making the once metaphorical sense of one term a new literal term altogether. An example of a dead metaphor is *culture* in example (12).

(12) A university is a *culture* of knowledge” (Bowdle & Gentner, 2005, p. 209).

While the word *culture* was once used metaphorically in relation to its literal sense of “growth of bacteria or cells for scientific purposes", it is now interpreted on the basis of its other unrelated sense of “particular heritage or society”, born from the extension of the primary literal meaning of the word. The interpretation of dead metaphors does not require imaginative leaps or concept associations since they can be considered as independent literal terms.

1.2.2 Types of linguistic metaphors

In the context of Lakoff and Johnson's (1980, 1993) conceptual metaphor theory, the term *linguistic metaphor* refers to the representations of conceptual metaphors in everyday language. According to this definition, a single conceptual metaphor may have multiple linguistic expressions.

Following Rai and Chakraverty (2020) classification of linguistic metaphors, the two main kinds of metaphorical expressions in language are *contracted* and *extended* metaphors. Contracted metaphors are those metaphors whose effect only extends to the sentence or phrase they are part of, whereas extended metaphors may affect multiple parts of their contextual discourse and are usually found in literary texts.

Contracted metaphors composed of one word are defined as *lexical metaphors*, while *multi-word* metaphorical expressions are composed of multiple words. Lexical metaphors are further subdivided into four types according to the linguistic form of the metaphor, a useful distinction in linguistically driven research. The four types include the following:

- *Type-I* or *nominal metaphors* are usually composed of a subject (the target domain) and an object (the source domain) joined together by a copular verb. An example of type-I metaphor is "time is *money*", here the conceptual mapping between the target domain and source domain is explicit - the abstract concept of *time* is associated with the more concrete concept of *money* in order to metaphorically emphasize certain shared attributes such as *value* and *importance*.

- *Type-II* or *Subject-Verb-Object (SVO) metaphors* refer to sentences containing metaphorical verbs. Differently from nominal metaphors, in SVO metaphors the conceptual mapping is implicit, this is illustrated in example (1) where the verb *drink* is used metaphorically to convey an excessive consumption actualizing the conceptual mapping CONSUMPTION IS DRINKING.

- *Type-III* or *Adjective-Noun (AN) metaphors* consist of a metaphorical adjective accompanied by a noun. The phrase "*sweet* child" (Rai & Chakraverty, 2020, p. 12) is a good example of an AN metaphor, here the metaphorical adjective *sweet* describes the child's nature, the conceptual mapping in this case being NATURE IS TASTE.

Type-IV or *Adverb-Verb (AV) metaphors* refer to metaphorical adverb-verb combinations such as "Ram speaks fluidly" (Rai & Chakraverty, 2020, p. 12): here the adverb *fluidly*, from the source domain LIQUID is used metaphorically to convey an effective way of speech, from the target domain COMMUNICATION.

1.3 Corpus-based approaches to metaphor

Corpus-based studies on the subject of metaphor generally focus on metaphor identification and annotation in corpora and on metaphor analysis based on quantitative corpus data. This section will cover a brief introduction to corpus linguistics methodology and give an overview of the insights that were gathered from corpus-driven research on metaphor.

1.3.1 Corpus linguistics

Corpus linguistics is challenging to define because, while it is a field of research that goes back to at least a hundred years, only in recent years has it become more relevant as a linguistic methodology. A precise definition of corpus linguistics has also proved elusive because linguistics itself is a generally heterogenous discipline, as such, any chosen methodological framework will vary accordingly.

In broad terms, corpus linguistics can be defined "as any form of linguistic inquiry based on data derived from [...] a corpus" (Stefanowitsch, 2020, p. 1), a corpus being a collection composed of "samples of language produced in genuine communicative situations" (Stefanowitsch, 2020, p. 1). Stefanowitsch further advanced a more elaborate definition that characterizes corpus linguistics as an instance of the scientific methodology and is illustrated as follows:

> Corpus linguistics is the investigation of linguistic research questions that have been framed in terms of the *conditional distribution of linguistic phenomena* in a *linguistic corpus*. (Stefanowitsch, 2020, p. 56)

The definition above highlights the importance of investigating the object of linguistic research under different conditions, as it is standard in any scientific method. For a more detailed overview of corpus linguistics as an instance of the scientific method and a guideline for conducting corpus linguistic research, refer to Stefanowitsch (2006, 2020).

Corpus linguistics studies have been conducted within different areas of linguistics research including that of metaphor.

1.3.2 Corpus linguistics research on metaphor

Generally, corpus oriented research on metaphor is carried out following two core steps: the first and usually more problematic is that of metaphor extraction, this is followed by a second phase of analysis of the data gathered depending on the purpose of the study. Most corpus-based approaches to metaphor use Lakoff and Johnson's (1980, 1993) conceptual mappings as a theoretical basis and, accordingly, metaphor extraction and analysis methods focus on conceptual mappings, source domains and target domains.

1.3.2.1 Metaphor extraction and annotation in corpora

There are different methods to extract metaphors within corpus linguistics. Most approaches are either based on manually annotating corpora or rely on particular lexical searches using corpus software, since it is not possible to use part-of-speech-tagging or grammatically annotated corpora to identify metaphors in view of the fact that metaphors, as seen in section 1.2.2, do not have unique linguistic forms compared to literal expressions, making metaphors and non-metaphorical expressions often linguistically indistinguishable.

Manual searching for metaphor identification consists in following specific annotation guidelines to manually annotate corpora for metaphors. An example of metaphor annotation guideline is the Metaphor Identification Procedure (MIP) designed by the Pragglejaz Group (2007). This procedure consists in annotating corpora for metaphors at the word level: each lexical unit in the corpus text analyzed is tagged by the annotators as either metaphorical or literal following the guideline provided. MIP provided the basis for the creation of the VU Amsterdam Metaphor Corpus (VUA) (Steen, et al., 2010), which is a sample of 117 documents from the British National Corpus Baby annotated for metaphors.

Other metaphor extraction methods that are based on specifically annotated corpora rely on either a corpus annotated for semantic fields (Semino, 2005) or a corpus with annotation based on conceptual mappings. Both types of annotated corpora are not exhaustive and most are limited in size. While manual metaphor identification can be highly accurate when following sophisticated annotation procedures such as the MIP, it is not feasible for corpora with large sizes, as it is a time consuming and labor intensive task. As of now, the largest corpus annotated for metaphors is the above mentioned VU Amsterdam Metaphor Corpus (VUA) (Steen, et al., 2010).

Lexically based searches for metaphor on the other hand either focus on keywords from the source domain vocabulary (Deignan, 2005) or from the target domain vocabulary (Koivisto-Alanko 2000; Tissari 2003). A number of studies (Martin, 2007; Stefanowitsch, 2006) have also searched for terms from both the source and the target domain. The vocabularies are either composed of individual words or of specific sets of terms from the relevant domains, and they can be chosen a priori or after a keyword analysis of the texts. It should be noted that since the vocabularies used cannot be exhaustive, additional manual annotation is required for a comprehensive search. Finally, metaphors can be extracted according to a method introduced by Goatly (1997) based on the "markers of metaphor", those metalinguistic expressions such as *metaphorically/figuratively speaking* or *in more than one sense* that can precede metaphorical expressions to introduce their non-literalness. Other metaphor markers are what are defined as "mimetic terms" (terms such as *image* or *likeness*), intensifiers such as *literally* and *actually*, and orthographic devices such as quotation marks.

1.3.2.2 Metaphor analysis in corpus linguistics

As regards to the corpus-based analyses carried out after metaphor extraction, several have proved to be more exhaustive and systematic than other non-corpus-based approaches, while also reviewing parts of the conceptual metaphor theory (Lakoff & Johnson, 1980, 1993) by focusing more on metaphors on a quantitative linguistic standpoint rather than a cognitive one (Stefanowitsch, 2006). Furthermore, Stefanowitsch reports that the use of corpus linguistics as a methodology has led to new insights into the underlying patterns of metaphor as a linguistic phenomenon.

In particular, approaching conceptual metaphors from a quantitative point of view lead Semino (2005) to reassess how conceptual mappings are established, and to emphasize the importance of specific mappings such as the already mentioned ARGUMENT IS WAR mapping, which proportionally appeared more than any other communication metaphor in the corpus examined.

Frequency data gathered in corpus-based studies further allows to explore which target domains are more strongly related to a specific source domain, as well as to discover broader sets of target domains in general when focusing on a source domain in particular. Correspondingly, studies that focus on the target domains isolate the source domains more strongly associated with the target domain in question.

Moreover, while metaphors do not manifest in unique linguistic forms, corpus-based research has shown that some metaphors can be associated with specific syntagmatic patterns. Hanks (2004), for instance, observed a common pattern that involves the partitive or quantifying *of* construction, as in “a storm of protest, a torrent of abuse, a mountain of paperwork [...]” (Hanks, 2004, p. 18).

Employing corpus-driven methods further allows to explore the textual and contextual properties of metaphor. For example, Martin (2007) observed that by analyzing the various contexts of a given type of metaphor it was possible to predict the likelihood of occurrence of the same type of metaphor in the immediately following discourse. When contexts contain target concepts or metaphorical expressions of a given metaphor, the likelihood of occurrence of the same type of metaphor in the subsequent discourse increases, while the likelihood of a literal reference to the source domain lowers.

Furthermore, other researchers used a corpus-based approach to focus their analyses on the degree of metaphoricity or conventionality that the extracted linguistic metaphors may convey. Hanks (2004), for instance, observed that the metaphorical mapping between two concepts (one of which is interpreted in the terms of the other) is more metaphorical when the two concepts have less semantic properties in common. Hanks considered the term *oasis*, and reported that while a desert oasis is the typically used form, an oasis in the Antarctic, either literal or metaphorical, is also possible and when the Antarctic oasis is confronted with an oasis in the city, the latter is interpreted as more metaphorical than the former; while an oasis in the mind is considered to be more metaphorical than both Antarctic and city oases.

Finally, Shutova (2011) was able to quantify the overall pervasiveness of metaphor in language by manually annotating a set of text of different genres from the British National Corpus (BNC) (Burnard, 2007) and then calculating the frequency of metaphor by genre. Shutova reported that on average metaphor occurs in every third sentence, proving that metaphor is a ubiquitous phenomenon in language.

1.4 Computational approaches to metaphor

While corpus linguistics research on metaphor has proved successful in small-scale studies and in analyses that employ already annotated corpora, metaphor extraction is still a problematic step in studies that address extensive not annotated corpora. For this reason, it is relevant to consider computational linguistics' role in the study of metaphor and more specifically in the advancements made to build computational models that automatically extract metaphors from corpora.

One of the main tasks addressed by computational approaches to metaphor is metaphor identification, also known as metaphor recognition. This section will present a brief introduction to computational linguistics methodology and give an overview of previous research that was conducted on metaphor identification.

1.4.1 Computational linguistics

Computational linguistics, alternatively referred to as Human Language Technology or Natural Language Processing (NLP), is an interdisciplinary field whose goal is "to get computers to perform useful tasks involving human language" (Jurafsky & Martin, 2008, p. 1). NLP has numerous applications in various areas of language including the domains of machine translation, virtual assistants, and text mining.

The ubiquity of metaphors in language together with the popularization of Lakoff and Johnson's (1980) CMT prompted an increased interest in NLP research around the theme of metaphors.

The NLP research area of computational metaphor processing focuses on two main sub-tasks as it develops techniques for both the identification and the interpretation of metaphors. Metaphor identification consists of automatically extracting metaphors from text, while metaphor interpretation mainly focuses on decoding the meaning of metaphors by usually paraphrasing metaphorical expressions into literal ones (Shutova, 2011).

In view of the fact that this thesis only adopts a computational method for metaphor identification, the subtask of metaphor interpretation will not be further developed. For a more detailed overview of automated metaphor interpretation refer to Rai and Chakraverty's survey of computational metaphor processing (2020).

1.4.2 Computational metaphor identification

There are three main methodological approaches adopted in machine learning research for metaphor identification tasks.

The first type of approach is characterized by the use of manually crafted rules and can be mostly found in the early work on metaphor detection. For instance, the met* (Fass, 1991) method was one of the first successful attempts at metaphor detection and it took inspiration from the Wilks (1978) *anomaly view* of metaphor presented in section 1.1.1. The purpose of the met* was to determine whether an expression was literal or figurative by detecting the violation of selectional preferences. In a second step the non-literal expressions were categorized as either *metonymic*, using hand-crafted patterns, or *metaphoric*, referring to a manually constructed database of analogies. Mason (2004) also drew inspiration from the principle of violation of selectional preference and proposed the CorMet, a system that automatically determines source and target domain mappings using domain-specific selectional preferences drawn from Internet corpora. Both the met* and the CorMet have limitations related to the violation of selectional preference, more specifically, the fact that the systems have issues with conventional metaphors that do not manifest preference violation, a problem that has been already mentioned in critical approaches to the anomaly view. Furthermore, it should be noted that constructing hand-coded rules is extremely costly.

The second type of methodological approach to metaphor detection includes corpus-based methods that employ statistical algorithms to detect metaphors in text. In particular, this kind of methods employ topic-modeling, word embeddings, and explore the use of different linguistic features. Heintz et al. (2013), for instance, used topic modeling to map source and target domains and, in a second step, identified sentences with words from both domains as metaphorical. Klebanov et al. (2014), on the other hand, used unigram embedding for metaphor detection by creating a supervised model that employed all content words from the training data (the unigrams) as features. Various linguistic features have been used in research including conceptual features such as word concreteness and abstractness (Tsvetkov et al., 2014; Köper and im Walde, 2017), sensory features (Shutova et al., 2016) and WordNet features like hypernyms and synonyms (Mao et al., 2018).

The third and last type of methods have been developed more recently following the resurgence of neural networks in NLP research. Most neural models address the metaphor identification task as either a sequence labeling task or a classification task. Sequence labeling systems output label sequences of input sentences, labeling the metaphoricity of each token in the sentences. Classification models on the other hand, take a known target word as input and then assign the target word a metaphorical or literal class. The vast majority of classification or sequence labeling systems are based on shallow neural networks such as Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) (Su et al., 2020). For instance, Wu et al. (2018) proposed a metaphor detection model based on CNN and BiLSTM, to extract both local and long-range (sentence) contextual information to sequence label metaphorical information.

Other structures of neural networks have been explored to better encode semantic information by detecting metaphorical patterns (Leong et al., 2018). In particular, Transformers based methods such as Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018), RoBERTa (Liu et al., 2019) and XL-Net have also been employed in token-level metaphor detection tasks in order to incorporate contextual information beyond local and sentence contexts. Dankers et al. (2020), for example, fine-tuned a BERT model to encode broader discourse properties in addition to lexico-syntactic and sentence contexts, yielding good performances in metaphor detection tasks. Su et al.'s (2020) DeepMet model, which was employed for this thesis, similarly used a RoBERTa based network to further encode contextual information. A more detailed account of this model is given in section 2.2.

Using Transformers based models like BERT and its variant is considered advantageous because this type of models are trained on large amounts of textual data, as such, they have more representational power than other models that were trained on smaller task-specific data. A greater representation power to encode semantic and contextual information seems to be crucial for metaphor identification tasks (Gong et al., 2020).

The models by Dankers et al. (2020) and Su et al. (2020) mentioned above were all designed in response to a shared task on metaphor identification that was conducted as a part of the ACL 2020 Workshop on Processing Figurative Language (Klebanov et al., 2020). The various models were all tested on the VUA Metaphor Corpus (Steen, et al., 2010) and on a subset of the TOEFL (Beigman Klebanov, Leong, & Flor, 2018). In the report of the task, Leong et al. (2020) note how more than half of the models used deep learning architectures based on BERT and how, compared to the 2018 shared task on the same VUA dataset (also reported by Leong et al.), the performance of the best 2020 models is evidence of general improvement. In particular, the best system in the first share task performed at F1 = 0.651, while the best performance in 2020 - Su et al.'s DeepMet model, was more than 10 points better at F1 = 0.769.

While the results of the two shared tasks on metaphor identification conducted in the last few years show how computational models for metaphor identification are constantly improving, the systems are far from perfect. Neidlein et al. (2020), for instance, note how recent models show gaps from a linguistic standpoint, on the grounds that most of the new models do not perform well in identifying novel metaphors compared to conventional metaphors. The authors conducted a linguistic study to draw an extensive comparison between a series of recent models, including a number of models from the 2020 shared task (Leong et al., 2020), to see their performance on non-conventional metaphor identification. The results showed how all models, with variation, perform substantially worse on novel metaphors and rarer word types in comparison to conventional metaphors. From these results, Neidlein et al. (2020) gathered that recent models are mainly optimizing the word sense disambiguation for conventional metaphors rather than generalizing metaphor properties, since they perform excellently with frequently seen word types. The models in question, however, show better generalization abilities, beyond word sense disambiguation, when the training input includes morphological variations or synonyms.

The fact that most recent models do not perform well in identifying novel metaphors may be related the fact that the majority of these models are trained on the VUA Metaphor Corpus (Steen et al., 2010) but they generally do not account for the fact that the VUA corpus was not originally annotated for computational processing. As Shutova (2011) points out, since Steen and colleagues were interested in historical aspects of metaphor, a vast portion of the metaphors annotated in the VUA corpus are highly conventional metaphors, dead metaphors, and borderline cases. For example, many instances of conventional polysemy such as the verb *show* in the sentence “They want to *show* you that they trust you” are annotated as metaphorical in the VUA corpus, therefore leading the model to optimize word sense disambiguation rather than creative metaphor identification.

1.5 Metaphor in political discourse

While metaphor is pervasive in all areas of human communication, it plays a critical role in specific types of linguistic varieties such as literary language and political discourse. Since this thesis analyzes metaphor in a corpus composed of American presidential speeches, an excursus on the state of the art on the topic of metaphor in political discourse and on political discourse itself is fundamental. The following subsections 1.5.1 and 1.5.2 will briefly outline the main views on political discourse and its relations to metaphor, while sections 1.5.3 and 1.5.4 will discuss the main findings on metaphor in political discourse from the research areas of corpus linguistics and computational linguistics, respectively.

