CHAPTER 6

Language Analysis in Political Psychology

Kayla N. Jordan

■ INTRODUCTION

Language is pervasive in politics, and advances in text analytic methods can not only expand the scope of political psychology but also allow a richer, deeper study of how people interact with political processes in the real world and their daily lives. While much of the work on language and politics has happened outside the realm of psychology, work in other disciplines (e.g., communications and political science) can inform how these methods may be applied to psychological questions. This chapter brings these diverse perspectives together to give guidance to political psychologists seeking to add language analysis to their research projects.

In my own work, I use multiple sources of politically relevant text data to explore individuals' psychological processes within the political domain. The main thrust of my work has been to consider political questions generally asked by political scientists, historians, and communication scholars from a psychological perspective using language analysis. Using texts from leaders, political candidates, voters, and the media, my research addresses the question how and why leaders are selected across time and context. In this chapter, I review the sources of text data I and other scholars have used to an-

swer political psychological questions, while also highlighting the contributions of scholars from other disciplines to this area. Additionally, the chapter highlights the types of questions that can be answered with text data and the common pitfalls to using text as data.

The rest of the chapter is organized around a few central ideas. First, I discuss how the work of other politically relevant disciplines can contribute to the use of language analysis in political psychology, as well as explore what language reveals about politics psychologically. Second, I review the different sources and ways in which language can be used to study political questions. Finally, I end the chapter with a discussion of a few areas of research in which there has been less work, and the questions and topics likely to prove interesting to the field in the future.

■ EXTENDING LANGUAGE ANALYSIS INTO POLITICAL PSYCHOLOGY

While language analysis still is not common in the field of political psychology, other disciplines have a long history of using text as data to study politics. Political scientists have studies how word choices in political ads and speeches impact voters' percep-

tions (Jessee, 2010; Markus, 1992; Petrocik, 1996). Political historians use political texts and qualitative text analysis to explore political leaders and processes of the past (Greenberg, 2016; Milkis & Nelson, 2011). Political communication scholars have used both qualitative and quantitative methods to understand how political figures use language to persuade and inform (Hart, Childers, & Lind, 2013; McKinney & Carlin, 2004). The realm of politics is complex, and it is vital to consider the contribution of these and other disciplines to understand how institutions, history, and rhetoric may impact psychological processes within the political domain.

Communication is the discipline with the longest history of using text as data in the realm of politics. The discipline has also found many ways to quantify text data to understand how political language is used and received. Indeed, much of the work cited in this chapter comes from communication scholars. That said, a small group of political psychologists has used qualitative and quantitative text analytic methods to investigate psychological phenomena, such as motivation (Winter, 2011), leadership (Simonton, 2003), and cognitive complexity (Conway, Conway, Gornick, & Houck, 2014) in political figures across time. As text data have become more available and text analytic methods become easy to implement (as evidenced by other chapters in this volume), new avenues to research have opened up for political psychologists to use political language to understand real-world political processes in ways that were difficult or impossible in years past. My purpose in the rest of this chapter is to give readers a glimpse of some of these new avenues that allow political psychologists to join scholars of other disciplines to more fully understand the functions and implications of political language.

■ TEXT AS DATA IN POLITICS

The first challenge in working with text as data is to determine where and how to collect the text. Political psychology encompasses many different potential populations, including voters/citizens, political figures, and institutions such as the media. While traditional psychology research methods can often be used to collect data about voters, other methods are generally necessary to study political figures and other actors in the political realm. In this section, I outline the common sources of data that can be collected to capture language in the political domain.

Reactions to Political Texts

It is commonly believed that word choice matters in politics. Consider the partisan fights over naming legislative proposals, such as Obamacare versus Affordable Care Act or estate tax versus death tax. Political operatives must believe that, in some cases, a single word can impact voters' attitudes and behaviors, but can they? One way in which political psychologists contribute to our understanding of politics is by testing whether words matter in political messages experimentally (or quasi-experimentally). Such tests can rest on the use of a single word or phrase or test broader stylistic or discourse-level features of language.

Such experimental manipulations are perhaps familiar in the realm of psychology, but they have been used in the political domain by many disciplines to understand a variety of linguistic features on numerous politically relevant outcomes, including vote intention, attitudes, and fundraising (see Table 6.1 for examples of manipulations). A few examples of such studies include the impact of complexity on political persuasion (Amsalem, 2019), loss frames (Arceneaux, 2012) and metaphor (Hartman, 2012) on issue opinions, storytelling on candidate perceptions (Gooch, 2018), verb choice on voting intention (Hauser & Schwartz, 2018), and verbal aggression on perceptions of credibility (Nau & Stewart, 2014).

While experimental manipulation is the only method for truly determining the causality of language manipulations, quasi-experimental and observational methods can also be used to begin exploring the effects of political language. For example, in my own work, I often use debates and other campaign events to explore the types of linguistic styles that may make a successful candidate. In a study currently in progress, I used presidential primary debates to explore whether

