Title

The Way EU Make Me Feel: Measuring Anxiety in the Brexit Negotiations Using Text and Audio

Author

Kappos, Cybele

Publication Date

2024

UNIVERSITY OF CALIFORNIA

Los Angeles

The Way EU Make Me Feel:
Measuring Anxiety in the Brexit Negotiations Using Text and Audio

A dissertation submitted in partial satisfaction
of the requirements for the degree Doctor of Philosophy
in Political Science

by

Cybele Kappos

2024

ABSTRACT OF THE DISSERTATION

The Way EU Make Me Feel:

Measuring Anxiety in the Brexit Negotiations Using Text and Audio

by

Cybele Kappos

Doctor of Philosophy in Political Science

University of California, Los Angeles, 2024

Professor Margaret E. Peters, Co-Chair

Professor Jeffrey B. Lewis, Co-Chair

Politics make us anxious. People experience anxiety – defined here as a fearful uncertainty about the future course of events – when thinking about political issues such as climate change and the economy, ahead of important elections, and even when hearing the voices of certain political figures. Yet, scholars have devoted little attention to the study of this critical emotion. We know that politics make us feel anxious, but what does the language of political anxiety *look* and *sound* like? By understanding the language of anxiety, we may someday be able to understand how political elites use anxiety as a tool to persuade colleagues and constituents alike. Using the Brexit negotiations (2016-2020) as a case study, I develop a methodological tool that measures anxiety and emotional intensity in elite political rhetoric using text and audio data. I develop a dictionary of anxiety that scores speeches based on the semantic similarity to a sample of highly anxious words. With this dictionary, I am able to study how anxiety varies with party affiliation and the topic of the speech. I then examine a large subset of the data in audio format using pitch as a proxy for emotional intensity. This novel approach combines the measure of anxiety in text and emotional intensity in audio, constructing a fuller picture of the speaker's emotional state. The composite measure reveals how expressions of emotion differ according to the role of the speaker in a meeting and their party affiliation.

The dissertation of Cybele Kappos is approved.

Georgia C. Kernell

Kathleen Bawn

Margaret E. Peters, Committee Co-Chair

Jeffrey B. Lewis, Committee Co-chair

University of California, Los Angeles

2024

To James, Alfie, and Toby, of course

Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

Graduate school has not been an easy journey. When I first entered the program in Fall 2018, I was one of two political theorists in my cohort. After taking Kathy Bawn's course on probability (200A) and Erin Hartman's course on linear regression (200B), I sensed a change of heart about my subfield. I spent many (many) long (long) days poring over the 200B problem sets which, over the course of the quarter, felt less impossible to complete. Despite being pushed to my limits in this course, I had discovered a new passion. R. I could do so many new things with this language including... measure emotion in text?!

After my first year, I spent the whole summer thinking about my place in the program. In August, I decided to email Davide Panagia, letting him know I would be switching over to comparative politics and methods. Davide responded to my email with such graciousness, expressing joy for my having found my path. To Davide I will always be immensely grateful for being the first person in the department to greet me with a beaming smile and for always listening.

The story of my time at UCLA would not be complete without mentioning the three women who have been my cheerleaders since the beginning; Jieun, Olivia, and Chelsea. Jieun, our paths have been interwoven from the very start. Together, we have survived the constant challenges and inequities that have stood in our way of this program. I am so grateful that you have always understood me. I could not have survived this program without you and am bound to you forever. Olivia and Chelsea, thank you for always reminding me that I am capable of many great things. Our many joyful moments together have been a source of light as we lived through our first years in the program, a global pandemic, and many trials in our lives. I am immensely grateful for your friendship.

I want to express my gratitude to Erin Hartman, who helped shaped me as a scholar. Erin's course on linear regression opened up a whole new world of political science for me. Erin always champions young methodologists, especially young women. She was one of the first people I brought the idea of doing text analysis to and she immediately set me on a path

to explore the many possibilities of text. Erin also mentored me as a I began to consider career paths and think about what I wanted to do. As I explore the non-academic path, Erin has helped me build confidence and connections as I embark on this journey. Thank you, Erin.

Kathy, thank you for teaching me how to teach. Teaching for Kathy Bawn was an immense joy. I feel lucky to have witnessed how to get a course of 120 students engaged with game theory. Kathy has also been such a steady believer in me over the years and often saw thing clearer than I could. I am so grateful to have had a mentor like you and so happy you were a part of my committee. The dissertation is a winding road and you always kept me focused on the bigger picture, the story, and the contribution of the work. Thank you, Kathy.

I would also like to thank Georgia Kernell, a crucial member of my dissertation committee. Georgia taught me how to write an academic paper. It is a strange thing to learn and relearn how to write over the course of one's life. In a pursuit of clarity and linearity in my academic writing, Georgia always gave me granular feedback (down to a missing period in a footnote) that helped me direct my dissertation and find my voice. I deeply appreciate how Georgia was always excited for my work, especially when I was too tired to see the importance of the work. Thank you, Georgia.

I must, of course, thank Jeff Lewis for being part of my life for the past six years. Jeff was my co-chair and second advisor. Jeff can probably talk about my dissertation better than I can and can effortlessly piece together what to me, at times, seems like chaos. Whether stopping by Jeff's office for a quick chat (and always a laugh) or to ask him to solve a coding problem that I had been struggling with for days (only for Jeff to solve it in one hour), I always appreciated having you there. Thank you, Jeff.

To the professor who has done the most for me, Maggie Peters. I cannot thank you enough. Maggie agreed to take me on as an advisee at the start of my second year after I had switched subfields. She took my international student status seriously, helping me to

get on track to advance to candidacy by the end of my third summer. What always struck me with Maggie was how calm she seemed about any big worry I expressed, whether it was a conference presentation, my dissertation prospectus, or my immense workload in my final year. Always responding with “You’ll be fine” or “You can do it”, I was at first skeptical, but then realized that indeed I *was* fine and I *could* do it. Maggie hired me as a research assistant when I had no summer funding in 2021, and has since had me on her team the whole time. Maggie, I sincerely appreciate how steady you’ve been all these years. This program would not have been the same without you. Thank you so much, Maggie.

I want to acknowledge my family, Natalia, Makis, Serap, and James, without whom I would not be where I am today. Thank you for always, always believing in me. I have had so much doubt over the years and very little faith that I was in the right place. With your support, it all turned out okay. Picking up my countless phonecalls, often with tears, has made the impossible possible. I love you so much and take so much strength in seeing how highly you think of me. Finally, I must thank James (and Alfie and Toby), without whom this final year simply would not have been possible. You have done everything for me, from emotional support, to taking care of our home, and always making me laugh. I am forever grateful for what you’ve done for me this past year.

I want to acknowledge all others who have made this journey possible over the years including Esther Blair (I’m so sorry for how many emails I’ve sent over the years), Belinda Sunnu and Barbara Wakasa (thank you both for always being so cheerful), Parker and Nicole (my reprieve from grad school), Rachel (who gave me strength), and so many others I am currently forgetting to mention. Thank you all. This was a group effort.

VITA

2017 B.A. (Politics), Pomona College

2021 M.A. (Political Science), University of California - Los Angeles

2024 M.S. (Statistics), University of California - Los Angeles

1 Introduction

The 2020 presidential debate was a stressful spectacle of Trump and Biden interrupting and yelling at each other. One journalist described the debate as "the most emotional in history", with little to no "rational" discussion (Filipovic, 2020). In an empty room, Eric Garcetti, the former mayor of Los Angeles, publicly announced that the city will need to furlough thousands of city workers because of the economic impact of the COVID-19 pandemic. "'We are bowed and we are worn down. We are grieving our dead,' the mayor continued, choking back tears" (Zahniser, Smith, and Reyes, 2020). In an altercation between Rep. Ted Yoho and Rep. Alexandria Ocasio-Cortez, Yoho confronted Ocasio-Cortez over a discussion about the connection between poverty and crime rates. He used offensive and profane language when speaking to his colleague. AOC made a floor speech condemning Yoho's behavior and the use of sexist, violent language (Sprunt, 2020). Yoho responded by saying "I cannot apologize for my passion" (ibid.). Politics are emotional.

Theories about the role of emotion in political communication span the history of political theory from the works of thinkers like Plato and Aristotle to more modern theorists like Hume and St. Augustine. Some of the guiding questions of these philosophical inquiries are: How should emotions inform our attitudes toward politics? How do politicians elicit emotional reactions in their audience? Which emotions activate political engagement? In Aristotle's *Rhetoric* (2001), emotions are discussed as a mode of political persuasion. Rhetoric has three modes of persuasion, (1) the personal character of the speaker, (2) putting the audience into a certain frame of mind, and (3) the apparent proof of the words of the speech itself. The second mode, putting the audience into a frame of mind, refers to the ability of the speaker to stir the emotions of the audience. The audience's emotional reaction to the rhetoric as well as their emotional predisposition both affect whether they will be persuaded by the argument.[1] Aristotle's theory shows an understanding that emotions color

[1] Aristotle delves into a few key emotions such as anger, calm, and fear. Although he does not develop a taxonomy as such, Aristotle still identifies key emotions that are associated with a negative (fear) or positive

every aspect of the process of political communication, from the disposition of the listener, to the intentions of the speaker, to the delivery of the speech.

Contemporary political scientists have continued to pursue a social scientific understanding of emotions. In an effort to understand emotion from a quantitative standpoint, scholars have constructed tools to measure the emotional content of political communication (Young and Soroka, 2012; Erisen and Villalobos, 2014; Tarr, Hwang, and Imai, 2023; Rheault and Cochrane, 2020). Much of this scholarship uses text, which is a rich source of data. Text data includes sources such as presidential speeches, campaign advertisements, floor speeches and manifestos. Scholars synthesize quantitative metrics like word count and word co-occurrence to capture the importance of words, topics, and concepts within the set of texts. Measuring emotion in text, often referred to as sentiment analysis, has most often been achieved through the creation of dictionaries (Cochrane et al., 2022; Rheault, 2016; Rheault et al., 2016; Osnabrügge, Hobolt, and Rodon, 2021). Dictionaries are a set of words or terms that are related to a given emotion (e.g. a dictionary of negative terms may include the terms "aggression", "hateful", and "anxious"). Using these types of tools, scholars can understand which emotions are present in a set of texts and, using additional data such as the speaker's political affiliation, explain when and why certain emotions are present in the data.

More recently, scholars have started to explore the potential of audio data for the study of emotions in political communication (Dietrich, Enos, and Sen, 2019; Dietrich, Hayes, and O'Brien, 2019; Dietrich, Schultz, and Jaquith, 2018; Knox and Lucas, 2021). This data is distinct from, but related to, text. Consider political campaigns, floor speeches, or the state of the union addresses. Each of these has an audio component and a text component, both of which are integral to how a listener listens to the speech. Incorporating audio contributes to a more complete understanding of political communication. While the data is notably more complex and computationally expensive, measures as simple as pitch have been shown to be

(pleasure) evaluation of a situation.

related to a speaker's emotional attachment to a political issue. Currently, political science has a very limited sense of how to computationally analyze audio data and, as a result, the promises of audio data have yet to be fully demonstrated. Borrowing methodologies from fields such as linguistics can allow us to expand the uses of audio data in political science.

In this dissertation, I propose an approach to measuring emotion in elite political rhetoric that examines both audio and text data. The central research questions are **How can we capture emotion in elite political rhetoric?** and **What can audio data contribute to our understanding of emotion that text cannot?** This dissertation pursues an understanding of the dynamics of emotion and more specifically anxiety through the use of natural language processing and audio analysis tools.

We can represent political speech through three dimensions, (1) text (2) audio and (3) visuals. Dimensions (1) and (2) capture different aspects of speech, namely *what* a person says and *how* they say it, respectively. By incorporating audio, which is a lesser used source of data in political communication, I hope to produce a more complete measurement of the level of emotion in political speech. This dissertation contributes to the areas of political communication and natural language processing methods in two main ways; 1) it advances the nascent area of political audio analysis and relates the findings back to text analysis methods, 2) it contributes to our understanding of how anxiety, an emotion which has not received as much attention in the field, operates in politics at the elite level. In doing so, this project tackles the fickle topic of emotions in politics while using multiple types of data and new methodological approaches.

1.1 What is emotion and why study it?

By understanding emotion to be at odds with reason, political scientists have not always known how to explain how emotion factors into political decision making. Traditionally, rationality has been the focus on political science, most often within the framework of Rational Choice Theory (RCT). RCT understands political behavior as "purposive action, consistent preferences, and utility maximization" (MacDonald, 2003). Political actors evaluate a set of actions and choose the courses of action that maximize their utility. RCT is not an indisputable tenet of the field; the validity of the model and the assumptions scholars make are debated, challenged, and criticized (Friedman, 1996; Brown et al., 2000; Goldthorpe, 1998; Smelser, 1992).

In recent decades, scholars have developed and tested important, influential theories about the role of emotion in political behavior. These theories of emotion complement, rather than replace, RCT. There is no single theory of emotion that is universally agreed upon; scholars debate which theory offers the most accurate and comprehensive taxonomy of emotions (Marcus, 2023). Despite these disagreements among scholars, emotions are universally treated as a response to a given situation. An emotional response is defined as a physiological and mental appraisal of a situation. Subsequently, mental processes move to the conscious level, where we deliberate, assess, and evaluate options. In other words, emotion precedes and informs reason. The implications of emotion on political behavior include responses to political rhetoric (Bakker, Schumacher, and Rooduijn, 2021), emotional connections to issues (e.g. anger towards the European Union) and subsequent voting behavior (Vasilopoulou and Wagner, 2017), and strength of partisanship (Huddy, Mason, and Aarøe, 2015). Anger, fear, surprise, joy, and anxiety are some of the key emotions that factor into our experience and understanding of politics.

Experiencing anxiety and remaining engaged in politics go hand-in-hand. To be even moderately informed on political events likely comes with a baseline level of anxiety. Contemporary politics – wars, climate change, housing crises, genocide, gun violence, threats to healthcare, rolling back protective rights, increasing polarization, the rising cost of living – leave many of us feeling an overwhelming loss of control.

1.1.1 Defining and Understanding Anxiety: Affective Intelligence Theory

The term "anxiety" has various meanings in today's society, one of which is a diagnosed mental condition. In this dissertation, I refer to anxiety as an emotional state, therefore emphasizing its transience. Describing a politician's rhetoric as anxious does not imply that they are an anxious politician generally speaking. My understanding of anxiety is derived from Affective Intelligence Theory, developed by political scientist George E. Marcus in the eighties (Marcus, 1988). This theory was expanded in close connection to the field of neuroscience and has since gained significant traction and credibility in political science (Gray, 1985; Vasilopoulou and Wagner, 2017; Marcus, 2023; Huddy, Mason, and Aarøe, 2015; Huddy, Feldman, and Cassese, 2007).

Marcus's (2000) definition of anxiety exists in relation to the broader framework of the disposition system and the surveillance system. These two systems characterize how a person interacts with their environment and determines their course of action. The *disposition system* is a comparison system. Its job is to ensure that an individual's actions in a situation are going according to plan. A course of action is executed and the disposition system compares the expected outcome to the status of the situation. The individual intakes somatosensory information; information about where the body is, its position, and its status. Based on this information, the system evaluates the success of the plan. Habitual behavior occurs when conditions are familiar to a person and their course of action is successful.

The *surveillance system's* job is to monitor threatening or unusual circumstances. Is the plan being executed in a successful fashion or is the plan failing? Can a course of action be decided? When we can no longer rely on habitual behavior because there is no predetermined action that corresponds to the event or because the intended course of action is failing, this signals to our mind that something threatening or unusual is happening. Anxiety identifies the times when we should cease operating out of habit and engage our minds to evaluate the situation. The surveillance system is responsible for interrupting habitual behavior.

Marcus's framework is used to describe how individuals operate in a given situation, including in a political context. The disposition system can be used to understand that habitual behavior is an important part of how we exist in a political space. Examples of political habits are voting out of partisan allegiance and remaining committed to opinions on issues like the climate crisis or reproductive rights. The surveillance system relates to how we respond to threats or unusual activity in our political surroundings, such as unconventional candidates for president like Donald Trump, or unanticipated developments like the overturning of Roe v Wade.

Affective Intelligence Theory has important implications including political attitudes, information processing, and voting. Huddy, Feldman, and Weber (2007) operationalize Marcus's theory, examining the effects of anger and anxiety on attitudes towards the Iraq war. They find that anxious individuals experience a heightened sensitivity and attention to threat, tend to overestimate risks, and engage in more careful information processing. Anxious individuals are therefore more risk-averse, since they feel a lack of control over the situation causing anxiety. In a different empirical study, Vasilopoulou and Wagner (2017) examine how emotions like anger and anxiety related to voting behavior during Brexit. They found that angry voters were more likely to vote to leave the EU. Anxious voters, as more risk-averse and fearful of the future, were more likely to vote to remain in the EU.

This dissertation relies on Affective Intelligence Theory to argue that the context of Brexit was anxiety-inducing. Affective Intelligence Theory tells us that anxiety arises when a situation is characterized by (1) uncertainty, (2) risks, and (3) a lack of control over the situation. The Brexit negotiations meet these criteria. The process of exiting the EU was lengthy and fraught. Three extensions of the exit deadline (March 2019, April 2019 and October 2019) underscore the non-linearity and complexity of these negotiations. In uncharted territory, governments now had to renegotiate many important political and economic issues.

The negotiation process involved a great deal of uncertainty and risk. How key issues like the border between the UK and Northern Ireland, the economic impact on sectors like agriculture, and the masses of people who would no longer be able to reside in the UK under free movement laws would be resolved was unclear from the start. The risks involved were enormous. The relationship between Northern Ireland and the UK is fragile. British agriculture would no longer have access to funding provided by the Common Agricultural Policy (CAP) and to the free movement of goods (Commission, 2024). Other sectors (e.g., academia), too, would suffer from the inability to use EU funding. Immigration, one of the key issues of the Leave campaign, was likely to impact sectors like healthcare which depend on a flow of EU workers who, under Brexit, could work in the UK without restrictions like visas. Changes to migration laws also, however, meant a change to emigration laws, since UK citizens would no longer be able to freely live and work abroad.

The UK did not appear to be in control of the negotiation process. Under three Prime Ministers (David Cameron, Theresa May, and Boris Johnson), numerous failed votes on Brexit proposals displayed the inability of parliament to work together. To this day, the impact of Brexit reverberates in British society and in their economy. This historical event changed the course of the future of the UK and the EU.

1.2 Outline of dissertation

In chapter 2, I examine Brexit as a suitable case study for research on anxiety. Providing historical context, I discuss the history of Euroskeptic movements in the UK and some of the key issues that motivated the two campaigns of the Brexit referendum, the Leave campaign and the Remain campaign. I give an overview of how the study of Brexit relates to political science research, focusing on voter and politician behavior.

In chapter 3, I begin with a general overview of text analysis to introduce any readers who are not familiar to the technical terms and fundamentals. This is to ensure that readers who do not have prior familiarity with the methods have a basic grasp of the tenets of text analysis. I discuss the text analysis methods used in this dissertation – dictionary methods and word embeddings – in more detail. This chapter includes a review of scholarly literature that has employed these methods and additionally, a technical explanation of the algorithm that word embeddings use.

Chapter 4 discusses each dataset used in this dissertation. I explain the importance of using Committee data, as opposed to the conventional Hansard dataset which only includes parliamentary debates. I go through the text analysis methodology in detail, explaining how to produce a custom lexicon using a dataset of speeches.

Chapter 5 validates the measurement tool. I examine random samples of the vocabulary in depth to give the reader a sense of how the scale has been calibrated. I also provide a few representative examples of full speeches and compare anxiety scores to human annotations of the relative anxiety of the speeches. The bulk of this chapter details the application of the anxiety lexicon to the Brexit dataset. I provide a series of descriptive findings that include temporal analyses and results disaggregated by party. These findings reveal that the role of anxiety in British parliamentary rhetoric is associated with the topic being discussed and the party affiliation of the speaker.

In chapter 6, I transition to the audio methods and analysis. This chapter first answers why pitch is a good candidate for measuring emotions. Next, the chapter details the methods I used to transform raw audio files into timestamped transcripts. This step, although tedious, is what enables researchers to compare measurements of emotion in audio to measurements of emotion in text.

In chapter 7, I develop an approach to compare text analysis to audio analysis. These two measures capture a distinct aspect of speech. First, I conduct a qualitative analysis of one 10-minute segment of a committee meeting. This helps explore the relationship between pitch and measurements of anxiety in text. Section 7.5 applies some of the insights from the qualitative analysis to the full dataset. These analyses confirm many of the qualitative findings on the patterns of emotional arousal in committee meetings. Emotional arousal is related to the party affiliation of the speaker as well as whether the speaker is a witness or MP.

In chapter 8, I conclude the dissertation with a discussion of the findings of this novel methodological approach to studying emotion. I briefly discuss the limitations of the research and future directions for audio and text.

2 Brexit as a case study

This dissertation uses Brexit as a case study of emotion, discourse, and rhetoric. Brexit is an appropriate choice not only given the salience of emotions in the context, but of anxiety in particular.

Brexit was an unprecedented event in European Union history. While waves of Euroscepticism have appeared in several EU member states – gaining traction in many places such as Hungary (Fidesz) and Greece (Golden Dawn and its outgrowth, the Greek Solution) – the Eurosceptic movement has never succeeded as much as it did in 2016. In June 2016, when the majority of British voters voted to leave the EU, the UK, including many of those who had voted to Leave, was shocked (Sundby, 2016).

2.0.1 A brief history of Brexit and British Euroskepticism

Although a member state had never voted to leave the EU, the United Kingdom and the EU (and all the iterations of the governing body that preceded its modern-day form) had a long and tumultuous relationship. Post-WWII Britain, struggling to recover from the devastation of the war, was not immediately accepted into the European Community (the European Community later evolved into the European Union). The resistance to UK membership was led by the influential French politician, Charles de Gaulle, who vetoed Britain's attempt to gain membership. The UK was admitted into the EC in 1972 when de Gaulle was replaced by his successor, Georges Pompidou (Haag, 2017).

In the following decades, the UK would see Euroskepticism represented in mainstream politics. Although Brexit is now often associated with the Conservative party, Euroskepticism has historically existed in both of the two largest parties – the Labour party and the Conservative party – albeit for different reasons. Labor unions were an important part of the Labour constituency. Unions viewed the Common Market as unfavorable to workers and were concerned about the effect of the Common Market on employment rates. There was also concern that Britain's domestic market would suffer due to European imports (News, 1975). Labour's anti-EU sentiment even resulted in a referendum on EC membership in 1975, where 67% voted to remain (Fitzgerald, Beadle, and Rowan, 2022).

These grievances grew as the European Community expanded its powers and the UK maintained its sovereignty in some areas. The UK retained the British pound as its currency when the Euro was introduced in the 1990s. The UK also opted out of the Schengen agreement (abolishing borders within the Schengen area) thereby maintaining its national borders (Gelatt, 2005). A defining moment in the EU's history was the Lisbon Treaty, signed in 2007, where the European Community became the European Union. This was a constitutional treaty that centralized much of the power of the Union, giving bodies like the European Parliament more power in the legislative process (Parliament, 2022). Importantly, the Lisbon Treaty contained Article 50, which was a clause that established a procedure for members to leave the EU.

Several crises in the decade preceding the 2016 referendum shook the British public's faith in the EU. The 2008 financial crisis, and the debt crises that followed in countries like Greece and Italy generated a notion that EU membership was a net loss for Britain; many felt like the immense sums of money the UK spent on the EU – totalling 8.5 million pounds in 2015 – could be spent at a national level rather than bailing out weak economies (Treasury, 2015). The 2015 refugee crisis resulted in widespread fears among the public about refugees flooding in from the French border.

Figure 1 shows why the refugee crisis affected every member state of the EU. Within the Schengen area, travelers are not subjected to checks and refugees, once in Schengen territory, could migrate across Europe with relative Ease. In this figure, black lines demarcate how border fences were erected and checks implements to limit the flow of these refugees by land. Although the UK was not a member of the Schengen area, once in France, the journey to the UK via the English Channel is short. The flow of migrants illegally crossing the channel was a major concern for British immigration policy (Makortoff, 2015).

Figure 1: Refugee flows in Europe (Tasch, 2016)

From a cultural standpoint, UK citizens felt different from their fellow Europeans. A Eurobarometer survey[2] from 2015 showed that on average, UK citizens felt less like a EU citizen (11 percentage points less) than citizens from all other surveyed nations (Commission, 2015). The same survey also hints to the other main points of contention; respondents thought that immigration, the economic situation and unemployment were the top issues facing the EU at the moment (ibid.).

The 2016 referendum marked the culmination of distrust in the EU. The UK was already formally distanced from the EU in many ways (the pound, the free movement of people) and with the vote to Leave, marked its desire for independence and sovereignty.

[2]The Eurobarometer is a survey of EU member states conducted by the European Commission several times a year on public opinions on a variety of topics and issues.

2.0.2 The Brexit campaigns

In this section, I include some examples of campaign materials to demonstrate the type of rhetoric that was characteristic of the campaign period. The materials are drawn from the official websites[3] of each campaign and other legitimate and prominent groups and organizations.

Figure 2a is a leaflet from Labour Leave (2016), a campaign group that consisted of Labour Party members who supported Leaving the EU. Figures 2b and 2c are from Vote Leave (2016), a coalition of politicians (including support from the UK Independence Party, the Labour Party and the Conservative Party) and interest groups (Vote Leave, 2016). Vote Leave was considered the official Leave campaign by the Electoral Commission.

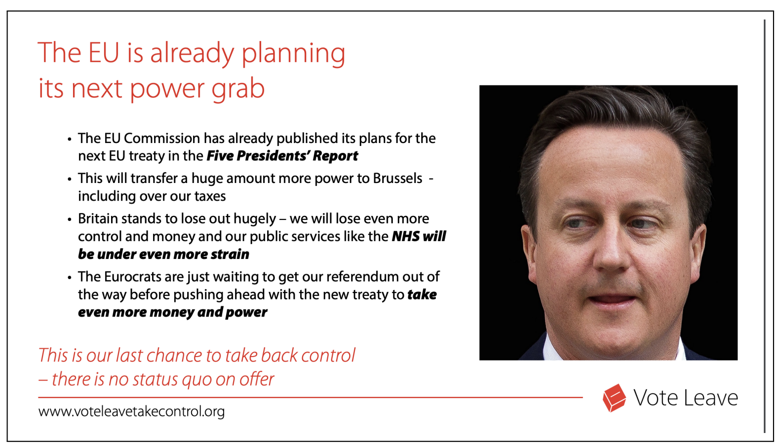
Despite being part of the same campaign, there is a visible difference between the content of and style of Labour Leave and Vote Leave's rhetoric. Labour Leave's pamphlet (figure 2a) emphasizes jobs, unfair labor contracts, the expanding power of the EU, and the freedom to be in the global market. The aesthetics of this pamphlet are relatively tame, compared to Vote Leave's materials, where the rhetoric is much coarser and sensationalist. The map in figure 2b became notorious for fear mongering and its conspicuous Islamophobia. The only two labeled countries on the map are Syria and Iraq (majority Muslim countries). Two of the listed countries, Turkey and Albania, are also majority-Muslim. Macedonia and Montenegro have a large Muslim population. What is implied in the campaign material by the large arrow that points directly from this region to the UK is a possible influx of immigrants, but more importantly, Muslim immigrants.

Figure 2c focuses on the expanding powers of the EU and the fiscal cost of EU membership, which highlight the loss of sovereignty. The tone reads quite differently to the similar issues raised in Labour Leave's pamphlet. The larger font size of terms like "power grab" and "take back control" are bold, assertive, and inflammatory.

[3] There were several websites for both campaigns, but only one of these was declared the official campaign website by the Electoral Commission

(a) Labour Leave leaflet

(b) Vote Leave leaflet



(c) Vote Leave leaflet

Figure 2: Leave campaign materials (Sources: Labour Leave campaign and Vote Leave campaign)

The Remain campaign, representing a vote to stay in the EU, is represented in the campaign materials in figure 3, as a vote for stability and economic strength as a member of the EU (Britain Stronger In Europe, 2016). The organization that produced these materials, Britain Stronger in Europe, was the official Remain campaign according to the Electoral Commission.

The materials in figures 3a, 3b, 3c convey the Remain campaign's perspective on issues of the economy, trade, jobs, and healthcare. The sources of the claims made in the materials are shown, lending legitimacy to the claims. Moreover, sources like business interest groups (CBI), government ministries (HM Treasury), and prestigious economic advisory firms give the materials an air of severity and truth.

The Leave campaign will be remembered as the more emotional of the two campaigns. This is partly due to the populist politics of the campaign, led by Nigel Farage and partly due to the blatant falsehoods of the campaign.[4] However, the Remain campaign also made emotional appeals. The image of the person in the wheelchair in 3b, the sheer number of jobs at stake, and the financial strength that is tied to the EU are all meant to play into the audience's sense of fear and empathy.

The referendum was a pivotal moment for the United Kingdom. Emotionally-charged campaign materials and political rhetoric elicited feelings of anger, frustration, enthusiasm for Britain's future. These efforts to elicit emotional reactions engaged voters, making Brexit an event that was riddled with emotions.

[4]For example, Brexit campaign politicians later admitted that the claim that the money sent to the EU would go to the NHS instead (see figure 4) was a lie (Helm, 2016).

(a)

(b)

(c)

Figure 3: Remain Campaign Materials (Source: Britain Stronger in Europe campaign)

2.1 The study of emotion in Brexit

Scholars have studied various aspects of the Brexit era, such as voter behavior and political strategy. In this section, I provide a brief overview of how political scientists have studied how emotion was important for voters and how emotions were a part of politicians' rhetoric and campaign rhetoric.