1.5.1 Politics and political discourse

The definition of politics can considerably vary according to the research area and topic of one's interest. Within the field of study of political discourse, Chilton (2004) defines politics as a general "struggle for power" which can be viewed on two different levels of interpretation: the first is a "macro" level, which entails the political institutions of a state and their aim to resolve conflicts of interest over different areas such as money, liberty, and influence, in order to assert the dominance of a specific group or individual; the second is a "micro" level and it involves the struggle for power and attempts of cooperation between individuals or specific social groups.

On the basis of this view of politics, Chilton defines political discourse as the set of linguistic actions that are employed on both macro and micro level to accomplish the above mentioned goals of dominance or cooperation. Such linguistic actions constitute at the macro level specific types of discourse such as parliamentary debates, broadcast interviews, and written laws; at the micro level, on the other hand, political discourse includes techniques such as persuasion, rational argument, or entreaties, among many others. At either level, it is evident that political power heavily relies on communication and that, as Chilton (2004, p. 4) emphasizes, there is ''a linguistic, discursive and communicative dimension" to politics that cannot be disregarded.

A similar and more concise definition of political discourse is Baranov and Kazakevich's (1991, p. 64), according to which political discourse is "the totality of all speech acts used in political discussions, as well as rules of public policy, sanctified by tradition and proven by experience".

However politics and political discourse are defined, it is generally agreed that the main aim of any analysis of political discourse is “to investigate the ways in which language is manipulated for achieving different political goals and shaping political reality” (Berberović and Delibegović Džanić, 2021, p. 1).

The analysis of language in politics belongs to a long tradition of studies that date back to the classical Greek-Roman study of rhetoric. The western classical tradition of rhetoric revolved around the delineation of methods to cultivate social and political competence in order to achieve certain goals. The set of communication techniques employed to accomplish specific political goals continued to be of similar interest through time and, while modern studies of rhetoric developed numerous different schools of analysis over the years, the focus remained constant in the literature.

Modern work in the field of rhetoric is extensive and often interdisciplinary in nature as it was developed concurrently within and across the fields of communication science, historical construction, social theory, and political science. In the early 1980s and 1990s political discourse became a subject of interest for linguistics as well, as the field of linguistics started to expand its focus on discourse in general (Wilson, 2005).

Linguistic analysis of political discourse, and more specifically of American political discourse, has been conducted on various subtypes of political discourse such as press conferences (Howe, 1988), political speeches and manifestos (Charteris-Black, 2004), and political news (Berberović and Delibegović Džanić, 2021). Most of these studies emphasize the role of persuasion and rhetoric in politics and how they are actualized through linguistic devices in political discourse in order to achieve the speaker’s specific goals. One of the linguistic devices that have been analyzed as a means of persuasion in political discourse is metaphor.

1.5.2 Metaphor in political discourse

Various critical discourse analyses of metaphor in political discourse show that metaphor in this type of textual context displays at least two additional functions beyond its primary function of describing one concept in terms of another.

The first function is the already mentioned use of metaphor as a persuasion device to achieve specific rhetoric goals. Charteris-Black (2011) in fact includes a good command of metaphor as part of a politician's persuasion skills, together with fine rhetoric skills and attention to personal style. Within conceptual metaphor theory, Lakoff and Johnson (1980) state that metaphors can be used to shape political reality on account of their power to either highlight or conceal specific aspects of reality. Musolff (2004) further considers metaphor to have an affective power, allowing the speaker to give sentences a particular connotation to influence listeners. Overall, political discourse literature regards metaphor as a powerful persuasive tool.

The second function that has been found for metaphor in political discourse is that of discourse structuring. Different analyses of metaphor patterns in political discourse revealed that metaphor is often used, either consciously or unconsciously, to organize discourse by establishing intratextual or intertextual coherence. Mulsoff (2004) explains how the same metaphor can be used in multiple texts in order to show one's position of either agreement or disagreement with a particular view. Semino (2008) additionally notes that some metaphors can link multiple texts on the same subject and that one way to establish this kind of intertextual relationship is to employ the same metaphor at different historical periods. Furthermore, Semino provides an overview of the different metaphor patterns that manifest these kinds of intertextual or intratextual relations. These patterns include metaphor repetition, clustering, extension, combination and mixing, and literal-metaphorical opposition.

As regards to the type of metaphors that are used in political discourse, Berberović and Delibegović Džanić (2021) point out that, similarly to other areas of research on metaphor, the majority of discourse studies primarily focus on conventional metaphors while mostly ignoring novel metaphors. This lack of research on creative metaphors is in part due to the predominance of CMT studies which focus on relatively fixed conceptual mappings rather than creative ones. The few quantitative studies that do analyze the use of creative metaphor in discourse prove that novel metaphor is not confined to literature and poetry as one might believe. Goatly (1997) in fact reports that, while novel metaphors occur with more frequence in modern lyric poetry and modern novels (56% and 28%, respectively), creative use of metaphor is also consistently found in other genres such as magazine advertising, popular science and national news reports, albeit in lower frequencies. Creative use of metaphor is found in political discourse as well and its function has been discussed in comparison to that of conventional metaphor. Semino (2008) for instance proposes that, similarly to conventional metaphors, novel metaphors are used in political discourse to achieve specific rhetoric goals, however, differently from conventional metaphors, they are mostly employed to support a specific argument that is related to a specific context. Semino further argues that creative metaphors are employed for discourse structuring analogously to conventional metaphors but also have a more memorable effect than conventional metaphors as a result of their novel nature. Whatever role novel metaphor plays in political discourse, it is evident that its use is as central as that of conventional metaphor, if not more.

Both corpus linguistics and computational linguistics studies have been conducted on metaphor patterns and metaphor as a persuasive device in political discourse. The remainder of this section will outline the main findings from these two areas of research.

1.5.3 Corpus-based research on metaphor in political discourse

Corpus linguistics methods have been employed to quantitatively and qualitatively analyze metaphors in various types of political discourse. The majority of these corpus based studies adopts CMT as a theoretical basis, meaning that they center around a specific set of conceptual metaphors and their recurrent linguistic manifestations in the form of conventional metaphors.

Howe (1988), for instance, conducted a study on metaphor in American political discourse in the New York Times newspaper and a number of periodicals to identify the most common source domains used in the corpora. He manually determined that the most used conceptual metaphors were related to the domains of war and sports. According to Howe, the use of metaphorical expressions from these domains aimed at establishing a common ground between the speaker and the voter that is based on domain areas that are known by both parties.

Charteris-Black (2004), on the other hand, adopted a corpus based and critical discourse analysis based approach, i.e. Critical Metaphor Analysis, to analyze and compare the conceptual metaphors in two different types of political discourse: British party political manifestos and American presidential speeches. Focusing on the American presidential speeches, Charteris-Black (2004, p. 87) defines them as “a very distinct type of political discourse because their purpose is to offer an idealized 'vision’ of the social world”. As such, the use of metaphor in this type of speeches is usually systematic and directed at conveying this particular vision. He further notes that, since political speeches usually follow a script written by a team of ghost writers, they exhibit a higher degree of structural planning than spontaneous speeches, and thus the use of metaphor is also directed at improving discourse structuring.

The corpus analyzed by Charteris-Black (2004) consisted of 51 inaugural speeches of US Presidents from George Washington to Bill Clinton. The first step of the Charteris-Black’s Critical Metaphor Analysis approach consisted in metaphor extraction; this step was carried out by manually searching keywords from source domains of interest in a sample of speeches and then throughout the corpus. For example, terms such as *step* and *path* were searched and labeled as metaphorical when associated with the conceptual mapping LIFE IS A JOURNEY instead of being used literally, this was determined with reference to the keywords’ contexts of use. Charteris-Black additionally compared the prevalence of different source domains in the corpus by measuring their “resonance”, that is, by multiplying the sum of tokens and the sum of types of metaphors from the same source domain and comparing the result to that of the other domains. The three most resonant source domains reported in Charteris-Black's (2004) research are those of conflict (36%), journey (16%) and buildings (14%).

The second step of the Critical Metaphor Analysis approach consisted in interpreting the keywords that were identified as metaphors in the first step. According Charteris-Black, metaphor interpretation consists in "establishing a relationship between metaphors and the cognitive and pragmatic factors that determine them" (2004, p. 37), which consists in conceptual metaphor identification.

The third and final step of the Critical Metaphor Analysis approach consisted in metaphor explanation, that is, in determining how metaphors are employed for persuasion purposes with reference to the textual context in which they occur.

Overall, the most common source domains extracted by Charteris-Black (2004) revolved around either familiar concepts, personal experiences or social activities. More specifically, the set of source domains, in order of resonance, consisted of the source domains of: conflict, as in the conceptual mapping POLITICS IS CONFLICT; journey, associated with the concept of progress towards a purposeful social activity or a more generic goal; fire and light, as in HOPE IS LIGHT, relating to sentiments of altruism and ambition; homes and buildings, as in SOCIETY IS A BUILDING; physical environment, as in CIRCUMSTANCES ARE WEATHER; and religion, as in the conceptual metaphor POLITICS IS RELIGION. These types of conceptual metaphors are often instances of reification, meaning that they convey abstract concepts as if they were physical ones. Charteris-Black (2004), similarly to Howe (1988), attributes the familiarity and physicality of the source domains to a rhetorical purpose. In particular, since the corpus is composed of inaugural political speeches, he argues that most metaphors drawn from familiar domains are employed to convey the value of abstract social ideals such as peace and justice and to give speeches a more coherent and intelligible structure.

1.5.4 Computational research on metaphor in political discourse

As previously stated in section 1.3, one major drawback of corpus linguistics research on metaphor is the fact that it cannot be extended to large corpora that are not specifically annotated. While small-scale studies that focus on specific conceptual metaphors are feasible and have proven effective, larger-scale studies cannot rely on hand-coded searches to analyze broader metaphor patterns. This issue persists in research on metaphor in political discourse, making large-scale domain agnostic studies on the use of metaphors in political discourse problematic.

Prabhakaran et al. (2021) addressed this problem by using two metaphor classifiers to automatically detect metaphors in a corpus composed of over 85K posts made by 412 US politicians in their Facebook public pages. The classifiers were based on Rei et al.'s (2017) architecture, which takes as input word pairs such as *(cure, crime)* and outputs a score that indicates the corresponding phrase as either metaphorical or literal. The aim of Prabhakaran et al.'s research was to investigate the general effectiveness of metaphor as a persuasive tool, identifying the broader patterns that allow metaphor to be used as a persuasive device in political discourse without only focusing on a limited set of source domains.

The effectiveness of metaphor was measured by comparing reader engagement in posts that contained metaphorical or literal use of the same words. Prabhakaran et al.'s analysis showed that metaphor usage positively affects engagement in political discourse: the engagement, measured according to shares, comments and reactions data, was greater in posts containing metaphorical expressions.

Prabhakaran and colleagues' success in employing computational linguistics methods to tackle the subject of metaphor in a large-scale political corpus proves that further computational research could help provide additional insight in the use and function of metaphor in political discourse.

2. Materials and methods

Having discussed the relevant background on metaphor in political discourse within the fields of corpus linguistics and computational linguistics research, this chapter establishes the materials and the methods selected and employed in this thesis. In particular, subsection 2.1 covers the construction, classification and summary statistics of the ad-hoc corpus of political speeches that was constructed for the purpose of being computationally annotated for metaphors. Subsection 2.2 presents the computational model that was selected for metaphor identification in the corpus and the preprocessing steps performed to set up the corpus for metaphor identification. Subsection 2.3 covers the methods employed for the analysis of the metaphors that were computationally identified in the corpus.

2.1 The American Presidential Speeches Corpus

The ad-hoc corpus of political speeches that was built for computational metaphor identification and metaphor analysis is composed of 1721 speeches of American Presidents from Dwight D. Eisenhower to Joe Biden. The corpus will be referred to as the American Presidential Speeches (APS) corpus.

While there already exist corpora composed of US presidential speeches such as the "Corpus of Political Speeches" (Kathleen, 2015) and the "United States Presidential Speeches" database (Lilleberg, 2019), an ad-hoc corpus was built for this thesis to specifically include campaign speeches together with presidential speeches, with the intent to explore the results of the metaphor identification task in this type of political speech as well. Furthermore, unlike other existing corpora, the speeches of the APS corpus were not automatically extracted from a single database, but rather manually selected from multiple databases in order to ensure balance and representativity of the corpus regarding the number of speeches selected per presidential term.

2.1.1 Corpus construction

The APS corpus was constructed between October 2021 and March 2022. The 1721 speeches in the corpus have been delivered by thirteen Presidents, from Dwight D.

Eisenhower to Joe Biden, in a period that spans from 1952 to 2021. The speech transcripts were downloaded from four main digital repositories which are illustrated in the list below.

- **The American Presidency Project**:[3] online repository of US presidential documents hosted at the University of California, Santa Barbara
- **American Rhetoric Online Speech Bank**:[4] database of full text, audio, and video versions of public speeches, including presidential speeches, compiled by University of Texas professor, Michael Eidenmuller
- **National Archives Presidential Libraries**:[5] presidential libraries administered by the US National Archives; the libraries are repositories of speeches, papers, records and historical materials of each President
- **Presidential Speech Archive, Miller Center**:[6] online repository of US presidential speeches hosted by the Scripps Library affiliated with the University of Virginia in cooperation with various presidential libraries

To ensure balance and representativity of the corpus, a number between 87 and 182 speeches were collected for each president, with an average of about 86 speeches for each one of the 20 individual presidential terms considered. A subset of the corpus of 80 speeches (about four per presidential term) that were delivered during the Presidents' presidential campaigns. This subset is analyzed separately and will be referred to as "Campaign Speeches" opposed to the "Presidential Speeches" subset, which represents the majority of the corpus.

Text retrieval from the selected online databases was performed with the help of the BootCaT web-scraping toolkit (Baroni and Bernardini, 2005). First, a list of URLs corresponding to each document page was compiled manually. The URL list was then fed into BootCat to automatically retrieve the corresponding webpage for each URL and convert them to plain text to build a corpus composed of 1721 speeches in *txt* format.

---

[3] https://presidency.ucsb.edu
[4] https://americanrhetoric.com
[5] https://archives.gov/presidential-libraries.html
[6] https://millercenter.org

The process employed to further prepare and filter the APS corpus for metaphor analysis is broadly based on the preprocessing procedure employed by Ficcadenti (2019) for a corpus of US presidential speeches transcripts downloaded from the Miller Center database.

Firstly, the transcripts were assessed for typos that could obstruct the text tokenization, that is, the automatic identification of individual words in the texts. In particular, misuses of punctuation such as in "*you.Therefore*" or "*be--and*", and missed blanks between words such as in "*thePresident*" and "*30,000of*", would impede an accurate text tokenization. These typos were corrected with the use of regular expressions, which is a method for searching specific patterns into strings with the use of pseudo-coding languages.

In the second step, the scraped texts were assessed for non-content data that is not relevant for the research, namely text metadata such as titles, authors, dates, speech context notes, and other website specific annotations such as indicators of applause or laughter. These metadata elements were also removed with regular expressions.

Following Ficcadenti's (2019) cleaning procedure, transcripts of public events such as press conferences were further stripped of any question from the reporters to isolate the Presidents' statements, since these are the sole focus of this thesis. The reporters' questions in the transcripts are signified by a "Q" followed by punctuation or space. A *for* loop was employed to identify the questions in the transcripts and remove all text after the first question.

In the next step, interactions with the public signified by speaker markers such as "THE AUDIENCE" or "Audience member" were removed with regular expressions, isolating the words pronounced by the Presidents.

The corpus was subsequently computationally preprocessed for statistical analysis with the Python NLP library *spaCy* (Honnibal and Montani, 2017). In this initial stage of preprocessing the text was split into sentences and tokenized (each word was automatically identified and separated).

In the final step, the tokenized documents were assessed with regards to their length to assess whether the modifications performed in the previous steps did not make them not suitable for a consistent analysis. In particular, speeches whose resulting number of tokens was less than 100 were filtered from the original corpus and omitted from the analysis. Overall, 87 speech transcripts were filtered from the original corpus. In addition, 37 speeches classified as presidential debates in the digital repositories were also omitted from the analysis, since debates are structured as conversations between the two candidates based on specific questions, and only considering the President's answers would be problematic. In total, 124 speeches in the original corpus were omitted from the analysis reported in Chapter 3.

2.1.2 Corpus classification

The speeches in the constructed APS corpus were classified by president, term of presidency, President political affiliation (i.e. Democratic or Republican), and document genre.

The genre of each Presidential Speech in the APS corpus was determined according to how the speech was classified or described in its source archive. The list below illustrates the set of genres for speech classification selected for this thesis:

- ***Address to the United Nations***: presidential address delivered before the General Assembly of the United Nations (UN)
- ***Farewell Address***: address delivered upon leaving the presidency
- ***Inaugural Address***: address delivered at the inaugural ceremony for the presidency
- ***Press Conference***: statement given by the President to reporters before answering questions from them
- ***Remarks***: address delivered by the President on current affairs or to frame an event, usually shorter than a speech, as an example, refer to Obama's speech: "Remarks on Gun Violence"[7], delivered on January 5, 2016
- ***Speech***: formal address delivered by the President to an audience on various subjects, as an example, refer to Kennedy's speech: "Address on the Space Effort"[8], delivered on September 12, 1962
- ***State of the Union Address***: annual speech delivered by US Presidents to a joint session of the United States Congress regarding the current condition of the nation
- ***Statement***: a brief communication setting forth particulars or facts, i.e. see Trump's "Statement on the Coronavirus"[9] delivered on March 11, 2020
- ***Victory speech***: speech delivered by the US president-elects after winning the presidential election

---

7 https://www.presidency.ucsb.edu/documents/remarks-gun-violence-0
8 https://millercenter.org/the-presidency/presidential-speeches/september-12-1962-address-space-effort
9 https://millercenter.org/the-presidency/presidential-speeches/march-11-2020-statement-coronavirus

- ***Weekly address***: weekly speech by the US President to the nation
- ***University Speech***: speech delivered in universities, air force, naval and military academies in the US and abroad, including commencement speeches and remarks to graduating classes
- ***Others***: class for genres represented by very few speeches such as concession speeches, memorandum, messages, proclamations, and proposals

Originally the number of genre classes selected for the Presidential Speeches subset was higher than the 12 reported above, but for ease of analysis, categories with less than 10 documents in the entire corpus were collapsed into the "others" category.