TABLE 6.1. Examples of Language Manipulation

Changed phrasing	"This is all backwards. We should do hearings first and then figure out what makes sense, and then do the legislation. What we are doing is we're doing legislation and then maybe hearings, maybe not, and then we're not even given an alternative. It just strikes me as unbelievable that after all we heard about openness and full discussion, we are rushing this to the floor in an emergency Rules Committee meeting and we're all agreeing it should be a closed process, so where's the openness and where's the discussion?" [Control]
	versus
	"**You obviously don't know what you are doing** [competence attack] because this is all backwards. We should do hearings first and then figure out what makes sense, and then do the legislation. What we are doing is we're doing legislation and then maybe hearings, maybe not, and then we're not even given an alternative. It just strikes me as unbelievable and **hypocritical** [character attack] that after all we heard about openness and full discussion, that we are rushing this to the floor in an emergency Rules Committee meeting and we're all agreeing it should be a closed process, so where's the openness and where's the discussion?" [Aggressive Condition]
	(Nau & Stewart, 2014)
Changed focus	"In his past term, Governor Steve Williams introduced regulations that **caused** extreme changes to the budget that were felt by many. He is the incumbent candidate running against local politician, Joshua Bayer." [Negative Valence]
	versus
	"In his past term, Governor Steve Williams introduced the regulations that **produced** extreme changes to the budget that were felt by many. He is the incumbent candidate running against local politician, Joshua Bayer." [Neutral Valence]
	(Hauser & Schwarz, 2018)
Changed style	"Would you favor or oppose a proposal by Vice President Joe Biden to raise the Social Security contribution rate for high-income individuals, or haven't you thought much about this?" [Elite cue only]
	versus
	"When talking about Social Security, Vice President Joe Biden said the following: 'I recently met with seniors. Many have worked for almost 40 years, caring for others. They're 70 years old now and happily retired. But like most seniors nowadays, they cannot afford medication, medical supplies, and nursing care from their Social Security benefits. Day-to-day expenses are impossible for them to meet. They worked hard their whole life and deserve more from an outdated Social Security program. Would you favor or oppose a proposal to raise the Social Security contribution rate for high-income individuals, or haven't you thought much about this?' " [Elite cue with impersonal story]
	(Gooch, 2018)
Complexity	"Solving the problem of illegal immigration to the country is a top priority for me. I favor allowing illegal immigrants who are otherwise law-abiding a path to full citizenship. The reason we need to naturalize illegal immigrants is that it will make our economy grow. I have no doubt that allowing illegal immigrants legal status is the right way to go—all other solutions to this problem just don't make sense." [Low Complexity]
	versus
	"Solving the problem of illegal immigration to the country is a top priority for me. I favor allowing illegal immigrants who are otherwise law-abiding a path to full citizenship. We need to naturalize illegal immigrants because it will make our economy grow. However, we must be careful: Allowing everyone to stay may encourage more illegal immigrants to come here, and we don't want that. My plan is to balance these two goals." [High Complexity]
	(Amsalem, 2019)

(continued)

TABLE 6.1. *(continued)*

Metaphor	"Congressman Alan Davidson, who specializes in technology issues, supports Network Neutrality legislation. He recently told reporters: 'Telecoms want to set up **toll booths** on the Internet to stand between content providers and their customers. Network Neutrality would prevent this from happening. It would ensure that we don't have a system where some companies have **access to an express lane**, while the rest are stuck **waiting in line at the toll booth**.'" [Metaphor]
	versus
	"Congressman Alan Davidson, who specializes in technology issues, supports Network Neutrality legislation. He recently told reporters: 'Telecoms want to **charge fees** on the Internet to connect content providers to their customers. Network Neutrality would prevent this from happening. It would ensure that we don't have a system where some companies have **access to fast services**, while the rest are **left with slower connections**.'" [Literal]
	(Hartman, 2012)
Framing	"The traditional family is the bedrock of our nation, and these groups are part of an agenda to encourage impressionable teenagers to adopt alternative lifestyles. School administrators and parents should be able to ban these groups so that we **do not lose** our connection to traditional family values" [Loss Frame]
	versus
	"The traditional family is the bedrock of our nation, and these groups are part of an agenda to encourage impressionable teenagers to adopt alternative lifestyles. School administrators and parents should be able to ban these groups so that we **can strengthen** our connection to traditional family values" [Gain Frame]
	(Arceneaux, 2012)

different language styles are associated with getting a party's nomination. Using dictionary-based algorithms developed using the Linguistic Inquiry and Word Count (LIWC) program, I compared winning versus losing candidates on three linguistic styles: authenticity, optimism, and confidence. These algorithms, which were developed by combining language analysis with more traditional methods, allow researchers to measure psychological constructs using natural language. For example, the authenticity metric was developed by comparing the language used when people are telling the truth versus lying. Multiple studies (Bond et al., 2017; Hancock, Curry, Goorha, & Woodworth, 2007; Newman, Pennebaker, Berry, & Richards, 2003) revealed that when people are being open and honest, they tend to use more I pronouns, present-tense verbs, and relativity markers (e.g., "new," "far," "here"). On the other hand, people who are being evasive tend to use more "he"/"she" pronouns and discrepancy markers (e.g., "should," "could"). Table 6.2 shows how differences between political figures can be realized in political language that can then be used to understand why some politicians are more successful and/or liked than others.