2.1.1 Voters' emotions

The Brexit campaigns were designed to play on voters' emotions by evoking fear and anger. In an article that empirically tests of Affective Intelligence theory, Vasilopoulou and Wagner (2017) argue that the Leave campaign was anger-inducing, while the Remain cam-

Figure 4: Brexit campaign bus (Source: Getty Images)

paign was anxiety-inducing. The Leave campaign framed EU integration (the process of political, legal, economic, and social integration of European member states) as a threat. Integration was equated to the ability of the EU to override national policy and law. Moreover, Leave portrayed the UK's membership in the EU as a fiscal burden as a result of the membership contribution scheme. One of the most famous ads for the Leave campaign was plastered across a bus, reading "We send £350 million a week to the EU, let's fund our NHS instead" (see figure 4). In contrast, the Remain campaign framed EU integration as a display of unity and strength.

Their theory states that angry citizens are less risk-averse, more confident and more aggressive. Anxious citizens are more risk-averse and rely on careful processing of information to deal with a threat. Vasilopoulou and Wagner (ibid.) hypothesize that angry citizens are more likely to vote to leave the EU and therefore eliminate the "threat" of the EU. The authors also hypothesize that anxious citizens are more likely to prefer re-negotiating the relationship to the EU and therefore vote to stay. In this way, remain voters wanted to avoid the risk of exiting the EU.

Using survey data, the authors find evidence in support of their hypotheses. Fear increases a voter's willingness to renegotiate the relationship to the EU, whereas anger is associated with a desire to leave the EU. Angry voter's stances on EU membership are associated with their underlying attitudes toward the EU, compared to fearful and enthusiastic voters. Their findings have implications for considering the emotions that campaigns elicit in order to achieve their objectives. Campaigns can evoke fear, enthusiasm, or anger, to mobilize their voters or to reinforce existing attitudes. This also speaks to the broader attitudes citizens harbor toward society and politics (e.g. the state of the economy, trust in the UK government, opinions on climate change) and how these emotions color their political perspectives.

The attitudes of voters towards the referendum extended to larger societal issues. A survey conducted by Ashcroft (2016) suggests that the majority of Leave voters had a substantially different outlook on the general economic and social climate of the time. The survey shows that Leave voters were more pessimistic than Remain voters about the quality of life in the UK compared to how it was 30 years ago. For many Leave voters, this negative outlook included viewing important sociopolitical forces such as multiculturalism, the green movement, and globalization as forces for ill, as compared to Remain voters.

Colantone and Stanig (2018) show that areas that were hit by economic globalization showed higher levels of support for the Leave campaign. Economic globalization displaced many regions' manufacturing activities, leading to a climate of change and uncertainty. The authors test this hypothesis against a competing hypothesis that higher immigration was associated with higher shares of Leave votes (Beauchamp, 2016). They find no evidence to support the competing hypothesis.

2.1.2 Politicians' emotions

Unlike most campaigns, Brexit was not explicitly partisan. With the exception of the UK Independence Party (UKIP), members of the same party – the Labour party, the Conservative party, as well as Liberal Democrats and Greens – held different positions on the referendum. As a result, voters could not simply rely on partisanship as a cue to vote.

Tolvanen, Tremewan, and Wagner (2021) argue that the Remain campaign ultimately positioned itself as centrist to win voters. Voting to Remain meant a vote for the status quo. Their theoretical framework posits that noncentrist campaigns form in response to centrist campaigns to attract voters. Noncentrist campaigns, however, are ambiguous, meaning they synthesize issues both from right-wing politics and left-wing politics.

The Leave campaign implemented this ambiguity in its agenda, making it unclear whether voting to Leave meant a shift to the right (e.g. fewer worker protections) or a shift to the left (e.g. less state intervention). There was a historical precedent for Euroskepticism not being a partisan topic (see section 2.0.1). From a right-wing perspective, EU regulations were oppressive and violated the nation's sovereignty. From a left-wing perspective, EU regulations reduced labor protections and oversight was too loose. Tolvanen, Tremewan, and Wagner (ibid.) argue that the Brexit campaign strategically employed the use of uncertainty. They test their theory in a lab setting, where they measure participants' voting preferences. They find that voters often prefer ambiguous, noncentrist platforms, especially when the opposition is a known centrist.

While the political positioning of the Leave campaign was ambiguous, the emotional charge of the campaign was clear. Cap uses the concept "strategic simulation of affect" to characterize the strategic motives of politicians to shape public discourse (Cap, 2016, 2). The concept describes how politicians evoke emotional reactions in the public in order to shape emotional associations with key issues. The Leave campaign pushed issues such as national sovereignty, British exceptionalism, and the lack of democratic accountability in the EU to tug at the British public's sense of patriotism and national self. Moreover, they strategically portrayed migration as an unstoppable, overwhelming flow into the UK. This was tied to issues like the pressure migration would create on the housing market and the welfare state. By emphasizing these key issues, and curating the emotional associations with the issues, Cap argues that politicians shaped the course of public discourse around Brexit.

Many scholars place emphasis on the populist rhetoric used by the Leave campaign, especially by key figures like Nigel Farage (Hughes, 2019). The following two quotes are provided to give the reader a sense of the rhetoric that was characteristic of the Leave campaign.

> “We are being sold that this is all about trade and that the single market is soft and cuddly and lovely like a baby puppy. But actually it is a smokescreen for the real, simply proposition of this referendum. It’s actually rather simple: do you wish us to be a self-governing, independent, democratic nation or part of a bigger, broader, European Union?” (News, 2016b)

> “[Brexit] will be a victory for ordinary people, for decent people.” (News, 2016a)

Despite the vast cultural cleavages of the locations where it emerges, populist rhetoric exhibits patterns. One of its notable characteristics is that it evokes the idea of “the people” who represent the “true essence” of a nation. The Leave campaign’s anti-immigrant rhetoric (see figure 2b) employed this tactic by eliciting a fear of immigrants who threaten the “people” of Britain because of their religion and cultural background. Populism is a salient phenomenon of contemporary politics at a global level. It is a highly successful strategy of political parties and politicians (e.g., Erdoğan in Turkey, Bolsonaro in Brazil, the Greek party Syriza, Orbán in Hungary, Alternative für Deutschland in Germany) and seems to emerge in contexts where large segments of the population exhibit some sort of disillusionment. Populist rhetoric is inherently affective insofar as it appeals to and amplifies divides in society, and by playing into fears of groups of people, like immigrants (Jagers and Walgrave, 2007).

3 Text as a source of data

In this chapter, I briefly review the basic terminology of text analysis methods and key parts of this type of methodological approach. This helps familiarize readers who are not acquainted with these methods. After giving a general overview of the classes of text analysis methods, I delve deeper into the primary text analysis methods used in this dissertation, namely dictionary methods and word embeddings. This is a more granular discussion of how these methods work and what they are used for.

3.1 Text analysis terminology and fundamentals

Text is a very rich source of data, but qualitative approaches cannot keep up with the enormous amount of text data that exists and continues to grow. Quantitative approaches tackle the problem of scale by computationally processing thousands of documents in very short periods of time. They can also help identify large-scale patterns in the data. However, automated methods do not and *cannot* fully replace qualitative research. Therefore, quantitative analyses need to be validated when employed and the application of any given method should be considered very carefully for a given research question.

A collection of texts is referred to as a “corpus”. Each text is often referred to as a “document.” These documents may not be literal documents, they can be tweets, speeches, press releases or any number of other things. “Tokens” are the words in the documents. In survey analysis the data consists of observations as the entries of each row with each column capturing some measurement of a variable for that given observation. In text analysis, the same principle applies. Each row is typically a document other columns contain metadata about the corresponding document. Metadata such as the date or author are necessary for most analyses.

How do we statistically analyze words? Texts need to be translated into numbers. Typically, unique tokens are identified, where words with the same root but different suffixes (e.g. run, ran, running) are considered the same token. In other analyses, the part of speech that the token is (e.g. noun, verb) is important to the analysis so the different forms of the token are considered separate tokens. While individual words are often interesting, scholars consider bigrams (pairs of words), trigrams, and more broadly, n-grams for analysis.

Most text analysis methods use the bag-of-words assumption, which states that the order of words does not matter, only the presence of words in a given document. Although this assumption is wrong, it massively simplifies the computational cost. Researchers who use text analysis also show that this assumption still allows for some interesting and valuable findings.

After this step, the document is represented as some sort of matrix. Examples include document-feature matrix (abbreviated as dfm) where the rows are the document IDs, and the columns are the unique features. Each row is then a series of mostly 0's and 1's (or maybe larger integers) which symbolize the count or absence of a token in a document. Another representation is the feature co-occurrence matrix (abbreviated as fcm) which includes the features as rows and as columns. Each cell is a count of how many times the corresponding row-token and column-token co-occur in a document. Several other formats exist.

3.1.1 An overview of text analysis methods

This numerical representation of text is the input of the statistical analysis. Text analysis has been used to approach a number of different questions. Grimmer and Stewart (2013) outline the general organization of text analysis methods. The two broad categories they delineate are classification and scaling. Each of these categories can be subdivided into supervised and unsupervised methods.

Supervised machine learning methods are relevant when the researcher is interested in categorizing the data into known categories. For example, if the researcher is interested in specific agenda topics (e.g. topics on fiscal spending and migration laws in a corpus of Congress speeches), then supervised learning methods would be applicable. The researcher can use a subset of the data (the training set) to train the algorithm to identify the topics and then apply this knowledge to test a sample of the data. Unsupervised methods are useful when the researcher does not know the categories ahead of time or when there are things to discover about the texts that they cannot pre-specify. For example, if there are some latent features about the texts, unsupervised learning methods may help classify the texts according using this feature.

Text analysis methods can be used in many different contexts and for different purposes. Ansolabehere and Iyengar (1996) use dictionary methods to show that negative advertising drives down voter turnout. This is a relatively simple use of dictionary methods. Beauchamp (2011) uses legislators' written and spoken text for ideological scaling. This method goes beyond the use of voting behavior to capture ideology. Grimmer (2010) uses a bayesian hierarchical topic model on press releases to measure political actors' attention to different topics. These are just a few examples to show how expansive the uses of text analysis are.

3.2 Dictionary methods

Dictionary methods are, as Grimmer and Stewart note, perhaps the most intuitive of the text analysis methods. Broadly speaking, dictionary methods entail classified documents as belonging to one or several categories. This is done with some metric using the count of words in each category in the text. Scholars have developed an array of diverse and interesting lexicons. These include the NRC lexicon used to classify emotions and sentiments (Liu, 2015; Mohammad and Turney, 2010, 2013). Lexicons like the Lexicoder Sentiment Dictionary are used to extract positive or negative valence of rhetoric (Nielsen, 2011). Other lexicons include more bizarre categories such as litigiousness and uncertainty (Bodnaruk, Loughran, and McDonald, 2015; Loughran and Mcdonald, 2016, 2011). In more recent years, efforts have been made to create lexicons that are context-specific given the limitations of using off-the-shelf dictionaries in any given context (Rheault et al., 2016).

Research that uses dictionary methods goes beyond simple word counts, however. Kellstedt (2000) uses Newsweek articles to detect articles that capture policy preferences on racial equality. The author uses dictionaries to classify the attitudes in the articles as liberal or conservative. This research produces longitudinal data on policy attitudes that cannot be captured by survey data. Survey data is both less frequent and items are not easily comparable across time. Other scholars emphasize the importance of dictionary methods in automating the process of coding documents (Laver and Garry, 2000; Young and Soroka, 2012). In another application, Burden and Sanberg (2003) take a more nuanced approach and capture both the volume and tone of budget-related rhetoric in presidential campaign speeches. Dictionary methods, although conceptually simple, can tackle fundamental questions of political science.

A drawback of using dictionary methods is that existing dictionaries may not always fit the context of study. For example, an emotion like enthusiasm may look different in an elite political context than it does on twitter. Another limitation is that there is not a dictionary for every research question. Dictionary methods are useful once a dictionary is ready but can require a lot of research of the backend to identify to a suitable dictionary or to create a custom one.

3.3 Word embeddings

Word embeddings are a scaling method used to capture semantic similarity of tokens. Words are represented as vectors and words that are close in meaning or used in similar contexts are expected to be close in the vector space. Figure 5 is a visualisation of a simple example of the method. This method was developed at Google by Tomas Mikolov et al. at Google in 2013, released in the `word2vec` package (Mikolov et al., 2013).

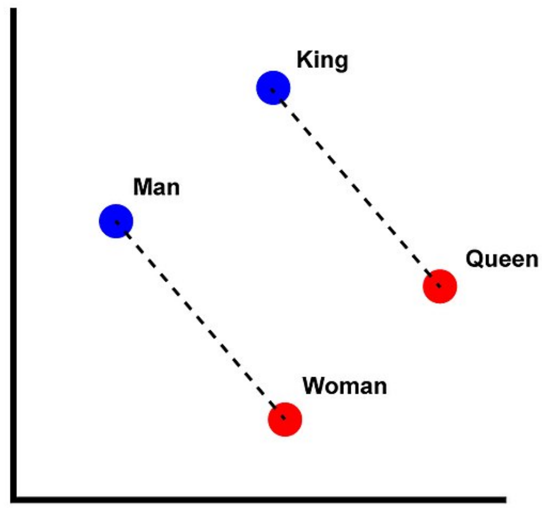


Figure 5: Visualisation of word embeddings

For the analyses in this paper, I used the `text2vec` package, which uses the Global vector for Word Representations (GloVe) algorithm. The GloVe algorithm consists of the following steps.

Algorithm 1: Glove algorithm

---

1. Translate corpus into a word-co-occurrence matrix X. Each cell in matrix X, denoted by X_{ij}, represents how often word i occurs in the context of word j. The width of the window is defined by a *window_size* parameter, which is some *window_size* number of terms before word i and *window_size* number of terms after word i. Typically, less weight is given to more distant words using this formula

$$decay = \frac{1}{offset}$$

2. Define soft constraints for each word pair

$$w_i^T w_j + b_i + b_j = log(X_{ij})$$

w_i is the vector for word i and w_j is the vector for the context word j and b_i, b_j are scalar biases for the main and context words.

3. Define a cost function

$$J = \sum_{i=1}^{V} \sum_{j=1}^{V} f(X_{ij})(w_i^T w_j + b_i + b_j - logX_{ij})^2$$

V is the size of the vocabulary. f is a weighting function to prevent learning only from extremely common word pairs. This function is defined as:

$$f(X_{ij}) = \begin{cases} (\frac{X_{ij}}{x_{max}}^{\alpha}) & \text{if } X_{ij} < x_{max}. \\ 1 & \text{otherwise.} \end{cases} \tag{1}$$

The cutoff x_{max} is set to 100 according to the authors of GloVe and $\alpha = 3/4$

Word embeddings have many applications, including predictive models (e.g. fill-in the-blank tasks) and building custom dictionaries. It is a powerful tool because it allows the researcher to observe changes in languages across many different kinds of contexts, such as geographic or temporal. Kim et al. (2014) trained the model on the Google Books Ngram corpus to track changes in the meaning of words over time. In another use of the method, Rheault and Cochrane (2020) capture ideology, a latent feature of text, by combining a model that uses word embeddings and metadata accompanying the text. Word embeddings allow for the dynamic analysis of language, doing away with the assumption that meaning is static across time and contexts.

3.3.1 Contribution

This dissertation's contribution is an in-depth examination of the importance of anxiety in political rhetoric and insight into anxiety operates in a political context. My custom anxiety lexicon measures how anxiety in expressed in the British House of Commons. To my knowledge, an off-the-shelf dictionary of anxiety does not exist. Even if an anxiety lexicon did exist, custom dictionaries are a much more powerful tool that capture how a specific emotion is expressed in a certain context. We do not expect a politician to convey anxiety in the same way a voter does. The anxious issues and the manner in which anxiety is conveyed can be identified using the approach of a custom dictionary.

Anxiety, according to Affective Intelligence Theory (see chapter 1), is one of the primary emotions that explain how we react to our environment. This emotion has important political implications such as information processing, voting behavior and political opinions. I argue that due to the environment of uncertainty during the Brexit negotiations, anxiety was a highly salient emotion. This study can help elucidate the state of political discourse at the time by measuring the levels of anxiety in parliamentary speeches.

4 Data & Methods

In this chapter, I discuss the datasets used for training machine learning models and for analysis. I also explain why these are the best datasets given the set of questions I ask. These datasets are a combination of existing data and data that I scraped from the British parliament's official website. I go on to discuss the methodological approach for the text analysis section of the analysis.

4.1 Design

What does anxiety look like in the context of political rhetoric? This is the core inquiry of chapters, 5, 7, and 7.5. These chapters show a comprehensive set of descriptive analyses, including measures of anxiety and emotional arousal by time period, party and committee. These descriptive findings show how anxiety and emotional arousal relate to these variables of interest. In chapter 5, I go beyond the timeframe of the Brexit negotiations to study how party politics factor into anxiety levels over a 20-year period.

This inquiry requires a corpus of elite political rhetoric. The unit of analysis is a speech, i.e. when a speaker speaks in turn. Metadata should at the very least include the speaker's name, the date and the topic of discussion.

The methodological approach to measuring anxiety in text involves the creation of a context-specific dictionary. We do not expect elites to show anxiety in the same way as voters, for example. Therefore, off-the-shelf dictionaries are not an appropriate choice for this inquiry.

The methodological approach to the audio analysis requires a good amount of pre-processing speeches to format them for analysis. The starting point is a collection of audio files, where each file is a meeting of a Special Committee or the main chamber. The resulting dataframe must include the metrics of emotional arousal (e.g., the average pitch) for each speech in the meeting.

4.2 Data

4.2.1 Hansard data

Initially, the Hansard dataset seemed like an appropriate choice for this project. It is a compilation of all House of Commons debates dating back to the 1988/89 session and is still being updated. The dataset has been used widely in scholarly work on discourse and rhetoric (Archer, 2017, 2018; Foxlee, 2018; Kruger, van Rooy, and Smith, 2019; Mair, 2019; Slembrouck, 1992). The data lends itself to text analysis because of its size; the timespan exceeds 40 years. Since many text analysis methods are computationally demanding, this makes the dataset a good choice. Moreover, the data is quite clean, not requiring much more than the usual pre-processing steps usually undertaken in text analysis (removing punctuation, lemmatizing etc.). Nevertheless, the dataset has been criticized for its deficiencies, such as the accuracy of the transcription (Mollin, 2007). For example, filler sounds and sounds like "um" or pauses are not captured in the transcript. This is arguably an important limitation of the data, and hinders the ability to conduct certain types of analyses (e.g. a linguistic analysis on cadence or clarity) on the data. The unit of analysis is a speech by a Member of Parliament (MP) and is accompanied by metadata including speech class (division, procedural, speech, table, or upper division), speaker name, speaker ID, constituency, the debate's minor and major heading, and the URL to the debate.

This data is interesting but does not fully capture the parliamentary dynamics of the negotiation period. The data does not include House of Commons Select Committees' oral evidence sessions, a less studied source of parliamentary rhetoric. In this calmer setting, small groups of MPs, tasked with investigating an issue or policy area, gather evidence and bring in expert witnesses to offer recommendations or proposals to parliament. Including this dataset allows us to more fully observe how Brexit transpired in debate and in Committee meetings.

4.2.2 House of Commons Select Committees

Select Committees exist in both the House of Commons and the House of Lords. The latter is not in the purview of this dissertation. Committees are formed in the main chamber and their members consist of backbencher MPs. There are a minimum of 11 members on each committee. There is a committee for each government department and the committees examine three aspects: spending, policies, and administration (UK Parliament).

Some committees have inter-departmental roles and others examine specific topics such as environmental policy and legislation. In some cases, committees are involved in investigations, such as allegations about an MP's conduct. Committees might be tasked, for example, with improving a given bill. They have the power to appoint specialists (e.g. academics or senior officials) or stakeholders (e.g. industry representatives or interest groups) to participate. Once the committees gather the oral and written evidence on their inquiry, they produce a report. Upon submission of the report, the government usually has 60 days to reply to the committee's recommendations. They can fully, partially, or not implement the recommendations.

Some scholars question the importance of the Select Committees on government deliberation, but in a 2013 study, Benton and Russell find evidence that challenges this assumption. The authors conduct a comprehensive examination of the impact of committee reports. They find that committees have an important capacity for oversight of government decision-making. Committees create exposure on a given issue, act as brokers between various government bodies, and therefore act as an accountability mechanism. Importantly, as committees gather evidence and tap into expertise on a given issue, they influence policy debate. Committee evidence sessions often receive more media coverage than main chamber debates; coverage of the latter has been declining over time (Kubala, 2011).

For this dissertation, I focus only on the oral evidence of committees, since it allows for the audio-text comparison. I conducted a search with the keyword "Brexit" from 2016 to 2020. There are 168 results in the time period, 121 of which are committee meetings, and the rest of which are main chamber debates. All transcripts were parsed using the `rvest` package. I extracted the speaker's name, the committee's title, the date, and the topic of the committee. I use the full dataset for the text analysis chapter, chapter 5, because scraping and parsing is relatively computationally efficient. For the audio analysis chapters, 7 and 7.5, I use a subset of this data. This is due to the computational cost of preparing this data for analysis. Given the resource constraints for this dissertation, I was only able to analyze the audio of around 50 Committee meetings.

4.2.3 Summary of the data

Dataset	Count of speeches	Count of sentences	Purpose
Hansard dataset 1999-2000	1,299,323	3,236,912	Training word embeddings
Committees dataset	38,718	113,058	Training word embeddings + Audio analysis
Committees + subset of Hansard = Brexit dataset	38,718 + 8398 = 47,116	113, 058 + 50,237 = 163,295	Analysis

Table 1: Summary statistics of data

I use several datasets in this dissertation for different purposes. Table 1 shows the size of each dataset and the various purposes of each. What I refer to as the Brexit dataset (the dataset used for the main analyses) includes a subset of the Hansard dataset. The Hansard data was filtered for parts of the debates that were related to Brexit. Filtering was done by

Year	Count of Speeches
2016	1469
2017	9377
2018	13144
2019	9556
2020	2822

Table 2: Count of Brexit speeches by year

using the minor heading (included in the metadata the dataset) which describes the topic of the debate at the time. This was necessary because main chamber debates cover a multitude of different topics in a single session.

The Hansard data from 1999-2000 combined with the Committees data is used to train the word embeddings model. Word embeddings require a lot of data to be trained and the Brexit dataset is not large enough.

The analyses shown in chapter 5 use the full Brexit dataset. This is, to my knowledge, a comprehensive dataset of all parliamentary debates and meetings about Brexit.

For the audio analysis, I use a sample of the Committees data. Using the audio files, I generate timestamped transcripts. Further details are discussed in chapter 6.

4.3 Methods for text analysis

The methodological design for the text analysis in this dissertation is based on a paper by Rheault (2016). Rheault's paper measures anxiety in the Canadian House of Commons. To my knowledge, this is the only political science research that quantitatively captures anxiety. Rheault's method uses word embeddings to scale the tokens of a corpus of Canadian parliamentary data. The output is a lexicon of tokens and their score of anxiety that can be used for dictionary analysis. The author then conducts analyses using the dictionary and some additional metadata, including the speaker's name and title annotations. The author explores the relationship between the class of the document and the topic of the discussion, producing some interesting results about how anxiety varies across topics (e.g. immigration appears to be a more anxious topic than constitutional affairs).

4.3.1 Training the model

The Brexit dataset I scraped is not large enough to train a word embeddings model. One of the limitations of word embeddings is the large size of the training data. To address this, I supplement the Brexit data with the Hansard data. I use a 20-year period of the

Hansard data (1999 until 2019). I chose these dates because at the time I was conducting the analysis, 2019 was the final year of complete data. This time period is both sufficient data-wise and also includes the majority of the timeline of the Brexit negotiations (2016-2020).

The data used to train the model is a superset of the data used for analysis but this should not theoretically affect the results. This claim is based on the assumption that rhetorical expressions of anxiety, meaning the words and phrases that British MPs use to express anxiety, do not change significantly over time. In other words, the way anxiety is expressed in British political elites' rhetoric does not change between discussing Brexit and discussing some other topic. There are several software packages for word embeddings, including `word2vec` and `text2vec`. I chose to use the `textmodel` package, because it allowed me to include seedwords easily in the process. Seedwords can be used in word embeddings to calibrate the scale of the output. I use the seedword list used by Rheault (2016) which consists of 30 words, 15 of which are anxious (e.g. risk, threat, and stress) and 15 of which are non-anxious (e.g. optimism, calm, guaranteed). The full list of seedwords is included in the appendix.

After some basic pre-processing (removing stopwords, numbers, punctuation and extra whitespace), I annotate the corpus using the `cleanNLP` package (Arnold, 2017). This process includes lemmatizing, tokenizing and attaching each token's part of speech. Lemmatization is the process of reducing the different forms of a word into a single form. For example, "working", "worked" and "works" are reduced to "work". These steps are explained in further detail in section 3.1. I then concatenate the lemma to the party of speech (PoS) with an underscore "_" (lemma_PoS) and by grouping together lemma_PoS by document id, I recreate the speeches as in the corpus originally. For example, a segment of speech that was originally "plenty of people vote" becomes "plenty_a of_p people_n vote_v" (v = verb, p = preposition, n = noun, a = adjective). This step allows for a more nuanced analysis of the output, because we can observe the anxiety by part of speech. Examples of this will be shown in the following section.

The output of the model includes the vector representation of every token in the dataset (in lemma_PoS format) and additionally, the `textmodel` package scales each token on a range of -1 to 1 with the seedwords on each respective end (anxious seedwords are scored as 1 and non-anxious as -1). The scaling is conducted by the software package using cosine similarity. The position of each token on a scale is its score. This produces a metric of the anxiety of each token.

Algorithm 2 details the workflow for the creation of the anxiety lexicon.

Algorithm 2: Creation of Anxiety Lexicon

For a corpus of document D of size N where each document is denoted as d_n where $n \in 1, \ldots N$

1. Pre-process all documents d_n
   (a) Remove numbers
   (b) Remove words less than 3 letters long
   (c) Remove punctuation
   (d) Remove common stopwords as well as procedural stopworks e.g. minister, hon, members
   (e) Remove excess whitespace
2. Remove speeches with fewer than 15 characters
3. Annotate data using `cleanNLP` package
4. Format and filter annotated data
   (a) Group together adjectives (comparative, superlative)

(b) Group together nouns (singular, plural)

(c) Group together adverbs (comparative superlative)

(d) Group together verbs (past tense, present participle, gerund, past participle, singular present)

(e) Filter annotated dataset to just adjectives, nouns, adverbs, interjections, and verbs

(f) Paste lemma to part of speech information (lemma_PoS)

5. Reconstruct speeches

6. Convert to document-feature matrix object

    (a) Remove special symbols

    (b) Include terms that appear more than 200 times

7. Convert to feature-co-occurrence matrix object with window of 15 tokens

8. Feed fcm object into word embeddings algorithm with seedwords (algorithm details discussed in Algorithm 1)

9. Extract cosine similarity output

4.3.2 A note on terminology

Instead of "non-anxiety", I will use the term "confidence". Confidence conveys a sentiment of security. Anxiety is about uncertainty or fear about one's course of action. It has a connotation of a lack of the ability to foresee or anticipate the outcome of events. Confidence, on the other hand, is an emotion of security. A confident person is more sure of their actions and the expected outcomes of their actions. These two terms may not be antonyms but they capture the two ends of the scale I use in the analyses. “Confident” is not as meaningful as “non-anxious”, the latter of which is a rather empty term.

5 Patterns of anxiety in the House of Commons: Findings from the custom anxiety lexicon

In the previous chapter, I outlined the methods and data used in the methodological approach. In this chapter, I show results from the custom anxiety lexicon and applications of the lexicon on the Brexit data. First, this chapter will explore the face validity of the lexicon. Face validity checks examine subsets or samples of the lexicon and evaluate whether these parts of the lexicon appears to measure what it is supposed to measure, namely anxiety. This helps establish the credibility of using this lexicon as a measurement for anxiety.

The second section of this chapter applies the lexicon to the Brexit dataset (the collection of parliamentary debates and committee meetings). I measure the anxiety levels in this dataset, disaggregating the results by party, date, and committee. I also compare these results to a sample of speeches in the same time period that are not about Brexit. Finally, I extend the findings about the relationship between anxiety and politics to a broader timeframe of 20 years. This final analysis, although slightly outside the scope of this dissertation, elucidates the dynamics of anxiety as an emotion in political speech.

To briefly summarize the methodology that produces this anxiety lexicon, I use a corpus of parliamentary speeches to train a word embeddings model. This model scales every unique word in the corpus according to a set of seed words which anchor the scale from least anxious to most anxious.

5.1 Validation of the anxiety lexicon

A simple but important step in checking the validity of the dictionary is to qualitatively examine a sample of the lexicon. Using the score output by the word embeddings model, I examine a random sample of the top 10% most anxious words and the top 10% least anxious words. The results are shown in tables 3 and 4. The purpose of including these

Lemma	Cosine similarity score
postbrexit	0.085
immigration	0.085
frustrate	0.086
ethnicminority	0.095
misjudgment	0.088
uncooperative	0.112
scaremongering	0.014
dismiss	0.033
debilitate	0.048
exploitative	0.023
intimidate	0.017
disapprove	0.045

Table 3: Sample of anxious words from top 10%

Lemma	Cosine Similarity Score
augment	-0.059
secure	-0.060
predictability	-0.085
assure	-0.110
anticipate	-0.113
facilitate	-0.140
ambitious	-0.171
empower	-0.108
restoring	-0.100
determined	-0.138
seamlessly	-0.119
streamline	-0.099

Table 4: Sample of confident words from top 10%

samples is to (1) provide evidence that the sentiment words on either side of the scale can be differentiated and (2) argue that these words can be used to express anxiety. I discuss a subset of the words included in the tables.