The Campaign Speeches subset, which pertains to each president's presidential campaign, or campaigns in the case of second terms, was categorized into the two genres of:

- ***Campaign announcement speech***: formal speech delivered to publicly launch presidential campaigns
- ***Campaign speech***: persuasive speech delivered by politicians running for President to advertise their platform
- ***Debates***: presidential debate held between the Democratic nominee and the Republican nominee for President, as mentioned in the previous subsection 2.1.1, the debates in the the original corpus were not considered for analysis as their structure resembles that of a conversation rather than a rhetoric speech

2.1.3 Corpus statistics

The resulting corpus after the filtering and cleaning process described in section 2.1.1 is composed of 1597 cleaned speech transcripts. For the data regarding the original corpus before the filtering process was performed, refer to the tables in Appendix A. Overall, the filtered APS corpus is composed of 2,677,469 word tokens, and has a vocabulary of 36,317 word types. Table 2.1 below shows the types and tokens distributions in the two subsets of the APS corpus: the Campaign Speeches subset and the Presidential Speeches subset.

Table 2.1

Distribution of tokens and types in the Campaign and Presidential Speeches subsets

	# Tokens	# Types
Campaign speeches	92925	1933
Presidential speeches	2584544	35797
Total	2677469	36317

Table 2.2 summarizes the number of Presidential Speeches and their average size in the number of tokens for each of the 20 individual presidential terms specified by year chronologically. The speech per President quantities were balanced according to individual presidential terms since not all the Presidents selected served for two terms, namely, John F. Kennedy, Gerald R. Ford, Jimmy Carter, George H. W. Bush, Donald Trump, and Joe Biden as of 2022. Each presidential term occupies about 5% of the Presidential Speeches with the exception of Richard Nixon's second term at 3%. This is because Nixon resigned after two years into his second term of presidency, thus limiting the amount of speeches delivered during this term that was supposed to last two more years.

The average Presidential Speech size in tokens tends to be larger for second terms with the exception of Ronald Reagan and Richard Nixon speeches, which are longer in the first term. Gerald Ford and Jimmy Carter's speeches are the shortest, with an average of about 1500 tokens, while George W. Bush and Bill Clinton's second term speeches are the longest with an average of about 2600 and 2400 tokens respectively.

Table 2.2

Distribution of presidential terms in the Campaign and Presidential Speeches subsets

	term	# camp docs	ave size	# term docs	ave size
Dwight Eisenhower	1953-1957 1957-1961	2 2	2065 3138	60 65	1569 1906
John F. Kennedy	1961-1963	4	1362	87	1913
Lyndon B. Johnson	1963-1965 1965-1969	1 2	1494 3454	77 73	1814 2114
Richard Nixon	1969-1973 1973-1974	2 1	526 4385	71 48	2022 1941
Gerald R. Ford	1974-1977	2	1632	79	1481
Jimmy Carter	1977-1981	3	2910	82	1499
Ronald Reagan	1981-1985 1985-1989	2 2	4161 1134	82 88	2226 1960
George H. W. Bush	1989-1993	3	3702	78	1554
Bill Clinton	1993-1996 1996-2001	3 --	2688 --	90 77	2073 2427
George W. Bush	2001-2004 2004-2009	2 1	3040 5125	80 93	2055 2602
Barack Obama	2009-2012 2013-2017	3 2	2895 4114	79 80	1708 1802
Donald Trump	2017-2021	2	1755	86	2061
Joe Biden	2021-2022	4	1660	79	2290
Total		43	2562	1554	1951

Note. There are no speeches for Bill Clinton's second presidential campaign because the two speeches selected in the non-filtered APS corpus for this period were debates, which, as already explained, were not taken into account for the analysis.

Table 2.3 and 2.4 below shows the distribution of genre types in the filtered APS corpus with a summary of the number and the percentage of documents and their average size in the number of tokens for each genre type. Table 2.3 summarizes the genre types distribution in the Presidential Speeches subset, and Table 2.4 in the Campaign Speeches subset.

Table 2.3

Distribution of genre types in the Presidential Speeches subset

	# docs	*percentage*	*ave size*
Address to the United Nations	48	3.1%	3157
Farewell address	11	0.7%	2982
Inaugural address	19	1.2%	1802
Press conference	331	21.3%	1050
Remarks	567	36.5%	1758
Speech	267	17.2%	2602
State of the Union address	65	4.2%	5490
Statement	26	1.7%	767
University/Academy speech	96	6.2%	2730
Victory speech	17	1.1%	1102
Weekly address	64	4.1%	682
Other	43	2.8%	2195
Total	1554	100%	4270

Presidential speeches classified as Remarks and Press conference have the highest percentages, occupying 36.5% and 21.3% of the Presidential Speeches in the filtered APS corpus.

The longest speeches are in the State of the Union Address and Address to the United Nations genres, with an average size of 5490 and 3157, respectively. Statements and Weekly addresses are the shortest speeches with an average size of 767 and 682 tokens, respectively.

As regards to the distribution of political affiliation in the APS corpus, as illustrated in detail in Table 2.5 below, a total of 22 Campaign Speeches were delivered by Democratic Presidents and the remaining 21 by Republican Presidents. A total of 830 Presidential Speeches were delivered by Republican Presidents and 724 by Democratic Presidents, respectively 53% and 47% of the Presidential speeches subset. The average sizes of the speeches delivered by Republican and Democratic Presidents are very similar.

Table 2.4

Distribution of genre types in the Campaign Speeches subset

	# docs	percentage	ave size
Campaign announcement	16	37.2%	1650
Campaign speech	27	62.8%	3086
Total	43	100%	2368

Table 2.5

Distribution of party affiliation in the Campaign and Presidential Speeches subsets

	# camp docs	percentage	ave size	# term docs	percentage	ave size
Democratic	22	51%	2463	724	47%	1957
Republican	21	49%	2644	830	53%	1965
Total	43	100%	2553	1554	100%	1961

2.2 The DeepMet model

After the construction, classification and preliminary analysis of the APS corpus, a computational model was selected in order to automatically identify metaphors in the presidential speeches transcripts.

The DeepMet model (Su et al., 2020) was selected for this thesis out of the models presented in the Second Shared Task on Metaphor Detection (Leong et al., 2020), which involved the most recent models designed for metaphor detection to date. The shared task was conducted as a part of the ACL 2020 Workshop on Processing Figurative Language (Klebanov et al., 2020) to further the research in the field of verb and part-of-speech (POS) metaphor detection technology.

The shared task provided two evaluation datasets: the VU Amsterdam Metaphor Corpus (VUA) (Steen, 2010) and the TOEFL database (Beigman & Flor, 2018). The VUA dataset is a sample of 117 documents from the British National Corpus, classified in the four genres of: Academic, Conversation, Fiction, and News. The VUA documents are annotated for metaphors according to the MIPVU procedure, which has a strong interannotator reliability of $\kappa > 0.8$ (Steen et al., 2010).

The TOEFL database on the other hand is a sample of the ETS Corpus of Non-Native Written English (Klebanov et al., 2018). It is composed of 240 argumentative essay responses written by non-native English speakers with medium and high proficiency levels. The essays were annotated for metaphors according to the procedure in Beigman Klebanov and Flor (2013), with an average inter-annotator agreement of $\kappa = 0.56\text{-}0.62$ (Beigman & Flor, 2018).

The databases provided by the shared task were both already divided into a training set and a test set. Table 2.6 below summarizes the number of documents and tokens for both the training sets and test set of the VUA and TOEFL databases.

Since the Second Shared Task on Metaphor Detection (Leong et al., 2020) aimed at developing new models for the identification of both verb metaphors and all part-of-speech (POS) metaphors, the evaluation dataset training sets and test sets were further divided for verb metaphor detection and all POS metaphor detection. The VUA train sets annotated for verb and all POS metaphors have metaphor percentages of 29% of 17240 tokens and 18% of 72611 tokens, respectively. While the TOEFL train sets have metaphor percentages of 13% of 7016 verb tokens and 7% of 26737 all POS tokens. Leong et al. (2020) report that thirteen teams took part in the shared task with varying focus on the four tracks. The track most utilized for evaluation was the VUA all POS track.

Table 2.6

Number of documents and tokens for the VUA and TOEFL datasets

	VUA		TOEFL	
	Training set	Test set	Training set	Test set
# docs	90	27	180	60
# tokens	12123	4081	2741	968

Note. From “A Report on the 2018 VUA Metaphor Detection Shared Task” by Leong, C. W. (Ben), Beigman Klebanov, B., & Shutova, E., 2018. *Proceedings of the Workshop on Figurative Language Processing*, 56–66, p. 2.

Regarding the characteristics and performance of the models presented for the task, Leong et al. (2020) report how more than half of the models used deep learning architectures based on BERT and how, compared to the 2018 shared task on the same VUA dataset (also reported by Leong et al., 2020), the performance of the best 2020 models is evidence of general improvement. In particular, the best system in the first share task performed at F1 = 0.651, while the best performance in 2020 - Su et al.’s DeepMet model, was more than 10 points better at F1 = 0.769. Overall, all teams performed better on the VUA data than on the TOEFL data. With the VUA database, the models performed substantially better on Academic and News documents than Fiction and Conversation documents.

The DeepMet model (Su et al., 2020) was specifically chosen to detect metaphors in the APS corpus because it achieved the highest scores out of all the other models participating in the task, performing at F1 = 0.804 and 0.749 for verb metaphor detection in the VUA corpus and the TOEFL database, respectively; and at F1 = 0.769 (VUA) and 0.715 (TOEFL) for POS metaphor identification. Su and colleagues made the DeepMet source code available online[10] for use.

As regards to how the DeepMet model works, Su et al. (2020) approached the metaphor detection task as a reading comprehension task. Machine reading comprehension tasks estimate a machine’s ability to read and understand Natural Language based on how it responds to text/document-based questions through knowledge and logic based inferences (Hermann et al., 2015). McCann et al. (2018) showed that

[10] https://github.com/YU-NLPLab/DeepMet

various NLP tasks such as machine translation or question answering can be successfully framed as reading comprehension tasks. Inspired by McCann et al.'s (2018) research, Su et al. (2020) framed their metaphor detection paradigm as a reading comprehension task. Put simply, Su and colleagues' reading comprehension metaphor detection task estimates the machine's ability to answer the question of whether a specific word is used metaphorically or literally based on its linguistic context. The DeepMet model labels query words as *literal* or *metaphorical* with consideration to their surrounding contextual linguistic information.

Su and colleagues define their reading comprehension paradigm for metaphor detection as triple (s, q_j, y_j) $(S, q_j \in Q, y_j \in Y)$, where S is a sentence, q_j is a sequence of query words from the query word vocabulary Q within the sentence S, and $y_j \in \{1,0\}$ is the predicted label for q_j, with 1 indicating a metaphor and 0 a literal query word; Y is the label sequence composed of each predicted y_j. The main goal of the DeepMet is to return the conditional probability $P(Y|S, Q)$.

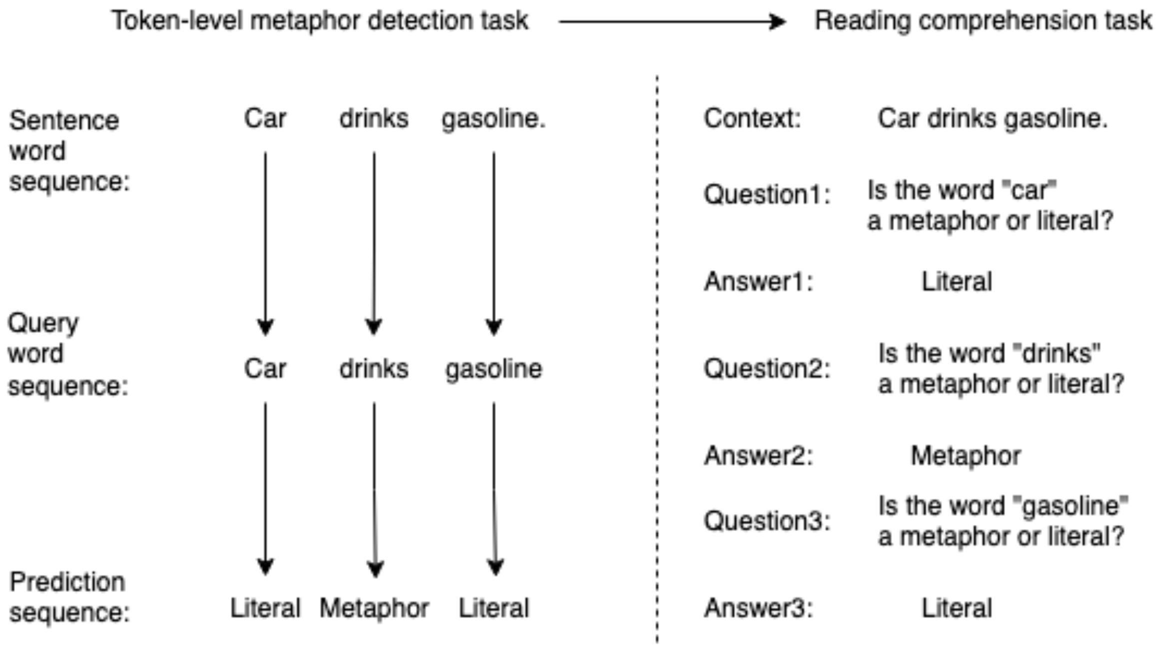
Su et al. (2020) illustrate the reading comprehension paradigm by giving as example the context "car drinks gasoline", in this case if the query word is *car*, with the question being "Is the word *car* a metaphor or literal?", the correct answer would be the label 0, which indicates a literal use of the word *car*. However, when the query word is "drink", the correct predicted label would be 1, as the verb *drinks* would be used metaphorically in this context. Figure 2.1 illustrates the schematic diagram of the metaphor detection task designed by Su and colleagues in more detail.

Based on this reading comprehension paradigm, Su et al. (2020) designed an end-to-end metaphor detection model. The overall architecture of the model is shown in Figure 2.2. The DeepMet uses a RoBERTa (Liu et al., 2019) based embedding layer of that, together with global text context information (full sentences), local context information (the short sentence fragments containing the query words) and the query word sequences, it also encodes both general and fine grained POS auxiliary features, isolating verbs, nouns, adjectives, and adverbs. Parts of speech such as punctuation, prepositions and conjunctions are ignored as they are not likely to trigger metaphors.

The embedded features are subsequently processed into Transformer stacks and ensemble for inference. The Transformer encoder layer has a siamese architecture that employs two Transformer encoder layers A and B to process global text features and local text features, respectively. The other features - the query word information and general and fine grained POS features - are shared by the two Transformer encoder layers.

Figure 2.1

Schematic diagram of metaphor detection task translated into reading comprehension task



Note. From "DeepMet: A Reading Comprehension Paradigm for Token-level Metaphor Detection" by Su, C., Fukumoto, F., Huang, X., Li, J., Wang, R., & Chen, Z., 2020, *Proceedings of the Second Workshop on Figurative Language Processing, 30–39*, p. 32.

Overall, the DeepMet achieved the highest scores out of all the other models participating in the task on both the VUA and TOEFL datasets, showing that building a metaphor detection model based on a reading comprehension can model the nature of metaphor comprehension successfully.

Su et al. (2020) performed an ablation experiment to establish the most influential features for metaphor detection, and observed that fine grained POS and global text features are the most helpful for the task. They also performed ablation experiments to analyze the model architecture and found that the Transformer encoder layer A has a greater influence on the model than the Transformer econder B, meaning that the global text information extracted by the first layer is better than the local text information extracted by the second.

Figure 2.2

The overall architecture of the DeepMet model

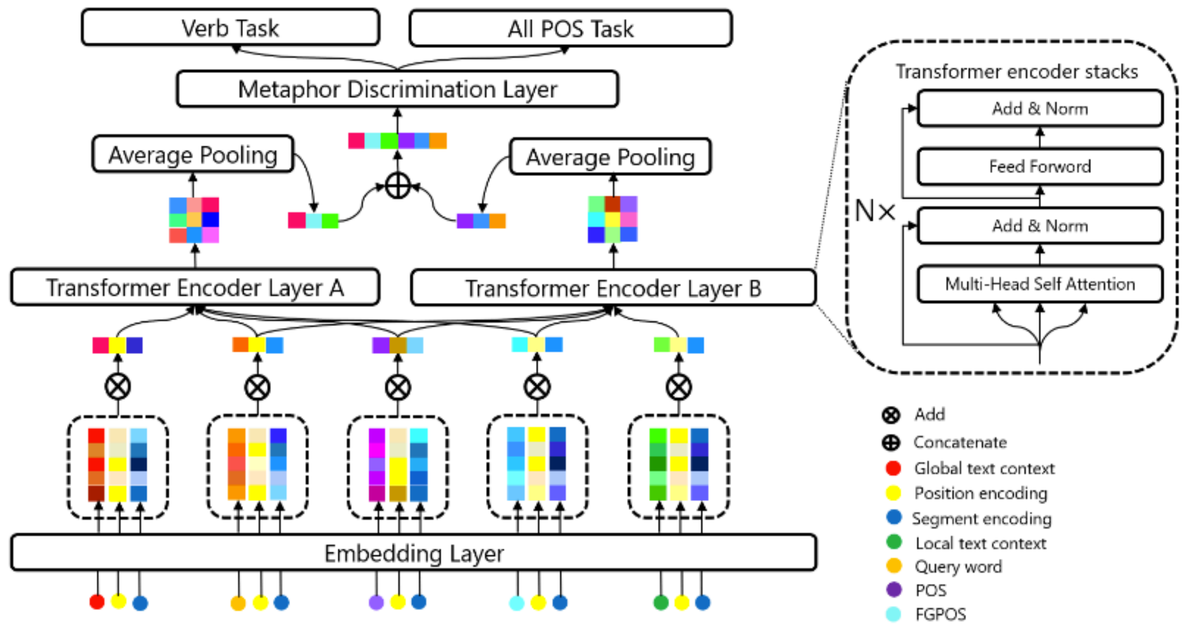


Figure 2: The overall architecture of our model (DeepMet).

Note. From “DeepMet: A Reading Comprehension Paradigm for Token-level Metaphor Detection” by Su, C., Fukumoto, F., Huang, X., Li, J., Wang, R., & Chen, Z., 2020, *Proceedings of the Second Workshop on Figurative Language Processing, 30–39*, p. 32.