The graph in Figure 6.1 shows how these differences in linguistic styles predicting success in presidential primaries (defined here as receiving the nomination) across eight elections. On average, across the primary season, the candidate who ultimately prevailed tended to use language in the debates that reflected a straightforward, optimistic approach compared to the losing candidates, who tended to display greater confidence. Winning candidates who did not conform to this pattern generally went on to lose the general election in November. This study is just one example of how language analysis can be combined with outcomes, such as elections and public opinion, to understand how psychological traits and processes impact the political world.
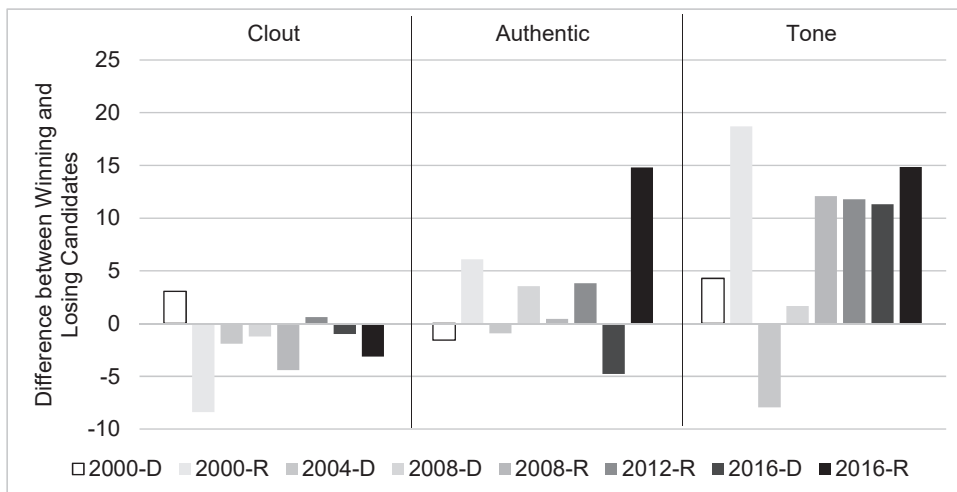
Measuring Political Language

In addition to measuring reactions to political language, the language itself can also provide valuable insights into what and how people think and feel about politics. The

TABLE 6.2. Example Responses from 2016 Republican Primary Debates

LIWC category	Low	High
Clout	I agree 100%, by the way, with Carly on the fact that the Democrats do not want to solve this problem, for the obvious reasons, but they do not. But I believe that a reading of the 14th Amendment allows you to have an interpretation where this is not legal and where it can't be done. I've seen both sides, but some of the greatest scholars agree with me, without having to go through Congress. (Donald Trump)	The fact is that we don't want to hear about your careers, back and forth, and volleying back and forth about who did well and who did poorly. You're both successful people. Congratulations. You know who's not successful? The middle class in this country who's getting plowed over by Barack Obama and Hillary Clinton. Let's start talking about those issues tonight and stop this childish back-and-forth between the two of you. (Chris Christie)
Authenticity	John Roberts has made some really good decisions, for sure, but he did not have a proven, extensive record that would have made the clarity the important thing, and that's what we need to do. And, I'm willing to fight for those nominees to make sure that they get passed. You can't do it the politically expedient way anymore. This is the culture in Washington. You have to fight hard for these appointments. This is perhaps the most important thing that the next president will do. (Jeb Bush)	But, you know, the fact of the matter is, we have extremely well-documented proof that there's no autism associated with vaccinations. But it is true that we are probably giving way too many in too short a period of time. And a lot of pediatricians now recognize that, and, I think, are cutting down on the number and the proximity in which those are done, and I think that's appropriate. (Ben Carson)
Tone	Well, let me tell you, Jake, the single biggest national security threat facing America right now is the threat of a nuclear Iran. We've seen six and a half years of President Obama leading from behind. Weakness is provocative, and this Iranian nuclear deal is nothing short of catastrophic. This deal, on its face, will send over $100 billion to the Ayatollah Khamenei, making the Obama administration the world's leading financier of radical Islamic terrorism. (Ted Cruz)	So, here's the deal. My wife is a Mexican American. She's an American by choice. She loves this country as much as anybody in this room, and she wants a secure border. But she wants to embrace the traditional American values that make us special and make us unique. We're at a crossroads right now. Are we going to take the Reagan approach, the hopeful optimistic approach, the approach that says that, you come to our country legally, you pursue your dreams with a vengeance, you create opportunities for all of us? (Jeb Bush)

basic premise behind psychological text analysis is that the words people choose have meaning beyond mere semantics. The words people choose to use can reveal important psychological processes. The most common method for analyzing text data in this way is to count words associated with some higher-level construct. For example, if you were interested in the motivations of a set of political candidates, you could count how often they used words associated with being power-motivated (e.g., "leader," "weakness," "demand") or affiliation-motivated (e.g., "help," "ally," "we") or achievement-motivated (e.g., "win," "earn," "excellent") (Jordan, 2019b) and maybe even determine whether those differences are related to some other variable (e.g., party-affiliation, election outcomes). In the rest of this section I focus mainly on the types of text data available in political contexts, but before jumping into the data, researchers must define the construct of interest, then determine what are the words (or phrases or structures) people use in language that capture what it means to be high or low in that construct.

FIGURE 6.1. Socioemotional traits of winning versus losing presidential primary candidates.

One method to capture political language is through open-ended survey questions. These questions can be added to traditional studies and can be used to (1) clarify or expand on answers to closed-format questions or (2) probe attitudes where it is unclear what people would naturally think about. One example is the American National Elections Studies (ANES, 2020) which is a series of nationally representative surveys conducted every election cycle stretching back to 1948. Though the majority of the questions asked in the ANES surveys are close-ended, the surveys are extremely comprehensive, and the open-ended questions probe more deeply into people's attitudes (e.g., "What do you dislike about the Democratic (or Republican) party?"). While developed from a political science perspective, the scope and design of the ANES can be instructive for political psychologists using survey designs. An important caveat for those incorporating open-ended survey questions is that it can be quite difficult to get participants to answer them satisfactorily. Such questions must be designed so that it is clear to participants that they need to provide an answer, and there is some incentive for participants to write a sufficiently lengthy, on-topic response.