I will also examine the dictionary subset by part of speech (nouns, adjectives, interjections, adverbs and verbs) to evaluate whether the algorithm was better at scaling certain parts of speech over others. While these explanations are speculative to an extent, my objective is to elucidate the inner workings of the word embedding algorithm since algorithms can so often behave as a black box – impervious to a deeper understanding.

As seen in tables 3 and 4, these words pass a face validity check. The term *frustrate* could appear in speech that is anxious. Frustration may convey exasperation with an issue, or a lack of control over a situation. Interestingly, the term *post-Brexit* emerges in this sample. The corpus used to train the word embeddings model included a dataset spanning 20 years and the Brexit negotiations are only a small fraction of this time period. Nevertheless, the algorithm determined that the term is highly charged with anxiety – it is frequently mentioned alongside other words that are associated with anxiety – and lands in the top 10% of the scale. Terms like *immigration* and *ethnicminority* [sic] may reflect ongoing issues with

British politics. As discussed in chapter 1, section 1.4.1, immigration has almost always been a contentious issue for the UK. The refugee crisis fueled the Brexit debates. One of the Leave Campaign's main selling points was that the borders would close to immigrants and refugees coming from non-European countries. The politics of ethnic minorities (treatment, housing, discrimination) is embedded into the sociopolitical framework of the UK. *Misjudgment* could appear in the context of a situation where issues have been mishandled or are believed to be currently mishandles, signaling anxiety again. Finally, a use of the term *uncooperative* might describe tense relations between members of parliament, government agencies, or other political bodies.

The terms from the sample of the bottom 10% of the scale also pass the face validity check. *Augment* can be used in many situations but it generally has positive connotations since it is a defined as making something better. Positivity is not equivalent to confidence but a speaker who uses *augment* might be discussing how things are improving, thereby conveying a level of optimism in their actions. Likewise, *secured* and *assured* are related to a sense of confident actions. *Predictability* stands in contrast to the fear of the future that anxious individuals experience. Therefore, it intuitively belongs on this side of the scale. A term like *anticipate* could be used in a context to convey confidence. For example, when a speaker is discussing expected challenges with an issue. To anticipate is to brace for or be prepared for a challenge.

The scores in parentheses should give the reader a sense of the range of the scale. While the range may not seem very large, the scores are best interpreted as a relative quantity. The relative distance between anxious and confident words (variance across categories) as well the variance within groups of anxious words and groups of confident words (variance within categories) gives the reader a better idea of the distribution of the anxiety scores of words.

Of note are words that invoke neither anxiety nor confidence at face value. For example, medical terms like *osteoporosis* and *encephalopathy* are included in the top 10% of anxious words. These may seem to discredit the face validity of the lexicon and the resulting analyses. What could explain this sort of result is that the algorithm is imperfect and these terms should in fact be scaled as neutral. Alternatively, the speaker could be discussing a topic that involves terms like "osteoporosis" and is anxious about the issue.

5.1.1 Results sorted by part of speech

Examining the results sorted by speech allows me to see the granularity of the analyses while also highlighting the imperfection of the annotation software (the `cleanNLP` package). These examples of the sorted results may provide some intuition about why a researcher might choose to include this information in their analysis. For example, a researcher may find that certain parts of speech are, on average, scaled higher than other parts of speech. Subsetting the lexicon to include only certain parts of speech can deliver a clearer, more definitive measure of anxiety.

Examples of anxious nouns: *insensitivity (0.145)*, *psychopath (0.166)*, *intoxication (0.168)*, *crudity (0.156)*, *disorientation (0.169)*
Examples of confident nouns: *milestone (-0.196)*, *ambition (-0.178)*, *transformation (-0.165)*, *future (-0.154)*, *endeavour (-0.148)*

Examples of anxious adjectives: *thoughtless (0.152)*, *unfaithful (0.168)*, *terrified (0.173)*, *troublesome (0.153)*, *helpless (0.147)*
Examples of confident adjectives: *smooth (-0.185)*, *ambitious (-0.181)*, *committed (-0.173)*, *inclusive (-0.158)*, *secure (-0.145)*

Examples of anxious adverbs: *terrifyingly (0.137)*, *inaccurately (0.141)*, *disturbingly (0.143)*, *lethally (0.146)*, *offensively (0.153)*
Examples of confident adverbs: *formally (-0.150)*, *constructively (-0.140)*, *diligently (-*

0.133), *warmly (-0.142)*, *firmly (-0.131)*

Examples of anxious verbs: *intimidate (0.144)*, *inflame (0.144)*, *invade (0.150)*, *denigrate (0.157)*, *implode (0.153)*
Examples of confident verbs: *deliver (-0.180)*, *finalize (-0.160)*, *ensure (-0.155)*, *reaffirm (-0.153)*, *strengthen (-0.148)*

These samples of nouns, adjectives, adverbs and verbs help show how anxiety is measured in the British House of Commons. They can also demonstrate how anxiety might operate differently based on the part of speech.

There is a clear difference in the valence of anxious words compared to confident words. The anxious words are often insulting or critical (e.g., *insensitivity* and *thoughtless*), whereas the confident words carry a sense of security and calm (e.g., *firmly* and *strengthen*). These example translate into the notion of a scale; these words can be placed on a spectrum that ranges from one concept (anxiety) to another concept (confidence). Scaling demonstrates that speech is more appropriately placed on a distribution of an emotion where words lean anxious, lean confident, or are more neutral. Assigning words a binary value of 1 (anxious) or 0 (confident) greatly simplifies language and reduces the nuances that scaling allows for.

The words that are scaled as anxious may not always appear as intuitive to the reader as the ones conveying confidence. Words such as *unfaithful*, *invade* or *psychopath* have an unequivocal negative valence, but are not necessarily thought of as anxious. The anxiety score of these words may be an indication of animosity that arises in anxious situations where there is a lack of direction and order. They may also be capturing the speaker's sentiment toward the situation (e.g. *troublesome* or *disturbingly*). Reviewing results such as these can help the reader expand the concept of anxiety beyond the traditional understanding as an emotion that captures fear of the future. This analysis shows how anxiety may be more concentrated around certain issues or topics, such as drug use or immigration (as in table

3).

Anxiety scores differ by part of speech. In figure 6, I plot the point estimate of the mean score of a given part of speech (e.g. adverb) and a confidence interval capturing 95% of the distribution of the levels of anxiety of that part of speech. This figure shows that there is a great amount of variation across parts of speech, with interjections carrying the highest levels of anxiety, followed by verbs, adverbs, nouns, and adjectives. The distributions of all parts of speech except interjections is rather narrow, as shown by the narrow confidence bands. The confidence intervals of the distributions of every part of speech interestingly do not overlap. Each part of speech appears to cover a distinct segment of the anxiety scale.

The annotation software is imperfect. For certain parts of speech, the results are relatively noisy (i.e. the part of speech assigned to a word is incorrect, with words such as *aleppo* identified as an interjection, *commitment* as a verb, or *autumn* as an adverb. Annotation (detecting the part of speech of a given token) is an interesting step in text analysis, but is not a conventional choice to make. Exploring how different parts of speech operate could be a fruitful path to take for the area of research if it is informed by a field such a linguistics, but it goes beyond the scope of most political scientists' focus. I do not explore the differences in the results by filtering based on different parts of speech at length in this dissertation. The primary reason I chose to include the step of annotation was to filter down the vocabulary of the dataset fed to the word embeddings algorithm to a simpler set of parts of speech (verbs, adverbs, nouns, adjectives, and interjections).

5.1.2 Comparing samples

A conventional validation exercise involves taking a sample of speeches and manually assigning them a label (or binary score) based on the classification task. For example, if a researcher is interested in classifying anger in a corpus, they would take a sample of the corpus and assign a label to each speech in the sample, denoting whether they consider the speech to be angry or not. They would then compare their manual annotations to the score

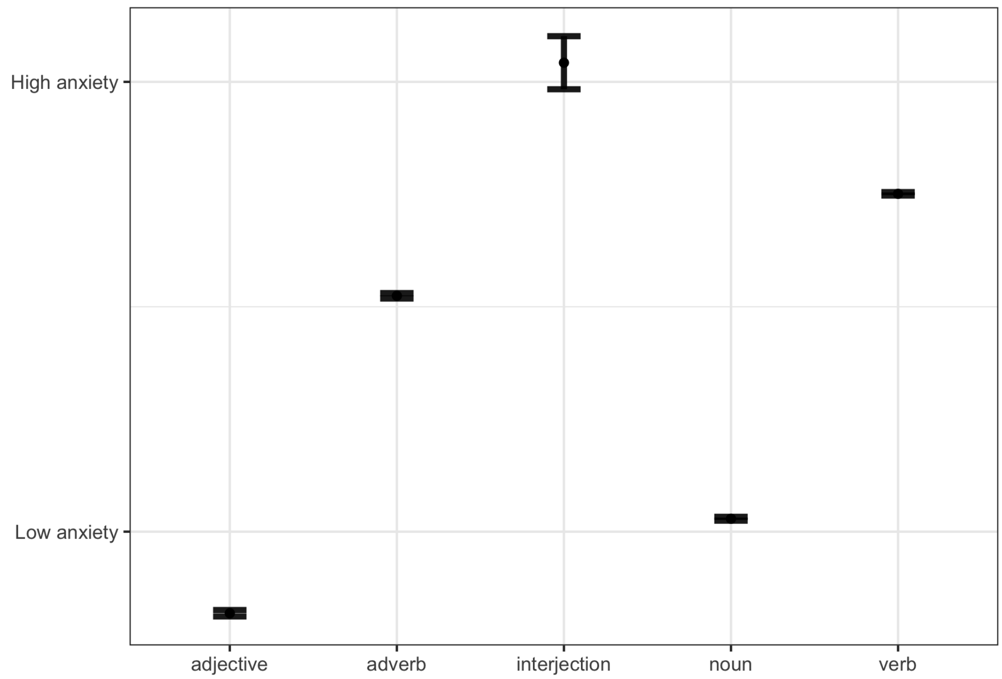


Figure 6: Arousal by part of speech. Mean anxiety score with 95% confidence intervals.

assigned by the dictionary analysis and evaluate the similarity of the manual scores using a few different metrics (usually the metrics are recall, precision, and accuracy). However, since this is a scaling task, it is difficult to annotate a sample of texts using a binary assignment (1 for anxious, 0 for confident) since texts are often not discretely anxious or confident. A more suitable test of the validity of the lexicon is to take pairs of texts and manually annotate which text the reader considers to be more anxious. Then, one can compare these texts again using the scores produced by the dictionary analysis. The scores from the sentiment analysis should ideally confirm that one text is more anxious than the other, matching the manual annotation.

I include two examples of pairs of texts that pass this test in tables 5 and 6. The left-hand side texts convey certainty and optimism for the state of the negotiations. The right-hand side texts convey uncertainty and fear about the possible negative consequences of Brexit. The right-hand texts are annotated as being more anxious than the left-hand text. Note: the right-hand texts happen to be questions in the selected examples but punctuation is not a part of the lexicon. Pre-processing the texts (see chapter 4) includes removing all punctuation.

In the left column of 5, the speaker is discussing the state of trade negotiations for two British islands, Guernsey and Jersey. They mention the World Trade Organization (WTO) and the Department for International Trade (DIT) in the speech, which are two important institutions in the context of trade. The speaker identifies an issue, namely how the future of the UK's trade might impact smaller, more remote places like the British Isles. They go on to emphasize that progress is happening, verified by references from representatives from the Isles and the speaker's colleague's contact at DIT. This speech, therefore, conveys a sense of security about the future of trade for areas of the UK.

In the right-hand column of table 5, the speaker is asking a question about the future of the British financial services sector. They identify the European Union as a possible threat to the sector. This is a good example of anxiety in speech. Based on the text, we cannot assess whether the speaker themselves is anxious about this issue, but the *content* of what they say carries an anxious tone. The threat is the possibility of an attack on the financial services sector from the European Union. The speaker does not identify a way to deal with this potential threat which conveys the uncertainty of the issue at hand. Instead, they ask who may be responsible for dealing with this threat. There is both fear (of a threat) and uncertainty (about who may be responsible with dealing with this threat).

The algorithm confirms that the left-hand speech is more anxious than the right-hand speech in table 5.

In the left-hand column of 6, the speaker is discussing trade in Northern Ireland after COVID. They argue that trade is integral to the peaceful relations between Northern Ireland and the United Kingdom. They identify an issue, namely the Northern Irish economy. They identify a problem: the impact of COVID on the Northern Irish economy. They also identify a solution: free trade as a driver of economic growth.

In the right-hand column of 6 the speaker is discussing the state of relations between

Confident text	Anxious text
It is important to say that we recognise that WTO membership is a top priority for Guernsey and Jersey going forward. In our regular engagement with them, they have made that absolutely clear to us. It is something we are very conscious of. We continue to work with other Departments. One of my colleagues has a very good contact at DIT, and he makes sure those discussions are happening.	There is concern that, beyond Brexit, whatever that place may look like, the financial services sector in particular may be susceptible to attack from the European Union. Beyond Brexit, to whom do you see that responsibility falling?

Table 5: Example 1 of an anxious and a confident speech

Northern Ireland and the British government. They cite on several occasions on which the British government stated that they intend to break international law relating to Northern Ireland. They bring up issues of Irish trust in the British government and how the violations of trust create an environment of instability and anxiety. The speaker identifies an issue, which is the state of relations between the UK and Northern Ireland. They identify a problem: violations of the Stormont House Agreement[5]. They do not identify a solution to the problem, instead they discuss their agitation and anger about how these violations of Stormont House will harm peaceful relations between the UK and Northern Ireland. This example is perhaps a more intuitive display of anxiety. In contrast to table 5 the speaker themselves seems to experience anxiety.

In this second example too, the algorithm confirmed that the left-hand text is more confident than the left-hand text.

5.2 Dictionary analysis of Brexit dataset

This project was designed to be an open-ended exploration of the text and audio data. Nonetheless, there are a few basic questions and loose hypotheses that helped focus

[5]The Stormont House Agreement was an agreement that was made in 2014 between the British and Irish governments to foster closer relations. It included several different topics including welfare, fiscal policies, and the recognition of the Irish language.

Confident text	Anxious text
Free trade can drive prosperity and economic growth, which are factors in ensuring peace around the world. Obviously, we all want to make sure that we continue to deliver on the huge gains that Northern Ireland has seen through the peace process. I want to ensure that Northern Ireland's economy has the ability to come out of covid with a turbo-boost in growth, to the benefit of all people in all communities in Northern Ireland. Being able to have good, positive trade is part of that.	Secretary of State, twice in the last six months, last week and on 18 March in relation to treaties arising from Stormont House, your Government have announced that they intend to break international law as regards Northern Ireland. [...] Do you understand why people cannot trust your Government? Do you have any handle on the anxiety, the unease and the instability your Government is creating every day in Northern Ireland, which is a fragile society?

Table 6: Example 2 of an anxious and a confident speech

and guide the analysis. One hypothesis is that anxiety was higher in the Brexit negotiations than it was in non-Brexit discussions. This is based on the understanding that Brexit was an unprecedented event; a member-state had never before left the European Union. The decision of the referendum impacted all areas of British society including migration, trade, and foreign relations. This contributed to a general atmosphere of uncertainty regarding the future of the UK, its place in the global economy, and the impact of Brexit on British socioeconomics.

The custom dictionary was merged with the Brexit dataset, giving each document (speech) in the corpus an anxiety score. To examine temporal variation, documents were grouped by month and year, which smooths out some of the noise of grouping the data by day, month and year.[6] Figure 7 shows the levels of anxiety over the time period of the Brexit negotiations (2016 to 2020). I smoothed the trends using a spline. The x-axis is time and the y-axis is anxiety score. The y-axis is a scale of low anxiety to high anxiety. This highlights the relative levels of anxiety in the texts.[7] Figure 7 shows that anxiety varies significantly

[6]When the analyses are too granular, the broad patterns of the data are difficult to discern

[7]The word embeddings that produced the dictionary used for these analyses were on a scale of -1 to +1, yet the minimum sentiment score for a given document is -0.16 and the maximum is +0.031, resulting in a total range of 0.191. This may not seem like a wide range and the anxiety scores themselves may seem insignificant, but this graph shows that the relative anxiety scores of these points in time reveal some

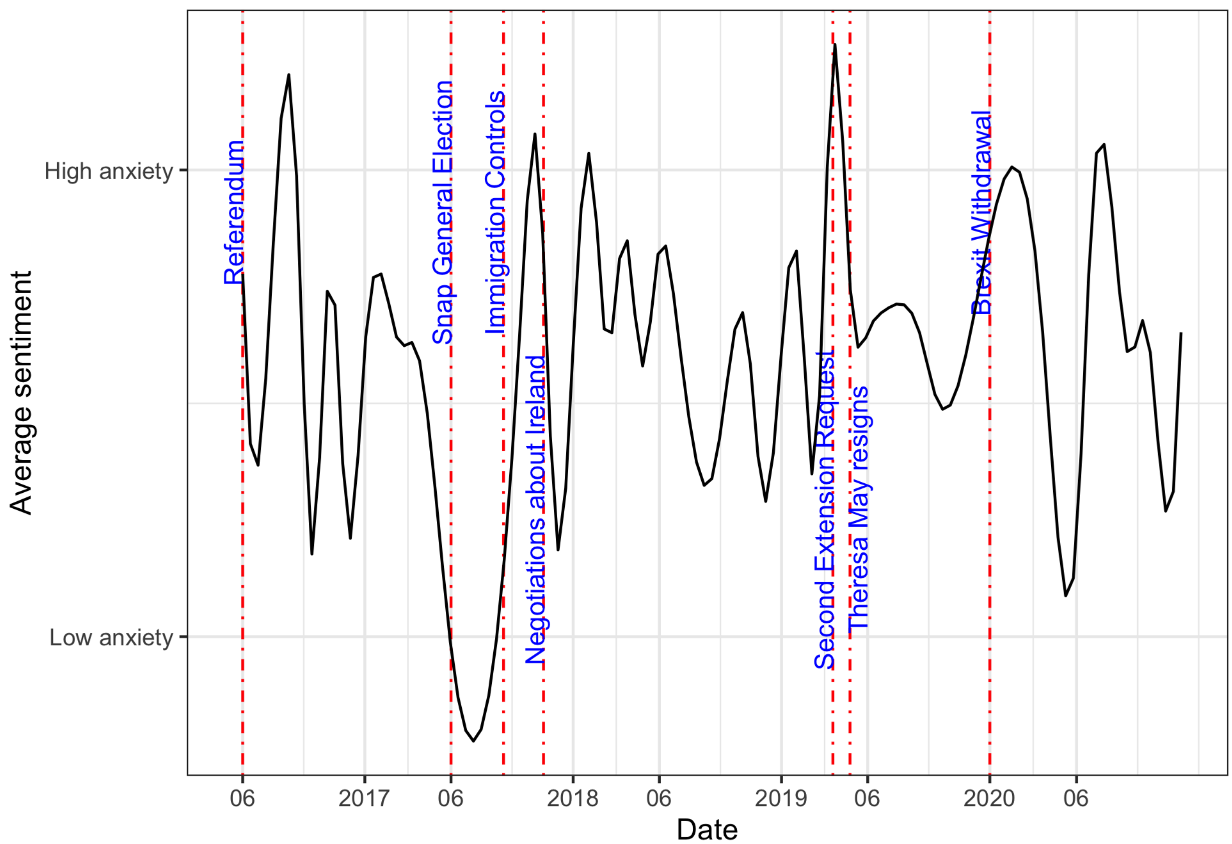


Figure 7: Anxiety over time. X-axis is time, y-axis is level of anxiety. Several important dates are marked with red, dashed vertical lines.

over time, with several sharp changes (both spikes and troughs).

To provide some insight into why these sharp changes in anxiety may have occurred, I overlayed a few important dates onto the graph, denoted with red, vertical, dashed lines. The Brexit negotiations timeline includes a great number of important dates, so the dates included are not comprehensive but may be insightful. I cannot claim at this stage that there is anything beyond an association between the dates and the trough or peak.

Some of the notable troughs occur in mid-2017 and mid-2020. Notable spikes include late 2017 and early in 2019. An important event in June 2017 was the snap general election, called by the Conservative government at the time. Conservatives at the time the election was called had a majority of seats. The election resulted in the Conservatives forming a

interesting patterns.

minority government with slightly fewer seats. Labour gained 30 seats.

A basic tenet of political science tells us that incumbents want to retain their power. In order to retain this position, they need to communicate that they have been successful in office. Politicians have a strategic motive to speak confidently and convey a message of security and stability. Voters do not want their representatives to be uncertain – anxious – about their decision-making. This downward trend in mid 2017 in figure 7 suggests something about the nature of the rhetoric leading up to the election. Politicians, especially members of the Conservative party, have an incentive to lower the levels of anxiety in their speech. Further evidence of this will be presented in the following section.

The trough in mid-2017 is followed by a sharp spike in anxiety in late 2017. This may be explained by the content of the negotiations during that period. An especially important issue following the election in September 2017 was the future of EU citizens in the UK.[8] Free movement –the ability to live and work in any EU member state – is one of the most important principles of the European Union. While legislators were keen to continue the free movement of goods and services between the UK and the EU, the continuation of the free movement of people did not seem as desirable (Home Affairs Committee, June 2018).

The promise to make changes to immigration law and tighten border control were central to the Brexit campaign (see figure 2b). Anti-immigrant and anti-refugee rhetoric and sentiment were rife. Although Muslim refugees were the primary target of the campaign, it was hard to foresee how the post-Brexit government would treat refugees and immigrants from non-EU countries compared to immigrants from European countries. The spike in late 2017 may reflect the high level of uncertainty about the issue of migration law reform.

A notable spike happens in April 2019 right before the second of three extensions was requested. The first extension was requested in March of the same year to extend the Brexit until April. In April an extension request was submitted to extend the exit period until the

[8] During the negotiations there were many issues of focus during a given time period but issues had varying levels of importance. Therefore, although we cannot say for certain that this particular issue is the cause of anxiety in this period, we do know that the immigration status of EU citizens was a very important issue.

31st of October, 2019. The third and final request happened in October 2019 for the exit to be delayed until the 21st of January, 2020. The draft exit bill failed three votes in January, early March and late March, after debates stretched on for days from the end of 2018 into 2019. The spike in anxiety may be associated with the increasing pressure the government was facing from EU leadership and the British public to pass a bill.

Shortly after this event, Prime Minister Theresa May resigned. This was due to her failure to actually reach an agreement and take Britain out of the EU. Although there may arguably be some anxiety around the choice of the new prime minister (anxiety levels are not notably low after she left office), there may also be a decrease in anxiety because Theresa May's popularity was dwindling.

The trough close to the middle of 2020 may be associated with the formal withdrawal from the EU. Negotiations had concluded and the withdrawal agreement had passed through parliament. There were (and continue to be) many ongoing discussions about the relation of the UK to the EU and the impact of Brexit, but the most uncertain period of the negotiations –i.e. negotiating with the EU and getting British parliament to approve of the bill– had passed.

5.2.1 Brexit rhetoric vs non-Brexit rhetoric

One of the central claims of this dissertation is that the Brexit negotiations created an atmosphere of high anxiety, stemming from the uncertainty and lack of protocol for the situation. To support the claim that discussions about Brexit were more anxious than discussions about other topics, I compare the levels of anxiety in Brexit speeches to a sample of speeches that were not about Brexit. The non-Brexit sample includes any other topics that may be discussed in a parliamentary setting (n = 85,243 speeches). Since the non-Brexit data is sampled from the Hansard dataset, which consists exclusively of parliamentary debates, I account for fixed effects of the setting of the speeches (committee meeting versus main chamber debate). The results are shown in figure 8 I measure the average level of anxiety by year from 2016 to 2020.[9] The red point estimates capture the mean anxiety of all speeches in that year and include a 95% confidence interval. The blue points capture the anxiety of a sample of non-Brexit speeches from the same time period. Green denotes the anxiety levels of Brexit speeches from the main chamber.

We observe that the average anxiety of the full dataset of Brexit speeches is significantly higher than the sample of non-Brexit speeches for every year between 2016 and 2019. This is seen in the difference between the red and blue data points and the separation (lack of overlap) of the estimates' confidence intervals.

Comparing speeches in the main chamber that are about Brexit to speeches not about Brexit, discussions about Brexit are higher in average anxiety than the sample of non-Brexit speeches for the years 2016, 2018 and 2019. The difference in means in statistically significant.

An exception is 2017, where the average anxiety of speeches about Brexit is lower than the anxiety of the non-Brexit sample. This may be related to the findings in figure 7, where 2017 exhibited a sharp fall in anxiety, possible associated with the general elections during this year. This analysis lacks a level of granularity as the results are grouped by year, which is a coarse unit of time. However, the results still seem to be significantly affected by the rhetoric of the elections.

To account for the possible bias of the sample of speeches not about Brexit, the same analysis is run on a second sample. Results are shown in the appendix, figure 26.

[9] While conducting this analysis, I did not have the Hansard data for the year 2020.

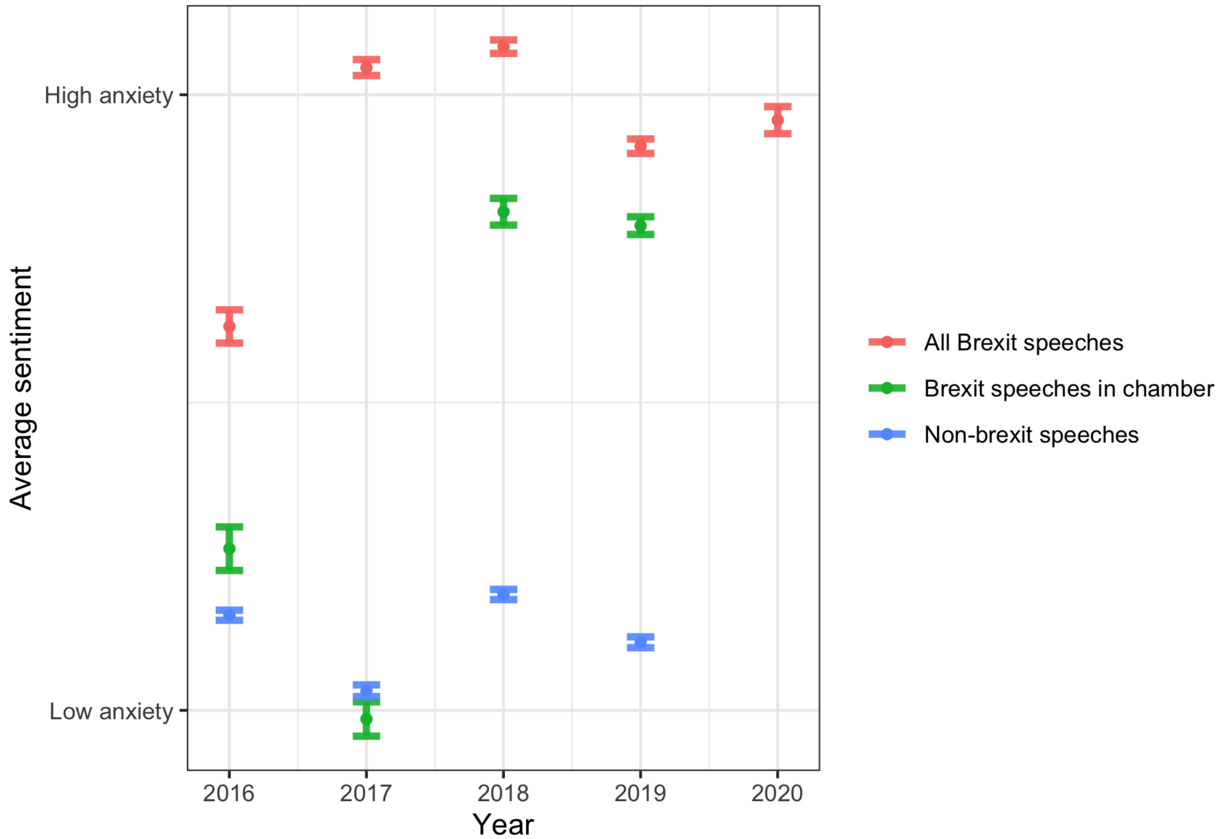


Figure 8: Average sentiment by year of all Brexit speeches (red), Brexit speeches in chamber (green) and sample of non-brexit speeches (blue). Mean anxiety with a 95% confidence intervals.