2.2.1 Limits of the model

In the error analysis Su et al. (2020) reported that the model predicts incorrectly when annotation in the datasets is ambiguous. The example reported is “The Health Secretary accused the unions of ‘posturing and pretending’ to run a 999 service yesterday” (VUA ID: a7w-fragment01 29), the underlined words are labeled as metaphors in the VUA corpus. In this example the DeeMet erroneously identified *accused* as literal, however, labeling *accused* as either metaphorical or literal in this context would be difficult even for human judgment. Furthermore, since the model works one word at a time at token level, metaphors that are triggered by multiple words are also difficult to detect. For example in “I stared at Jackson Chatterton, and at last sensed the drama that lay behind his big calm presence.” (VUA ID: ccwfragment04 2095) *big* is erroneously labeled as literal.

Su et al. (2020) finally mention a future implementation of linguistic theory into their framework to make their model more explanatory. A deep learning model such as the DeepMet cannot be considered very explanatory because, while it provides accurate predictions, it does not provide interpretable insights on why a word is identified as a metaphor or not. Having a linguistic theoretical basis could not only make predictions more accurate but also more interpretable from a linguistic point of view as they could account for different types of metaphors. As Neidlein et al. (2020) point out, most recent deep learning models, the DeepMet included, seem to be mainly getting better at optimizing the word sense disambiguation for conventional metaphors rather than actually generalizing metaphor properties. Neidlein et al. (2020) show this by comparing how selected models from the Second Shared task (Leong et al., 2020) perform on novel metaphors and conventional metaphors. They report that while the models perform excellently with frequently seen word types, they all perform substantially worse on novel metaphors and rarer word types. This could be related to the fact that the VUA Metaphor Corpus (Steen, et al., 2010) used for model training in the shared task was not originally annotated for computational processing and is largely composed of highly conventional metaphors, dead metaphors and borderline cases. Thus, an implementation of linguistic theory related to the distinction of conventional and novel metaphors, could also better a model’s abilities to generalize on novel metaphors.

2.2.2 Corpus preprocessing for metaphor identification

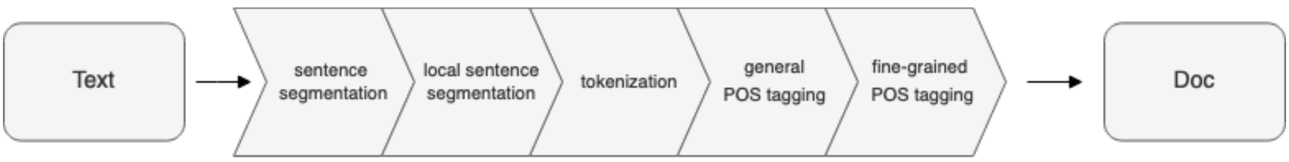
Before employing the DeepMet model (Su et al., 2020) for automatic metaphor identification in the APS corpus, the already cleaned and tokenized corpus text was further preprocessed with the preprocessing code provided by Su et al. (2020) in order for the corpus to be correctly encoded by the metaphor identification model.

Firstly, the corpus composed of multiple documents in *txt* format was converted into a single comma-separated values file (i.e. a *csv* file). The corpus data in the *csv* file was organized at token level, each row representing a single token in the corpus, or in Su et al.'s (2020) terms, a single query word. Each query word in the *csv* file is accompanied by the document id that specifies where the word is located, the global text context of the query word (the full sentence), the local context of the token (the short sentence fragment containing the token), and the token general and fine-grained part-of-speech (POS) features.

The sentence segmentation, tokenization and general and fine grained POS features were obtained for each token with the *spaCy* (Honnibal and Montani, 2017) framework. Figure 2.3 below illustrates a schematic representation of the spaCy corpus text preprocessing that was performed on a document level to create the *csv* document used as input for the DeepMet model (Su et al., 2020).

Figure 2.3

Corpus text preprocessing pipeline



Tokens tagged as punctuation, prepositions, determiners and conjunctions were stripped as they are less likely to trigger metaphors. The resulting annotated and preprocessed corpus was subsequently used as input for the DeepMet model (Su et al., 2020) for metaphor identification.

2.3 Methods

This section is devoted to the methods employed to analyze the results of the metaphor identification task performed by the DeepMet model (Su et al., 2020) on the APS corpus. In order to present a comprehensive analysis of the results of the metaphor identification task, both a quantitative and qualitative approach were adopted.

The distribution of the tokens identified as metaphors by the model was firstly analyzed quantitatively according to individual campaign and term of presidency, speech genre, and the president's political affiliation (i.e. Democratic or Republican). The methods employed to perform this quantitative analysis are covered in subsection 2.3.1.

The metaphors identified in the APS corpus were then analyzed following part of Charteris-Black's (2004) Critical Metaphor Analysis approach, classifying and analyzing a subset of the metaphorical tokens related to the set of source domains drawn up by Charteris-Black (2004) specifically for American presidential speeches. The methodology background and methods employed to perform this analysis are covered in subsection 2.3.2.

2.3.1 Quantitative approaches to metaphor analysis

A quantitative corpus-based approach was employed to explore the distribution of the tokens identified as metaphors by the DeepMet model (Su et al., 2020) in the APS corpus. The frequency of lemma tokens and types identified as metaphors was calculated on the *csv* file outputted by the model converted to a Pandas (McKinney et al., 2010) dataframe.

The distribution of metaphor tokens and types was calculated for the two subsets of the APS corpus, the Campaign Speeches subset and Presidential Speeches subset, in order to compare the percentage of metaphor use in these two types of speeches.

The distribution of metaphors in the APS corpus was then calculated according to the POS tag of the tokens identified as metaphors: this was performed as a means to gain insights into which parts of speech were more likely to be identified as metaphors by the model compared to the POS distribution of the literal tokens in the corpus. The results of the POS distribution analysis were also employed to filter function word tokens identified as metaphors and to analyze clusters of consecutive tokens tagged as metaphorical by the model (i.e. multi-word metaphorical expressions). Function word unigrams identified as metaphors were filtered and not considered in the final quantitative analysis since they are not relevant metaphorical expressions when isolated.

The percent frequency distribution of the filtered dataset of metaphors identified by the model was then computed based on: presidential campaign and presidential term, in order to get insights into the quantitative use of metaphor through time; genre of speech, to assess whether metaphor use changes based on the genre of the speech delivered; and the presidents' political affiliation, to determine whether quantitative use of metaphor changes if the president who delivered the speech is affiliated to the Republican or Democratic party.

2.3.2 Quantitative and qualitative approaches to source domain analysis

An exploratory quantitative and qualitative manual approach was employed to gain insights into the type of metaphors that were identified by the DeepMet model (Su et al., 2020) in the APS corpus.

Subsequently, a "top-down" analysis of the type of metaphors identified in the APS corpus was performed by way of searching keywords related to the same source domains and related conceptual metaphors drawn up by Charteris-Black (2004) for his corpus-based research on metaphor in American presidential speeches outlined in section 1.5.3. In his research, Charteris-Black (2004) manually classified the metaphors used in 51 American presidential inaugural speeches from George Washington to Bill Clinton into the source domains of *body parts*, *buildings*, *conflict*, *fire and light*, *journey*, *physical environment*, and *religion*. The keywords from the source domains of interest were first manually searched in a sample of speeches and then throughout the corpus.

The keyword search in this thesis was performed directly on the set of metaphors identified by the DeepMet model (Su et al., 2020) using an ad-hoc filter on the filtered metaphor set converted to a Pandas (McKinney et al., 2010) dataframe. The domain keywords were chosen a priori in lemma form to represent the seven selected source domains as accurately as possible. Table B1 in Appendix B shows in detail the keywords associated with each source domain.

Following the first step of Charteris-Black's (2004) Critical Metaphor Analysis approach, the prevalence of the different source domains in the corpus was compared by way of measuring their *resonance*, that is, by multiplying the sum of tokens and the sum of types of metaphors from the same source domain and comparing the result to that of the other domains. Charteris-Black defines resonance as "an indication of the extent to which metaphor source domains are found in a particular corpus and therefore is a measure of their productivity" (2004, p. 89).

In addition to computing the measure of resonance of the source domains in the APS corpus, further exploratory data analysis was performed to gain insights into the statistical association between the source domains and corpus metadata variables. In particular, the incidence of the source domains was firstly investigated in the two subsets in relation to the President's political party affiliation (i.e. Democratic or Republican). In order to present an overview of the types of metaphors used by the two political parties, a preliminary qualitative analysis based on Charteris-Black's (2004) findings was also conducted on the most relevant results of the quantitative analysis in the Campaign Speeches subset. In particular, following the second step of Charteris-Black's (2004) Critical Metaphor Analysis approach, relevant metaphors from the selected source domains were associated with their underlying conceptual bases. For ease of reference, the list below illustrates a summary of the set of source domains and related conceptual metaphors used in American presidential speeches reported by Charteris-Black (2004):

- ***Body parts***: Charteris-Black (2004) considers metaphors drawn from the source domain of the human body as combinations of metaphor and metonymy based on familiar relations between certain body parts and specific activities. These are metaphors such as *heart* in "We've touched the *heart* of the city" or *hands* in "The future is in your *hands"*
- ***Buildings***: the metaphors drawn from the source domain of buildings, similarly to journey metaphors, are employed to conceptualize aspirations toward political and social objectives with a positive connotation. The conceptual metaphors identified by Charteris-Black (2004) for this domain are: WORTHWHILE ACTIVITY IS BUILDING, as in "We will *build* a better future for ourselves" and SOCIETY IS A BUILDING, as in "Justice is one of the *pillars* of our society"
- ***Conflict***: according to Charteris-Black's (2004) analysis, metaphors drawn from the source domain of conflict are mostly related to the conceptual metaphor of POLITICS IS CONFLICT, which is employed to highlight the struggles and sacrifices that speakers describe as necessary to achieve abstract social goals, such

as freedom and rights, and to solve social issues, such as poverty and disease. Instances of *conflict* metaphors are found in sentences such as “We will *fight* the *war* against poverty and injustice”. The conflict metaphors analyzed by Charteris-Black tend to follow a specific rhetorical pattern that consists in the identification of an enemy, the assembling of allies, and a military struggle against the enemy that leads to victory and subsequent punishment of the enemy.

- ***Fire and light***: Charteris-Black (2004) combined the two source domains of fire and light because he found that they are employed similarly in the corpus he analyzed. In particular, both source domains are used to express positive meanings with conceptual metaphors such as SEEING IS UNDERSTANDING and HOPE IS LIGHT for the source domain of light, as in “We need to *shine a light* on this situation” and “The *light* of freedom will *shine* across the country”, and PURIFICATION IS FIRE for the source domain of fire, as in “Mill *fires were lighted at the funeral pile* of slavery” (Charteris-Black, 2004, p. 102)

- ***Journey***: metaphors drawn from the source domain of journey are employed to conceptualize political objectives as traveling destinations. Charteris-Black (2004) associates the progress towards a political or social goal is associated with a progress on a predetermined path, guided by a leader and is represented by the conceptual metaphor: PURPOSEFUL SOCIAL ACTIVITY IS TRAVELING ALONG A PATH TOWARDS A DESTINATION, as in “We’re *heading towards* a better future”

- ***Physical environment***: Charteris-Black (2004) grouped the two source subdomains of weather and natural geographical features under the same source domain of physical environment. The conceptual metaphor identified by Charteris-Black (2004) for weather domain is A SOCIAL CONDITION IS A WEATHER CONDITION, which is related to the positive or negative changes in social conditions conceptualized as changes in weather conditions as in “It’s been a *stormy* year for our country”, while metaphors related to the features of the landscape are associated to the conceptual metaphor of A SOCIAL CONDITION IS A FEATURE IN A LANDSCAPE, as in “There is hope on the *horizon* for this matter”.

- ***Religion***: the metaphors drawn from the source domain of religion are employed to conceptualize social and political objectives as spiritual aspirations. Charteris-Black (2004) associates this source domain of religion with the conceptual metaphor POLITICS IS RELIGION, which is expressed with linguistic metaphors such as “Our *mission* is *sacred*” and “I have a *vision* for this country”.

Source domain incidence was finally investigated chronologically for the Campaign Speeches subset to determine whether there are any significant changes in source domain distribution through time in this particular genre of speech. This analysis was performed by computing the relative frequency of each source domain for each year one or more campaign speeches were delivered.

3. Analysis

3.1 Metaphor identification statistics

Overall, a total amount of 452,393 tokens in the filtered APS corpus were labeled as metaphors by the DeepMet model (Su et al., 2020). The resulting metaphor vocabulary is composed of 7801 lemma types. About 17% of the tokens in the APS corpus are identified as metaphors. Table 3.1 below shows the distributions of metaphor tokens and lemma types in the Campaign Speeches and Presidential Speeches subsets of the APS corpus.

Table 3.1

Distribution of metaphors in the Campaign and Presidential speeches subsets

	# metaphor tokens	# metaphor types	% metaphors
Campaign Speeches	15884	1933	17.0%
Presidential Speeches	436509	7651	16.8%

Note. The percentage of metaphors is the ratio of tokens identified as metaphors to the total number of tokens in the two subsets. Tokens tagged as punctuation, determiners, adpositions, and conjunctions were filtered as they are less likely to trigger metaphors.

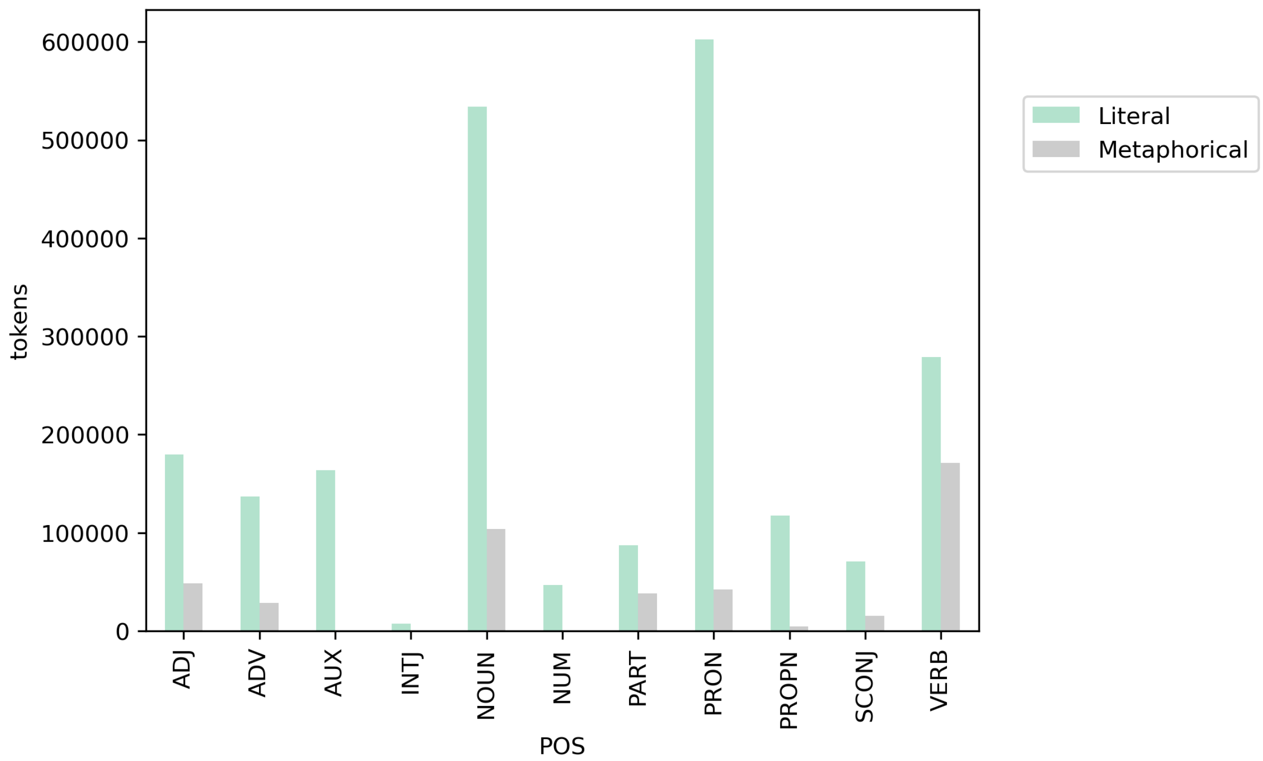
Figure 3.1 illustrates the distribution of the tokens labeled as literal and metaphorical according to their part-of-speech tags. The most likely POS tags of tokens that trigger metaphor in the corpus are content words tags, that is, tokens tagged as verbs (VERB), nouns (NOUN) and adjectives (ADJ).

The tokens tagged as metaphorical whose POS is other than verb, noun, adjective, and adverb are considered to be part of multi-word metaphorical expressions or model labeling errors. In order to confirm this, the tokens tagged as metaphorical are divided into two subsets: one composed of metaphorical unigrams, which consists of isolated tokens labeled as metaphors, and the other composed of clusters of consecutive tokens tagged as metaphorical, i.e. multi-word metaphorical expressions.

A total of 292,825 isolated tokens were labeled as metaphorical and, according to their POS tag, 19% of these unigrams are functions words such as pronouns (PRON), particles (PART), conjunctions (SCONJ) and auxiliaries (AUX). Since these function words

Figure 3.1

Distribution of literal and metaphorical tokens according to POS in the APS corpus



are isolated, they cannot be considered part of multi-word metaphorical expressions and were thus considered labeling errors and filtered from the results.

A total of 292,825 isolated tokens were labeled as metaphorical and, according to their POS tag, 19% of these unigrams are functions words such as pronouns (PRON), particles (PART), conjunctions (SCONJ) and auxiliaries (AUX). Since these function words are isolated, they cannot be considered part of multi-word metaphorical expressions and were thus considered labeling errors and filtered from the results.

The remaining metaphorical n-grams are analyzed on the basis of their POS tag combinations and were filtered accordingly. In particular, 19,846 out of 56,330 bigrams (i.e., two consecutive tokens) labeled as metaphorical are isolated and then joined as single strings to be counted for the summary statistics analysis along with the unigram counts. The bigrams selected are composed of phrasal verbs (e.g. *go forward*, *take place*), compound nouns (*front line*, *safety net*), adjective-noun combinations (*high standard*, *free world*), which can be considered as Adjective-Noun (Type-III) metaphors, and adverb-verb combinations (*strongly support*, *deeply move*), which can be considered as

Adverb-Verb (Type IV) metaphors. The remaining non-relevant bigram components are considered individually as unigrams and are filtered if they are function words. Trigrams are addressed similarly: the trigrams selected are composed of phrasal verbs (*look forward to, stand-up-to*) and multi-word expressions with different POS combinations such as *beacon of light and take concrete steps*. As with non-relevant bigrams, the remaining non-relevant trigrams and tetragrams components are considered individually as unigrams and are filtered if function words.