Another source of political language is social media and other online forums. In the last few decades, the Internet and social media platforms have provided new ways for people to not only find political information but also to communicate and discuss their political attitudes with each other and even political leaders. Social media and online forums provide an insight into people's political thoughts and attitudes in an unprompted manner. In traditional surveys, even with open-ended questions, people's responses may be influenced by the wording of the question or may elicit attitudes that are not strongly held/deeply considered. Online data such as Tweets or blog posts can better represent the issues that people truly think and care about in their daily lives.

In recent years, Twitter and Facebook have been highly visible (and criticized) platforms in political processes. Social media platforms give a novel look into how political figures themselves adapt and use these technologies for campaigning and governing. Unlike the past, when politicians communicated in relatively formal settings through traditional media outlets, social media and other emerging technologies are more informal communications by politicians themselves (or their social media teams). These platforms can also provide a real-time, unobtrusive look into how citizens are reacting and interacting with political events and processes. Numerous studies in the past decade have sought to both investigate traditional psychological questions in this new

domain (Sagi & Dehghani, 2014) as well as to understand how these platforms impact people's psychological processes (Grinberg, Joseph, Friedland, Swire-Thompson, & Lazer, 2019; Jordan, Pennebaker, & Ehrig, 2018; Siegel et al., 2021).

Twitter has been the platform most studied, likely due to the public nature of the platform and the relative ease of collecting Tweets. While the platform is constantly changing, presenting new challenges to researchers, Murphy (2017) provides a guide for using Twitter data in psychological research that is useful for researchers interested in this domain. In addition to general guidelines and advice for psychological research using Twitter, more technical guides are available to explain the mechanics of how to collect Twitter data (Helpful Tools for Researchers, n.d.). Facebook (as well as other social media platforms such as Instagram) have been studied (Bond et al., 2012; Matz, Kosinski, Nave, & Stillwell, 2017), but they present more of a challenge for research due to the private nature (i.e., most users have private accounts that can only be seen by their friends, whereas on Twitter, many users have public accounts) of the platforms and restriction on data collection (i.e., most platforms do not allow/heavily restrict what data researchers can collect, whereas platforms such as Twitter have application programming interfaces (APIs) allowing for automated retrieval of Tweets and public user profiles with an R/Python script). Collecting data from these platforms is not impossible, but it can involve incorporating more traditional research methods, such as using surveys to ask participants directly for the Facebook or Instagram profiles. Another consideration when looking to incorporate social media data into research is that these platforms and their user bases are constantly changing (with bots and fake accounts becoming an increasing problem), and new platforms emerge over time. Hence, it is vitally important to keep abreast of these changes and to carefully consider what aspects of these platforms may lead to robust, enduring insights in political psychology.

Another recent source of data in political psychology has been Google Trends, which aggregates searches made on its platform to reveal what people are most interested in

knowing about in a given time and/or location (Google, 2020). Online searches can be particularly informative about people's true psychological states as they are private behaviors not influenced by social pressures as other online behaviors generally are (e.g., Tweets, Facebook posts). For example, one study of Google Trends in the domain of political psychology examined the relative frequency of searches with racist language in U.S. states to predict decreased support for Barack Obama in 2008 and 2012 (Stephens-Davidowitz, 2014). Other studies have indicated that foreign trips by presidents increase searches about the president, potentially explaining increases in presidential approval following such trips (Cohen, 2016), and that following a Republican electoral victory, red states show more pornography searches, and as do blue states following a Democratic victory (Markey & Markey, 2011). As demonstrated by these examples, Google Trends can be useful for answering questions that may be difficult to answer using traditional methods. For example, Google Trends can show what information/traits people are looking for in a candidate or how people in different states search for different types of political information before an election.

However, using Google Trends data presents many challenges that researchers need to consider. First, Google Trends data are aggregated data. In psychology, we generally try to study patterns at an individual level, and the aggregated nature of Trends data makes this difficult and limits the conclusions that can be drawn from the results. Second, it is difficult to know if Google users at a given time or place are representative of the population of interest. While this is generally an issue with much real-world data, the aggregated nature of the data makes this concern more of an issue in this context. Finally, Google Trends data are reported as relative (not absolute) frequency of searches, which can make interpretation of the data difficult. For example, say the relative frequency of searches for "the Secretary of Education" doubled after some event. However, since the base rate is not known, searches may have doubled from 500,000 to 1,000,000 or they may have only doubled from 500 to 1,000, which would lead researchers to draw significantly different con-

clusions. Researchers using Google Trends should take care to collect the appropriate data for their given research question.

A third source of political language is online archives. Language analysis is particularly useful for studying political figures themselves, as text is often some of the only psychological data available about them. Some contexts have much more data available. For example, near-complete archives are available for the American presidency, but archives of local leaders such as mayors or archives of world leaders of smaller nations are more difficult to find and are often more limited. Depending on the research question and psychological construct under study, texts from different types of leaders and/or context may be most appropriate. In the following subsections, I detail some of the most common data available for different levels of leadership, as well as various national contexts.

As the most visible political figures, executive leaders (e.g., presidents, prime ministers, or governors) often yield the greatest amount of text data. From "stump" speeches to interviews to yearly addresses, executive leaders consistently produce political language that researchers can analyze to understand various psychological and other processes. However, the context of the text is important to consider, as different types of text may reveal different processes about the leader. For example, a prepared address reveals different information about a leader than an interview or debate. A prepared address is likely to be more formal and may reflect strategic decisions or a speechwriter's style. An interview or debate may still be strategic but is likely to be more informal and reflect the leader's own traits and style. Psychologically, prepared speeches and other written texts are likely to be more influenced by impression management considerations and likely reflect more about the psychology of the leader's team or party than the leader him- or herself.