5.2.2 The main chamber

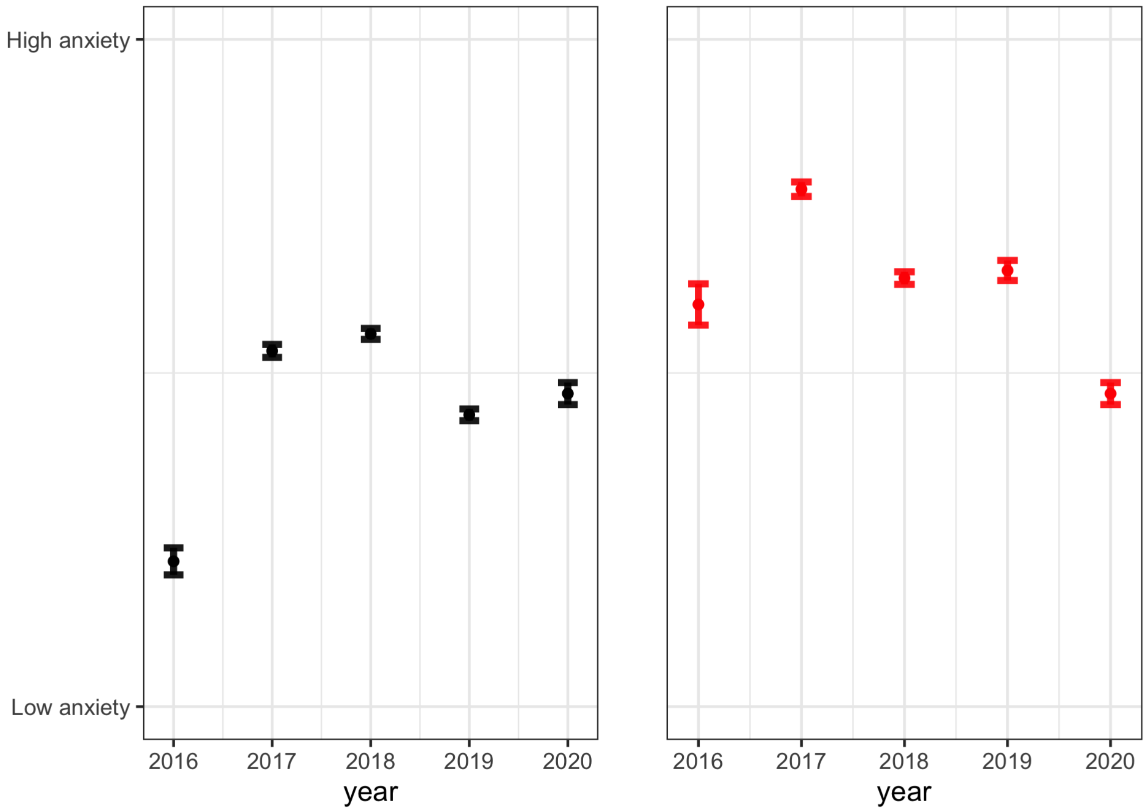


Figure 9: On the left, average sentiment is plotted by year for Brexit speeches including speeches in the main chamber. On the right, average anxiety is plotted for speeches in special committees (excluding the main chamber). Mean anxiety with 95% confidence interval.

Figure 8 sparked an interest in the unique trends of Brexit speeches in the main chamber, as compared to speeches in special committee meetings. To better understand the dynamics of the main chamber, I calculated the average annual level of anxiety in the full Brexit dataset that includes the main chamber (figure 9 [left]) and on the Brexit dataset excluding the main chamber (figure 9 [right]). The results shown in figure 9 include a 95% confidence interval.

As seen in the right panel of figure 9, when excluding the main chamber debates, the average anxiety of the data is high. This dataset consists of all special committee meetings. When including the main chamber speeches, this lowers the dataset's overall anxiety score.

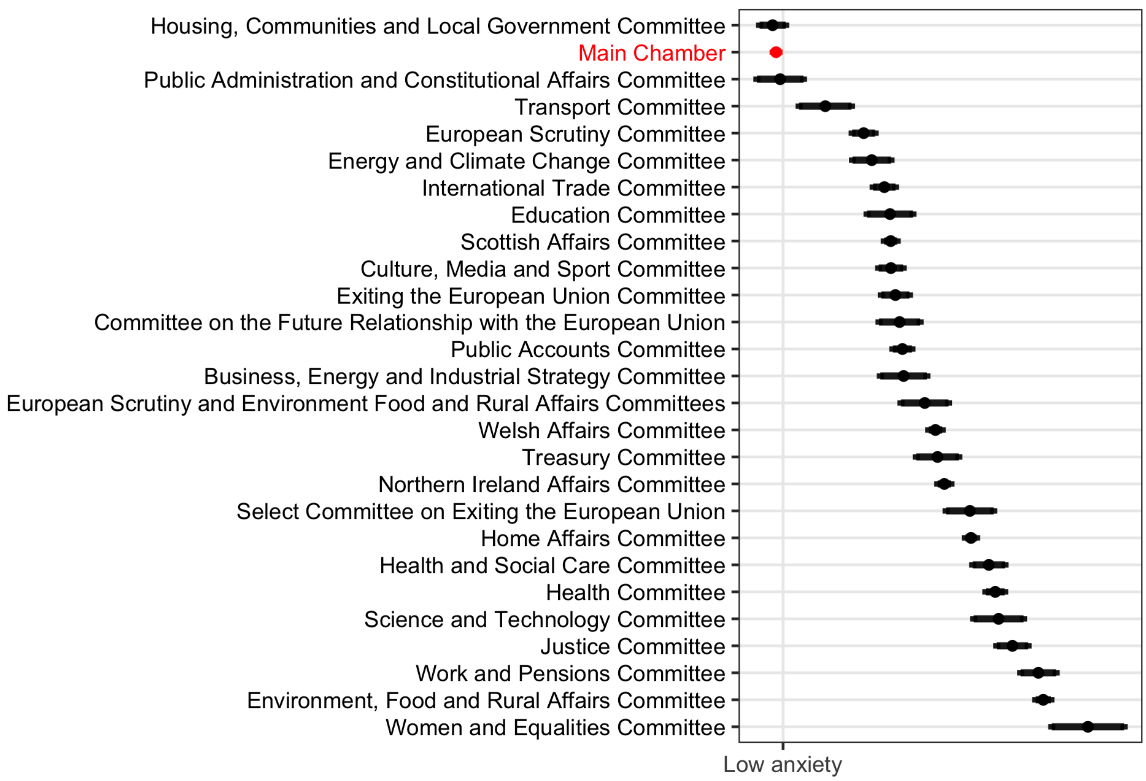


Figure 10: Average anxiety by committee. The y-axis shows the name of the committee, the x-axis shows the level of anxiety. Each estimate includes a 95% confidence interval.

There is a statistically significant difference between the average levels of anxiety every year except for 2020. As noted above, this year does not include main chamber debates.

The anxiety score of the special committees' meetings may be due to an averaging effect. There may be committees with higher anxiety than the main chamber and others with lower anxiety. To account for this, I calculate the average anxiety score of every committee. The results are shown in figure 10. I include a confidence interval for every committee. The confidence interval for the main chamber is very small compared to that of the other committees but is included in the figure.

The average score of the main chamber is statistically significantly lower for all but two of the committees; the Housing, Communities and Local Government Committee and the Public Administration and Constitutional Affairs Committee.

These findings show the differences in anxiety in the political rhetoric of the main chamber compared to the rhetoric in special committee meetings. The relatively lower levels of anxiety is a somewhat counterintuitive finding, due to the popular understanding of the dynamics of the main chamber.

The main chamber is where all MPs congregate to debate bills, legislation, and other parliamentary affairs. Despite the formal nature of this chamber, the main chamber often receives coverage for shouting, heckling, and jeering, as seen figure 11 (Martin, 2022; Reid, 2019; Demianyk, 2020; Barry, 2019). Despite emotions apparently running high, the anxiety analysis shows that the content of the rhetoric in parliament is lower than speeches in committees but higher than speeches about other topics.

This may be due to the performative nature of the main chamber. In this setting, politicians are attempting to argue for their stance on an issue. Committee meetings are not debates even though political opinions arise in the setting of a committee, too. In debates, MPs may be more inclined to either appear confident about the course of the government or, if they are a member of the opposition, to criticize the government's political choices. To explore whether the lower levels of anxiety in the main chamber are associated with party affiliation, the results are further disaggregated.

Figure 11: House Speaker John Bercow calling for order in the House of Commons with MPs in background (Source: New York Times)

5.2.3 Party politics during Brexit

The trends of anxiety in rhetoric in the main chamber evoke normative theories of party politics (government versus opposition). The party in power and opposition have conflicting strategic motives in the main chamber. The party in power's job is to demonstrate that they are doing a good job. We would expect that the rhetoric of government would be more confident since this communicates higher levels of certainty and security to the public.

In contrast, we would expect the opposition to show higher levels of anxiety since this communicates a message of uncertainty (about the government) to the public. This serves their interests since they want to discredit the party in power and secure more votes in the next election cycle.

I estimate the distribution of the anxiety in rhetoric of speakers of the two largest parties, the Conservative and the Labour party in the Brexit dataset. I limit the numbers of speakers to the top 20 speakers, ranking speakers by count of speeches. This is to filter out some of the noise when including the full dataset of Conservative and Labour speakers ($N = 352$). The estimated distribution of the full dataset of speakers is in the appendix, figure 27.

Figure 12 shows the estimated distribution. There is an observable distance between the two distributions, with the Labour speakers being on average more anxious than the Conservative speakers. By conducting a t-test of the means of distribution of Conservative speakers and the Labour speakers, I find that the difference in means is statistically significant at the $p = 0.05$ level.

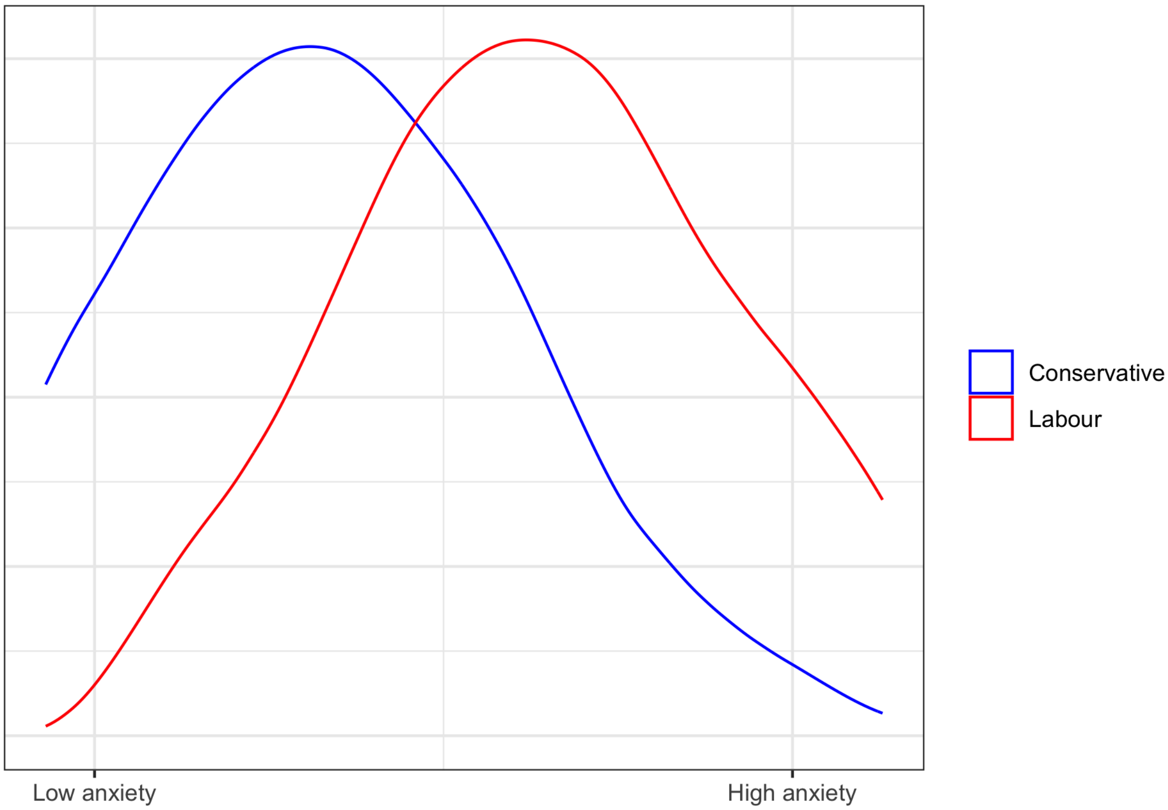


Figure 12: Estimated density of the top 20 Labour and Conservative speakers.

For a more granular analysis, I examined anxiety at the speaker level for all Brexit speeches. Each speaker's average anxiety is calculated as well as a 95% confidence interval. For the sake of visual legibility, I limit the analysis to the top 15 speakers of the Conservative party and the top 15 of the Labour party. The results are shown in figure 13.

Theresa May, leader of the conservative party from 2016 to 2019, has the lowest anxiety score in this subset of Conservative speakers. This estimate is statistically different than all other Conservative speakers in the sample. Similarly, Jeremy Corbyn, leader of the Labour party from 2015 to 2020, has the third lowest anxiety score of this subset of Labour speakers. Party leaders' rhetoric appears, therefore, to be less anxious. This aligns with the findings from earlier sections that government rhetoric is on average less anxious than opposition rhetoric. Party leaders speaking with lower anxiety than other party members may reveal a strategic motive to convey confidence to party members and voters. Jeremy

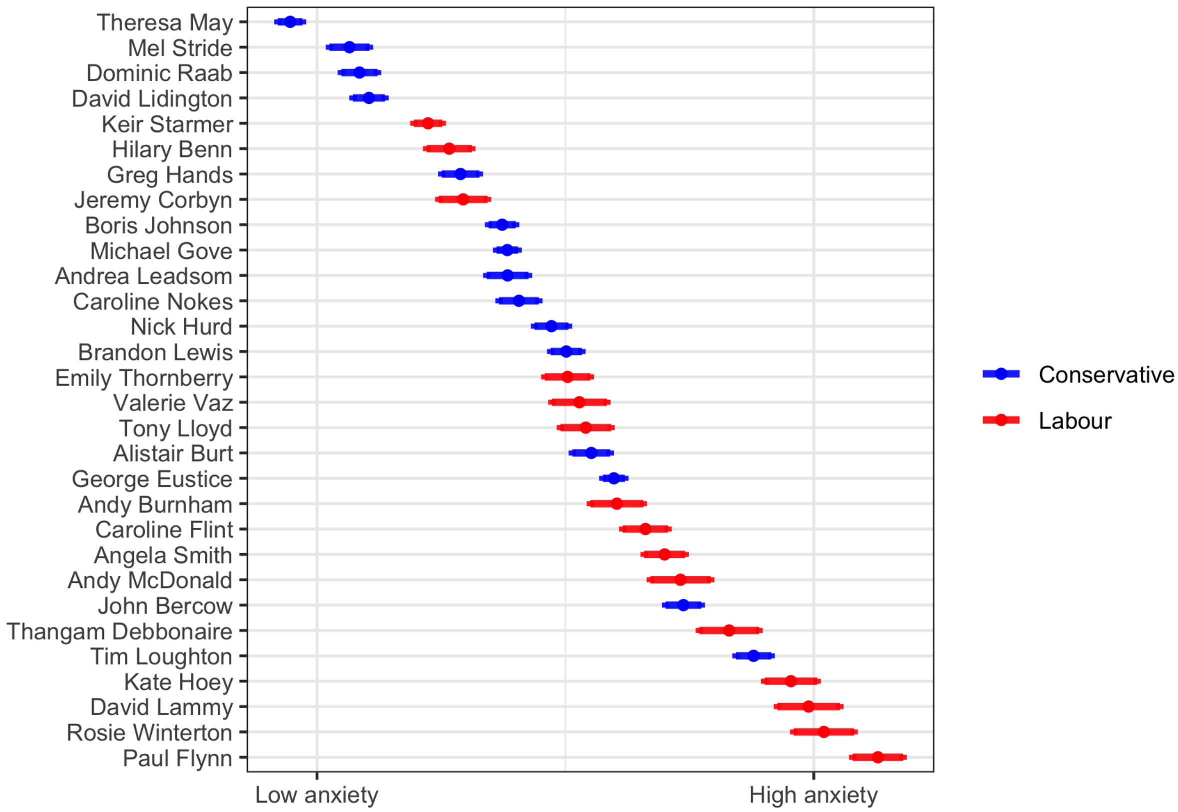


Figure 13: Point estimates of the top 15 Labour and Conservative speakers' anxiety including a 95% confidence interval

Corbyn, although he is a member of the opposition and therefore expected to speak with higher levels of anxiety, still has to keep the Labour party united.

Other notable figures include Boris Johnson, who did not become the Tory party leader until 2019. Johnson is at the middle of the scale. Keir Starmer, who succeeded Jeremy Corbyn as leader of the Labour party in 2020, is the least anxious of the Labour party members in the analysis.

More can be said about the distribution of individual MPs' anxiety scores but this is beyond the scope of this exploratory analysis.

To better understand the dynamics of party politics over time, I measure the average anxiety of Conservative and Labour speakers for the entire time period of the Brexit dataset (2016 through 2020). This analysis includes only speakers in the main chamber. The results are displayed in figure 14.

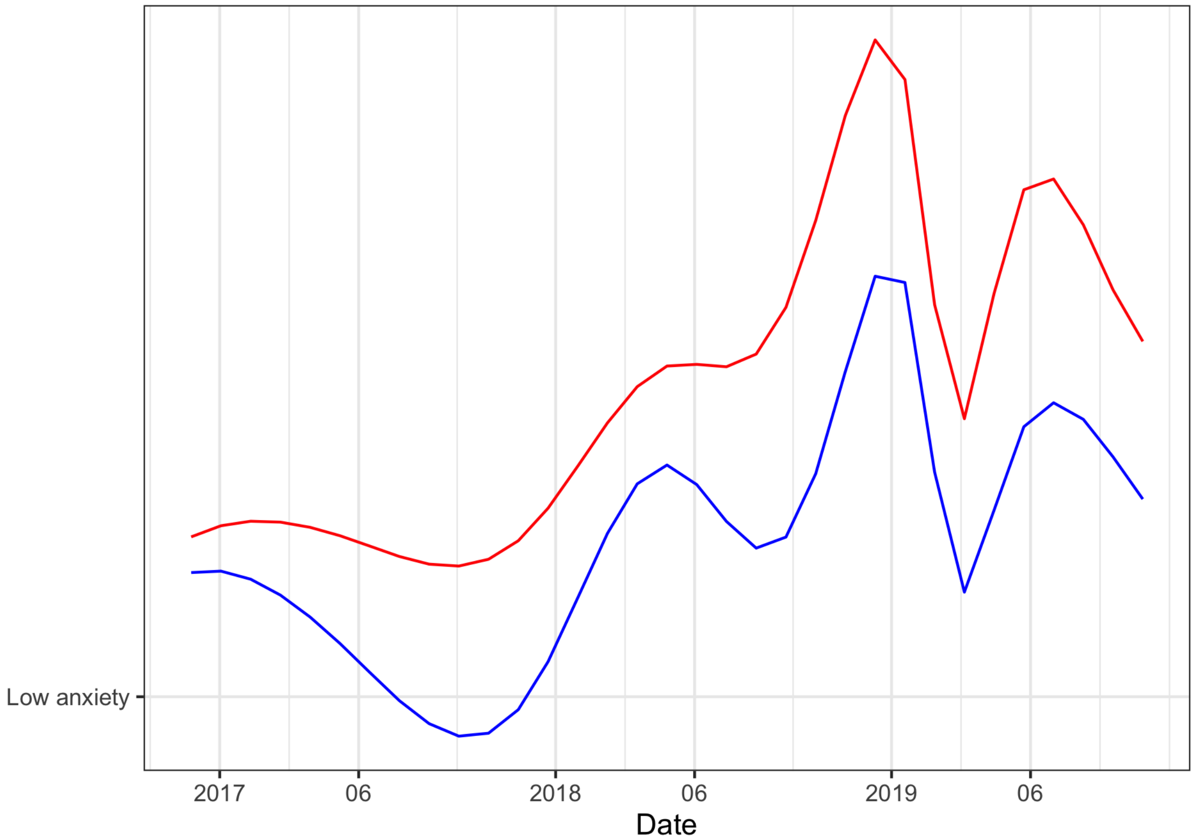


Figure 14: Anxiety over time for Labour (red) and Conservative speakers (blue) in the main chamber

The anxiety in the rhetoric of Conservative and Labour speakers reveals a distinct separation over the entire course of the negotiations. Conservative MPs consistently speak with lower anxiety while Labour MPs consistently speak with higher anxiety. The anxiety levels are continuously distinct while strikingly parallel. Both paries display a dip in anxiety in the latter half of 2017 (following the election). The anxiety levels are the overall lowest in late 2017. This is followed by a rise in anxiety in early 2018 and an even sharper rise at the end of 2018. In 2019, levels of anxiety plummet in early 2019.

The highest levels of anxiety for both Conservative and Labour MPs is in late 2018 going into early 2019. This is when the first withdrawal bill was published and debated. It was first voted on in early 2019 when it was voted down, with 202 ayes and 432 noes. One the reasons why the anxiety is so high during this period may be the long duration of the debates and the inability to pass an exit bill more than two years after the referendum.

The lowest level of anxiety in the main chamber for both parties is in late 2017. Following the snap general election in June 2017, the pace of negotiations with the EU intensified and significant progress was reported (Council, Council). The lower anxiety levels may reflect the lower levels of uncertainty about the state of the negotiations as well as the fact that the snap elections reinforced the Tories' power.

The rhetoric in the main chamber exhibits unique patterns along party lines. To understand whether party lines were generally a facet of parliamentary politics, I examined the levels of anxiety only in committee meetings disaggregated by party. The results are in figure 15.

There is a clear contrast to figure 14, as the two trend lines overlap most of the time and do not show clear separation on average. The anxiety of Conservative speakers exhibits higher levels of fluctuations compared to Labour speakers, with noticeable sharp dips in anxiety in late 2016 and the latter half of 2019. The overlap between these trend lines may reflect the composition of committee meetings, which is entirely backbencher MPs. Moreover, since the theatrics of the main chamber are not as relevant in the context of a committee meeting, the convergence of the levels of anxiety of MPs of these two parties may reflect a convergence in agenda. There may not be as many strategic motives to discredit political opposition.

5.2.4 Anxiety beyond Brexit

Is the ability to distinguish between the levels of anxiety in Conservative and Labour rhetoric unique to the context of Brexit? To answer this question, I analyzed a dataset that goes beyond the 4-year Brexit period. As discussed in chapter 4, the data used to train the word embeddings model was a set of British parliamentary speeches spanning 20 years. Since this data was already part of the analysis, it was a simple process to conduct a dictionary analysis using the same anxiety lexicon that has been used until this point.

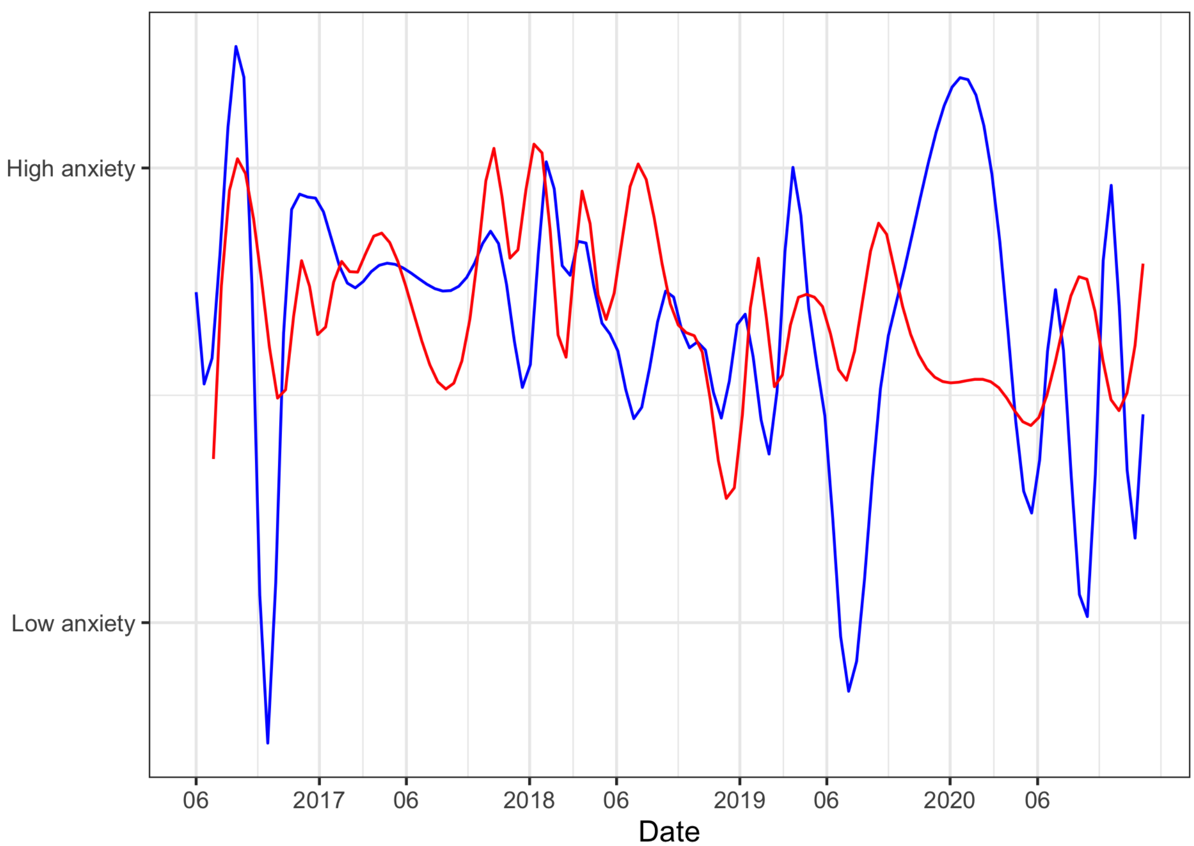


Figure 15: Anxiety over time for Labour (red) and Conservative speakers (blue) in committee meetings

For comparability to the analyses in the previous section, I examine only Conservative and Labour speakers in the larger Hansard dataset. Grouping by month and year, I calculate the average anxiety of the Conservative party and the Labour party. The results of estimating anxiety levels over time are in figure 16.

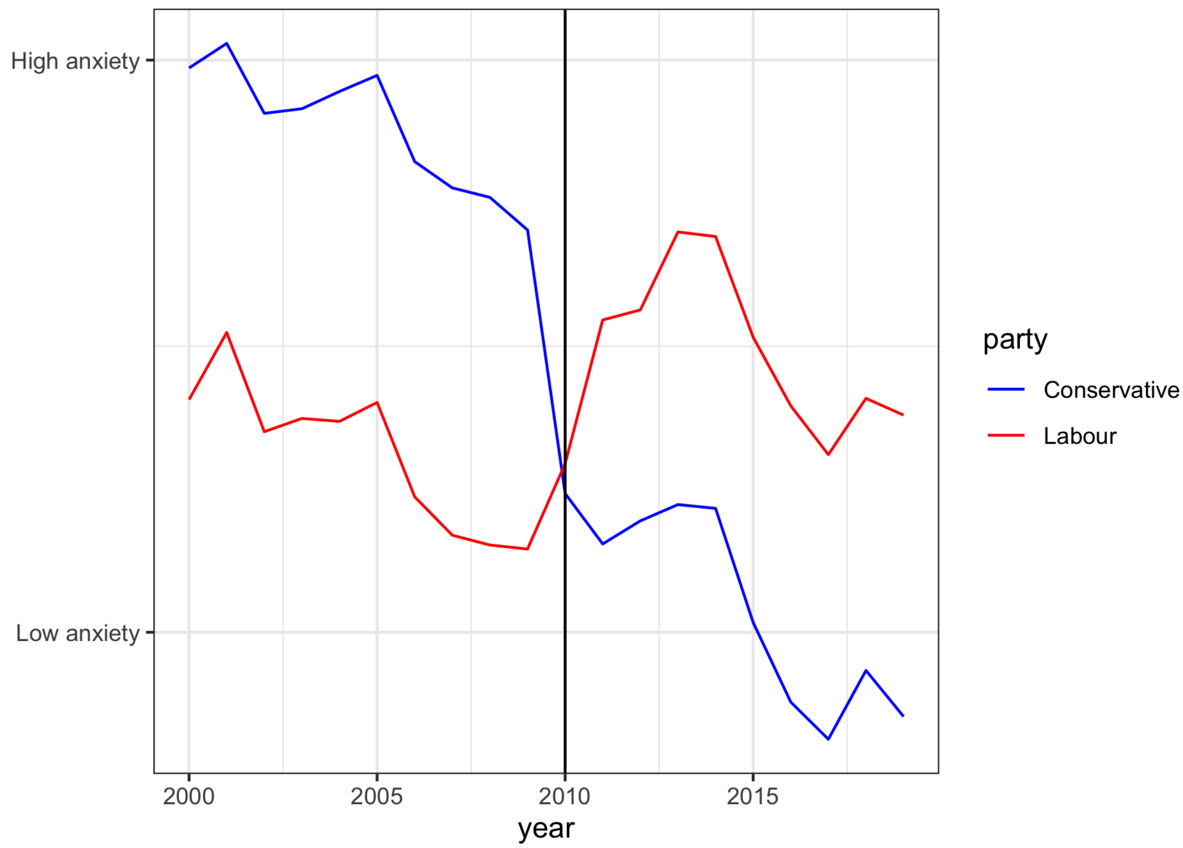


Figure 16: Average anxiety from 2000 to 2019 in the main chamber for the Conservative party (blue) and the Labour party (red). The vertical line in 2010 marks an election year where the government changed from Labour to Conservative.

This analysis reveals quite a distinct temporal trend. Prior to 2010, the Conservative party rhetoric is significantly more anxious than the rhetoric of the Labour party. In 2010 there is a vertical line marking an important change of government from the Labour to the Conservative party. After this point, the rhetoric of the Labour party abruptly switches to being significantly more anxious than the rhetoric of the Conservative party. The two parties also show parallel trends in these two time segments of 2000 to 2010 and 2010 to 2019. The levels of anxiety move in the same direction, mirroring the parallel trends of the rhetoric during the Brexit negotiations.

This analysis appears to show the importance of party membership in determining anxiety levels. This large-scale temporal analysis can be viewed through the perspective of the theory that opposition discredits the government. One way to discredit the government is through rhetoric. In the context of this research, anxiety conveys a message of uncertainty and fear, which can signal to the voters that the government is performing poorly. To synthesize the discussion of anxiety in rhetoric about Brexit, in the appendix, figure 29 includes two time trends of anxiety on the same graph. The results show that the Brexit negotiations displayed unusually high levels of anxiety.