A total of 305,429 metaphorical tokens composed of unigrams, bigrams and trigrams are considered in the statistical analysis. Table 3.2 presents the summary statistics for these tokens in each presidential campaign and presidential term in chronological order.

Focusing on the individual presidential campaigns, there is a low degree of variability among the metaphor relative frequencies: the highest percentage of metaphors is found in George W. Bush's first campaign period at 14% and the lowest in Nixon's first campaign period at 9%.

There is a low degree of variability among the metaphor relative frequencies in the individual presidential terms as well: the highest percentage of metaphors is found in both Eisenhower's terms and in Obama's first presidential term at 12.4% and the lowest is found in Joe Biden's first term at about 10%.

Table 3.2

Relative frequency of metaphors in the Campaign and Presidential Speeches subsets

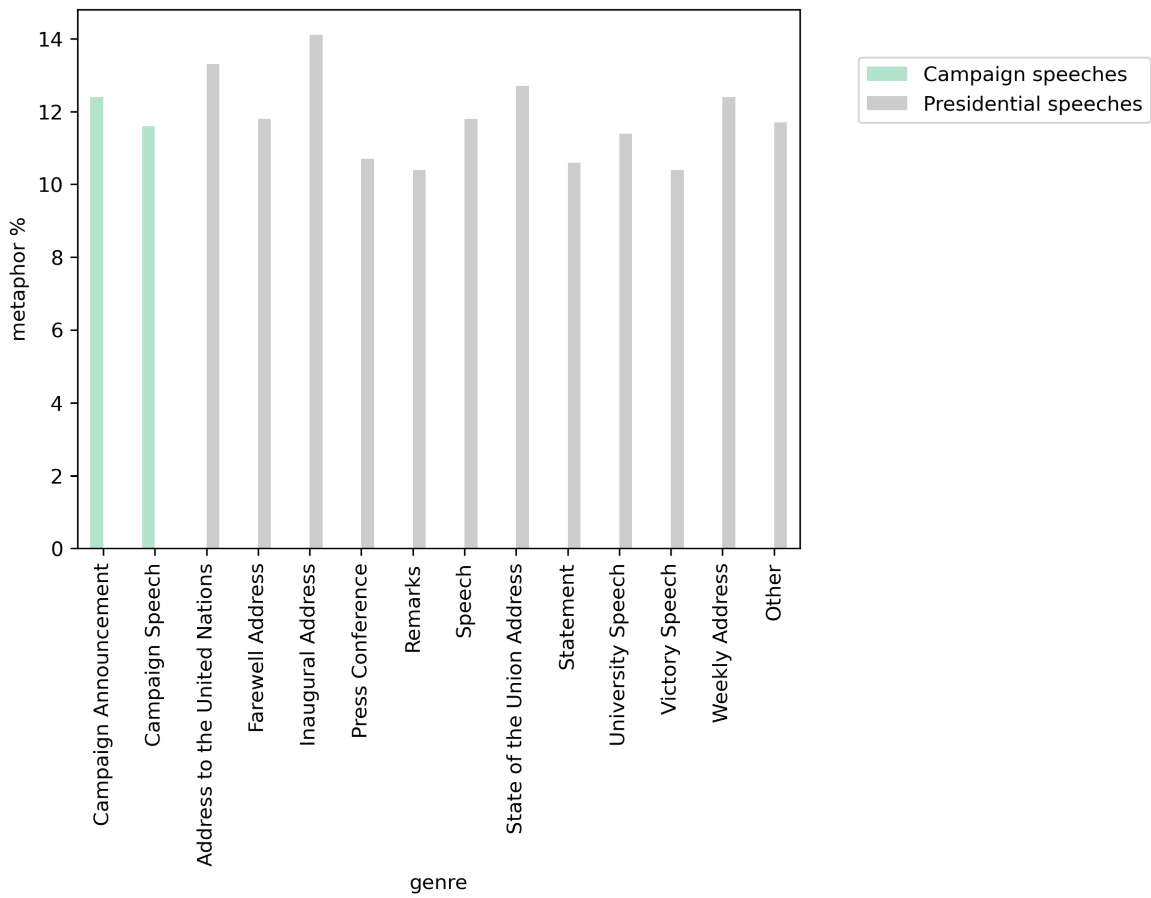
	term	% camp met	% term met
Dwight Eisenhower	1953-1957 1957-1961	13.5% 10.7%	12.4% 12.4%
John F. Kennedy	1961-1963	11.5%	12.2%
Lyndon B. Johnson	1963-1965 1965-1969	11.8% 10.0%	10.0% 10.3%
Richard Nixon	1969-1973 1973-1974	9.0% 10.4%	11.7% 10.3%
Gerald R. Ford	1974-1977	12.1%	11.0%
Jimmy Carter	1977-1981	10.7%	11.3%
Ronald Reagan	1981-1985 1985-1989	13.6% 13.5%	11.9% 11.8%
George H. W. Bush	1989-1993	12.2%	12.3%
Bill Clinton	1993-1996 1996-2001	11.6% --	11.1% 11.4%
George W. Bush	2001-2004 2004-2009	14.1% 13.2%	11.8% 11.7%
Barack Obama	2009-2012 2013-2017	11.1% 12.1%	12.4% 12.0%
Donald Trump	2017-2021	11.9%	10.3%
Joe Biden	2021-2022	10.7%	10.0%

Note. The values for Bill Clinton's second campaign are nil since the only speeches he delivered in the original APS corpus consisted of two debates, which, as already explained, were filtered for the metaphor identification task and were not considered in the subsequent analysis.

Figure 3.2 shows the summary statistics for the content words labeled as metaphorical in each genre of speech in the Campaign and Presidential Speeches subsets. The highest percentages of metaphors are found in the Inaugural Addresses at 14.1% and in the Addresses to the United Nations at 13.3%, while the lowest percentages of metaphor use are found in the Remarks and Victory Speeches, both at 10.4%.

Figure 3.2

Relative frequency of metaphors in Campaign and Presidential subsets genre types

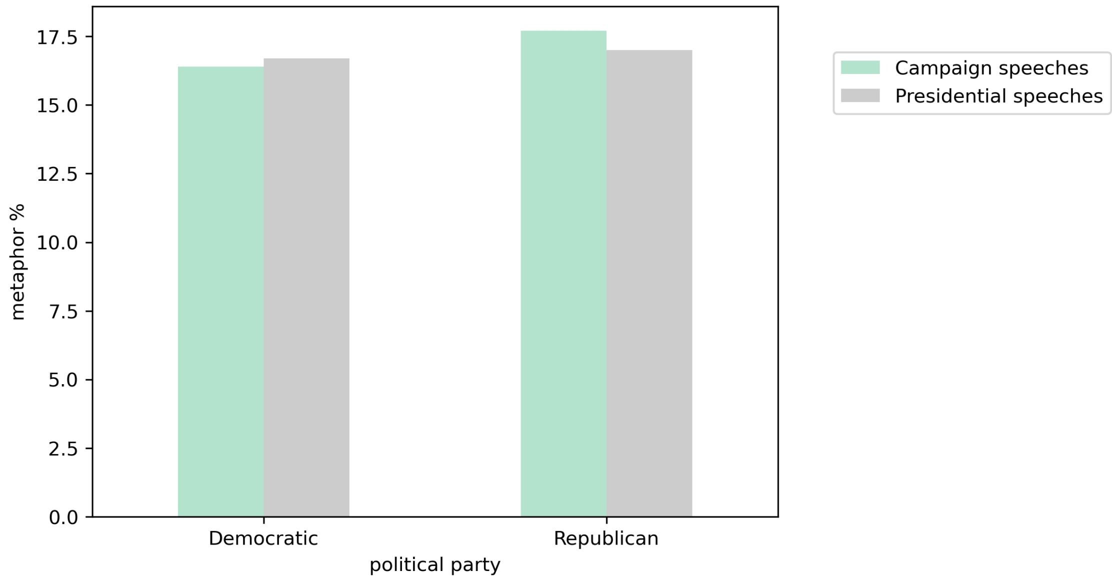


Note. The percentage of content words labeled as metaphorical in the individual speech genres is computed as the ratio of metaphors to content words per genre.

Regarding the distribution of metaphor based on political affiliation in the two subsets, as illustrated in Figure 3.3, the percentages of metaphor use are similar: ranging from 16.4% to 17.7%, metaphor frequencies are slightly higher in the Republican campaign subset. Moreover, metaphor use in both Campaign and Presidential Speeches is minimally higher in speeches delivered by Republican presidents. This is expected as Republican speeches are marginally higher in number than Democratic speeches.

Figure 3.3

Relative frequency of metaphors in Campaign and Presidential subsets party affiliation



3.2 Source domain analysis

The aim of the analysis described in this section is to gain insights into the type of metaphors that were identified by the DeepMet model (Su et al., 2020) in the APS corpus.

A global analysis of the results of the identification task shows that the vast majority of the tokens identified as metaphors by the model are highly conventional and dead metaphors. For instance, the verbs labeled as metaphorical with the highest frequencies are verbs such as *have*, *make*, *take* and *get*, and similarly, the metaphorical nouns with the highest frequencies are nouns such as *thing*, *system*, *plan* and *part*, while metaphorical adjectives with the highest frequencies are adjectives such as *great* and *high*. Polysemous words such as *face*, *pass* and *stand* are also labeled as metaphorical in high frequencies. These results were in part expected as the DeepMet model (Su et al., 2020) employed for metaphor identification in the APS corpus was trained on the VUA Metaphor Corpus (Steen et al. 2010) and, as mentioned in subsection 2.2.1, a large majority of the tokens annotated as metaphorical in the VUA Corpus are in fact highly conventional metaphors, dead metaphors, or borderline cases.

In order to narrow down tokens identified as metaphors that are more relevant for an accurate analysis of metaphor usage in American presidential speeches, a search is conducted for keywords from a pre-selected set of source domains. The source domains selected are those drawn up by Charteris-Black (2004) for his corpus-based research on metaphor in American presidential speeches and are the source domains of: *body parts*, *buildings*, *conflict*, *fire and light*, *journey*, *physical environment*, and *religion*. This set of source domains is considered exhaustive for the APS corpus after a comprehensive manual review of the tokens automatically identified as metaphorical.

It should be noted that, qualitatively speaking, the vast majority of the metaphors categorized by source domain are expected to be conventional metaphors, as are the metaphors analyzed by Charteris-Black (2004).

Overall, a total of 22,720 tokens and 374 lemma types identified as metaphors by the DeepMet model (Su et al., 2020) are keywords from the source domains manually identified by Charters-Black (2004). The metaphors drawn from the pre-selected source domains account for about 7% of the total of tokens identified as metaphors by the DeepMet model (Su et al., 2020).

3.2.1 Source domain resonance

The first step of the analysis of the identified metaphors categorized as keywords from the pre-selected source domains consists in measuring and comparing the incidence of the different source domains in the APS corpus. This is performed using Charteris-Black's (2004) proposed measure of *resonance*, which consists in multiplying the sum of tokens and the sum of types of the metaphors from the same source domain and comparing the result to that of the other domains.

Figure 3.4

Source domain resonance in the APS corpus and in Charteris-Black's US Inaugural corpus

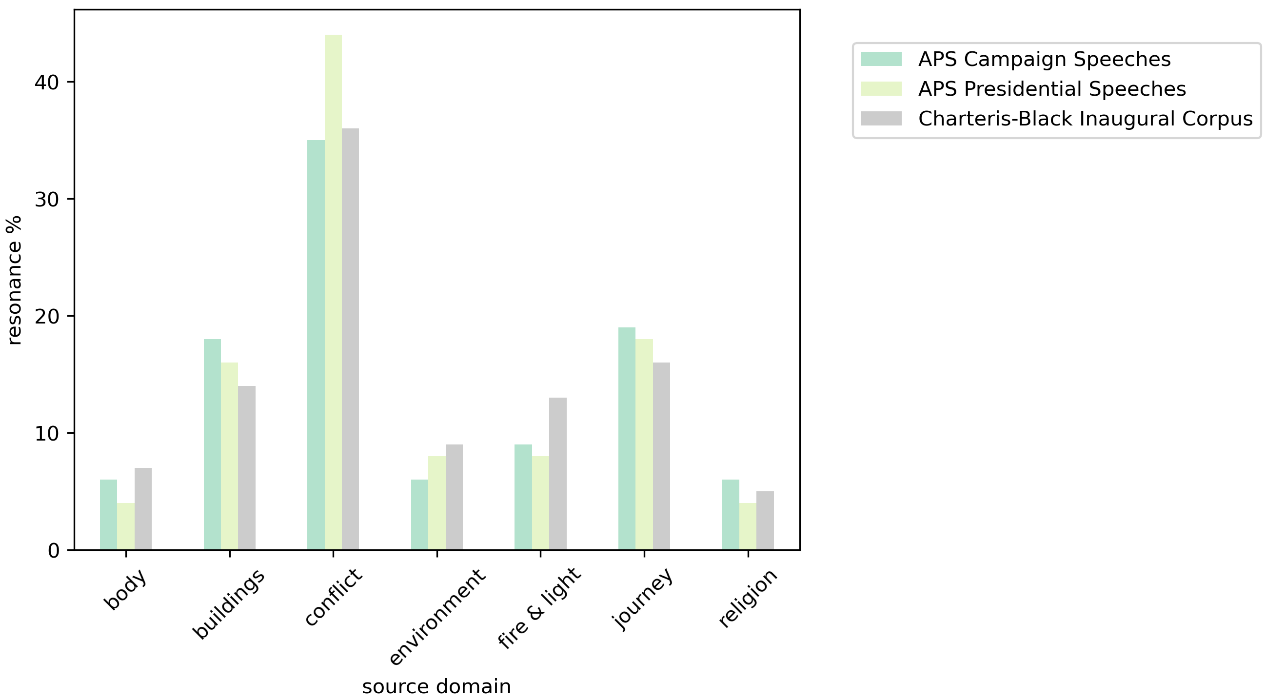


Figure 3.4 displays the resonance of the source domains in the two subsets of the APS corpus in comparison to Charteris Black's data. It should be reiterated that Charteris-Black's (2004) data refers to a corpus of 51 inaugural presidential speeches from George Washington to Bill Clinton, as such, it refers to a different time period and to a single genre of presidential speech compared to the APS corpus. However, the plot shows that the results of both analyses seem to converge to similar source domains measures of resonance. As can be seen in Figure 3.1, the three most resonant domains in both subsets

of the APS corpus and in Charteris-Black's (2004) Inaugural Corpus are those of *conflict*, *journey,* and *buildings*. Furthermore, metaphors drawn from the source domain of *conflict* are more than twice as resonant as any other group of metaphors, in both the APS corpus subsets and the Inaugural Corpus. The two least resonant source domains in the APS subsets, *religion,* and *body parts*, are also similarly resonant in Charteris-Black's (2004) corpus. Metaphors of *fire and light* are marginally more resonant in the Inaugural Corpus.

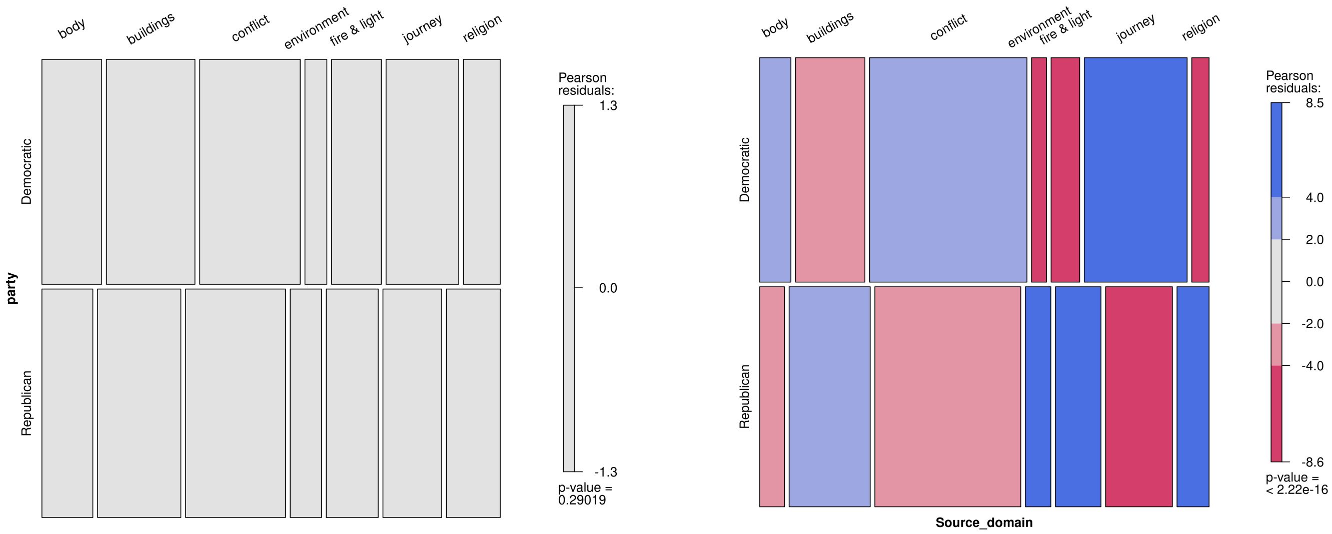
For a more detailed overview of the source domains distributions and resonance measures in the APS corpus and in Charteris-Black's Inaugural Corpus, refer to table C1 and C2 in Appendix C.

3.2.2 Source domains in Democratic and Republican speeches

The second step of the quantitative analysis of the source domains consists in exploring the statistical association between the variables of political party and source domain in the Campaign Speeches and Presidential Speeches subsets.

Figure 3.5

Source domain distribution and resonance in the Campaign Speeches subset



Note. Overall frequency distribution of the selected source domains for each political party (left) in the Campaign Speeches subset; source domain resonance distribution for each political party (right) in the Campaign Speeches subset.