One of the largest sources for political texts in the American Presidency Project compiled by Gerhard Peters and John Woolley (2019) at the University of California, Santa Barbara. This online repository contains thousands of texts written (or spoken) by U.S. presidents and other executive of-

ficials throughout the entire course of U.S. history (1789–present). The archive contains documents from a wide variety of contexts including written, formal declarations (e.g., letters, executive orders, written statements); spoken, formal addresses (e.g., State of the Union addresses, inaugural addresses, convention speeches); and spoken, off-the-cuff remarks (e.g., debates, interviews, press conferences). While generally not as extensive, archives of leaders from other nations are available online. Some archives include mainly speeches or texts from campaigns such as the Election Speeches Archive (Museum of Australian Democracy, 2016), while other repositories include speeches and texts from wider governing contexts such as the British Political Speech Archive (Finlayson & Atkins, 2017). Researchers interested in executive leadership have an extensive literature from which to draw, particularly in the U.S. context (Bond et al., 2017; Braun, van Swol, & Vang, 2015; Crew & Lewis, 2011; Dyson & Preston, 2006; Jarvis, 2004), but also in other national and international contexts (Blaxill, 2013; Dalvean, 2017; Kaufer & Hariman, 2008). One of the most promising future directions for researchers in this area is to compile more extensive archives of multinational leadership to grow the field beyond reliance on the American context.

While not as visible, legislators are important actors in most political systems. The main source of text data from legislators is congressional or parliamentary records. Nearly all legislative bodies in developed countries have an online repository of debates or speeches given during legislative sessions. The number of text available from legislators tends to be larger, but they are more limited in context, as they are generally speeches in legislative sessions (or committees) or formal texts, such as bills. Other types of texts are available, including campaign debates, stump speeches, and interviews, but they are not systematically available and typically are not stored online in a central repository or archive. Some corpora of legislative texts require some knowledge of web scraping to collect. However, many corpora are available in a downloadable format, including the U.S. Congressional Record and U.K. Parliament (Jordan, 2019a). Several studies have been conducted

using legislative texts for which data have been made available or could be requested (Grimmer, 2010; Holtzman, Kwong, & Baird, 2015; Kriner & Shen, 2014, Neiman, Gonzalez, Wilkinson, Smith, & Hibbing, 2016; Frimer, Aquino, Gebauer, Zhu, & Oakes, 2015). Unlike executive leaders whose language is useful for understanding individual-level processes, legislators' language reveals more group-level processes, as they must work together in larger groups on behalf of their region, party, and/or factions. As online archives at all levels of government become more extensive, political texts from diverse samples and context will further our understanding of political leaders beyond the executive and national levels. Though I have focused here on national-level leaders, another important area going forward is to develop archives and datasets looking at not only lower level political leaders but also judges, bureaucrats, and political interest groups. Research on such political actors is necessary to fill gaps in our understanding of political processes.

Often, what the average citizen learns about political figures and their actions is filtered and framed by media sources. Hence, an important set of research questions revolves around what the media say about politics and how it is framed. Such questions are often studied through surveys and experiments manipulating media messages to test their effects. However, another avenue to address this set of questions is to directly measure media messages through language analysis. Particularly when combined with additional data such as public opinion polls, language analysis of real media stories can provide a more ecologically valid look at the real-world impact of the media on political processes and attitudes. Many outlets such as the *New York Times* (NYT) and CNN maintain outlines archives of articles or transcripts that can be sampled manually or using Web scraping scripts. Several studies have been conducted using media articles or transcripts (Baum & Groeling, 2008; Fetterman, Boyd, & Robinson, 2015; Freedman, Fico, & Love, 2007; Neiman et al., 2016), which are useful resources for developing research questions and building corpora.

In one of my studies, I used two archives of two news sources, CNN and NYT, to investigate the relationship between media coverage of presidential candidates and election outcomes. Behind the project was a major research question: Are there words the media uses to describe presidential candidates that are associated with success or failure? While a long history of communication scholarship has studied this question (Gunther & Christen, 2002; Son & Weaver, 2005), I approached the question from a psychological perspective, considering how the media create mental representations of presidential candidates in the minds of voters by the words associated with the candidate (i.e., what were the 15 words used before and after the candidate's name?). Table 6.3 provides a snapshot of the data showing the types of issues associated with winning and losing candidates over several election cycles. The strongest finding was that candidates deemed by the media as more important (by mentioning their name more) are more likely to win, and candidates surrounded by scandal are likely to lose (Jordan, 2020). Beyond demonstrating how media data can be used, this study highlights the potential value of bringing a psychological perspective to the work of other disciplines.

The previous sections are not exhaustive lists of the potential corpora available in the context of political psychology. Advances in technology and the ongoing explosion of data availability continue to make larger or novel corpora available. From the digitizing of historical documents to new forums and technologies used in the course of the political process, text as data presents an ever-involving area for research in political psychology (see Table 6.4). The previous sections simply give a sense of the data that are now or may soon be available depending on the specific research questions.