5.3 Discussion

This section provided some interesting insight into how anxiety in political rhetoric operates in British parliament and into the idiosyncrasies of anxiety during the Brexit negotiations. Discussions about Brexit show higher levels of anxiety in general. When accounting for whether the speech is happening in the main chamber or a committee, the levels of anxiety are lower in the main chamber than the average anxiety of special committees. This finding holds when comparing the main chamber's average anxiety to each individual committee, with few exceptions.

The rhetoric in the main chamber shows unique patterns along party lines, where the Conservative party shows lower levels of anxiety than the Labour party. This appears to have to do with the division of government and opposition, since in an analysis that extends beyond the context of Brexit, the party that was in power consistently showed lower levels of anxiety.

This analysis focused only on the emotion in the content of political rhetoric in the House of Commons. In the following section, I introduce audio analysis as a methodological approach in political science. Subsequently, I analyze committee data, using emotional arousal (pitch) to capture the emotional state of the speaker. These findings supplement the insight of this methodological chapter and add to the complexity of measuring the emotional state of the speaker.

6 Audio as a source of data

There is not a large and established body of scholarship in political science on audio analysis. This is a somewhat surprising fact, considering how important audiovisual political content is in the political sphere. Consider, for example, debates, campaign speeches, media statements, or even chants at a picket line. We intake a lot of political information through auditory channels.

In recent years, there have been a number of seminal pieces that focus on the importance of audio data in political science (Dietrich, Hayes, and O'Brien, 2019; Knox and Lucas, 2021; Dietrich, Schultz, and Jaquith, 2018).[10] Other scholarly fields, such as linguistics and psychology, have established a tradition of audio analysis (Bezooijen, 1984; Schroeder and Epley, 2016). This scholarship is accompanied by sophisticated tools and software that automate many of the conventional steps in audio analysis. The methods and findings from in these fields should not be overlooked as political science seeks to establish its own tradition. Of course, questions that are asked by political scientists are different than those asked by linguists and scholars of psychology but they are not mutually exclusive (hence the development of subfields in political science like political psychology). In this section, I will only review political science scholarship examining the political implications of audio data. Introducing the methods used in this dissertation, I review psychology literature to inform a deeper understanding of the audio metrics used in this dissertation. This is an interdisciplinary dissertation that bridges psychology and political science.

In an article by Dietrich, Hayes, and O'Brien (2019), the authors analyze Congressional floor speeches to identify the emotional state of the speaker. They argue that vocal pitch is a reflection of true emotional state of the speaker. Pitch is an aspect of speech that is difficult for the speaker to control and therefore reflects their level of emotional engagement with the issue. They hypothesize that female representatives in Congress speak with

[10] In fact, many of the articles mentioned in this section inspired the conception of this dissertation

greater emotional intensity when speaking about issues that relate to women. They test their hypothesis using a dataset of text and audio data from 74,158 floor speeches in the US House between 2009 and 2014. To prepare the data for analysis, they use timestamps found in closed captioning to split proceedings into individual speeches. Using the software `Praat` to measure average vocal pitch in an individual speech, the authors scale this measure according to the individual speaker's baseline vocal pitch. Using dictionary analysis, they code whether a given speech is about women or not. They find that female Members of Congress deviate higher than their baseline pitch when talking about a women-related issue. Their finding that emotional attachment to the issue at hand is revealed by pitch is supported with additional analyses that show that more emotionally intense congresswomen are rated more highly by women's interest groups. This research has important implications for studies of the policy preferences of policymakers.

In another study, Dietrich, Schultz, and Jaquith (2018) show that pitch influences the likelihood a floor speech is aired on major news networks. Using a dataset of 145,706 hours of cable news broadcasts, they consider what types of speeches in Congress receive coverage. Emotionally intense speeches (speeches with a higher vocal pitch) and speeches with negative valence are more likely to elicit an emotional response from viewers. Since networks choose content according to what will increase their viewership (i.e., what is newsworthy), networks choose to cover speeches that will engage their audience through emotions. Media is often characterized as sensationalist. The authors develop a piece of software called `AudioNewser`, which identifies which floor speeches appear on cable news. This automates the important but tedious step of filtering through the data to find relevant content. They scale the texts by valence using the Google Cloud Natural Language Application Programming Interface. They find that an increase in vocal pitch from the speaker's baseline increases the likelihood that a floor speech is aired. Speeches with a negative valence are also more likely to be aired. This research gives a better understanding of how politicians strategically use emotions to gain media coverage and expand the reach of their political agendas.

An important study by Knox and Lucas (2021) identifies skepticism in justices' tone during judicial proceedings. Audio data, according to the authors, contains nuanced information that cannot be captured by text models. The metric of interest is the speaker's tone, which is a measure of uncertainty. The authors' definition of tone includes several features like the zero crossing rate, the loudness of the speech and the several elements of the spectrum of the audio segment, resulting in a composite measure of speech. They also look at the context of the speech, including dynamics between speakers, interruptions, and interjections. The authors created a new software package for their novel analysis of audio data. This complex analysis studies conversational flow in order to be able to predict the outcome of a legal case in court.

Knox and Lucas's composite measure of audio is more complex than that used by Dietrich and coauthors (deviation from baseline pitch). The complexity of the measure is arguably a double-edged sword, helping perhaps uncover more nuances of the speech that a single metric would allow while creating an analysis that is a black box, i.e. it is difficult to understand the role and the meaning of each facet of audio.

Studies have also measured voters' preferences for the vocal pitch of political candidates (Anderson and Klofstad, 2012; Klofstad, 2016). Men tend to have lower-pitched voices and women have higher-pitched voices. Anderson and Klofstad (2012) find that when female candidates are running, voters prefer female candidates with low-pitched voices. They also find that men prefer male candidates with low-pitched voices. Women do not distinguish between male candidates with high-pitched and low-pitched voices. These types of studies also indirectly convey findings about the gender dimensions of political decisions.

This dissertation builds on the audio analysis of articles such as the ones discussed above. However, the research bridges new audio analysis with the established techniques of text analysis, which can also be used to measure the emotional content of speech. Rather than rely on either audio or text analysis, I synthesize the findings from both types of analysis in order understand how the metrics relate to one another and provide a more complete picture of the speaker's emotional state.

6.1 Why measure pitch?

What is pitch? Pitch is also known as fundamental frequency or F_0. Pitch can be mathematically defined using the following equation:

$$F_0 = \frac{1}{2L}\sqrt{\frac{\sigma}{\rho}}$$

where L is the vocal fold (wave) length, σ is the longitudinal stress on the vocal folds and ρ is the vocal fold tissue density.

Owren and Bachorowski (2007)'s biological/physiological explanation of pitch is as follows:

> "For vowel sounds, source energy arises from air passing through the *glottis*, which is the opening between the *vocal folds* (known informally as the vocal cords). [...] All vowels are produced this way, and it is the regular vocal-fold vibration involved that gives these sounds their relatively clear, tonal quality. The basic rate of vocal-fold vibration is referred to as the *fundamental frequency (F_0)* of a voiced sound, and this physical proprety is also a primary correlate of the psychological experience of pitch." (page 244)

An increase in pitch is an increase in F_0, as vocal folds vibrate at a higher frequency. Both positive and negatively valenced states tend to show increased speech rate and changes in the source energy of the sounds, i.e. the mean F_0 and the F_0 range both increase.

There are many ways to try to measure of emotion. In a review piece, Mauss and Robinson (2010) classify four general approaches. (1) Self-reported emotions; for this approach, participants are simply asked to classify their own emotions. This approach is useful insofar as people can accurately capture their own emotions in the moment. The approach is not as accurate for predictions of emotional states or recollections of past emotional states.

There is uncertainty, however, about the extent to which people can identify their own emotions. Researchers have found, however, that dimensional approaches to emotion, as opposed to discrete emotions, are reliable models.

While this approach has its merits, there is of course no metadata including the speaker's self-reported emotional state.

(2) The second general approach is autonomic responses. "The autonomic nervous system (ANS) is a general-purpose physiological system responsible for modulating peripheral functions" (Mauss and Robinson, 2010, p. 102). The most common metrics of emotions using the ANS are electrodermal responses (skin conductance) and cardiovascular measures (including heart rate, blood pressure among other things). Research using ANS responses has not been able to clearly identify the autonomic responses that accompany distinct emotions. There is more evidence to suggest valence and arousal can be more easily identified using ANS responses.

This approach is also not feasible for this project since the only data I have is audio and text. Additionally, there are limitations to using this approach without proper lab equipment.

(3) Neuroimaging and EEG (electroencephalography) has revealed some interesting findings about the activity of the brain for certain emotional states. This is quite an advanced approach to research on emotions and not suited for this research.

Politicians cannot be placed in MRI machines while making political statements. This renders this approach impossible for the context of this study.

(4) Finally, there are a range of behavioral measures of emotion. Behaviors include facial responses, vocal characteristics, and body language. Vocal characteristics appear to be quite good at measuring arousal. Facial characteristics appear to be quite good at measuring the valence of a person's emotional state. Body posture has not received as much attention, but is a promising approach.

The association of pitch and emotional arousal has been consistently supported by new research. Moreover, pitch has been found to be an "honest" measure of emotional state. Ekman et al. (1991) conducted a study that showed that pitch was a good indicator for telling whether the participant was lying or not.

Although research on emotions has developed many metrics, given the data available for this project and the variation in validity of these many metrics, pitch is a simple and reliable approach to capturing emotional arousal.

Pitch is a good candidate for measuring emotion in the context because of the nature of the data. The political rhetoric has already transpired and features like autonomic responses or self-reported emotions have of course not been documented. The data is also occurring in the real world and lab-setting measurement such as neuroimaging are not feasible. Therefore, considering the information we do have, text and audio, vocal characteristics are the most appropriate. Facial responses could also be a possible candidate for sentiment analysis but the video formats are not consistent enough across meetings.

6.2 Methods for audio analysis

Using a subsample of audio recordings from the Committees dataset and machine learning, I was able to produce timestamped transcripts. The transcripts have effectively the same content as the scraped and parsed transcripts used in the text analysis chapters. However, these transcripts are more accurate, i.e. they are verbatim transcripts of the meetings as opposed to the parsed transcripts, which have been edited and cleaned.

In order for the researcher to be able to compare anxiety in text to emotional arousal in audio, they need to be able to extract or identify the time segments in the audio that correspond to the text. In an ideal case, the corpus the researcher is working with is accompanied by closed-captioning (CC) or subtitles. If that is the case, the researcher simply has to obtain the subtitle file (in .srt [SubRip Subtitle File] format, for example) or an analogous file with timestamps and text. Unfortunately for this project, this information was not available. Although the House of Commons has now introduced CC for parliamentary meetings, this feature is not consistently available and they have yet to retroactively enabled this feature in older videos.

Details of the process used to conduct audio analysis are provided in the following subsections. The first step is to identify different speakers' voices in the file. Subsequently, the full audio recording is split according to the speaker-level timestamps. The individual audio segments are transcribed and pieced together, allowing for the text analysis segment. Finally, using audio analysis software, the measurements of interest can be captured from the individual files.

To provide future researchers a better sense of the scope of audio analysis software available for similar tasks, I detail which software worked and which did not for my particular study.

6.2.1 Aligning text and audio

Aligning the audio to the transcript is the most tedious and complicated step in this process. In the case where the researcher is only using a few minutes or hours of data, this step can be done manually by listening through the audio or watching the video. There are several resources for an assortment of audio-analysis-related tasks such as transcription and the extraction of features like fundamental frequency and volume. Resources for forced alignment are fewer and often less straightforward to use. A list of the most frequently-used forced alignment tools can be found on github at `https://github.com/pettarin/forced-alignment-tools` by Alberto Pettarin. Pettarin is also the creator of `aeneas`, a forced alignment (audio alignment) package for python. This list is not maintained very regularly.

For this project, the desired output of forced alignment should look like this:

> **Chair:** I welcome the Immigration Minister and Ms Wilkinson, who are giving evidence to us today. We are very grateful for your time. We want to cover some of the issues around the Brexit proposals, proposals around the existing immigration

system operation, and Brook House, to feed into a series of different inquiries. Can I start by asking you some factual questions about the Brexit arrangements you have planned? When do you plan to start registering the 3 million EU citizens? => [00:00:00, 00:00:15]

Brandon Lewis: First, thank you for inviting us to do this, this morning; that works for us, too. It is a pleasure. I think this is my first Home Affairs Select Committee appearance since I have been in the Home Office, so it is really good to have a chance to come and have a conversation. => [00:00:16, 00:00:23]

Tim Loughton: No pressure. => [00:00:24, 00:00:25]

Brandon Lewis: Yes, absolutely: no pressure there. In terms of the settled status of European citizens who are here, obviously we are dependent on that stage of the negotiations completing so that we can move forward, but the intention is to do it later in 2018. => [00:00:26, 00:00:37]

Each speech is accompanied by an interval marking the beginning and end of the audio segment corresponding to the text.

Many of the tools listed on the link noted above are for alignment at the phoneme-level since phonemes are often the unit of analysis in linguistics. Since much of the software in forced alignment has been developed such analysis, the output of such tools is time alignment at the phoneme-level. Moreover, for many of the software packages, the use must input a transcript with timestamps for each utterance, which is what we need the output to be in this project. This renders all tools that perform phoneme-level alignment unsuitable for this research.

Narrowing down the list to tools that conduct alignment at the utterance level, I initially found `aeneas` to be the most suitable for the task. `Aeneas` is a relatively simple piece of software. Its main advantage of this software is that it only requires a transcript –what was *recorded* as being said– and an audio file –what was *actually* said– as input and outputs a file with timestamps for each utterance from the transcript. The granularity of the utterances (word-level, sentence-level, speech-level) is determined by the user. This software was not ultimately suitable for the audio files in the data because the transcripts and the audio were too dissimilar. This is the main disadvantage of this software, namely it requires that the transcript and audio be mostly the same. In my case, the transcripts that had been provided by the House of Commons had been edited and cleaned to reflect a clearer version of what the speaker actually said. This is a process that makes it easier for the general public to follow the meetings. Any mistakes speakers made (stuttering, saying the wrong word, repeating a word or phrase, correcting themselves while speaking) had been edited out. These are normal parts of human speech, we often misspeak or are unable to maintain a clear, linear train of thought. Yet when these normal parts of speech appear in a transcript, it can seem uninterpretable. Moreover, and perhaps more importantly, many files had been "purged" of outbursts by members of the committee or witnesses. `Aeneas` was therefore not suited for the task, since it requires a mostly accurate transcript.

Given that existing forced alignment software was not suitable for the data used in this dissertation, I decided to use a combination of (1) speaker diarization software and (2) transcription software instead of forced alignment software. This approach has the same result – a timestamped transcript – as forced alignment. When researchers do not have access to a transcript of an audio file, this is a roundabout way to obtain a transcript. Speaker diarization is the process of identifying when different speakers are speaking in an audio file. The software distinguishes between different speakers' voices and outputs a series of time segments, each delineating when a speaker is talking. The limitation of diarization is that it does not transcribe the speech. Transcription software takes an audio file as input and outputs a transcript of the audio. Restricting the list to free, open-source software, I ended up combining `pyannote` for speaker diarization and OpenAI's `whisper` for transcription.

6.2.2 Diarization and transcription

The full algorithm that details the steps for audio analysis are shown in Algorithm 3. The information I needed to extract from the audio was (1) who was speaking, (2) what they said, and (3) when they said it. This turned out to not be a straightforward task. Thankfully, other researchers had tackled a similar problem and by modifying existing code, I was able to produce a pipeline that efficiently takes an audio file and outputs a transcript.

The first step in the process is speaker diarization. This involves identifying when different speakers are speaking in an audio file. I chose the `pyannote` package for this task. The software only requires an audio file, but the user can choose to specify certain parameters such as the number of speakers, if this information is known ahead of time. I included this parameter to better control the process but found that not specifying this parameter still resulted in an accurate count of speakers. The diarization software is computationally expensive without a GPU. Before using the GPU, an audio file that was 1.5 hours long took about an hour to diarize on a 2022 MacBook Pro M1 chip. I decided to use a paid version of Google Colab's GPU because the GPU on the free version did not have enough RAM for the process. After upgrading to the Colab GPU, the process took about 5 minutes, which was a massive computational gain (12 times faster!). For files this long the RAM reached up to 15GB for some audio files. The computational cost is something to consider given the constraints of many people's computers, even with the speed and power of many modern machines. Therefore, cloud computing is a advisable step and is becoming exceedingly affordable.

Diarization outputs a timestamp and a speaker ID. The software performed well and was even able to distinguish speakers when they spoke at the same time. There were some issues, including not being able to distinguish speakers with similar voices when the audio quality was not good (e.g. technical issues with microphones during meetings). An example of the output is as follows:

[00:00:00.008 – 00:00:14.354] SPEAKER 03

[00:00:13.404 – 00:00:31.112] SPEAKER 02

[00:00:29.295 – 00:00:41.230] SPEAKER 03

The next step in the algorithm is to transcribe the audio. Using the timestamps from the diarization step, I could then loop through the audio file and transcribe the section of the audio delimited by the start timestamp and the end timestamp. The transcription was done using OpenAI's `whisper`. The software includes many pretrained models which perform very well. I chose the smallest available model due to time constraints. I did not have the resources to compare the performance of different models. I specified the language the audio was in although the software can autodetect this. I discovered that when speakers had an accent, e.g. Irish, the software would occasionally try to transcribe in Gaelic. There was also an instance where a speaker spoke briefly in Welsh, but I removed this segment from the analysis.

For this task, I used the M1 GPU to slightly speed up the process. Transcription of an 1.5-long audio file would take about 45 minutes to an hour. The accuracy of the results was quite good, and overall resulted in a more faithful transcript (with repetition, filler words and mistakes made by the speaker). However, the performance definitely did suffer with certain accents such as Scottish, Irish and Welsh accents.

The final step, which was quite a time-consuming manual task, was to go back to the original transcripts from the House of Commons and to match the speakers' names to the transcript. The transcript outputted by the script only labeled speakers by number, e.g. SPEAKER 03. Using a sample of a speaker's speeches, I was able to map the speaker's number ID to their name. This step was performed on each individual transcript.

6.2.3 Splitting the audio file into segments

Splitting the audio file was achieved using the software `Audacity`. `Audacity` is a free, open-source audio editor. The user can open the full audio file and choose to import objects called "labels." Labels are contained in a text file with three columns; the start timestamp of a segment, the end timestamp and a label for the segment. Based on this file, the software will split the full-length audio file into as many labeled segments as there are using the timestamps. Finally, the user can export all the individual audio files for analysis. I removed all segments that were less than 3 seconds long. The result was 18,335 audio files for analysis.

6.2.4 Extracting pitch of individual speakers

The fundamental frequency, or pitch, of each file was calculated using `Praat`. `Praat` is a standard tool for audio analysis. Developed by researchers at the University of Amsterdam, this free tool can be used to conduct a wide range of basic and advanced audio analysis tasks (Boersma and Weenink). Pitch is calculated by dividing the autocorrelation of a windowed signal by the autocorrelation of the window itself. I use the default settings for the selection of pitch window.

The measure of pitch for each speaker is then subtracted from the mean pitch of the speaker. Other researchers take this additional step to account for any individual-level characteristics such as gender (women speak on average at a higher pitch than men) (Dietrich, Hayes, and O'Brien, 2019). Standardizing pitch focuses on when a speaker deviates from their baseline and the relative magnitude of the change.

Algorithm 3: Audio processing and pitch extraction

For a set of audio files A of size M where each file is denoted a_m with $m \in 1 \dots M$

1. Diarize all files using audio analysis software (e.g. pyannote) (Optionally include the number of speakers)
   (a) Store output as text file with each line including start time stamp, end time stamp and speaker ID e.g. [00:00:00.008 → 00:00:13.421] A SPEAKER_04
2. Segment audio file using timestamps from previous step and transcribe using transcription software (e.g. OpenAI's whisper)
   (a) Remove all audio files shorter than 3000 milliseconds
   (b) Export transcriptions of audio segments into a single audio file
3. Using transcript generated in previous step, split full audio file into individual audio files by speech
4. Use speech analysis software, calculate pitch of individual audio files (e.g. Praat)

6.2.5 Pitch is not enough

Humans typically understand language by listening to speech, processing what is being said, and by looking at the speaker. In other words, there are three components to text; the verbal component, the audio component, and the visual component. Language involves all of these dimensions, so why should models of language limit themselves to just text, just audio or just video? In this project, I will examine how coupling audio and text metrics might help measure the quantity of interest, anxiety or more broadly, emotion.

In the following analysis, I conduct a qualitative comparison of two items. First, I gather a human interpretation of the anxiety and emotional content of speech, examining both the verbal content and the audio recording itself. I compare this to the computational estimation of the levels of anxiety and emotion of the speech in both the text and audio. Second, I assess how the text metrics of anxiety compare to the audio metrics of emotion. In the pursuit of understanding how we may better represent the speaker's emotional state, I assess the patterns of how these emotion in audio and emotion in text move. The findings show that the two metrics are not always correlated but reveal an interactive relationship.

7 Sentiment analysis using audio and text data

7.1 Qualitative analysis

In this section, I examine a 10-minute segment from a Home Affairs Committee on Home Office Preparations for Brexit. This sample was chosen because speakers exhibit emotional range, allowing for more insight to be drawn from the sample. A qualitative approach allows for a more meticulous and careful approach to subject matter – synthesizing audio and text analysis – that is relatively unfamiliar to researchers. By closely examining the movement of audio and text metrics, and how these correlate with the speaker's party identification and role in the meeting, I was able to begin to understand the relationship between these variables. Subsequently, I was able to examine whether these patterns occurred at the dataset level by conducting a quantitative analysis to confirm or challenge my qualitative insights.

Members of Parliament that are present in this segment include Tim Loughton (Conservative - East Worthing and Shoreham), Kate Green (Labour - Stretford and Urmston), and Yvette Cooper acting as chair (Labour - Normanton, Pontefract and Castleford). Witnesses in this segment include Shanker Singham (CEO, Competere and Chair of the Alternative Arrangements Commission Technical Panel) and Tony Smith (Panel member of the Alternative Arrangements Commission Technical Panel).

To contextualize their discussions, the Irish border – with Ireland being part of the EU and the UK leaving the EU – was a highly salient issue during the Brexit negotiations because of the fragility of Irish-British relations. The Alternative Arrangements Commission Technical Panel was formed in 2019 to "to develop credible and practical Alternative Arrangements relating to the Irish Border" (UK Parliament, 2019). In this snippet, the meeting attendees are discussing the transition from the single, effectively borderless, market to the creation of border where the movement of goods has to be monitored. The witnesses discuss technological solutions for addressing smuggling and criminal activity that might happen at the border. MPs also ask questions about the timeline for the implementation of these changes. All audio files for the quotes included in this section can be found in the supplementary materials section.

7.1.1 Procedures for annotation

I developed a systematic approach to annotating the data. To annotate the transcript, I classified a speech as anxious, confident or neutral based on the content of the speech itself. I did not rely on speculative assumptions about the underlying emotional state of the speaker. Rather, I evaluated the word choice, including key words that I considered anxious (e.g. "smuggling"), and other textual characteristics that conveyed anxiety (e.g. adjectives such as "urgently").

To annotate the corresponding audio segment, I had to follow a longer procedure. The main issue I faced was an inability to assess whether a speaker was deviating from their baseline pitch because I was not familiar with each speaker's baseline. To calibrate my annotation, I listened to the full meeting, which was slightly more than an hour's worth of audio. In this way, I familiarized myself with the speakers and their general speech patterns, including their inflections and pitch. This greatly improved my ability to make clear assessments of the audio segments. I annotated the audio similarly to how I annotated the text, using a classification of low/neutral/high emotional arousal.

I then compared the accuracy of my annotations to the scores given by the algorithms. For simplicity, I divided the sentiment scores of the machine into high anxiety or emotional arousal when the score exceeded one standard deviation above. Similarly, low anxiety or emotional arousal was considered anything below one standard deviation. Anything in between one standard deviation above and one below is considered neutral. I compared these buckets to the annotations I made.

7.1.2 Contributions

Comparing the accuracy of human annotations to machine annotations was not the only objective of this inquiry. I also wanted to learn which emotional states I could observe using a combination of text and audio measures. In other words, I wanted to formulate a more formal scheme of emotional states as a result of combining audio and text analysis. I was not only looking to see if a human listener made the same assessments as the algorithms, but to see *when* and *why* the assessments differed. I also wanted to better understand the relation between emotion in audio and text. The score calculated by the custom anxiety lexicon in the speech component is a fundamentally different measure of sentiment to the measure of pitch extracted. My primary research question was:

What emotional states can be captured when combining measures of speech and audio?

7.1.3 Results

A general overview of the comparisons of qualitative vs quantitative measurements revealed that I fully (both the levels of anxiety and emotional arousal were correctly identified) or partially (one of two dimensions was correctly identified) assessed the speeches 94% of the time. Completely accurate assessments were made 73% of the time. Audio annotations were incorrect more often than text annotations. Of course, since the sample size is so small (33 speeches) these measures of precision are not very meaningful. However, these findings can help guide future research that uses datasets of a larger scale.

The analysis revealed several different sets of patterns of dynamics between the audio and text measurements. There is a relationship between the two metrics, but it is not a linear relationship. Namely, an increase in pitch does not always correspond to an increase in anxiety in text. Rather, the two measures display an interactive relationship. This analysis shows how we can capture the dynamics of speech: how speakers react to other members of the meeting and what is being discussed.

7.2 Members of parliament

7.2.1 Arousal fluctuates while anxiety remains high

In the sample, two MPs are asking questions, Kate Green and Tim Loughton. Yvette Cooper was also present but was acting as Chair and did not ask as many questions as her colleagues. Links for all quotes that are included in this section are attached as additional media with the dissertation.

Kate Green displays the highest level of anxiety and emotional arousal across speakers, including witnesses. She either spoke with levels of emotional arousal above her baseline or below her baseline, but not at baseline. The range of her emotional arousal was quite wide, ranging from 1.34 s.d. below her baseline to 2.28 s.d. above her baseline. Her anxiety was high most of the time, ranging from 0.24 s.d. below the baseline to 2.75 s.d. above her baseline.

Kate Green was the only speaker for which I found anxiety and arousal increased simultaneously. An example of such a speech is as follows:

> "Our understanding from evidence that we have had, for example, from Katie Hayward, Dr. Katie Hayward, is that any check that is done away from the physical border where the goods actually cross does create a risk of those goods being swapped, exchanged, different goods actually being shipped off the border over the border. So how do you enforce between the assessment that is been made at the factory floor and what actually travels over the border?"
> (Quote 1: Kate Green, 00:54:30.3 - 00:54:51.9)

To me as a reader, the content of this speech is anxious. The speaker expresses uncertainty about the possibility of goods being swapped or exchanged. The danger of this is the inability of the UK to monitor the imports of goods at the border. Her tone is more animated than usual, signaling that she is more emotionally aroused than baseline (audio file quote_1_audio.wav).

The arousal of this speech was 1.26 s.d. above average and anxiety was 0.93 s.d. above average. The "hot" words (the words that had the highest anxiety scores in the speech) that contributed to the high anxiety score (+1.5 s.d.) of this speech include "evidence", "cross", "risk" and "enforce". For a visualization of the levels of anxiety and arousal in this speech compared to the sample as a whole, see SI figures 31 and 32.

The parallel increase in anxiety and arousal is an intuitive pattern. An increase in the level of anxiety of what a person is talking about may be accompanied by an increase in emotional intensity. However, anxiety and arousal do not appear to always have a linear relationship. In several speeches, high anxiety is paired with low arousal, as in the following example by the same speaker, MP Kate Green:

> "My understanding from the evidence we received from the Deputy Chief Constable of the Police Service for Northern Ireland recently was that there was a highly foreseeable likelihood of an increase in cross-border crime and an increase in the volume of smuggling. So why do you think an event of No Deal that we shouldn't be too worried about the risk of smuggling?"
> (Quote 2: Kate Green, 00:52:10.7 - 00:52:21.9)

Kate Green's speech is anxious according to the Affective Intelligence definition of anxiety. There is both uncertainty about how Brexit would affect trade and a strong fear about increases in smuggling and cross-border crime.

This speech had an anxiety score of 1.32 s.d. (see figure 18) above baseline and an emotional arousal score of 0.61 s.d. (see figure 17) below baseline. These figures show the speech's anxiety and arousal level, respectively, compared to the levels of anxiety and arousal in the rest of the sample. The speaker's lowering of her pitch seems to be at odds with a quite obviously anxious speech. It also stands in contrast to the speaker's general tendency to speak with high levels of anxiety and high arousal.