The mosaic plots in Figure 3.5 are created using the R statistical computing environment[11] and the vcd package (Meyer et al., 2006). In these plots, the widths of the cells in a row are proportional to the frequency distributions (left plot) and resonance measures (right plot) of each source domain in the Campaign Speeches delivered by Democrat and Republican Presidents. The height of the first row of cells is proportional to the overall frequency (left plot) and resonance (right plot) of the seven source domains in

[11] http://www.r-project.org

the speeches delivered by Democrats as compared to the speeches delivered by Republicans, displayed in the second row.

In total, there is about the same metaphor frequency in both Democratic and Republican campaign speeches. The total measure of resonance is also equal for the two political parties.

The color shades in the mosaic plot indicate the significance of the differences between the cells in a column. The degrees of the differences are calculated by comparing the relative frequencies of the metaphors or the relative resonance of a specific source domain between the two parties using a Pearson residual test (Meyer et al., 2006). Blue shading represents positive values and red shading represents negative values. The stronger the shade, the higher the absolute residual value and the more significant the deviation from the cross-party distribution.

Whereas the single source domains in Democratic and Republican campaign speeches have similar distributions (left plot), the distribution based on source domain resonance (right plot) presents some significant differences. For instance, while the proportion of *fire and light* metaphors is nearly identical among the two political parties, the resonance of the source domain of *fire and light* is higher in Republican speeches, meaning that Republican *fire and light* metaphors are more varied vocabulary-wise, since resonance is computed as the product of types and tokens. Furthermore, the resonance measures of the source domains of *religion* and *environment* in Republican campaign speeches are higher than expected, as is the source domain of *buildings* to a smaller degree. Conversely, the proportion of resonance of the source domain of *journey* is higher than expected in Democratic campaign speeches, as are the source domains of *conflict* and *body parts*, to a smaller degree.

From a preliminary qualitative analysis of the most resonant source domains, it is shown that not only do these domains have similar distributions in the two parties, but they are also employed similarly in Democratic and Republican campaign speeches. Keywords such as *fight*, *battle*, *protect* and *defend* are the most frequently used conflict metaphors in both parties and to convey the struggles and sacrifices deemed necessary to achieve abstract social goals, as illustrated in the speech extracts in Text 3.1 and Text 3.2.

Text 3.1

I have developed an image of America as fulfilling a noble and historic role as the <u>defender</u> of freedom in a time of maximum peril and of the American people as confident, courageous and persevering. (Kennedy, 1960, Democratic).

Text 3.2

The alternative to bureaucracy is not indifference. It is to put conservative values and conservative ideas into the thick of the fight for justice and opportunity. (Bush, 2000, Republican)

In Republican speeches, *conflict* metaphors are also employed, to a smaller degree, to describe the process of the presidential campaign in terms of conflict, as illustrated in Text 3.3.

Text 3.3

And I know that you will all fight even harder for the great victory our party is going to win in November because we're going to be together in that election campaign. (Nixon, 1968, Republican)

The most frequent keywords from the source domain of *journey* are also similar in the two parties: nouns such as *path* and *steps* are employed to conceptualize the political objectives presented in the campaigns as traveling destinations, as illustrated in Text 3.4 and 3.5.

Text 3.4

We're offering a better path – a future where we keep investing in wind and solar and clean coal [...] the path we offer may be harder, but it leads to a better place. (Obama, 2012, Democratic)

Text 3.5

Each step towards real unification of Europe is a major victory to the free world. (Eisenhower, 1952, Republican)

Regarding the source domain of *building*, the most frequent keywords identified as metaphorical in both parties are words such as *build*, *restore* and *foundation*, which, like *journey* metaphors, are also generally employed to conceptualize an aspiration toward a political or social objective presented in the campaign speech. Instances of this are illustrated in Text 3.6 and 3.7.

Text 3.6

Tonight I want to talk with you about my hope for the future, my faith in the American people and my vision of the kind of country we can build together. (Clinton, 1992, Democratic)

Text 3.7

We're here to lift the weak and to build the peace, and most important, we're here, as Dr. Warren said, to act today for the happiness and liberty of millions yet unborn, to seize the future so that every new child of this beloved Republic can dream heroic dreams. (Reagan, 1984, Republican).

In Democratic speeches, instances of *building* metaphors are also found for describing the process of the presidential campaign in terms of building a structure, as in Text 3.8.

Text 3.8

And that kind of campaign takes time to build. So even though I'm focused on the job you elected me to do, and the race may not reach full speed for a year or more, the work of laying the foundation for our campaign must start today. (Obama, 2012, Democratic)

Finally, it should be noted that from the preliminary qualitative analysis of the source domain keywords that were identified as metaphors, a non-insignificant number of labeling errors are detected. For instance, the verb *attack* in Text 3.9 was labeled as metaphorical by the DeepMet model (Su et al., 2020), while in the context of the sentence it is used literally.

Text 3.9

And I think about the young sailor I met at Walter Reed hospital, still recovering from a grenade attack that would cause him to have his leg amputated above the knee (Obama, 2012)

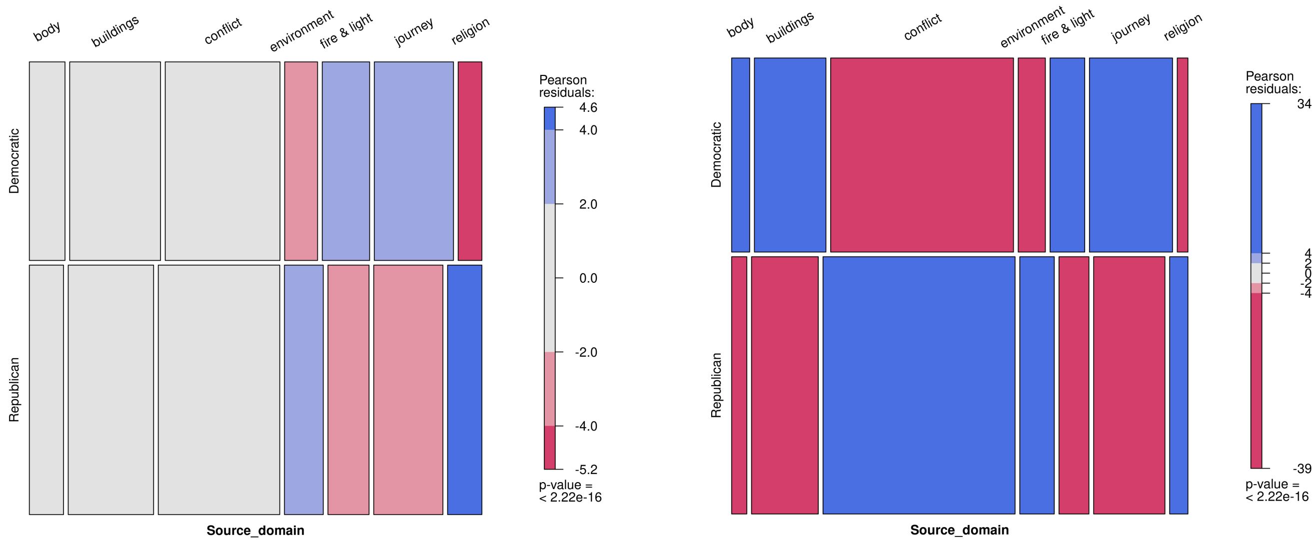
From the preliminary qualitative analysis a significant number of borderline cases are also detected. Cases such as the one reported in Text 3.10 would not be relevant to a more in depth political discourse analysis.

Text 3.10

In the same way, we have worked unceasingly for the promotion of effective steps in disarmament so that the labor of men could with confidence be devoted to their own improvement rather than wasted in the building of engines of destruction. (Eisenhower, 1956)

Figure 3.6

Source domain distribution according to party in the Presidential Speeches subset



Note. Overall token distribution of the selected source domains for each political party (left) in the Presidential Speeches subset; source domain resonance distribution for each political party (right) in the Presidential Speeches subset.

Moving on to consider the Presidential Speeches subset, Figure 3.6 displays on the left the frequency distribution of the selected source domains in the two parties, and on the right the source domain resonance distribution. Overall, there is a marginally greater number of metaphors from all the selected source domains in Republican Speeches, which is expected as there are more Republican presidential speeches in the APS corpus (53% to the Democratic speeches 47%). Similarly, the total measure of resonance is also higher for Republican speeches.

Focusing on the single source domain distributions, while the source domains of *conflict*, *buildings*, and *body parts* have all similar distributions in the two parties, the frequency of *religion* metaphors is significantly higher in Republican Speeches, as is the source domain of *environment* to a smaller degree.
Conversely, the source domains of *fire and light* and *journey* are marginally higher than expected in Democratic speeches. As in the Campaign Speeches subset, the resonance distribution in the Presidential Speeches subset presents more significant differences. The

source domain of *journey*, *fire and light*, *buildings* and *body parts* are significantly more resonant than expected in Democratic speeches. Conversely, the remaining source domains of *conflict*, *environment* and *religion* are significantly more resonant than expected in Republican speeches.

A global qualitative analysis shows that metaphors from the most resonant source domains (*conflict*, *journey* and *buildings*) are employed similarly to those reported for the Campaign Speeches. As such, the analysis prioritized the other source domains with significant differences in the two parties: the source domains of *religion* and *fire and light*.

In particular, metaphors drawn from the source domain of *religion* are mostly employed to conceptualize social and political objectives as spiritual aspirations in both parties, however, as already mentioned, they are more frequent in Republican speeches. Text 3.11 and 3.12 illustrate two examples of *religion* metaphors identified in the Presidential Speeches subset.

Text 3.11

There was a hunger in the land for a spiritual revival; if you will, a crusade for renewal. (1984, Reagan, Republican)

Text 3.12

My fellow leaders, this is a moment where we must prove ourselves the equals of those who have come before us, who with vision and values and determined faith in our collective future built our United Nations, broke the cycle of war and destruction, and laid the foundations for more than seven decades of relative peace and growing global prosperity. (2021, Biden, Democratic)

Regarding the source domain of *fire and light*, while *fire and light* metaphors are more resonant in Democratic speeches, the most frequent keywords from this source domain are similar in the two parties: words such as *light*, *bright*, *shine* and *fire* are found in high frequencies, together with their counterparts *dark* and *shadow*, in lower frequencies. A global analysis shows that this type of metaphors are employed to convey positive meanings, often related to hope, as illustrated in Text 3.13 and 3.14.

Text 3.13

Help us to open wide the doors of opportunity and invite all to come in, for when we have done this, it will one day be said of America that she was a burning and shining light in man's journey on earth. (Johnson, 1964, Democratic)

Text 3.14

Let us pledge together to make these next four years the best four years in America's history, so that on its 200 th birthday America will be as young and as vital as when it began, and as <u>bright</u> a <u>beacon</u> of hope for all the world. (Nixon, 1973, Republican)

Moreover, as was shown in the Campaign Speeches subset analysis, a non-insignificant number of labeling errors are detected in the Presidential Speeches subset as well, together with borderline cases and large frequencies of highly conventional metaphors.

In summary, in both the Campaign and Presidential Speeches subsets, the distribution of the seven pre-selected source domains is similar for Democratic and Republican speeches. The most common source domains are those of *conflict*, *journey* and *building*. More significant differences are found in the source domain resonance in the two parties, this is related to the use of more varied vocabularies for the relevant source domains. Qualitative analyses highlighted similar usage of metaphor in the two parties and confirmed the significant presence of highly conventional metaphors, borderline cases, and labeling errors.

Finally, it should be reiterated that the qualitative analyses presented for the Campaign and Presidential Speeches subsets were not in depth analyses, but rather preliminary analyses used to expound on the quantitative analyses. Therefore, the findings should be interpreted with caution and further research is needed to explore the underlying conceptual bases of the keywords that were automatically identified as metaphors.

3.2.3 Diachronic source domain distribution in campaign speeches

Source domain incidence was subsequently investigated chronologically for the Campaign Speeches subset, to determine whether there are any significant changes in source domain distribution through time in this particular genre of political speech.

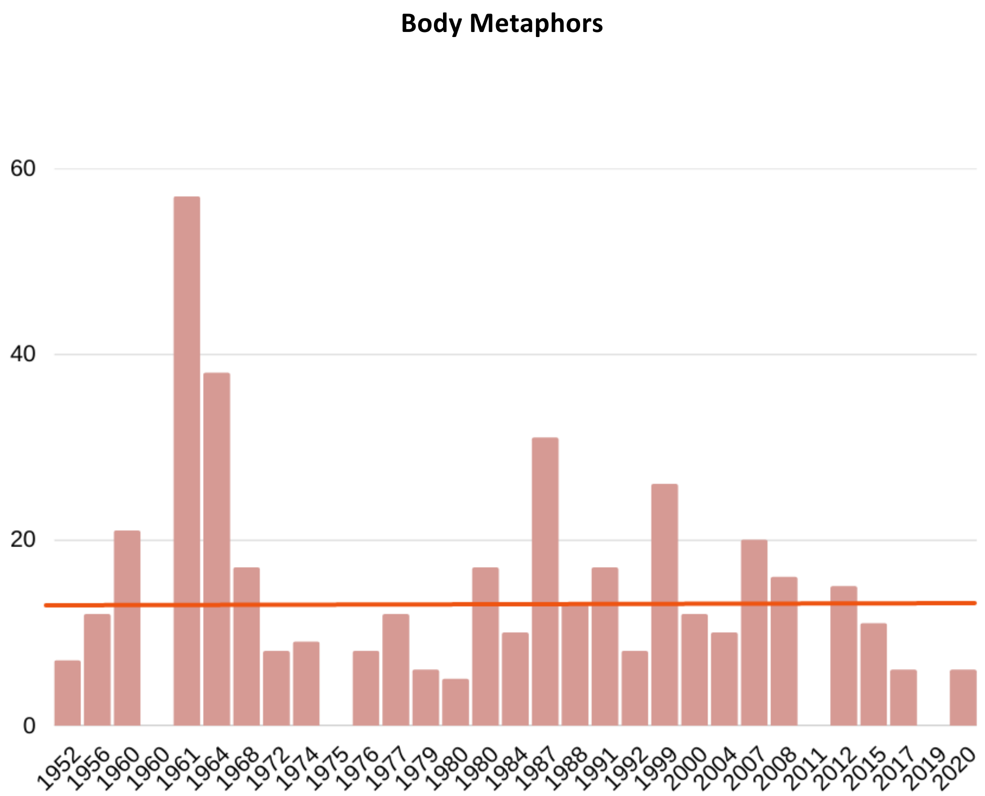
In Figure 3.7, seven plots display the relative frequency of the seven source domains with respect to the year of the campaign speech. A global analysis shows that the frequency of *conflict* and *buildings* metaphors is on average higher than any other source domain frequency. The other source domains, on the other hand, only have isolated peaks relative to one or a few years.

Upon closer examination, the highest peaks are associated with the use of fewer source domains per year, which corresponds in most cases to a lower absolute number of metaphors used in the relevant year. For instance, in Kennedy's 1961 campaign speech, only seven tokens were identified as metaphors, four of which are keywords from the source domain of *body parts*, while the other three are divided between the source domains of *buildings*, *environment* and *journey*. Similarly, in Nixon's 1968 campaign announcement six tokens were identified as metaphors, three of them being keywords from the source domain of *religion*, while the others are divided between the source domains of *body*, *conflict* and *environment*. The highest peaks for the *journey* and *religion* source domains correspond to Ford's 1975 campaign speech, where only two tokens were identified as metaphors, one from the source domain of *journey* ("I have found these leaders in Bo Callaway of Georgia [...] and many others from every State and from every <u>walk</u> of life who have volunteered to help") and one from the source domain of *religion* ("I want every delegate and every vote that I can get that can be won to my cause within the <u>spirit</u> and the letter of the law and without compromising the principles for which I have stood all of my political and public life"). The highest peak in the *environment* source domain corresponds to Carter's 1977 campaign speech, which presents eight metaphors, three of which are *environment* metaphors. The only peak in the source domain of *buildings* corresponds to Obama's 2011 campaign announcement speech, which presents metaphors from the only two source domains of *buildings* (67%) and *conflict* (33%). Conversely, the only peak in the source domain of *conflict* corresponds to Trump's 2015 campaign announcement which presents five metaphors of *conflict* out of nine metaphors in total.

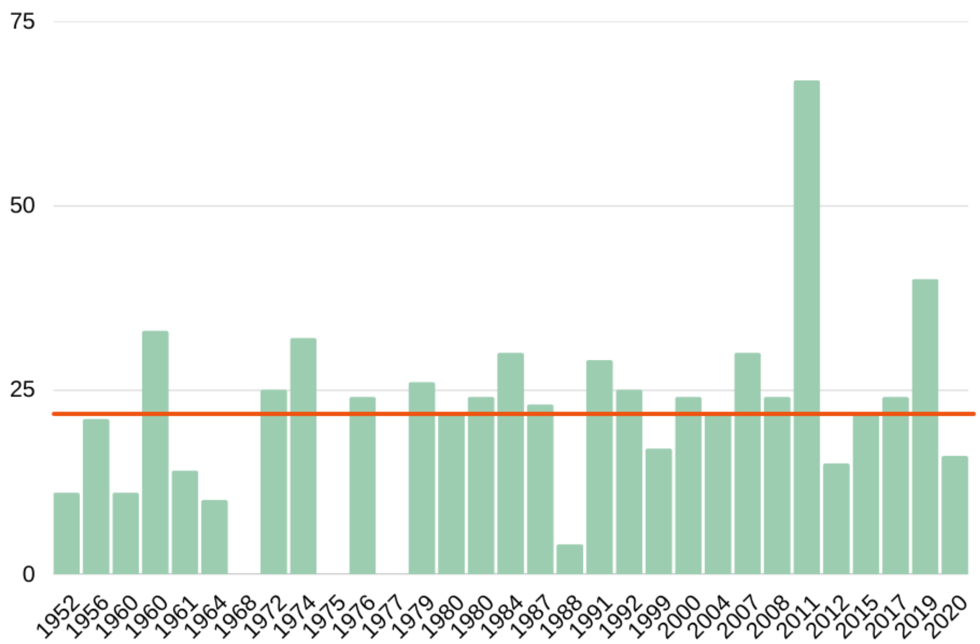
In summary, source domain incidence is mostly constant over time, with the exception of less common source domains that spike in certain years. These spikes coincide with fewer source domains and fewer metaphors overall in those years.