■ WHAT'S NEXT?: FUTURE DIRECTIONS IN POLITICAL PSYCHOLOGY AND LANGUAGE

A major criticism of psychological research in recent years has been the historical reliance on WEIRD (Western, educated, industrialized, rich, democratic) populations (Henrich, Heine, & Norenzayan, 2010). Text as data, particularly in the context of political psychology, is uniquely situated

TABLE 6.3. Topics (Words) Uniquely Associated with Each Candidate

	Winner	Loser
2000	Issues (tax, education, abortion, social security), Places (Texas, SC, Michigan, Philadelphia, Arizona)	Issues (economy, health care), Voters (union, black, labor), Fundraising, Events (speech, interview)
2004	Record (Iraq, administration, official, commission, budget)	Scandal (Vietnam, veteran, wife), Campaign (poll, debate, advisor, nomination)
2008	Race (White, Black), Support (endorse, young, together, people)	Issues (economy, war, military, policy, tax, security), Campaign Strategy (mate, nominee, advisor)
2012	Record (administration, official, Iran, military, Congress, court)	Scandal (Bain Capital, tax return), Campaign Strategy (mate, pick, plan, advisor)
2016	Controversial Statements (Twitter, Mexico, judge, rig), Controversial Ties (Putin, university, business, organization), Events/Media Coverage (said, Texas, Fox, interview, comment, news)	Scandal (e-mail, FBI, server, investigation, private, foundation), Campaign Strategy (aide, margin, poll, trail, battleground, Black, college)

as a method for branching out beyond traditional samples of American undergraduate students. As I indicated in the first section of this chapter, numerous data sets are available for bringing the study of political psychology outside the laboratory and even outside the American context. Though the majority of studies cited in this chapter have relied on American samples, several have not and have included studies of political contexts and figures in Australia (Dalvean, 2017), Canada (Jordan, Sterling, Pennebaker, & Boyd, 2019; Suedfeld, Conway, & Eichhorn, 2001), the Middle East (Conway, Suedfeld, & Clements, 2003; Pennebaker, 2011), and the UK (Blaxill, 2013; Spirling, 2016), to name a few.

The spread of Internet/online communication and online archiving of political texts in most nations makes it increasingly easy to expand studies beyond an American context. Additionally, recent advances in machine translation makes it easier to analyze non-English texts. Windsor, Cupit, and Windsor (2019) demonstrated that, in political contexts, using Google Translate on non-English texts yield comparable text analytic results to traditional translation. Expanding into samples outside Western contexts is necessary to gain a more comprehensive understanding of people in political contexts.

A second promising area in political psychology is cross-disciplinary work. The realm of politics encompasses more than psychological issues; there are also rhetoric, institutional, history, and information issues. In fact, much of the work cited in this chapter was not conducted solely by psychologists; it was also conducted by communication scholars, political scientists, and computer scientists. To truly understand people and political processes, numerous perspectives need to be considered. Going forward, political psychologists might work to seek out collaborations with persons in other disciplines, with expertise in other aspects of the political process. Specifically, political psychologists would likely benefit from collaborations with communication scholars and political scientists. Communication scholars have a much longer history of working with text as data and may therefore be able to contribute data, in the form of existing archives and corpora. Political scientists have a greater understanding of the institutional norms and procedures that may have important implications for political language. For example, the Congressional Record is a rich source of political text data, but the rules and regulations of how and when members of Congress can speak may influence the conclusions drawn from

TABLE 6.4. Types of Text Data Available for Political Psychology Questions

Population	Contexts	Example texts	Selection of available archives
Voters/ citizens	Surveys	Open-ended questions	American National Election Studies (*electionstudies.org*)
	Social media	Tweets, Reddit posts, Facebook groups	Twitter API (*developer.twitter.com*), Google BigQuery (Reddit, *cloud.google. com/bigquery/public-data*)
	Other Internet sources	Search terms	Google Trends (*trends.google.com*)
Political figures	Executive leaders (e.g., presidents, prime ministers, governors)	Formal addresses (e.g., Inaugural addresses, State of the Union address, Prime Minister's Questions, interviews, debates, campaign speeches, press conferences/releases	American Presidency Project (*presidency.ucsb.edu*), British Political Speech Archive (*britishpoliticalspeech. org*), Australian Prime Ministers Centre (primeministers.moadoph.gov.au), Legislative Reference Library of Texas (*lrl.texas.gov/legeleaders/governors*)
	Legislators (e.g., members of Congress or members of Parliament)	Floor speeches or debates, press releases, campaign communications (e.g., speeches, ads, social media)	Congressional Record (*gpo.gov*), Hansard (UK; *api.parliament.uk/ historic-hansard*), Open Parliament (Canada; *openparliament.ca*), EUSpeech (Schumacher et al., 2016)
	Judicial officials	Opinions	U.S. Supreme Court (*supremecourt.gov*)
	Local officials	Council meetings, mayoral addresses	
The media	Newspapers	*NYT, Wall Street Journal,* local newspapers	
	Cable/broadcast News	CNN, Fox News, MSNBC, CBS news programs	
	News websites	Daily Kos, The Blaze	
Institutions/ organizations	Activist groups	Nonprofit groups, political action committees	
	Political parties	Manifestos, party platforms, convention speeches	American Presidency Project

psychological text analysis. If a researcher wanted to study cross-party aggression using the Congressional Record, they might fail to find evidence of such aggression due to rules and norms restricting members of Congress from verbally attacking their colleagues. Politics is complex, encompassing people, institutions, history, and technology; more multidisciplinary work is necessary to truly have a holistic understanding of politics.