An interpretation of this speech is that the speaker may be posturing, i.e. she may be lowering her pitch voluntarily to appear composed about the issue or to elicit a certain

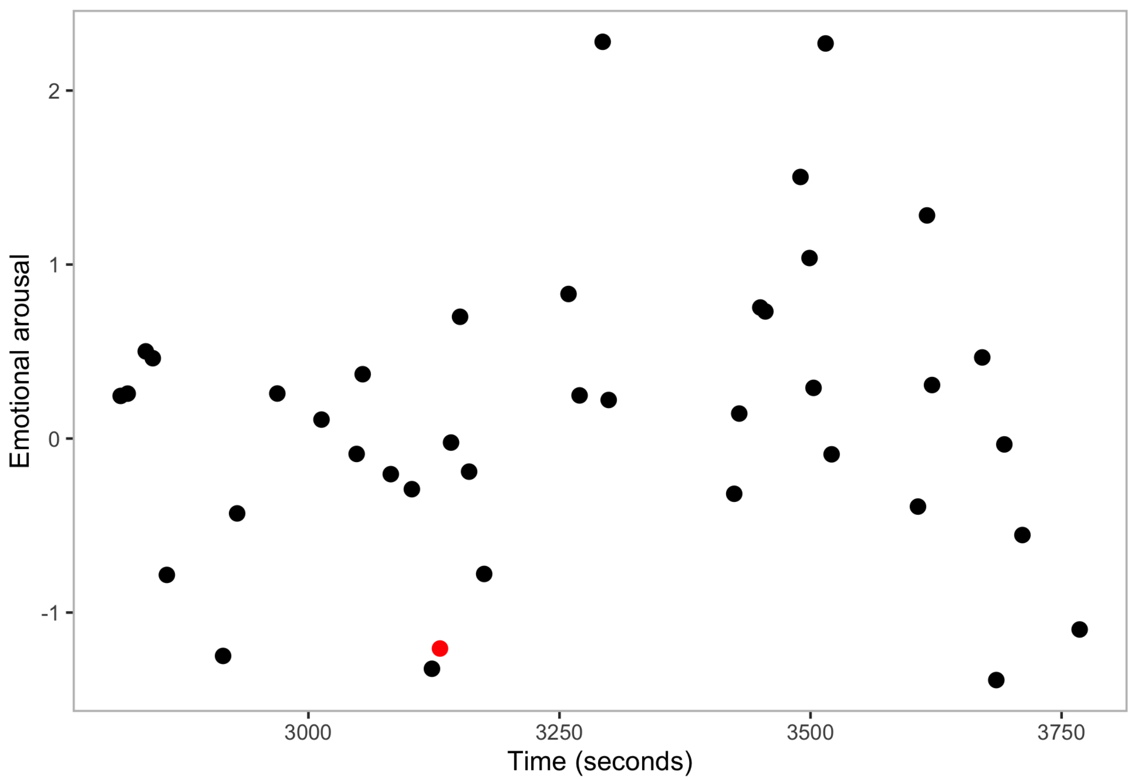


Figure 17: Emotional arousal for MP Katie Green's speech in quote 2. Speech data point is in red. The data points in black display arousal for the other speeches in the sample. The x-axis is time (seconds) and the y-axis is arousal in standard deviations.

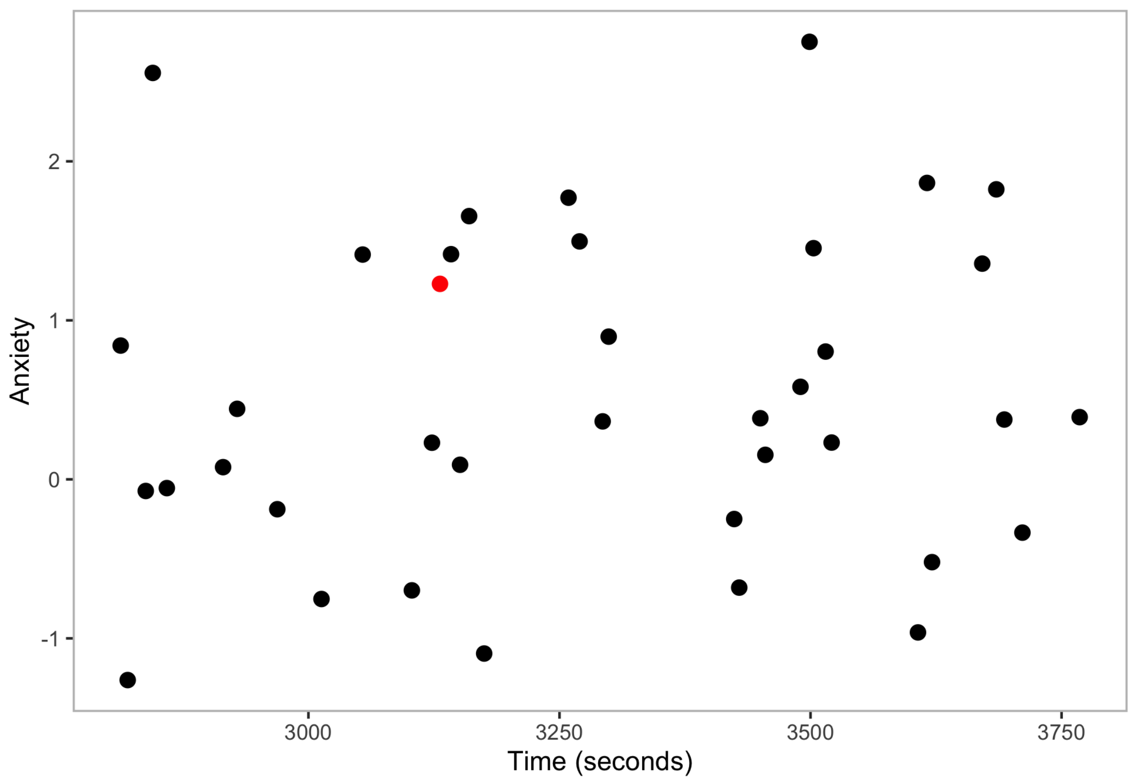


Figure 18: Anxiety score for MP Katie Green's speech in quote 2. Speech data point is in red. The data points in black display anxiety for the other speeches in the sample. The x-axis is time (seconds) and the y-axis is anxiety in standard deviations.

reaction. To explain this reasoning, we might consider real-world cases in which a person may lower their pitch despite speaking about an anxious issue. Consider a trivial example, such as if a teenager wants to tell their parents that they aren't going to college, against their parents' wishes. They may want to appear calm and confident in order to avoid an emotional outburst from their parents (whether they achieve this is of course dependent on many family-level factors). In a similar way, Kate Green may want to demonstrate composure, despite the content of her speech and perhaps despite actually herself feeling anxious about the issue.

Another interpretation is that the content of the speaker's speech should be treated separately from the speaker's tone. The content of the speech may very well be anxious, while the speaker themselves is not *feeling* any anxiety.

Future research should consider ways to measure whether deviations from baseline in pitch may be voluntary (as substantiated by existing literature [see or strategic (Dietrich, Hayes, and O'Brien, 2019)]). Additonally, it should consider the different strategic motivations between increases and decreases in pitch.

7.2.2 Neutral arousal and high anxiety

The second MP in the meeting does not display the same patterns of emotion in speech. For Tim Loughton, levels of emotional arousal are not associated with anxiety. Loughton's pitch generally remains neutral (around baseline), never increasing by more than half a standard deviation. His anxiety, on the other hand, increases by up to 2.5 standard deviations.

The mismatch between high levels of anxiety and low levels of arousal can be shown in the following example:

> "And the concept that more smuggling and criminal activity might happen because there might be more monitoring, in inverted commas, of a border than no monitoring now seems to me absurd"

(Quote 3: Tim Loughton, 00:47:24.8 - 00:47:38.7)

The content of the speech reads as anxious because of the possibility of increased criminal activity. Quite like the speeches of Kate Green shown in the examples in the previous section, members of the meeting are focused on this potential, negative consequence of Brexit, criminal activity at the Irish border. The issue is discussed with a level of uncertainty, which also contributed to my assessment. This is confirmed by the anxiety score determined by the scaling done with the anxiety lexicon, which is high at +2.5 s.d. Words like "criminal" and "smuggling" contribute to this score as well as words capturing uncertainty like "concept", "seem" and "happen". Anxiety is a future-oriented emotion, which helps explain the score of these words.

The affect in his pitch in this speech conveys a degree of confusion about how increased monitoring of a border would lead to an increase in criminal activity. However, the speaker's voice does not carry the same emotional charge about the issue that Kate Green's does when discussing the same issue. The difference between the two speakers may be attributed to the speaker's level of commitment to the issue, as discussed in the work of Dietrich, Hayes, and O'Brien (2019). It may be the case that Kate Green is more committed to the issue than Tim Loughton.

There do appear to be differences across speakers's emotional states that, when holding the issue constant, may be based on attitudes towards the given issue.

7.3 Witnesses

Witnesses generally exhibit lower levels of arousal and anxiety on average compared to the MPs. When witnesses' deviate from baseline, it appears to be a response to what the speaker before them said, elucidating the dynamics of speech.

7.3.1 Calm and confident

Shanker Singham exhibits the lowest levels of arousal and anxiety across all speakers. Across the span of this segment, Singham's anxiety never goes above baseline.

Singham's absence of emotional intensity and anxiety signals calmness and confidence. He responds well to anxiety-ridden issues, such as the assessment of risk of criminal activity at the Irish border. The following speech is an example of a response to a question about the risk of no deal:

> "I mean a lot of it depends on what we do and what the EU does in the event of no deal and as I say our proposals and our work is not for an no deal scenario, it is for a deal scenario. But I would think that in the event of no deal, you would have the UK government said we will not put up physical infrastructure and we will not... The no deal planning essentially allows on a temporary basis product to come from the EU into the UK without many checks at all. It depends what the EU does. If the EU puts up a hard border there then, yes, there will be problems. If the Irish government said we do not want to put up infrastructure on the Irish border then there will have to be a border between Ireland and the EU 26. There is no other way of doing it. The EU cannot have a completely open channel in the event of no deal to the single market because we have to have some checks. They will have to have their checks to protect the integrity of the single market and the customs union somewhere in the island of Ireland either on the border or between Ireland and the EU 26. And there is no other way of doing it. That is why we are very anxious to ensure there is a deal and that is why our proposals are designed to make sure we do get a deal and a transition period and all of these things can be rolled out."
> (Quote 4: Shanker Singham, 00:52:54.8 - 00:54:18.3)

To a reader, the content of this speech is confident. The speaker delineates all possible

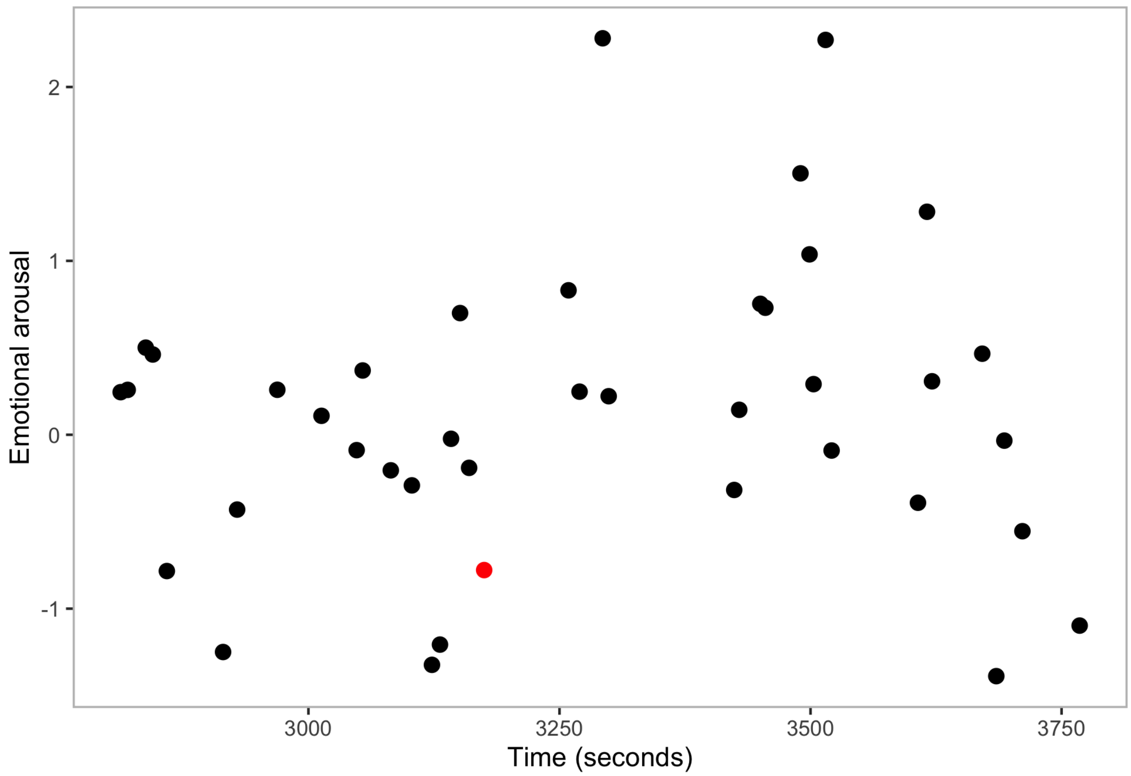


Figure 19: Emotional arousal for witness Shanker Singham's speech in quote 4. Speech data point is in red. The data points in black display arousal for the other speeches in the sample. The x-axis is time (seconds) and the y-axis is arousal in standard deviations.

permutations of the borders between Ireland, the UK, and the EU. He identifies the key issue, which is that the EU must protect the single market. Although there is mention of some uncertainty about the ability of the UK and the EU to reach a resolution at the beginning and end of the speech, Singham is not uncertain about what the outcomes of the deliberations could be.

In this speech, Singham's pitch is 0.78 s.d. (see figure 19) below baseline, and anxiety is at 1.1 s.d. (figure 20) below baseline. This confirms our expectations about the content of the speech. Similarly, for his tone, Singham displays composure. Unlike quote 2, Singham's low pitch does not seem to reflect underlying anxiety. Similar to Kate Green's quote, however, low pitch does seem to signal a desire to appear composed.

The deviation below the baseline for pitch may be attributed to the emotional tone of

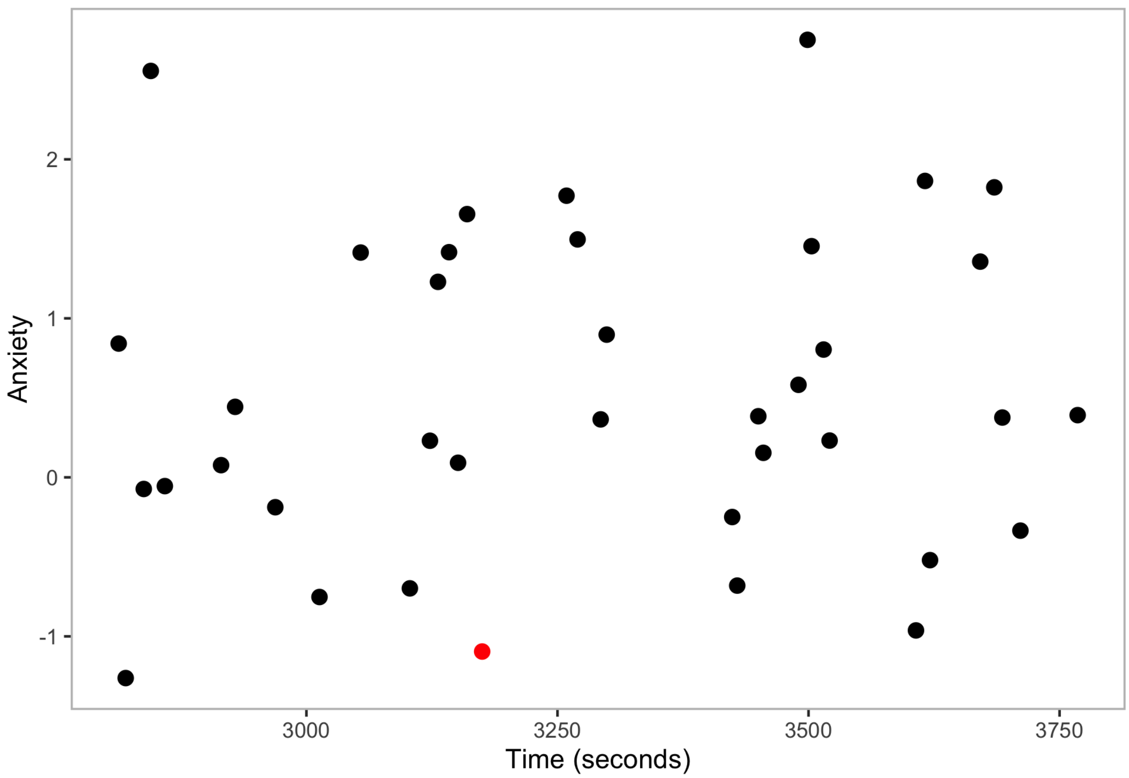


Figure 20: Anxiety score for witness Shanker Singham's speech in quote 4. Speech data point is in red. The data points in black display anxiety for the other speeches in the sample. The x-axis is time (seconds) and the y-axis is anxiety in standard deviations.

the previous question by Kate Green.[11] Green's question, despite not displaying a high level of emotional arousal (0.2 s.d below baseline), is highly anxious in content (1.66 s.d. above baseline).

The importance of dynamics appears in another speech by Singham, who again deviates from his baseline arousal. This example is a response to a question by MP Tim Loughton about the paradox that the existence of a new border would lead to increased levels of criminal activity.

> "Well I think we will have more data so actually the ability to enforce and the ability to monitor will be greater under the mechanisms that have been proposed than they are now. Right now there is no way of really monitoring and there is no way of really knowing who the small traders are who are moving across the border now. So I think the only concern the EU would have, to answer your question, is the quality of the checks that the UK- Do they trust the checks that the UK will be imposing at the ports and harbors of the Irish sea? I do not see any other reason to to object to these proposals other than if they truly are saying that they cannot be a customs border in the island of Ireland. Then it is very difficult to see how you solve this problem but we do not think a customs border means what people often think it means which is checks and checkpoints and all of those things."
>
> (Quote 5: Shanker Singham, 00:47:39.4 00:48:35.0)

Singham answers Tim Loughton's question from quote 3, which asks about whether the current low levels of criminal activity currently reflect a lack of monitoring. Singham answers the questions in a calm way. There are no signs of anxiety in the text. Similar to the levels of emotion in quote 3, Singham is quite composed and collected.

[11] The question reads: "So do you agree with the Deputy Chief Constable at an event of no deal the reason to restrict all cross border crime and smuggling potentially a significant risk? I think we were told by the Prime Minister and Secretary of the Department of Justice that was impossible really to quantify the risk."

In this speech, pitch is 0.8 s.d. below baseline and anxiety is 0.05 s.d. below baseline. The level of anxiety is neutral. Although the content in this example may resemble that of quote 3, the difference in anxiety score is attributed to overall distribution of anxiety of words in the speech. In this example, there are several words that are associated with confidence, such as "monitor" and "ability". There are happen to be several words that fall in the middle of the anxiety scale, such as "actually", "port" and "customs". These words drag the weight of the distribution of anxiety in the speech toward the center of the scale, making it a more neutral speech.

The arousal in this speech is below baseline, which is evidence in support of the hypothesis that levels of arousal can be partially explained by the anxiety in the content of the preceding speech. Singham lowers his pitch below baseline because he is aware of the anxious nature of the issue.

The other witness, Tony Smith exhibits similar levels of neutral or low emotional arousal and anxiety. One exception is a speech where Yvette Cooper, the chair, is pressuring him to answer a question about when criminal activity data will be in place. Cooper presses him to give a more concrete answer to which Smith responds:

> "I am afraid I really would not want to hazard a guess and mislead the commission because I really do not know. The answer to that I am sorry but I could not say. I mean I think you know there are different categories of import-export... certainly those are already import export outside of the EU as well. We can pretty well bank those ones but there is a huge, as you know, publicity campaign going out now to say these are the things you need to the first and get the ELRI number get them on our register."
>
> (Quote 6: Tony Smith 00:50:53.9 - 00:51:21.8)

Here the witness's arousal remains relatively low at 0.4 s.d. above baseline but Smith's anxiety leaps to 1.4 s.d. above baseline. The anxiety is in response to the timeline of implementing the proposed changes. Smith appears to be projecting fear about his own ability to carry out the project he is responsible for.

7.4 The dynamics of witnesses and MPs

The examples shown in the results section reveal that text and audio contain distinct information about emotions in political speech. We are able to measure emotion in text on a linear scale. The anxiety lexicon used in this case was able to measure how anxious the content was in the speech. As in quotes 1 and 3, the anxiety analysis captured that both of these texts, despite differences in the actors' emotional states, were anxious in nature. Both speakers expressed uncertainty and fear about the future of the Irish border.

Arousal, on the other hand, does not exhibit linearity. Increases in arousal revealed similar emotional states to decreases in arousal, as in examples 1 and 2. In both quotes the content of the speeches was anxious but arousal moved in opposite directions. This difference in arousal appears to show how different speakers displayed their levels of emotional attachment to the issue. Without considering the level of anxiety in the text, the direction of emotional arousal would not be interpretable. Vice versa, by only considering the levels of anxiety in each speech, the speeches would not appear different.

The lowering of pitch below baseline suggested strategic motives. MP Kate Green in quote 2, as well as witness Shanker Singham in quote 4, both sound very composed. In quote 2, Green's lowering of pitch appears to be a case of posing as composed despite the nature of what she is talking about. In quote 4, the content of Singham's speech is confident, which can be interpreted as a strategic decision to address the uncertainty of the Irish border. When interpreting the witnesses' behavior, it is important to consider the dynamics of the meeting segment as a whole.

The analysis reveals how sentiment in text and audio is a dynamic process, reflecting the speaker's reactions towards preceding speeches. In quotes 5 and 6, the witnesses reflect their emotional reaction to the preceding question. In quote 5 Singham responds to Loughton's remark in quote 3. He senses the anxiety in the content of the remark and despite Loughton's neutral pitch, he lowers his pitch and gives a neutral answer that makes the discussion calmer. In quote 6, Smith is unable to respond calmly to the Chair of the session pressing him to give exact dates. The audio and text together reflect his anxious response to the question.

7.5 Emotional arousal in Parliament: exploratory findings from a quantitative analysis

In this section, I show a set of descriptive, quantitative findings that are guided by the insights in the previous qualitative section. Taking each speech in the dataset of transcribed audio files (see chapter 6 for more details), I estimate the level of anxiety and arousal in each speech.

This is a comparative analysis of the data using information provided in the dataset, including party affiliation, speaker role information and date information.

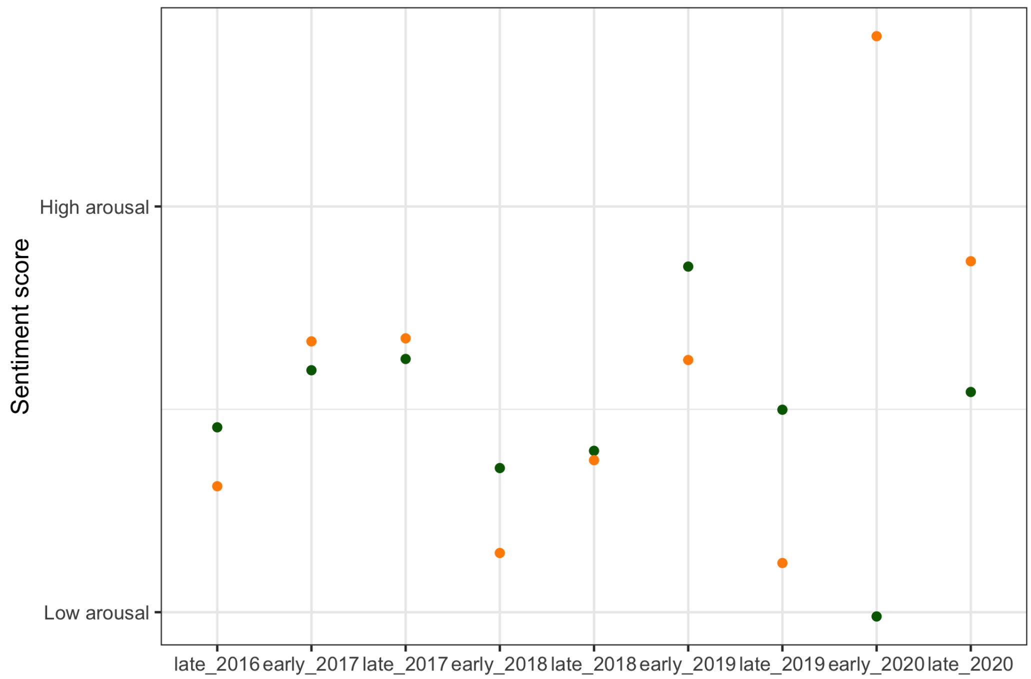


Figure 21: Average anxiety (green) and arousal (orange) of speeches over time. The x-axis shows time in buckets of six months. The y-axis is the level of sentiment in each time point.

I first study how anxiety and arousal in the speeches moves over time. Dates are grouped into buckets of 6 months for each year (early 2016, late 2016 etc.) to slightly smooth out the trends. The results are shown in figure 21. Both measures are standardized at the speaker level to be able to compare the measures. In the unstandardized format, arousal is measured in Hertz and anxiety is measured according to the scaling of the word embedding dictionary analysis (chapter 4). By standardizing the measures, we capture the deviation of each speaker from their baseline.

Based on this broad picture of the data, the two metrics of speech, anxiety and arousal, move in the same direction. Both anxiety and arousal increase from late 2016 to late 2017, followed by a decrease in early 2018. Both measures increase again in late 2018 and late 2019. In early 2020 there is a sharp split in the trend, where anxiety plummets whereas arousal sharply increases. In late 2020 the measures converge.

This analysis indicates that there is a relationship between the two measures. As was shown in the qualitative analysis section of this chapter, the relationship does not seem to be linear. In order, therefore, to better understand what factors determine how these two measures relate, I capture the average anxiety of speeches based on the role of the speaker. In section 7, MPs and witnesses displayed important differences in how arousal moved in relation to anxiety. The measures moved at times in opposite directions and at other times, they moved in tandem. In figures 22 and 23, I show the average anxiety and average arousal, respectively, of each type of speaker.

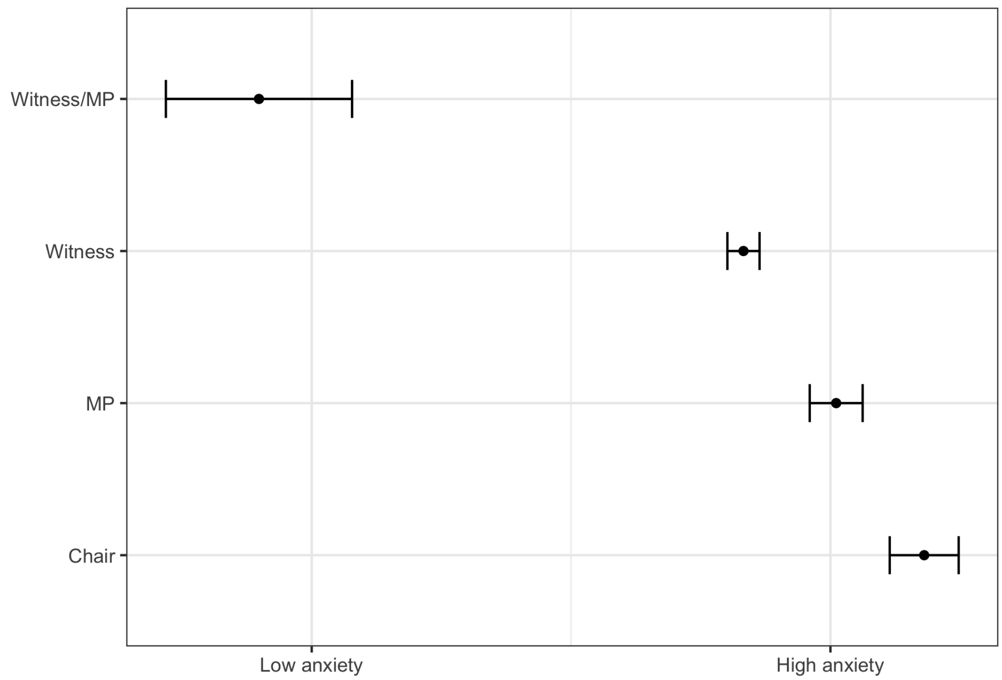


Figure 22: Average anxiety and 95% confidence interval by the role of the speaker

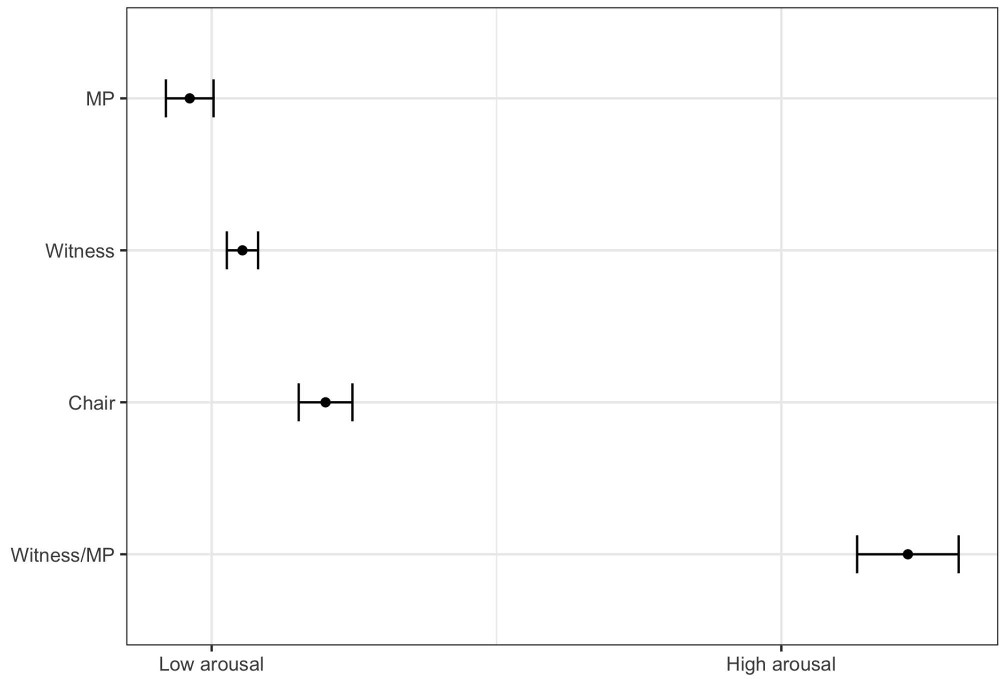


Figure 23: Average arousal and 95% confidence interval by the role of the speaker

A witness/MP is a Member of Parliament who is participating in a panel as a witness. Since this is a distinct role from participating as a MP and given that their involvement in parliament differentiates them from other witnesses, I classify this as a different category of speaker. Indeed, these speakers exhibit a different pattern of anxiety and arousal, with their anxiety being the lowest of the four categories and their arousal being the highest of the four. MPs have the second highest levels of anxiety but the lowest levels of arousal. Witnesses have the second lowest average level of anxiety and the second lowest average arousal. The Chair of committees shows the highest level of anxiety and the second highest level of arousal. The differences between all averages in both analyses are statistically significant.