Figure 3.7

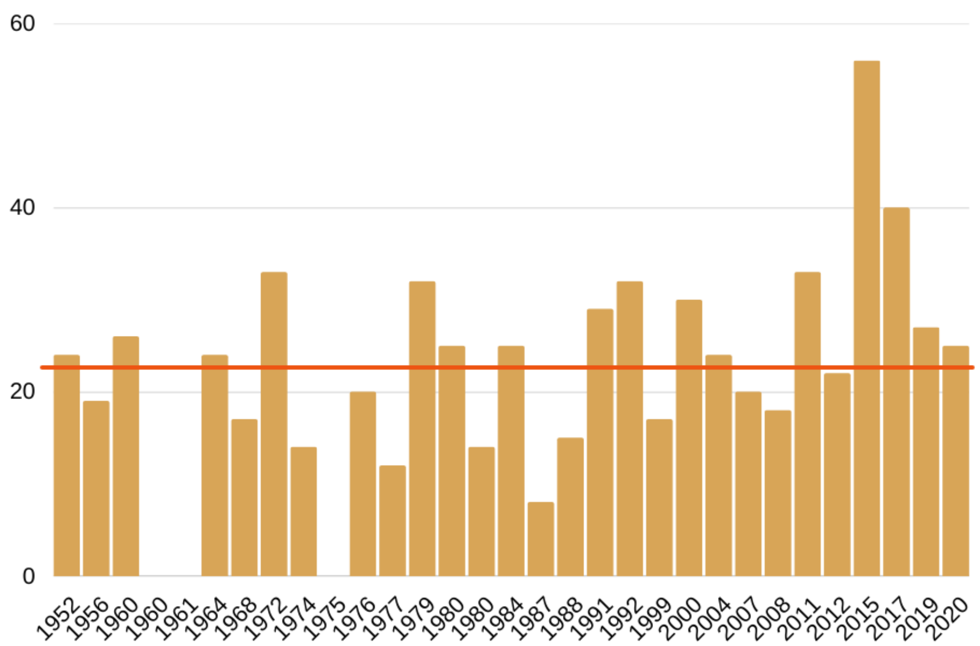
Diachronic source domain distribution in the Campaign Speeches subset



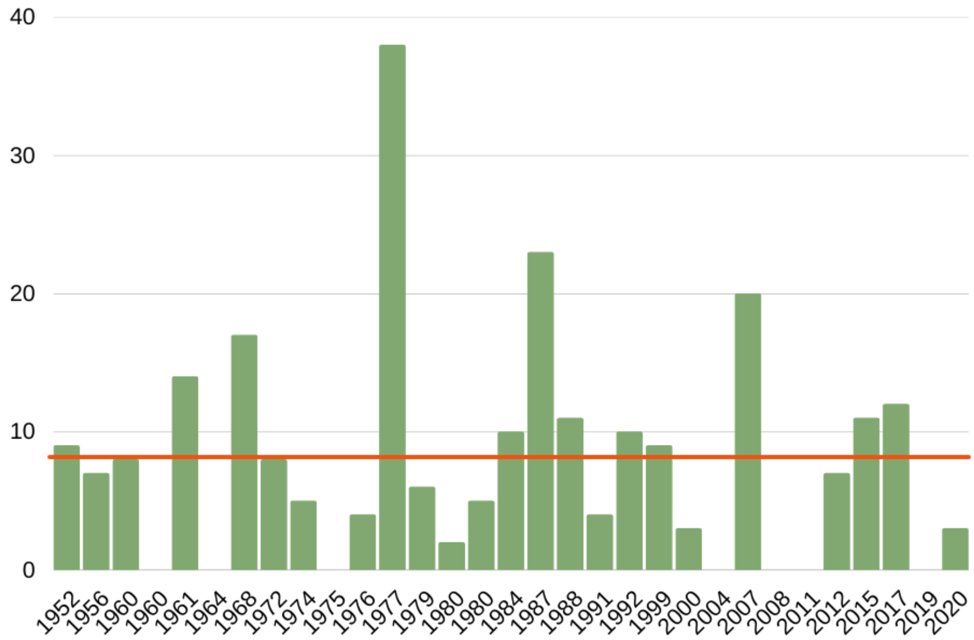
Building Metaphors



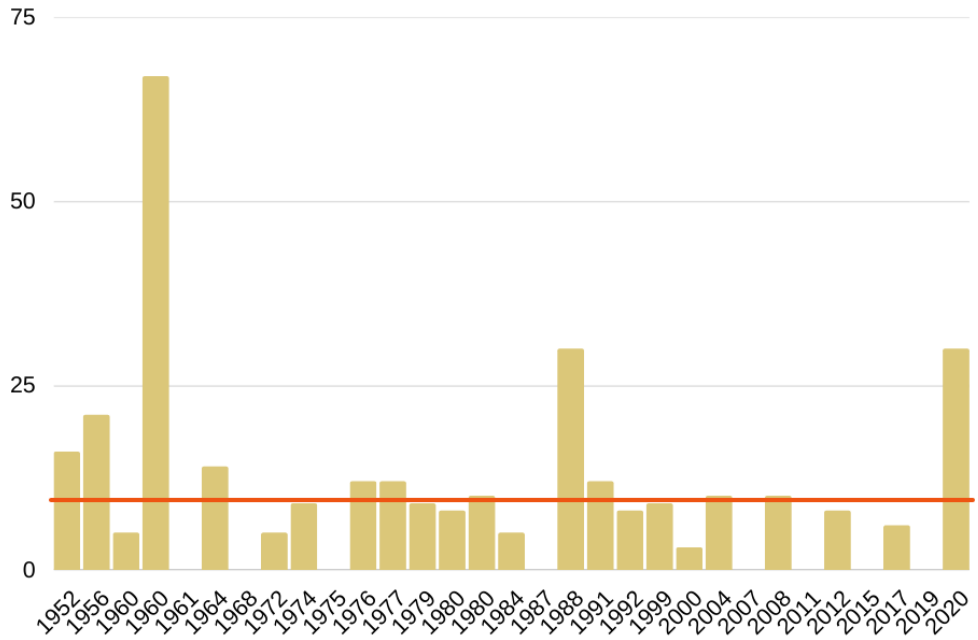
Conflict Metaphors



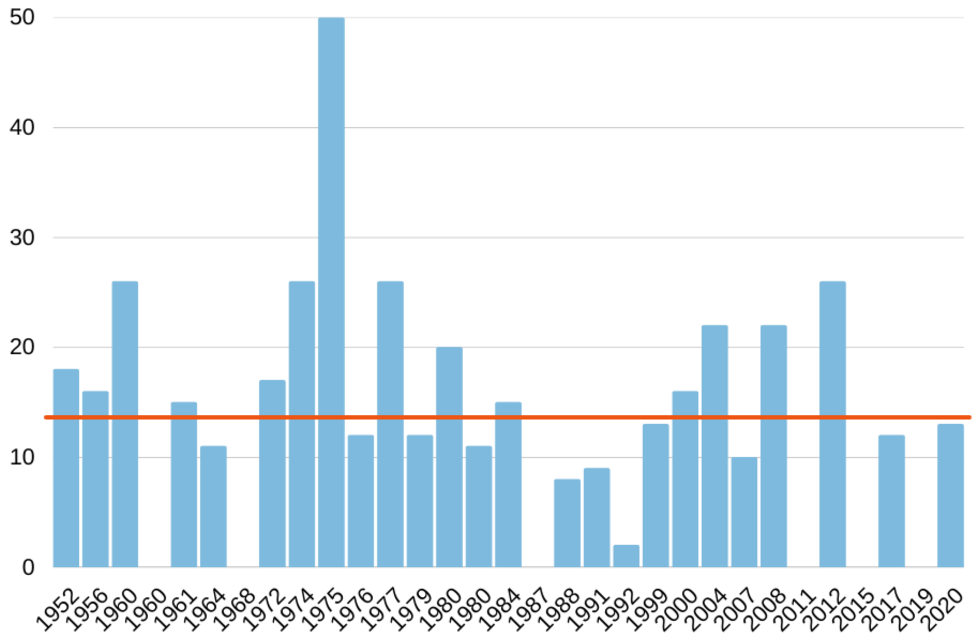
Environment Metaphors



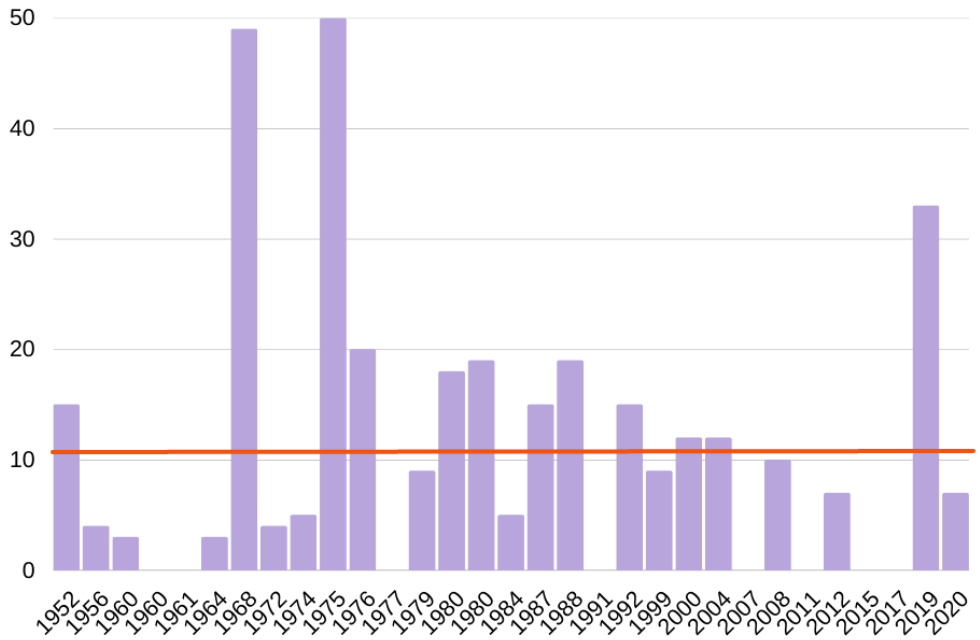
Fire and Light Metaphors



Journey Metaphors



Religion Metaphors



3.3 Discussion

The first aim of this thesis was to overcome the challenge of metaphor extraction in large corpora of political speeches by employing a computational model for automatic metaphor identification. The American Presidential Speeches corpus was constructed to be analyzed for metaphors and the DeepMet model (Su et al., 2020) was selected for the task of metaphor identification. Using an NLP-based metaphor identification approach allowed to perform a large-scale explorative study on metaphor use in a large number of American presidential speeches, which had not been possible in previous corpus-based research on this topic.

A limit that was encountered in using the DeepMet model (Su et al., 2020) for metaphor identification was the fact that this model does not perform well in identifying novel metaphors, as was reported by Neidlein et al. (2020) for most of the recent deep learning models designed for metaphor identification. For this reason, the analysis of metaphor use in the APS corpus was expected to be restricted to conventional metaphors, that is, words that have adapted to a metaphorical sense over time, to the point that the metaphorical sense in question becomes one of the word's possible meanings, rather than creative associations of concepts.

A global analysis of the results of the metaphor identification process performed on the APS corpus revealed that the vast majority of tokens identified as metaphors, were highly conventional metaphors, dead metaphors, and borderline cases. These findings confirmed the limitations of the model and further highlighted the need to construct corpora annotated specifically for novel metaphors to use them as training datasets if one's objective is to analyze creative figurative language usage in large-scale corpora.

The second aim of this thesis was to analyze the results of the metaphor identification process and to compare these findings to the results of a corpus-based study that does not rely on computational models for metaphor detection, namely, Charteris-Black's (2004) research on 51 US presidential inaugural speeches.

The quantitative analysis of the results of the metaphor identification process revealed that content words identified as metaphors were used in similar percentages across time periods, regardless of speech type (Campaign Speeches vs. Presidential Speeches), and political affiliation (Democratic vs. Republican). An implication of these findings is that conventional metaphor use seems to be consistent in American political presidential discourse, strengthening the idea that metaphor is ubiquitous in language regardless of genre (Shutova, 2010).

In order to isolate relevant conventional metaphors and compare their use to Charteris-Black's (2004) findings, a deductive approach based on source domain keyword search was adopted, isolating the same source domains drawn up by Charteris-Black. These source domains being those of *body parts*, *buildings*, *conflict*, *fire and light*, *physical environment*, and *religion*. The distribution and productivity of these source domains were analyzed across the two corpora and within the APS corpus according to speech type (Campaign Speeches vs. Presidential Speeches), political affiliation (Democratic vs. Republican). The main findings of this analysis are as follows:

- The resonance scores of the source domains were similar in both corpora regardless of genre, signifying that the small set of source domains drawn up by Charteris-Black (2004) for a specific genre of American presidential speech, the inaugural speech, is actually consistently employed in similar proportions in American presidential discourse at large, and regardless of genre, since similar source domain resonance scores were observed for the Campaign Speeches subset of the APS corpus.
- The most common and resonant source domains in both corpora are those of *conflict*, *journey* and *buildings*, which confirms that these are not only the most common source domains in American presidential inaugural speeches, but also in American presidential discourse at large.
- Comparing source domain distribution in Democratic and Republican campaign and presidential speeches revealed that source domain distribution does not vary significantly in the two parties in both subsets, with the exception of *religion* metaphors. *Religion* metaphors were, in fact, significantly more frequent in Republican campaign speeches and, to a lesser extent, in Republican presidential speeches. This result could be related to the party's more conservative ideological views, however a more in depth qualitative analysis should be performed to confirm this hypothesis.
- Comparing source domain resonance in Democratic and Republican campaign and presidential speeches revealed that source domain resonance varies for specific source domains according to party, even though the corresponding source domain

distributions are similar, this signifies that parties use more varied vocabularies for certain source domains.

- A preliminary qualitative analysis of the source domains also confirmed that the metaphors associated with these source domains were employed similarly to those reported by Charteris-Black (2004), specifically, they are employed to convey abstract social goals such as justice and peace in more concrete terms.
- The qualitative analysis also revealed a significant number of dead metaphors and borderline cases in the selected domain sources as well, further confirming the need to construct a corpus annotated for metaphor with computational analysis in mind.
- Lastly, a diachronic analysis on the Campaign Speeches subset revealed that source domain incidence in this genre of speech is mostly constant over time, with the exception of less common source domains such as those *body parts* and *fire and light*. These source domains presented high spikes of relative frequency in years that coincide with the presence of fewer source domains and overall fewer metaphors.

4. Conclusion

This thesis presented an explorative computational study of metaphor in an ad-hoc corpus, the APS corpus, which cointains 1721 American presidential speeches, from Dwight D. Eisenhower to Joe Biden. The APS corpus was constructed to include both campaign speeches and presidential speeches and was balanced on single terms of presidency.

The first aim of the study was to overcome the challenge of metaphor extraction in large corpora such as the one constructed for this thesis by employing a computational model for automatic metaphor identification, namely the DeepMet model (Su et al., 2020). The second aim of this thesis was to quantitatively and qualitatively analyze the results of the metaphor identification process to gain broader insights into the use of figurative language in American presidential political discourse.

The results were first analyzed quantitatively, comparing and contrasting the use of metaphor across different time periods (1952-2021), different types of speeches (Campaign Speeches vs. Presidential Speeches), and different political parties (Democrats vs. Republicans). The results were subsequently analyzed adopting a quantitative and qualitative approach to determine which types of metaphors are detected by the computational model and how they are distributed in the corpus.

The first analysis revealed that the DeepMet model (Su et al., 2020) detected metaphors in about 17% of the tokens in both the Campaign Speeches and Presidential Speeches subsets of the APS corpus. The most frequent parts of speech that were marked as metaphors were content words, that is, verbs, nouns and adjectives. The relative metaphor frequencies did not vary much across different campaign periods or presidential terms, indicating a consistent use of metaphor in the corpus. Moreover, metaphor use was also found to be quantitatively consistent across the two political parties. The highest percentages of metaphors were found in the Inaugural Addresses and in the Addresses to the United Nations presidential speech genres.

The second analysis revealed that the vast majority of metaphors detected by the DeepMet model were highly conventional metaphors, dead metaphors, and borderline cases. A deductive approach based on source domain keyword search was adopted to narrow down relevant metaphors. The source domains analyzed were the same seven domains drawn up by Charteris-Black (2004) in his study on the use of metaphor in a corpus of 51 US presidential inaugural speeches. These source domains being those of *body parts*, *buildings*, *conflict*, *fire and light*, *physical environment*, and *religion*. The analysis of the distribution and productivity of these source domains confirmed that the results reported by Charteris-Black (2004) for a smaller and genre-specific corpus of American presidential speeches are consistent in a larger corpus that contains different presidential speech genres and that covers a more recent timeline. Overall, the most common source domains are those of *conflict*, *journey* and *buildings*. A preliminary qualitative analysis showed that these source domains are typically employed to convey abstract social goals such as justice and peace in more concrete terms. Moreover, the qualitative analysis further confirmed the significant presence of highly conventional metaphors, borderline cases, and labeling errors in the keywords from the source domains as well. Lastly, a diachronic analysis on the Campaign Speeches subset revealed that source domain incidence in this genre of speech is mostly constant over time, with the exception of less common source domains such as those *body parts* and *fire and light*. Less common source domains presented high spikes of relative frequency in years that were revealed to coincide with the presence of fewer source domains and overall fewer metaphors in those years.

While this thesis presented promising results for the possibility to computationally study metaphor in large-scale corpora, it should be noted that these results are limited to conventional metaphors associated with well-established source domains. A major limitation of the DeepMet model (Su et al., 2020), and of other recent automatic metaphor identification models, is, in fact, its poor performance in detecting novel metaphors.

In order to collect data on the use of novel metaphors in American presidential speeches, future work may look further into constructing a corpus that is annotated specifically for novel metaphors, to use as a training dataset for deep learning models such as the DeepMet (Su et al., 2020).

Another direction of future work could be determining whether the limited set of source domains for conventional metaphors that has been found to be consistent in American presidential speeches is also representative of other types of American political discourse. Further research should also examine if these findings extend to political discourse in other cultures and languages.