■ CONCLUSION

Political scientists and communication scholars have long studied language in poli-

tics, but many questions remain unanswered when it comes to the psychology of political language. Advances in technology and the Internet continue to open new avenues for studying political psychology with text. Text data and text analytic methods provide political psychologists with an unprecedented opportunity to understand political processes across time and place, and to explore how politics impacts people in their everyday life. Language analysis can be challenging, but for those political psychologists interested in text data, this chapter (and its numerous references) should provide at least a starting place for doing this type of research.

■ REFERENCES

Abe, J. A. A. (2012). Cognitive–affective styles associated with position on war. *Journal of Language and Social Psychology, 31*(2), 212–222.

American National Election Studies. Retrieved January 27, 2020, from *www.electionstudies.org.*

Amsalem, E. (2019). How informative and persuasive is simple elite communication?: Effects on like-minded and polarized audiences. *Public Opinion Quarterly, 83*(1), 1–25.

Arceneaux, K. (2012). Cognitive biases and the strength of political arguments. *American Journal of Political Science, 56*(2), 271–285.

Baum, M. A., & Groeling, T. (2008). New media and the polarization of American political discourse. *Political Communication, 25*(4), 345–365.

Blaxill, L. (2013). Quantifying the language of British politics, 1880–1910. *Historical Research, 86*(232), 313–341.

Bond, G. D., Holman, R. D., Eggert, J. A. L., Speller, L. F., Garcia, O. N., Mejia, S. C., . . . Rustige, R. (2017). "Lyin' Ted," "Crooked Hillary," and "Deceptive Donald": Language of lies in the 2016 US presidential debates. *Applied Cognitive Psychology, 31*(6), 668–677.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature, 489*(7415), 295–298.

Braun, M. T., van Swol, L. M., & Vang, L. (2015). His lips are moving: Pinocchio effect and other lexical indicators of political deceptions. *Discourse Processes, 52*(1), 1–20.

Cohen, J. E. (2016). Presidential attention focusing in the global arena: The impact of international travel on foreign publics. *Presidential Studies Quarterly, 46*(1), 30–47.

Conway, L. G., Conway, K. R., Gornick, L. J., & Houck, S. C. (2014). Automated integrative complexity. *Political Psychology, 35*(5), 603–624.

Conway, L. G., III, Suedfeld, P., & Clements, S. M. (2003). Beyond the American reaction: Integrative complexity of Middle Eastern leaders during the 9/11 crisis. *Psicologia Politica, 27,* 93–103.

Crew, R. E., Jr., & Lewis, C. (2011). Verbal style, gubernatorial strategies, and legislative success. *Political Psychology, 32*(4), 623–642.

Dalvean, M. (2017). Changes in the style and content of Australian election campaign speeches from 1901 to 2016: A computational linguistic analysis. *ICAME Journal, 41*(1), 5–30.

Dyson, S. B., & Preston, T. (2006). Individual characteristics of political leaders and the use

of analogy in foreign policy decision making. *Political Psychology, 27*(2), 265–288.

Fetterman, A. K., Boyd, R. L., & Robinson, M. D. (2015). Power versus affiliation in political ideology: Robust linguistic evidence for distinct motivation-related signatures. *Personality and Social Psychology Bulletin, 41*(9), 1195–1206.

Finlayson, A., & Atkins, J. (2017). British Political Speech Archive. Retrieved from *www.britishpoliticalspeech.org/speech-archive.htm.*

Freedman, E., Fico, F., & Love, B. (2007). Male and female sources in newspaper coverage of male and female candidates in US senate races in 2004. *Journal of Women, Politics and Policy, 29*(1), 57–76.

Frimer, J. A., Aquino, K., Gebauer, J. E., Zhu, L. L., & Oakes, H. (2015). A decline in prosocial language helps explain public disapproval of the US Congress. *Proceedings of the National Academy of Sciences USA, 112*(21), 6591–6594.

Gooch, A. (2018). Ripping yarn: Experiments on storytelling by partisan elites. *Political Communication, 35*(2), 220–238.

Google. (2020). Google Trends. Retrieved from *https://trends.google.com/trends/?geo=US.*

Greenberg, D. (2016). *Republic of spin: An inside history of the American presidency.* Norton.

Grimmer, J. (2010). A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases. *Political Analysis, 18*(1), 1–35.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science, 363*(6425), 374–378.

Gunther, A. C., & Christen, C. T. (2002). Projection or persuasive press?: Contrary effects of personal opinion and perceived news coverage on estimates of public opinion. *Journal of Communication, 52*(1), 177–195.

Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes, 45*(1), 1–23.

Hart, R. P., Childers, J. P., & Lind, C. J. (2013). *Political tone: How leaders talk and why.* University of Chicago Press.

Hartman, T. K. (2012). Toll booths on the information superhighway?: Policy metaphors in the case of net neutrality. *Political Communication, 29*(3), 278–298.

Hauser, D. J., & Schwarz, N. (2018). How seemingly innocuous words can bias judgment: Semantic prosody and impression formation. *Journal of Experimental Social Psychology, 75,* 11–18.

Helpful tools for researchers. (n.d.). Retrieved June 18, 2020, from *https://developer.twitter.com/en/use-cases/academic-researchers/helpful-tools*.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*(7302), 29–29.

Holtzman, N. S., Kwong, S., & Baird, K. L. (2015). Exploring political ideologies of senators with semantic analysis tools: Further validation of CASS. *Journal of Language and Social Psychology, 34*(2), 200–212.

Jarvis, S. E. (2004). Partisan patterns in presidential campaign speeches, 1948–2000. *Communication Quarterly, 52*(4), 403–419.