For the MPs in the committee meetings, the high level of anxiety and low arousal evokes the results from section 7 where MP Tim Loughton spoke about anxious matters but with lower than average arousal. Similarly, the lower levels of anxiety and arousal for witnesses aligns with the qualitative analysis of witness Shanker Singham's speeches, which consistently had low levels of arousal and neutral or low levels of anxiety.

In the qualitative chapter there was evidence of differences for levels of arousal along party lines for the MPs in the sample. To examine this finding in the full dataset, I group by party membership and speaker role to capture the average anxiety and arousal. The results are in tables 24 and 25.

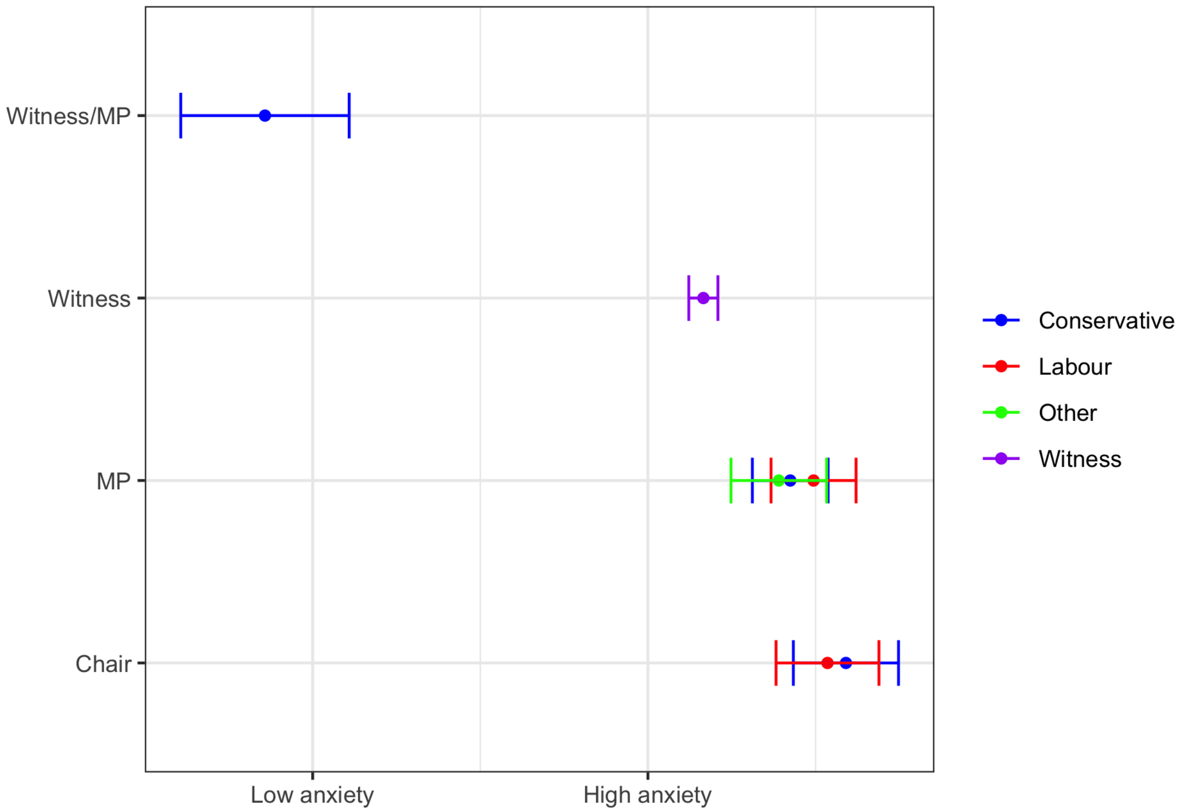


Figure 24: Average anxiety by role and party membership of the speaker

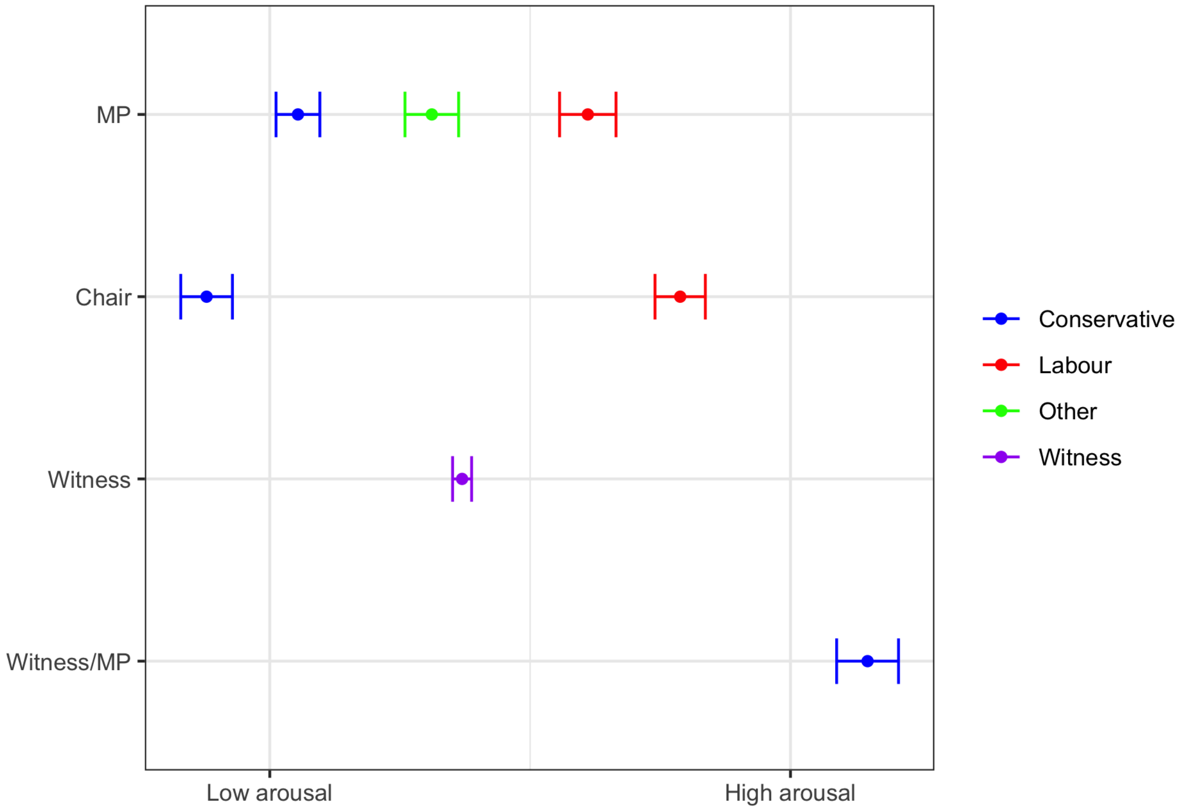


Figure 25: Average arousal by role and party membership of the speaker

For roles where there are speakers from more than one party, the differences in average anxiety are not statistically significant. This aligns with the findings from chapter 7.5. In figure 39, the temporal trends of anxiety in speeches from committee meetings showed a lot of overlap. Similarly, in this subset of the data, for both MPs and Chairs, participants have relatively high levels of anxiety.

In contrast, in figure 25, differences in the mean arousal of speakers of different parties are distinct. For MPs, Conservative party speakers have the lowest average arousal, followed by members of parties other than Labour and Conservative. Labour MPs speak with the highest levels of arousal. The average arousal of the Chair shows an even sharper split in average arousal. Conservative speakers speak with far lower levels of arousal than Labour members.

This confirms the findings from chapter 7. The qualitative analysis showed Labour

MP Kate Green speaking with higher levels of arousal, compared to MP Tim Loughton. The anxiety in the content of the two speakers was high, but the arousal was high for Green and neutral for Loughton.

In chapter 7.5, although there was overlap in the levels of anxiety across parties in the committee meetings, the main chamber showed clear patterns along party lines. The significant difference in means of levels of arousal for MPs and Chairs in the committees may be related to the difference in means of levels of MP's anxiety in the main chamber. By including arousal as a measure of emotion in speech, we are able to see better understand how party members speak differently in the context of committees based on their political affiliation.

7.6 Discussion

In this chapter, I presented a set of descriptive findings showing levels of arousal in parliamentary committee meetings. I compared the metrics of emotional arousal to measures of anxiety in speech in order to provide a more complete understanding of different speakers emotional states. Anxiety and emotional arousal do seem related as measures of emotion, as shown in figure 21. However, as initially shown in chapter 7, the two measures do not have a linear relationship.

Speakers showed significantly different levels of anxiety and arousal across groups (witness/MPs, witnesses, MPs and chairs). However, party affiliation could only help explain differences in levels of arousal. MPs and chairs of different political parties showed similar levels of anxiety. In section 7.5, differences in levels of anxiety by party only existed in the main chamber and not in committee meetings. The findings from this chapter show that differences in emotional states across parties can be captured using audio.

8 Conclusion

In this dissertation, I explored methodological approaches to measuring anxiety in text and audio data. Upon establishing the validity of the measurement tool for anxiety in text, I examined how anxiety operates in the British House of Commons. Focusing on the Brexit negotiations, which took place between 2016 and 2020, I demonstrated that levels of anxiety were relatively higher in speeches about Brexit than they were for speeches on other topics. Using audio analysis, I showed that audio data contains different but complementary information about the speaker's emotional state. Using a combination of audio and text, I was able to disentangle the nuances of different speakers' emotional profiles.

I selected Brexit as a case study for anxiety in text and audio data. The 2016 Brexit referendum was a turning point for British history and a historical moment for the European Union. The uncertainty of this time period makes it a good candidate for the study of anxiety because anxiety arises in unknown, uncertain situations where there is no predetermined course of action. The Brexit campaign is also remembered for its rhetoric, with slogans like "take back control" capturing the heightened emotions around the many key issues of the referendum. The importance of rhetoric in this time period lends itself to the study of emotion in text.

To study this time period, I compiled a dataset of parliamentary committee meetings about Brexit. This was supplemented with the Hansard dataset, which includes all speeches from parliamentary debates. The committee meetings are an understudied source of data in text analysis and provide insight into the unique setting of committee meetings which are less polarized.

After creating a custom anxiety lexicon, I analyzed the full Brexit dataset in chapter 5. Anxiety varies greatly over time with certain time periods displaying sharp peaks or troughs in anxiety. Anxiety was highest in the year 2018. Discussions about Brexit were on average higher in anxiety than discussions about all other topics. An important analysis showed that the main chamber, contrary to expectations, has the lowest levels of anxiety in text. This finding reminds us that measuring emotions in elite political speech is ultimately a measurement of the content of speech and not necessarily of the emotions of the speaker themselves. Committee meetings had higher levels of anxiety on average. Further analyses that incorporated the party affiliation of the speaker showed that, during the Brexit negotiations, Conservative MPs spoke with lower levels of anxiety than Labour MPs. A larger analysis of anxiety from 2000 to 2020 revealed a pattern of the party in power speaking with lower levels of anxiety than the opposition. In sum, the analysis of text shows that anxiety is associated with the topic of discussion, the party affiliation of the MP, and whether the discussion occurs in the main chamber on in a committee meeting.

Incorporating the audio data into the project was not as straightforward as hoped. The transcripts did not perfectly match the audio, which ultimately led to my needing to generate transcripts directly from the audio files. Details of the process are explained in chapter 6. For every meeting, the resulting data matched speeches to the audio segments in the meeting's audio file.

After conducting a qualitative analysis of a 10-minute segment of a meeting, I expanded the analysis to the full dataset of segmented audio files and speeches. This part of the project used pitch as a proxy for emotional arousal in audio and compared this metric to the measurement of anxiety in text. The qualitative analysis of pitch revealed variation along the dimensions of speaker role (MP or witness) and of party affiliation. These differences appeared to suggest information about the speakers' intentions and emotional attachment toward the issues. The quantitative analysis confirmed that there were statistically significant differences in pitch along the dimensions of interest. Measurements of emotion in audio and measurements of anxiety in text have an interactive, non-linear relationship that, when combined, can offer a fuller understanding of the emotional profile of speakers.

8.1 Limitations and future research

The main limitation of the audio analysis segment of this study is scale. The data used in the text analysis is comprehensive because it includes every speech in the House of Commons that discusses Brexit. The audio analysis only ended up including around 50 of the 130+ meetings used for text analysis. This was due to the computational limitations of the project. I transcribed all meetings alone with very limited additional computational resources (google Colab). It would take either a long time or a huge upgrade in computational power to be able to transcribe these meetings. The computational cost is, unfortunately, one of the challenges of audio analysis in general. Many of the inaccuracies in the data could be resolved by using a larger model in the transcription step (see chapter 6).

This project has expanded on the growing body of audio analysis in political science scholarship and showed the importance of anxiety in elite political speech. Future research should consider the particular role of anxiety in political communication and study the emotion in different political contexts. Additionally, future studies of anxiety should consider what other factors may be associated with levels of anxiety in speech such as gender, age, or voting record.

9 Appendix

List of custom stopwords removed in pre-processing

'also', 'hon', 'minister', 'debate', 'secretaries', 'secretar', 'governs', 'secretary', 'bill', 'friend', 'committee', 'committees', 'will', 'government', 'governing', 'must', 'party', 'get', 'go', 'want', 'say', 'one', 'govern', 'people', 'peoples', 'come', 'can', 'us', 'know', 'mr', 'said', 'refer', 'however', 'make', '£', 'member', 'house', 'think', 'take', 'see', 'look', 'even', 'thing', 'let', 'go', 'gentleman', 'gentlemen'*

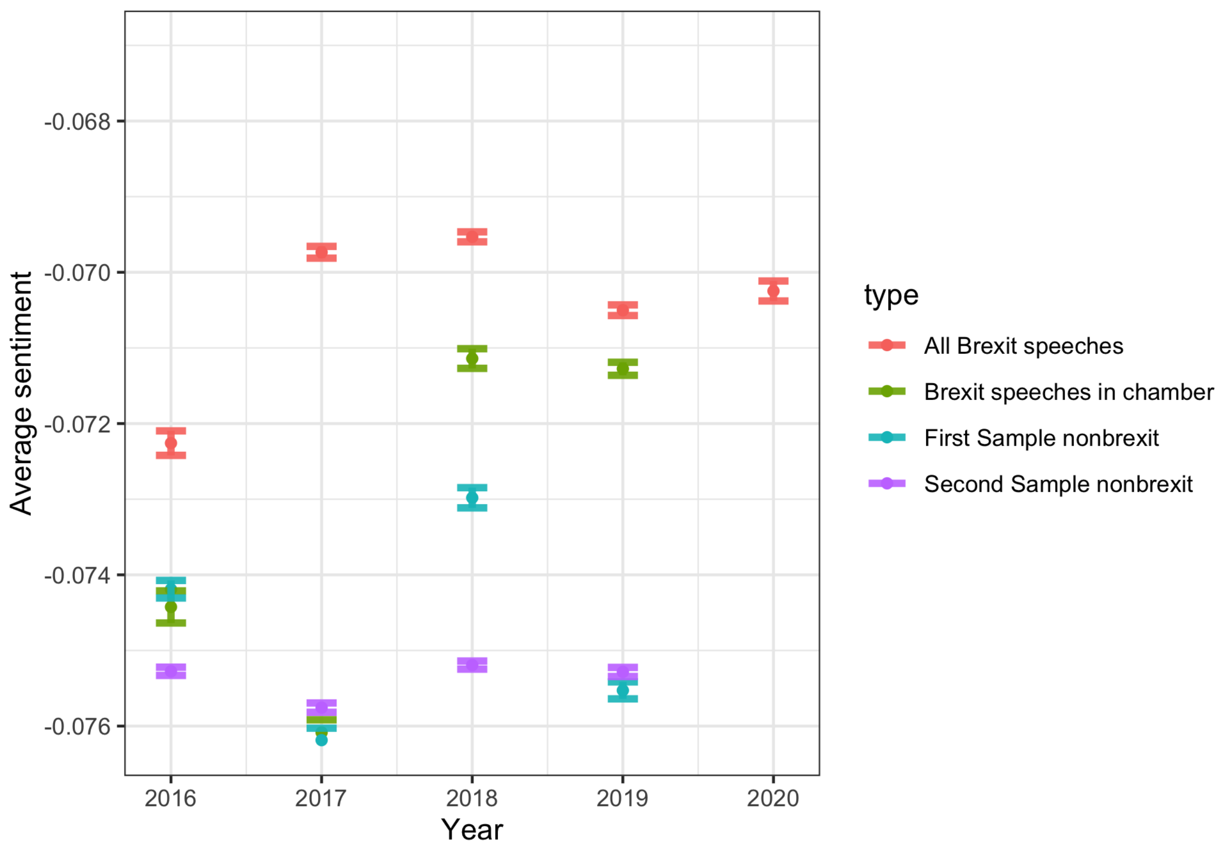


Figure 26: Anxiety by year. Second sample of non-Brexit speeches included.

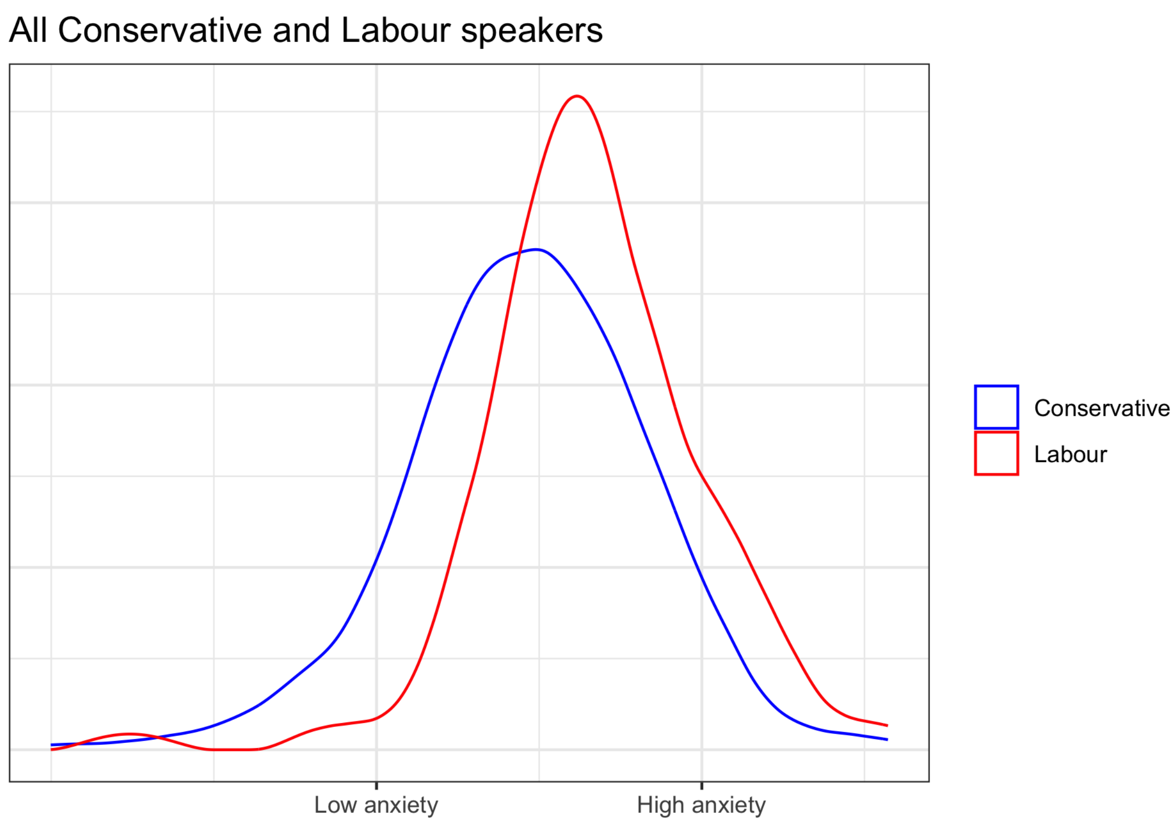


Figure 27: Estimated distribution of all Conservative and all Labour speakers

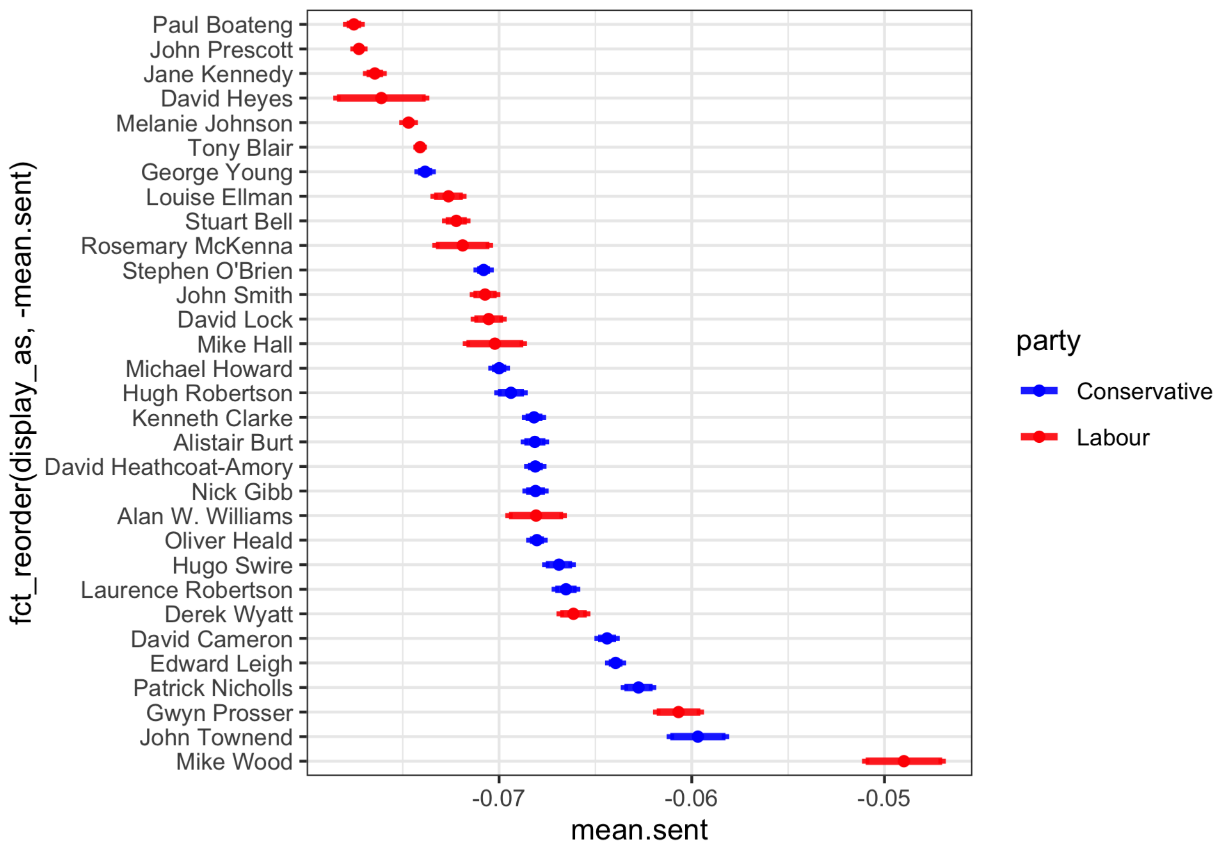


Figure 28: Second example of members of parliament mean anxiety score and confidence interval. Conservative speakers (blue) and Labour speakers (red)

Time Period	Number of Conservative Speakers	Number of Labour Speakers
2000-2004	195	455
2005-2009	218	436
2010-2014	342	403
2015	374	300

Table 7: Count of speakers in Conservative and Labour party for each time period studied outside the Brexit negotiations (2000-2015)

Figure 29: Average anxiety from 2000 to 2019 in the main chamber for the Conservative party (blue) and the Labour party (red). The vertical line in 2010 marks an election year where the government changed from Labour to Conservative. From 2016, there are two additional trend lines that are main chamber debates about Brexit.

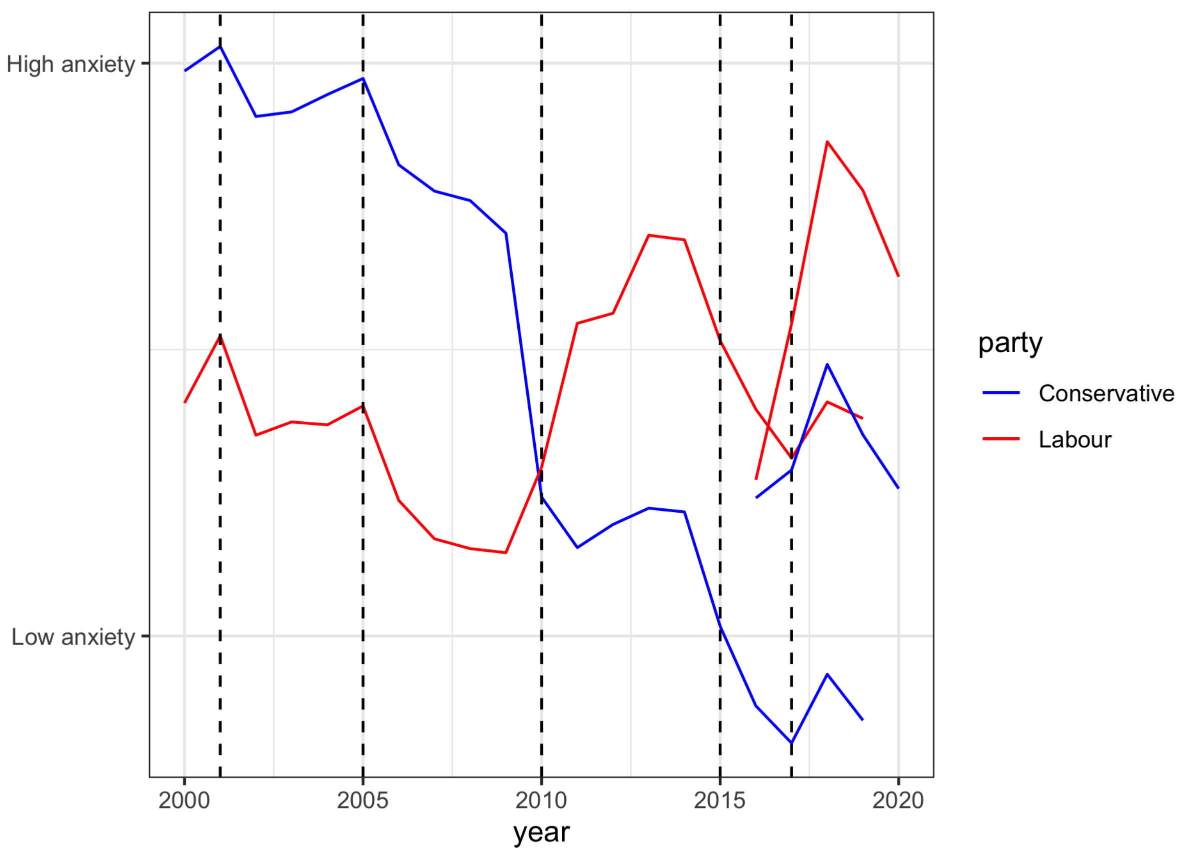


Figure 30: Average anxiety from 2000 to 2019 in the main chamber for the Conservative party (blue) and the Labour party (red). The vertical lines mark all national election years

Committee	Count of speeches
Communities, Local Government Committee	277
Culture, Media and Sport Committee	544
Environment, Food and Rural Affairs Committee	1895
European Scrutiny and Environment Food and Rural Affairs Committee	229
European Scrutiny Committee	1771
Exiting the European Union Committee	593
Health Committee	792
Home Affairs Committee	2895
Housing, Communities and Local Government Committee	330
International Trade Committee	755
Justice Committee	676
Liaison Committee	467
Northern Ireland Affairs Committee	4106
Public Accounts Committee	2585
Public Administration and Constitutional Affairs Committee	420