Appendix A

Table A1

Distribution of presidential terms in the Campaign and Presidential Speeches subsets

	term	# camp docs	ave size	# term docs	ave size
Dwight Eisenhower	1953-1957 1957-1961	2 2	1302 845	71 79	1707 2258
John F. Kennedy	1961-1963	8	4297	89	2715
Lyndon B. Johnson	1963-1965 1965-1969	1 2	1311	82 78	2707 2840
Richard Nixon	1969-1973 1973-1974	3 0	161 0	87 50	2592 2696
Gerald R. Ford	1974-1977	2	165	90	1596
Jimmy Carter	1977-1981	6	534	85	1887
Ronald Reagan	1981-1985 1985-1989	4 4	1017 1347	85 88	3282 2669
George H. W. Bush	1989-1993	5	4104	85	3047
Bill Clinton	1993-1996 1996-2001	8 2	1258 8076	91 78	2841 3626
George W. Bush	2001-2004 2004-2009	5 4	475 187	80 93	2650 3298
Barack Obama	2009-2012 2013-2017	6 5	215 235	79 81	2347 2638
Donald Trump	2017-2021	5	319	89	3596
Joe Biden	2021-2022	6	769	81	2306
Total		80	2047	1641	2665

Table A2

Distribution of tokens and types in Campaign and Presidential Speeches subsets

	# Tokens	# Types
Campaign speeches	111831	8501
Presidential speeches	2803352	41436
Total	2915183	42224

Table A3

Distribution of genre types in Presidential Speeches subset

	# docs	percentage	ave size
Address to the United Nations	48	2.9%	2700
Farewell address	11	0.7%	2842
Inaugural address	19	1.2%	1499
Press conference	399	24.3%	4604
Remarks	578	35.2%	1733
Speech	267	16.3%	2903
State of the Union address	65	4%	4755
Statement	30	1.8%	595
University/Academy speech	98	6%	2539
Victory speech	17	1%	981
Weekly address	64	3.9%	579
Other	45	2.7%	1747
Total	1641	100%	2290

Table A4

Distribution of genre types in Campaign Speeches subset

	# docs	percentage	ave size
Campaign announcement	16	20%	1479
Campaign speech	27	33.75%	2867
Debates	37	46.25%	11524
Total	80	100%	15870

Table A5

Distribution of party affiliation in Campaign and Presidential Speeches subsets

	# camp docs	ave size	# term docs	ave size
Democratic	44	1641	744	1661
Republican	36	1101	897	1766

Appendix B

Table B1

Metaphor source domains keywords

BODY PARTS	arm, artery, blood, body, bone, brain, cheek, eye, ear, finger, fingertip, fist, flesh, foot, footprint, gut, hand, head, heart, leg, mouth, shoulder, skin, stomach
BUILDINGS	architect, architecture, arena, backbone, base, basin, basis, bastion, bedrock, bridge, build, buildup, cathedral, ceiling, cement, construction, cornerstone, door, doorstep, edifice, entrance, exit, floor, fortress, foundation, foundation-stone, framework, house, interior, pillar, rebuild, restore, roof, room, structure, threshold, tower, underpin, wall, window
CONFLICT	advance, adversary, aggression, aid, ailment, alarm, alignment, alliance, annihilate, antagonist, arm, armor, army, arsenal, ascendancy, assault, assistance, at-peace, attack, battle, battle-cry, battlefield, battlefront, battleground, beach-head, beat, blast, blockade, bloodlust, bomb-heavy, bounty, bullet, bulwark, capitulation, captain, captive, carnage, citadel, combat, combat-equipment, command, conflict, conquer, contest, conversion, deescalation, defeat, defect, defend, deploy, detente, devotion, enemy, fight, firepower, flank, front-line, outflank, flashpoint, foe, loser, maneuver, overpower, overrun, protect, resistance, retreat, sergeant, shield, shoot, slaughter, strategy, strife, surrender, tactic, threaten, triumph, war, warfare, weapon
FIRE AND LIGHT	ashe, beacon, beacon-of-light, beam, blaze, blind, bonfire-light, bright, brilliant, burn, candle, clear, dark, dawn, dazzle, dim, eclipse, fire, flame, flare, glare, gleam, gloom, glow, heat, illuminate, kindle, light, pale, radiance, ray, see, shadow, shine, spark, star, sun, sunshine
JOURNEY	arrival, barrier, burden, bump, chart, course, crossroad, crossroads, captain, departure, destination, direction, drive, embark, emigrate, exodus, explore, follow, footstep, forward, frontier, go-forward, harbor, harness, immigrant, journey, march, milestone, move-forward, obstacle, pace, paddle, passenger, path, pave, plan, plank, rail, railroad, ride, road, sail, ship, sink, station, step, track, travel, trip, tour, voyage, walk, way
PHYSICAL ENVIRONMENT	abyss, air, atmosphere, avalanche, backdrop, bay, blizzard, border, breeze, cave, cavern, chasm, chill, climate, cloud, cold, continent, corner, crest, crosscurrent, current, deadlevel, desert, dry, eclipse, environment, field, firmament, flood, floodgate, fog, foothills, freeze, hail, hole, horizon, hot, jungle, land, nature, rain, storm, tide, warm, warmth, whirlwind, wilderness, wind
RELIGION	angel, apostle, church, church-support, creed, crusade, destiny, devil, faith-base, fate, hell, magic, miracle, miraculous, pray, prophet, sacred, saint, soul, specter, spirit, vision

Appendix C

Table C1

Summary of selected source domains and resonance in the APS corpus

Source domain	*Total types*	*Total tokens*	*Resonance*	*% of total resonance*
Conflict	86	6303	542058	41%
Journey	67	3922	262774	20%
Buildings	41	4695	192495	15%
Fire and light	44	2362	103928	8%
Physical environment	56	1918	107408	8%
Body parts	32	1904	60928	5%
Religion	32	1616	51712	4%
Total	374	22720	1321303	

Table C2

Summary of source domains and resonance in Charteris-Black's US Inaugural corpus

Source domain	*Total types*	*Total tokens*	*Resonance*	*% of total resonance*
Conflict	18	116	2088	36%
Journey	12	76	912	16%
Buildings	12	66	792	14%
Fire and light	15	51	765	13%
Physical environment	16	35	560	9%
Body parts	6	72	432	7%
Religion	4	76	304	5%
Total	77	492	5853	

Note. From "Corpus Approaches to Critical Metaphor Analysis" by Charteris-Black, J., 2004. *Proceedings of the Workshop on Figurative Language Processing*, p. 90. (2004).

References

Ahrens, K. (2015). *Corpus of Political Speeches*. Hong Kong Baptist University Library. Retrieved from https://digital.lib.hkbu.edu.hk/corpus/

Baranov, A. N., Kazakevich, E. G. (1991). Parliamentary debates: traditions and innovations, *Modern political language from ritual to metaphor*, 63-70. Moscow: Znanie.

Baroni, M., & Bernardini, S. (2005). BootCaT: Bootstrapping corpora and terms from the web. In M.T. Lino, M.F. Xavier, F. Ferreira, R. Costa & R. Silva (Eds.), *Proceedings of Fourth International Conference on Language Resources and Evaluation* (LREC 2004), 1313– 1316. Lisbon: ELDA.

Berberović, S., & Delibegović Džanić, N. (2021). *A cognitive linguistic study of the use of creative figurative language in American political discourse*. Livre de Lyon.

Black, M. (1962). *Models and metaphors: Studies in language and philosophy*. Cornell University Press.

Blei, D.M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3* (Jan), 993-1022.

Bowdle, B. F., & Gentner, D. (1995). The Career of Metaphor. *Psychological Review*, 112(1), 193–216.

Burnard, L. (2007) *Users' reference guide to the British National Corpus* (XML edition). Oxford: Oxford University Computing Services. Retrieved from http://www.natcorp.ox.ac.uk/XMLedition/URG/

Dalla Libera, C. (2017). Le metafore concettuali in un approccio comunicativo nell'apprendimento delle lingue straniere. *EL.LE*.

Danesi, M. (2001). *Lingua, metafora, concetto. Vico e la linguistica cognitiva*. Bari: Edizioni del Sud.

Dankers, V., Rei, M., Lewis, M., & Shutova, E. (2019). Modelling the interplay of metaphor and emotion through multitask learning. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2218–2229.

Davidson, D. (1978). What metaphors mean. *Critical Inquiry* 5(1), 31–47.

Deignan, A. (2005). *Metaphor and corpus linguistics*. J. Benjamins Pub.

Deignan, A. (2008). Corpus linguistics and metaphor. In R. Gibbs, Jr. (Ed.), *The Cambridge Handbook of Metaphor and Thought* (Cambridge Handbooks in Psychology, pp. 280-294). Cambridge: Cambridge University Press.

Devlin, J., Chang, M., Lee, K. & Toutanova, K. (2018), Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding, *CoRR*.

Fass, D. (1991). met*: A Method for Discriminating Metonymy and Metaphor by Computer. *Computational Linguistics*, 17(1).

Ficcadenti, V., Cerqueti, R., & Ausloos, M. (2019). A joint text mining-rank size investigation of the rhetoric structures of the US Presidents' speeches. *Expert Systems with Applications*, 123, 127–142.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy, *Cognitive Science*, 7(2), 155-170.

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

Gong, H., Gupta, K., Jain, A., & Bhat, S. (2020). IlliniMet: Illinois System for Metaphor Detection with Contextual and Linguistic Information. *Proceedings of the Second Workshop on Figurative Language Processing*, 146–153.

Guastini, D. (2004). Aristotele e la metafora: ovvero l'elogio dell'approssimazione. Istituto di Filosofia Arturo Massolo. Università di Urbino. Isonomia. Retrieved from http://www.isonomia.uniurb.it/vecchiaserie/guastini/guastini2004.pdf

Hanks, P. (2004). THE SYNTAGMATICS OF METAPHOR AND IDIOM. *International Journal of Lexicography*, 17(3), 245–274.

Heintz, I., Gabbard, R., Srinivasan, M., Barner, D., Black, D. S., Freedman, M., & Weischedel, R. (2013). Automatic Extraction of Linguistic Metaphor with LDA Topic Modeling. *Proceedings of the First Workshop on Metaphor in NLP*, 58–66.

Hermann, K. M., Kočiský, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., & Blunsom, P. (2015). *Teaching Machines to Read and Comprehend.*

Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*.

Jacobs, T., & Tschötschel, R. (2019). Topic models meet discourse analysis: A quantitative tool for a qualitative approach. *International Journal of Social Research Methodology*, *22*(5), 469–485.

Jurafsky, D., Martin, H. J. (2008). Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. Prentice Hall.

Köper, M., & Schulte im Walde, S. (2017). Improving Verb Metaphor Detection by Propagating Abstractness to Words, Phrases and Individual Senses. *Proceedings of the 1st Workshop on Sense, Concept and Entity Representations and Their Applications*, 24–30.

Keysar B., Shen Y., Glucksberg S., and Horton W. S. (2000). Conventional language: How metaphorical is it? *Journal of Memory and Language*, 43(4), 576–593.

Kirby, J. T. (1997). Aristotle on metaphor. *American Journal of Philology*, 118(4), 517–554.

Klebanov, B. B., Diermeier, D., & Beigman, E. (2008). Lexical Cohesion Analysis of Political Speech. *Political Analysis*, 16(4), 447–463.

Klebanov, B. B. & Flor, M. (2013). Argumentation-relevant metaphors in test-taker essays. *Proceedings of the First Workshop on Metaphor in NLP*, 11–20.

Koivisto-Alanko, P. & Tissari, H. (2007). Sense and sensibility: Rational thought versus emotion in metaphorical language. In A. Stefanowitsch & S. Gries (Ed.), *Corpus-Based Approaches to Metaphor and Metonymy*, 191-213. Berlin, New York: De Gruyter Mouton.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and Thought*, 2nd edition, 202–251. New York and Cambridge, U.K.: Cambridge University Press.

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

Leong, C. W. (Ben), Beigman Klebanov, B., Hamill, C., Stemle, E., Ubale, R., & Chen, X. (2020). A Report on the 2020 VUA and TOEFL Metaphor Detection Shared Task. *Proceedings of the Second Workshop on Figurative Language Processing*, 18–29.

Leong, C. W. (Ben), Beigman Klebanov, B., & Shutova, E. (2018). A Report on the 2018 VUA Metaphor Detection Shared Task. *Proceedings of the Workshop on Figurative Language Processing*, 56–66.

Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., & Stoyanov V. (2019). Roberta: A robustly optimized bert pretraining approach.

Manning, C., Raghavan, P., & Schuetze, H. (2009). *Introduction to Information Retrieval*. Prabhakar Raghavan.

Mao, R., Lin, C., & Guerin, F. (2018). Word Embedding and WordNet Based Metaphor Identification and Interpretation. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1222–1231.

Martin, J. (2007). A corpus-based analysis of context effects on metaphor comprehension. In A. Stefanowitsch & S. Gries (Ed.), *Corpus-Based Approaches to Metaphor and Metonymy*, 214-236. Berlin, New York: De Gruyter Mouton.

Martin, J. H. (1988). Representing regularities in the metaphoric lexicon. *Proceedings of the 12th Conference on Computational Linguistics*, 1, 396–401.

Mason, Z. J. (2004). CorMet: A Computational, Corpus-Based Conventional Metaphor Extraction System. *Computational Linguistics*, 30(1), 23–44.

McCann, B., Keskar, N. S., Xiong C., & Socher, R. (2018). The natural language decathlon: Multitask learning as question answering.

McKinney, W., & others. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, 445, 51–56.

Meyer, D., Zeileis, A., and Hornik, K. (2006). The strucplot framework: Visualizing multi-way contingency tables with vcd. *Journal of Statistical Software*, 17(3), 1–48.

Murphy G. L. (1996). On metaphoric representation. *Cognition*, 60(2), 173–204.

Musolff, A. (2004). *Metaphor and political discourse: Analogical reasoning in debates about Europe*. London: Palgrave Macmillan.

Neidlein, A., Wiesenbach, P., & Markert, K. (2020). An analysis of language models for metaphor recognition. *Proceedings of the 28th International Conference on Computational Linguistics*, 3722–3736.

Nguyen, H. L., Nguyen, T. D., Hwang, D., & Jung, J. J. (2015). KELabTeam: A Statistical Approach on Figurative Language Sentiment Analysis in Twitter. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 679–683.

Nunberg, G. (1987). Poetic and prosaic metaphors. *Proceedings of the 1987 Workshop on Theoretical Issues in Natural Language Processing*. Association for Computational Linguistics, Las Cruces, New Mexico, 198–201.

Prabhakaran, V., Rei, M., & Shutova, E. (2021). How Metaphors Impact Political Discourse: A Large-Scale Topic-Agnostic Study Using Neural Metaphor Detection. *Proceedings of the International AAAI Conference on Web and Social Media*, *15*, 503–512.

Pragglejaz Group. (2007) MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22, 1–39.

Rai, S., & Chakraverty, S. (2021). A Survey on Computational Metaphor Processing. *ACM Computing Surveys*, 53(2), 1–37.

Rei, M.; Bulat, L.; Kiela, D.; & Shutova, E. (2017). Grasping the Finer Point: A Supervised Similarity Network for Metaphor Detection. *Proceedings of EMNLP 2017*.

Rehurek, R., & Sojka, P. (2011). Gensim–python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, 3(2).

Semino, E. (2005) The metaphorical construction of complex domains: the case of speech activity in English. *Metaphor and Symbol* 20: 35–70. In A. Stefanowitsch & S. Gries (Ed.), *Corpus-Based Approaches to Metaphor and Metonymy*, 36-62. Berlin, New York: De Gruyter Mouton.

Semino, E. (2008). *Metaphor in discourse*. Cambridge: Cambridge University Press.

Shutova, E., Kiela, D., & Maillard, J. (2016). Black Holes and White Rabbits: Metaphor Identification with Visual Features. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 160–170.

Shutova, E. (2011). Computational Approaches to Figurative Language. Ph.D. thesis, University of Cambridge, UK.

Steen, G. (Ed.). (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins Pub. Co.

Stefanowitsch, A. (2006). Corpus-based approaches to metaphor and metonymy. In A. Stefanowitsch & S. Gries (Ed.), *Corpus-Based Approaches to Metaphor and Metonymy*, 1-16. Berlin, New York: De Gruyter Mouton.

Stefanowitsch, A. (2020). *Corpus linguistics: A guide to the methodology*. Zenodo.

Su, C., Fukumoto, F., Huang, X., Li, J., Wang, R., & Chen, Z. (2020). DeepMet: A Reading Comprehension Paradigm for Token-level Metaphor Detection. *Proceedings of the Second Workshop on Figurative Language Processing*, 30–39.

Thibodeau P. H., Boroditsky L. (2015) Measuring Effects of Metaphor in a Dynamic Opinion Landscape. *PLOS ONE*, 10(7): e0133939.

Tissari, H. (2011). English words for emotions and their metaphors. In M. Winters, H. Tissari & K. Allan (Ed.), *Historical Cognitive Linguistics*, 298-330. Berlin, New York: De Gruyter Mouton.

Tsvetkov, Y., Boytsov, L., Gershman, A., Nyberg, E., & Dyer, C. (2014). Metaphor Detection with Cross-Lingual Model Transfer. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 248–258.

Wilks, Y. (1975). A preferential, pattern-seeking, semantics for natural language inference. *Artificial Intelligence*, 6, 53–74.

Wilks, Y. (1978). Making preferences more active. *Artificial Intelligence*, 11, 197–223.

Wilson, J. (2005). *Political discourse. The Handbook of Discourse Analysis*, 398-415.

Winner, E. (1988). *The point of words: Children's understanding of metaphor and irony*. Cambridge, MA: Harvard University Press.

Wu, C., Wu, F., Chen, Y., Wu, S., Yuan, Z., & Huang, Y. (2018). Neural Metaphor Detecting with CNN-LSTM Model. *Proceedings of the Workshop of Figurative Language Processing*, 110-14. New Orleans, Louisiana. Association for Computational Linguistics.

Sitography

Eidenmuller M., E. (2023, May 15). American Rhetoric. https://americanrhetoric.com

The R Foundation. (2023, May 15). The R Project for Statistical Computing. https://www.r-project.org/

The U.S. National Archives and Records Administration. (2023, May 15). Presidential Libraries. https://archives.gov/presidential-libraries.html

University of Virginia. (2023, May 15). Miller Center. https://millercenter.org

Woolley, J. & Peter, G. (2023, May 15). The American Presidency Project. https://presidency.ucsb.edu