Jessee, S. A. (2010). Partisan bias, political information and spatial voting in the 2008 presidential election. *Journal of Politics, 72*(2), 327–340.

Jordan, K. N. (2019a). Long-term trends in politics and culture. Retrieved from *https://osf.io/xhtf9*.

Jordan, K. N. (2019b). Power, friends, and success: What drives the Democratic front-runners? [Blog post]. Retrieved from *https://wordwatchers.wordpress.com/2019/08/01/power-friends-and-success-what-drives-the-democratic-front-runners*.

Jordan, K. N. (2020). *Understanding political leaders and political processes through language.* Unpublished doctoral dissertation, University of Texas at Austin.

Jordan, K. N., Pennebaker, J. W., & Ehrig, C. (2018). The 2016 US presidential candidates and how people tweeted about them. *SAGE Open, 8*(3). [Epub ahead of print] *https://journals.sagepub.com/doi/10.1177/2158244018791218, 2158244018791218*.

Jordan, K. N., Sterling, J., Pennebaker, J. W., & Boyd, R. L. (2019). Examining long-term trends in politics and culture through language of political leaders and cultural institutions. *Proceedings of the National Academy of Sciences USA 116*(9), 3476–3481.

Kaufer, D., & Hariman, R. (2008). Discriminating political styles as genres: A corpus study exploring Hariman's theory of political style. *Text & Talk: An Interdisciplinary Journal of Language, Discourse Communication Studies, 28*(4), 475–500.

Kriner, D., & Shen, F. (2014). Responding to war on Capitol Hill: Battlefield casualties, congressional response, and public support for the war in Iraq. *American Journal of Political Science, 58*(1), 157–174.

Markey, P., & Markey, C. (2011). Pornography-seeking behaviors following midterm political elections in the United States: A replication of the challenge hypothesis. *Computers in Human Behavior, 27*(3), 1262–1264.

Markus, G. B. (1992). The impact of personal and national economic conditions on presidential voting, 1956–1988. *American Journal of Political Science, 36*, 829–834.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences USA, 114*(48), 12714–12719.

McKinney, M. S., & Carlin, D. B. (2004). Political campaign debates. In L. L. Kaid (Ed.), *Handbook of political communication* (pp. 203–234). Erlbaum.

Milkis, S. M., & Nelson, M. (2011). *The American presidency: Origins and development, 1776–2011.* CQ Press.

Murphy, S. C. (2017). A hands-on guide to conducting psychological research on Twitter. *Social Psychological and Personality Science, 8*(4), 396–412.

Museum of Australian Democracy. (2016). Election speeches archive. Retrieved from *https://electionspeeches.moadoph.gov.au*.

Nau, C., & Stewart, C. O. (2014). Effects of verbal aggression and party identification bias on perceptions of political speakers. *Journal of Language and Social Psychology, 33*(5), 526–536.

Neiman, J. L., Gonzalez, F. J., Wilkinson, K., Smith, K. B., & Hibbing, J. R. (2016). Speaking different languages or reading from the same script?: Word usage of Democratic and Republican politicians. *Political Communication, 33*(2), 212–240.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665–675.

Pennebaker, J. W. (2011). Using computer analyses to identify language style and aggressive intent: The secret life of function words. *Dynamics of Asymmetric Conflict, 4*(2), 92–102.

Peters, G., & Wooley J. T. (2019). The American Presidency Project. Retrieved from *www.presidency.ucsb.edu/index.php*.

Petrocik, J. R. (1996). Issue ownership in presidential elections, with a 1980 case study. *American Journal of Political Science, 40*, 825–850.

Sagi, E., & Dehghani, M. (2014). Measuring moral rhetoric in text. *Social Science Computer Review, 32*(3).

Schumacher, G., Schoonvelde, M., Traber, D., Dahiya, T., & De Vries, E. (2016). EUSpeech: A New Dataset of EU Elite Speeches. In *Proceedings of the International Conference on the Advances in Computational Analysis of*

Political Text (pp. 75–80). University of Zagreb.

Siegel, A., Nikitin, E., Barberá, P., Sterling, J., Pullen, B., Bonneau, R., . . . Tucker, J. A. (2021). Trumping hate on Twitter?: Online hate in the 2016 U.S. election and its aftermath. *Quarterly Journal of Political Science, 16*(1), 71–104.

Simonton, D. K. (2003). Qualitative and quantitative analyses of historical data. *Annual Review of Psychology, 54*(1), 617–640.

Son, Y. J., & Weaver, D. H. (2005). Another look at what moves public opinion: Media agenda setting and polls in the 2000 U.S. election. *International Journal of Public Opinion Research, 18*(2), 174–197.

Spirling, A. (2016). Democratization and linguistic complexity: The effect of franchise extension on parliamentary discourse, 1832–1915. *Journal of Politics, 78*(1), 120–136.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics, 118,* 26–40.

Suedfeld, P., Conway, L. G., III, & Eichhorn, D. (2001). Studying Canadian leaders at a distance. In O. Feldman & L. O. Valenty (Eds.), *Profiling political leaders: Cross-cultural studies of personality and behavior* (pp. 3–19). Praeger/Greenwood.

U.S. Supreme Court (2020, January). Supreme Court of the United States. Retrieved from *supremecourt.gov*.

Windsor, L., Cupit, G., & Windsor, A. (2019). Automated content analysis across six languages. *PLOS ONE, 14*(11), e0224425.

Winter, D. G. (2011). Philosopher-king or polarizing politician?: A personality profile of Barack Obama. *Political Psychology, 32*(6), 1059–1081.