Table 8: Count of speeches by committee for audio analysis

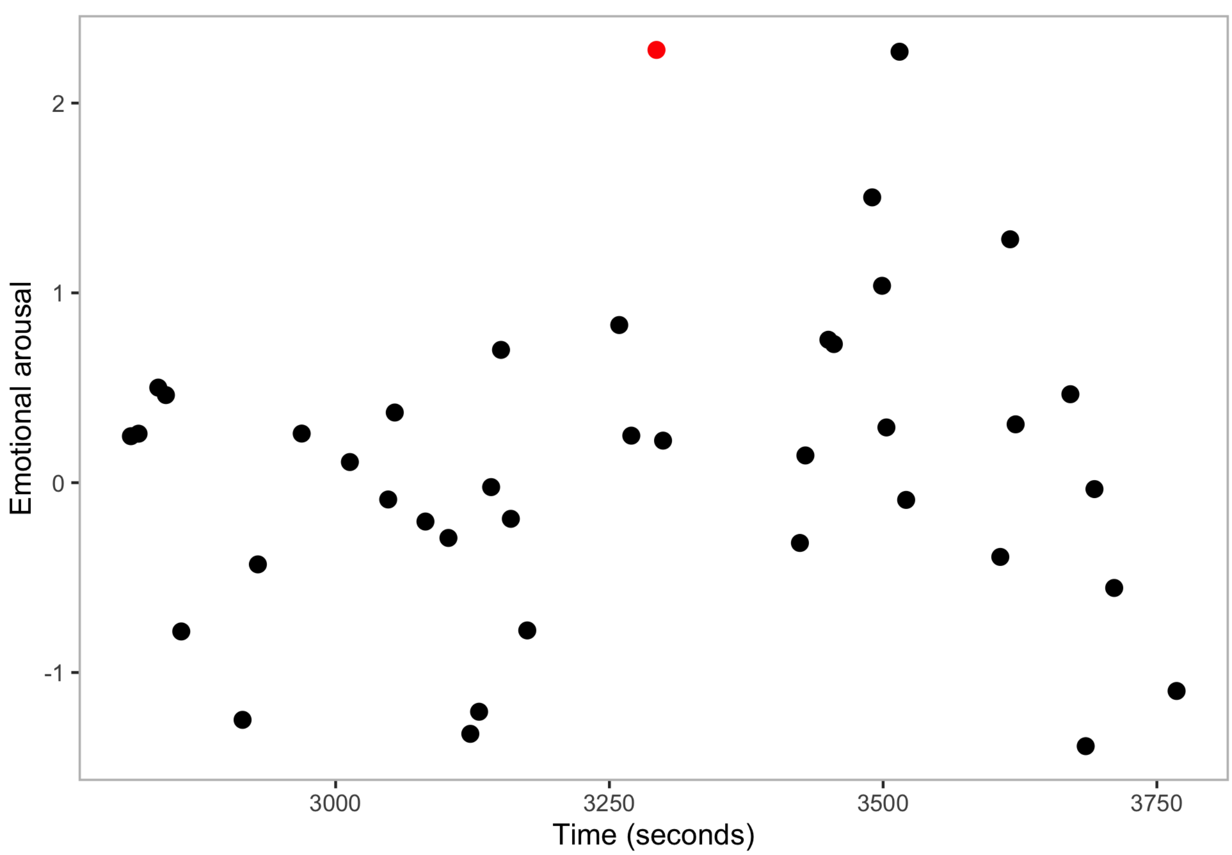


Figure 31: Emotional arousal of Kate Green's speech in quote 1. Speech data point is in red

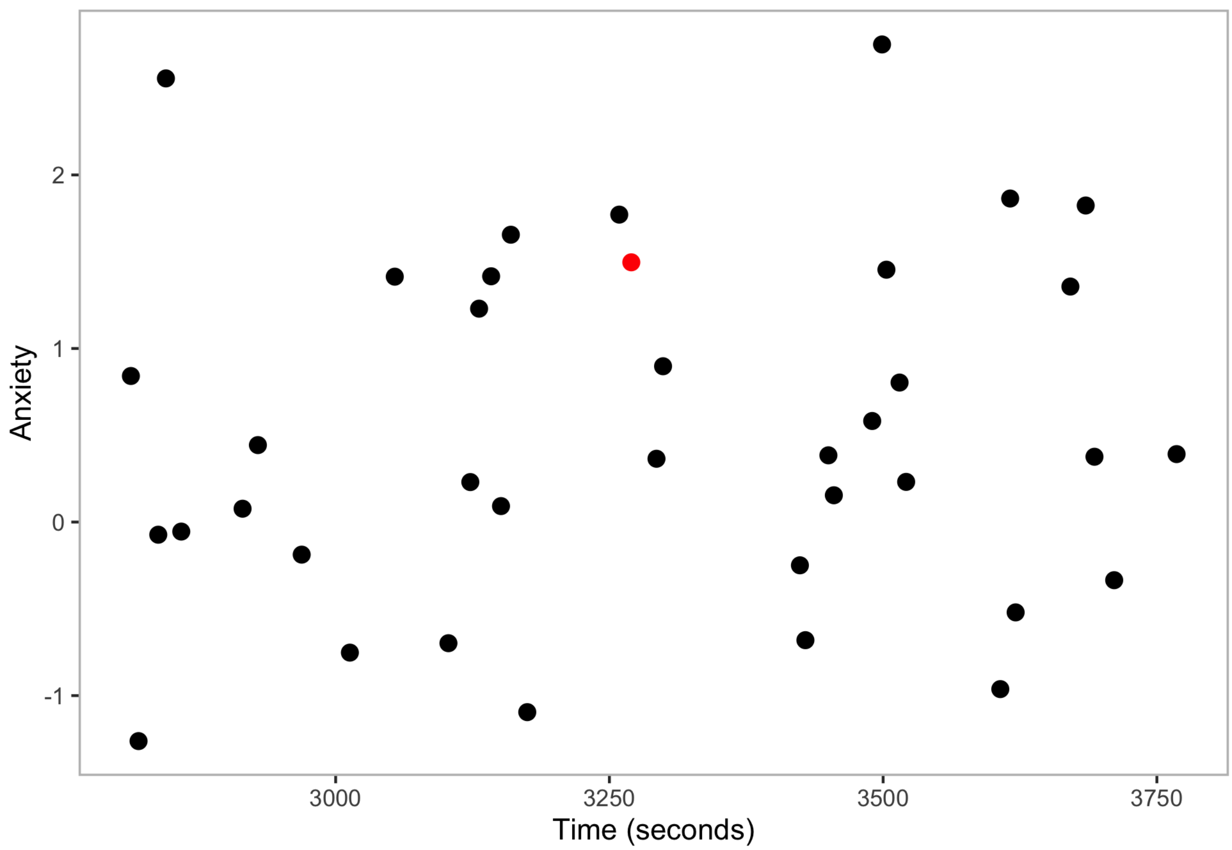


Figure 32: Anxiety score of Kate Green's speech in quote 1. Speech data point is in red

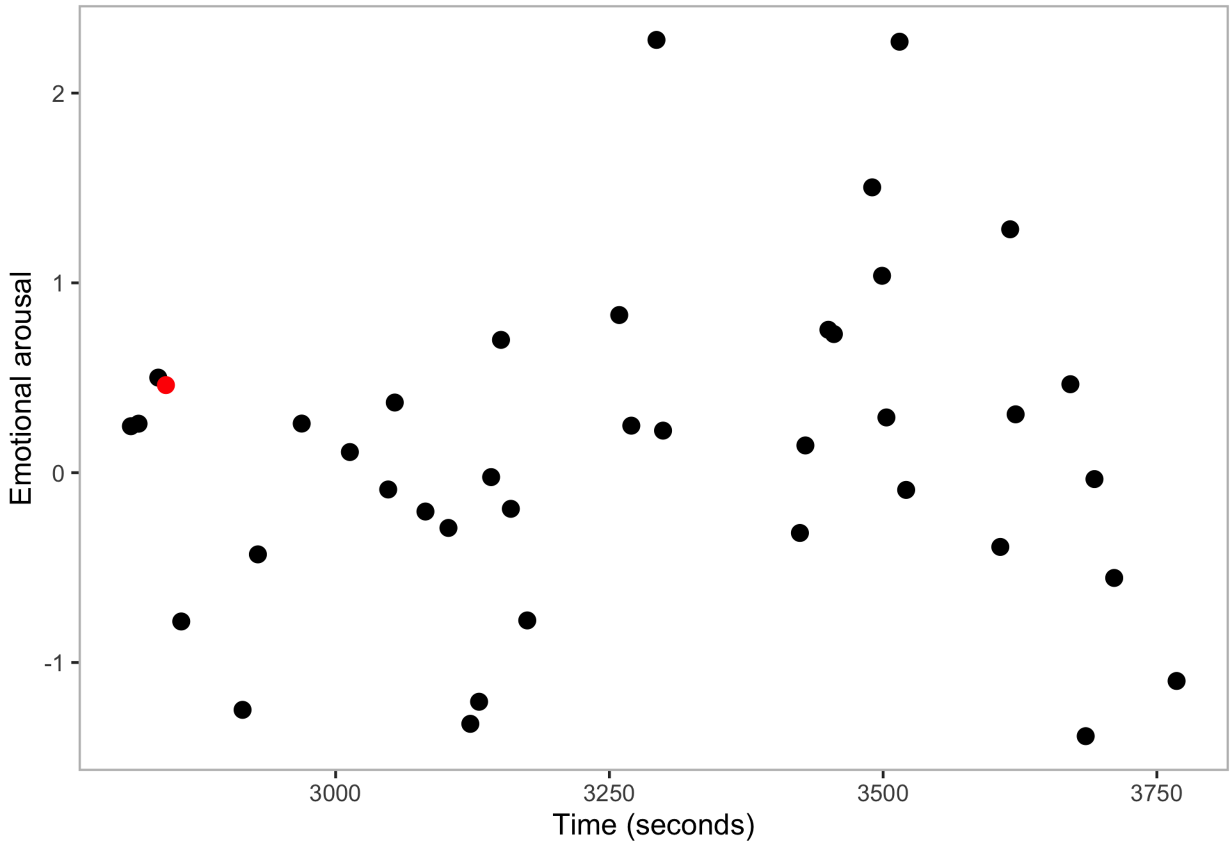


Figure 33: Emotional arousal of Tim Loughton's speech in quote 3. Speech data point is in red

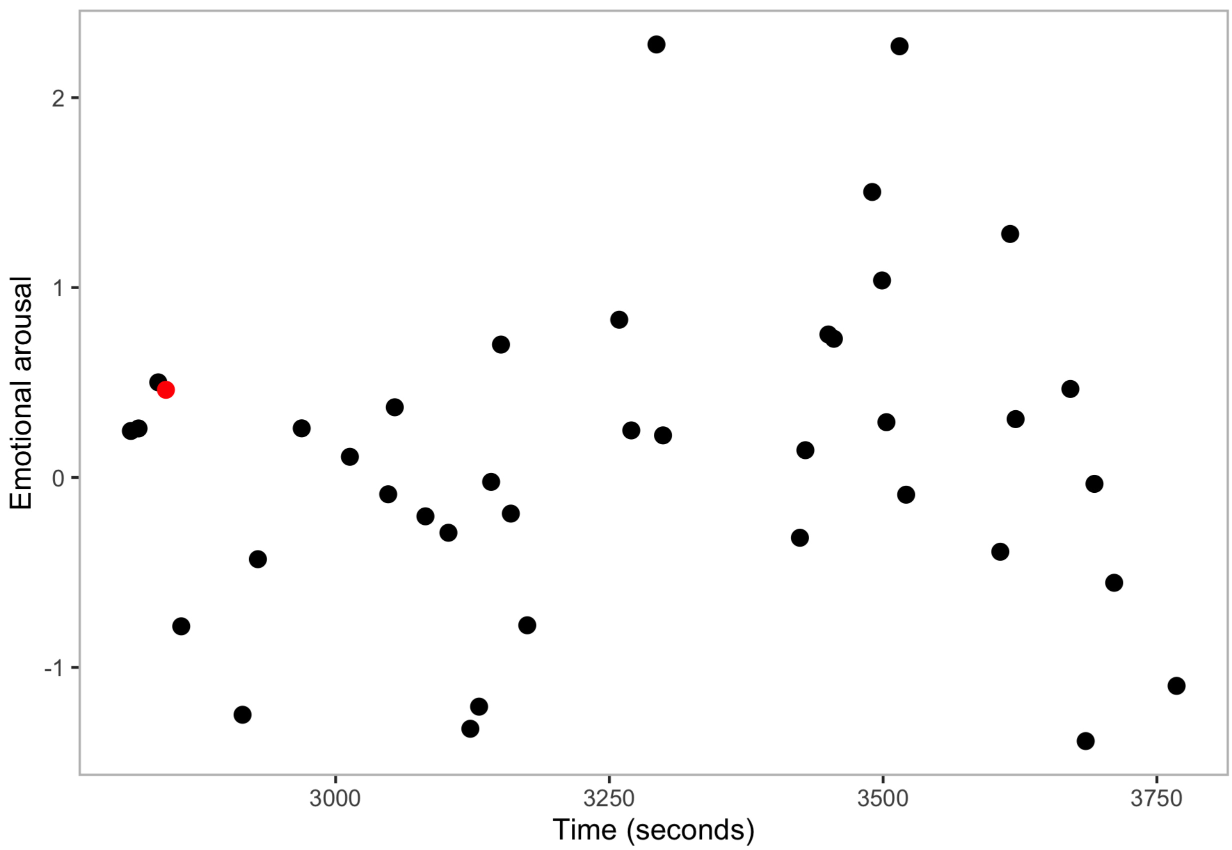


Figure 34: Anxiety score of Tim Loughton's speech in quote 3. Speech data point is in red

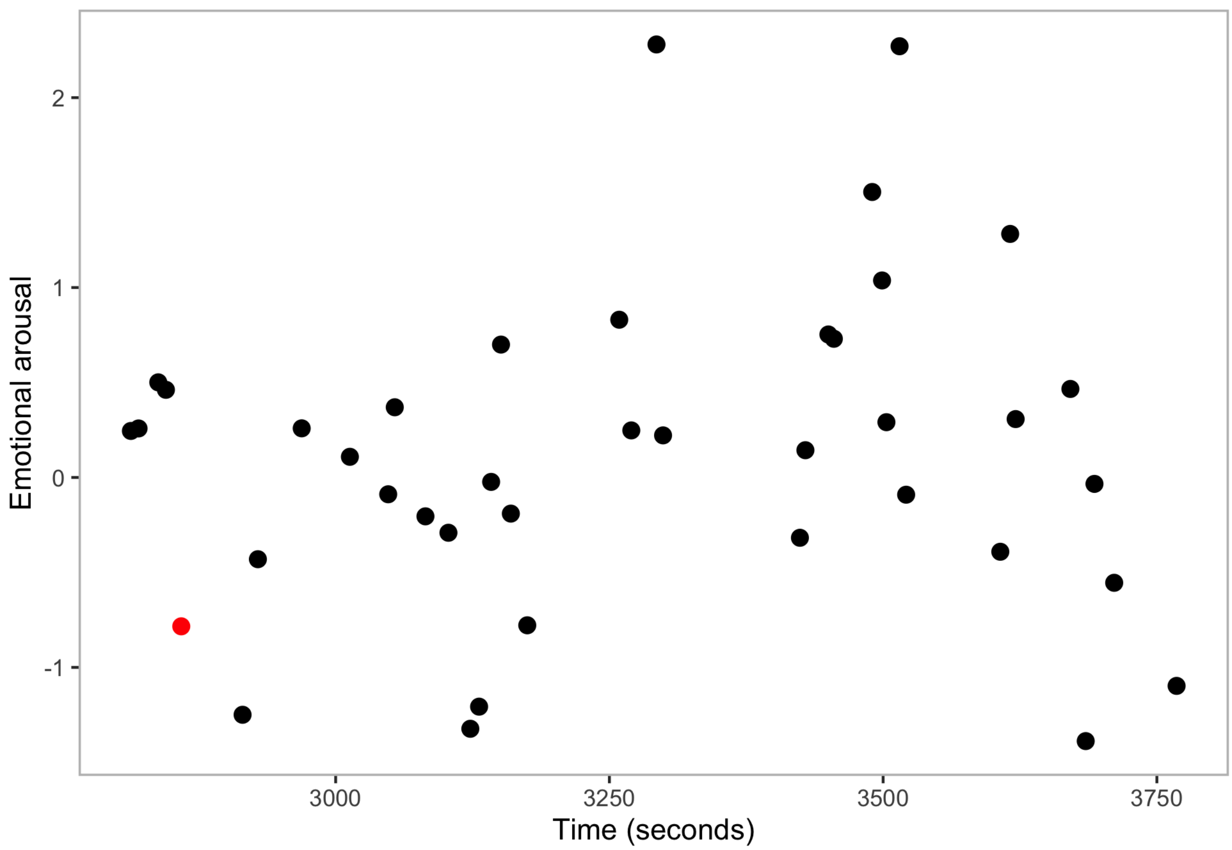


Figure 35: Emotional arousal of Shanker Singham's speech in quote 5. Speech data point is in red

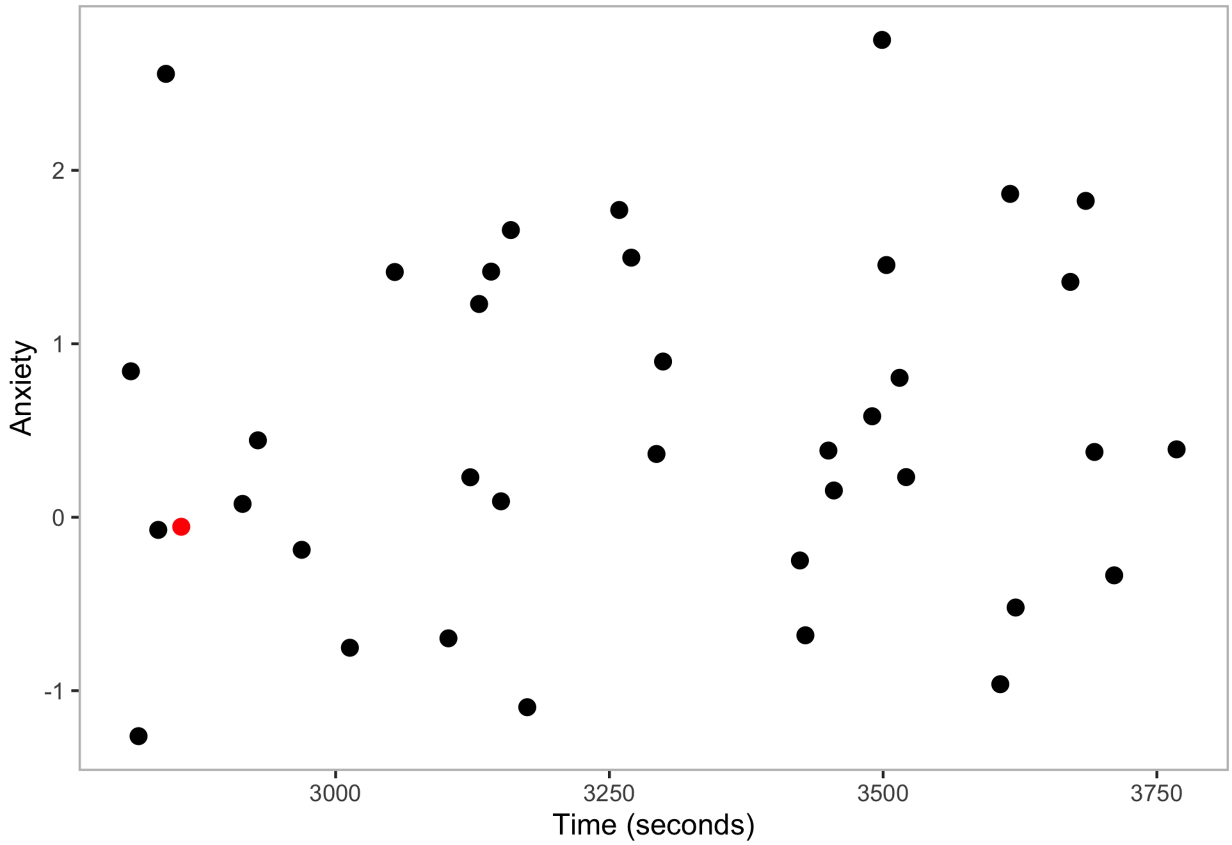


Figure 36: Anxiety score of Shanker Singham's speech in quote 5. Speech data point is in red

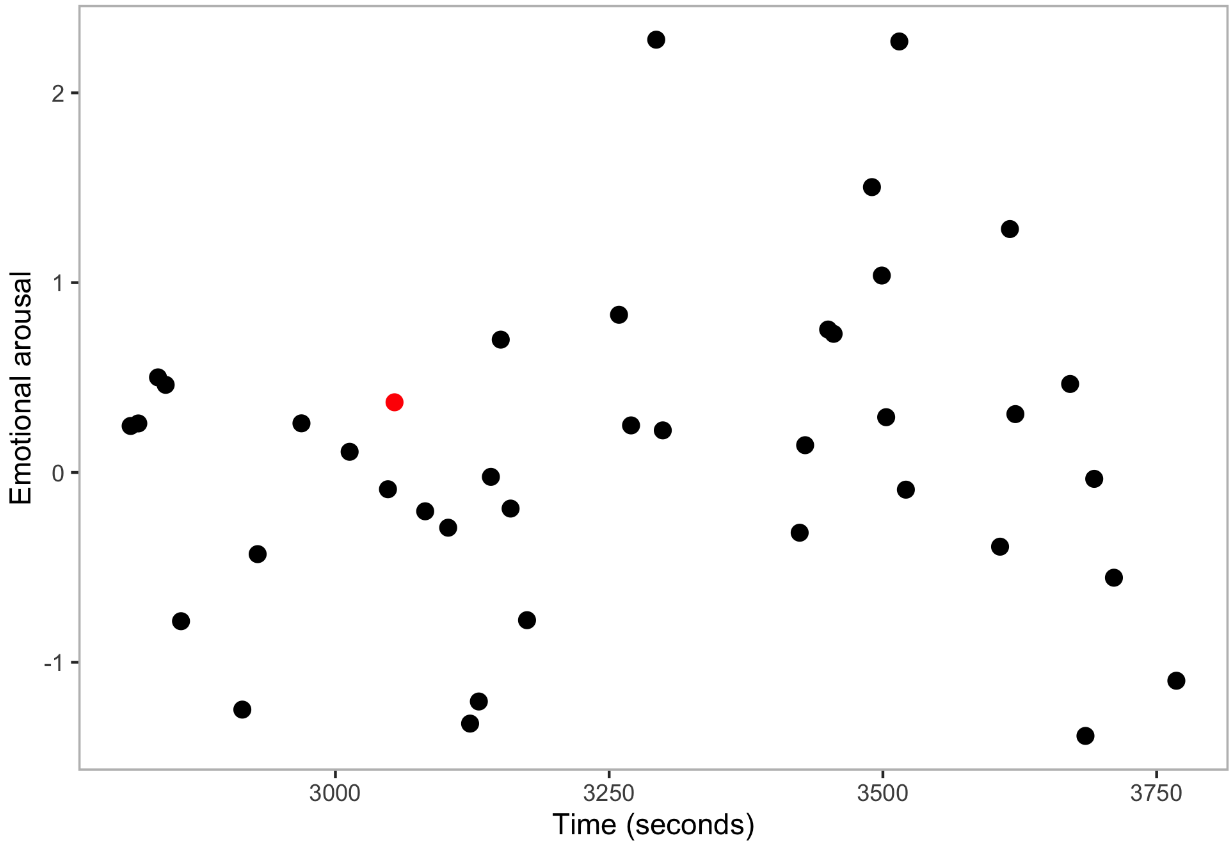


Figure 37: Emotional arousal of Tony Smith's speech in quote 6. Speech data point is in red

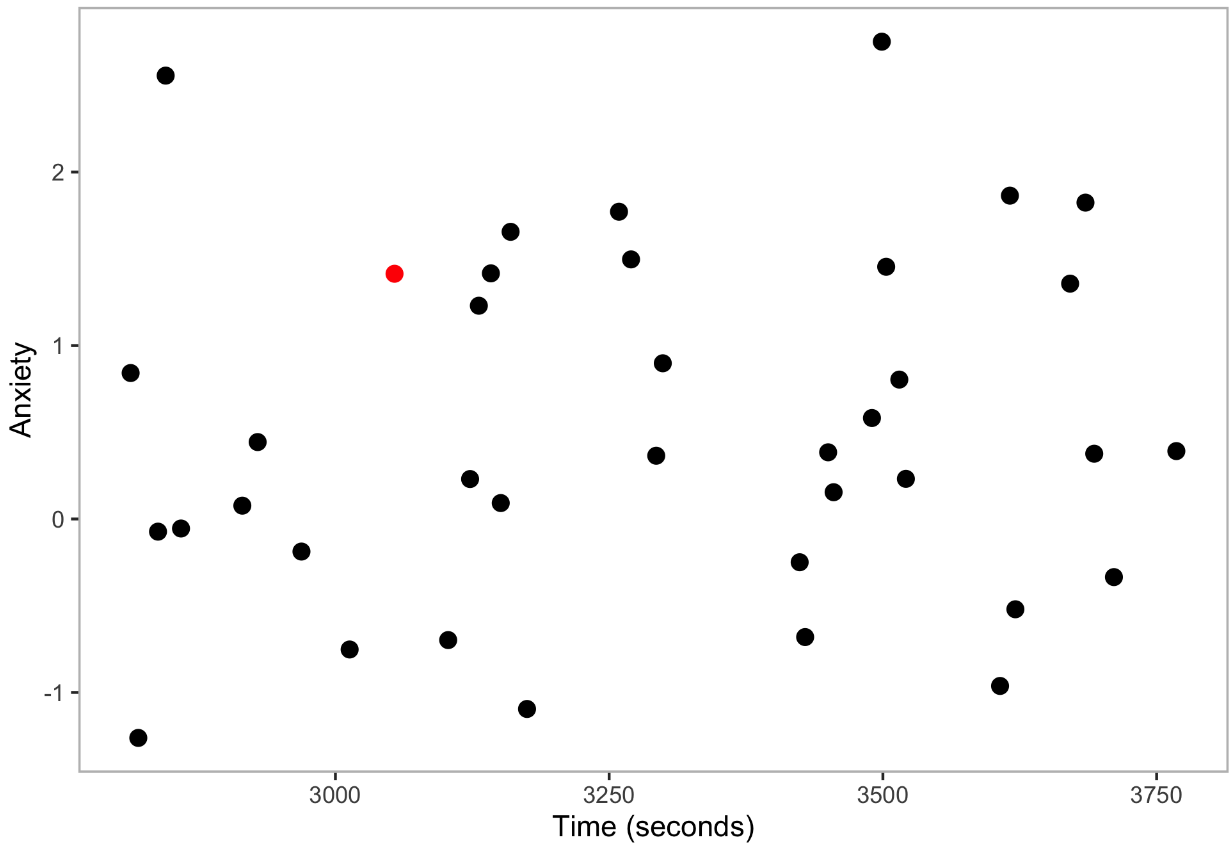


Figure 38: Anxiety score of Tony Smith's speech in quote 6. Speech data point is in red

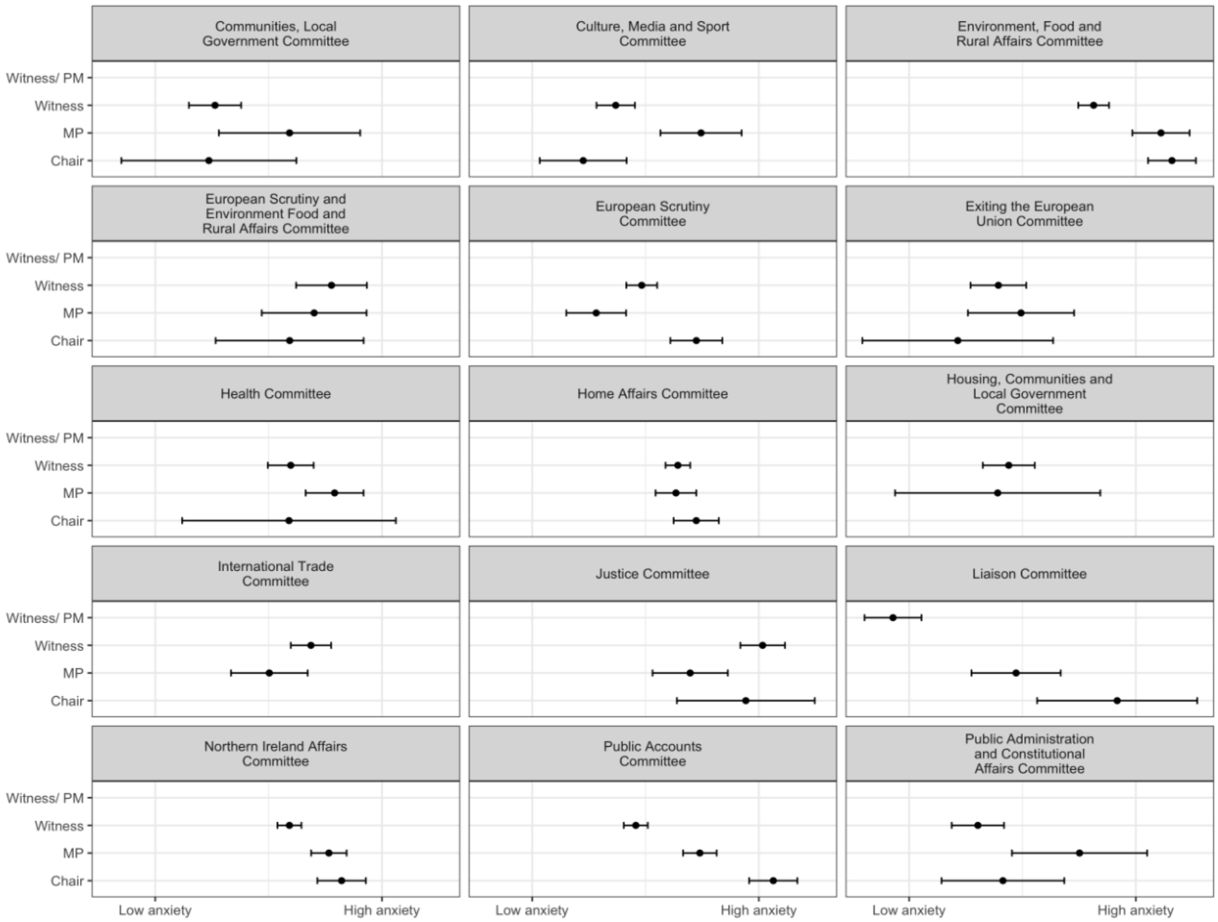


Figure 39: Anxiety by role of speaker for each committee

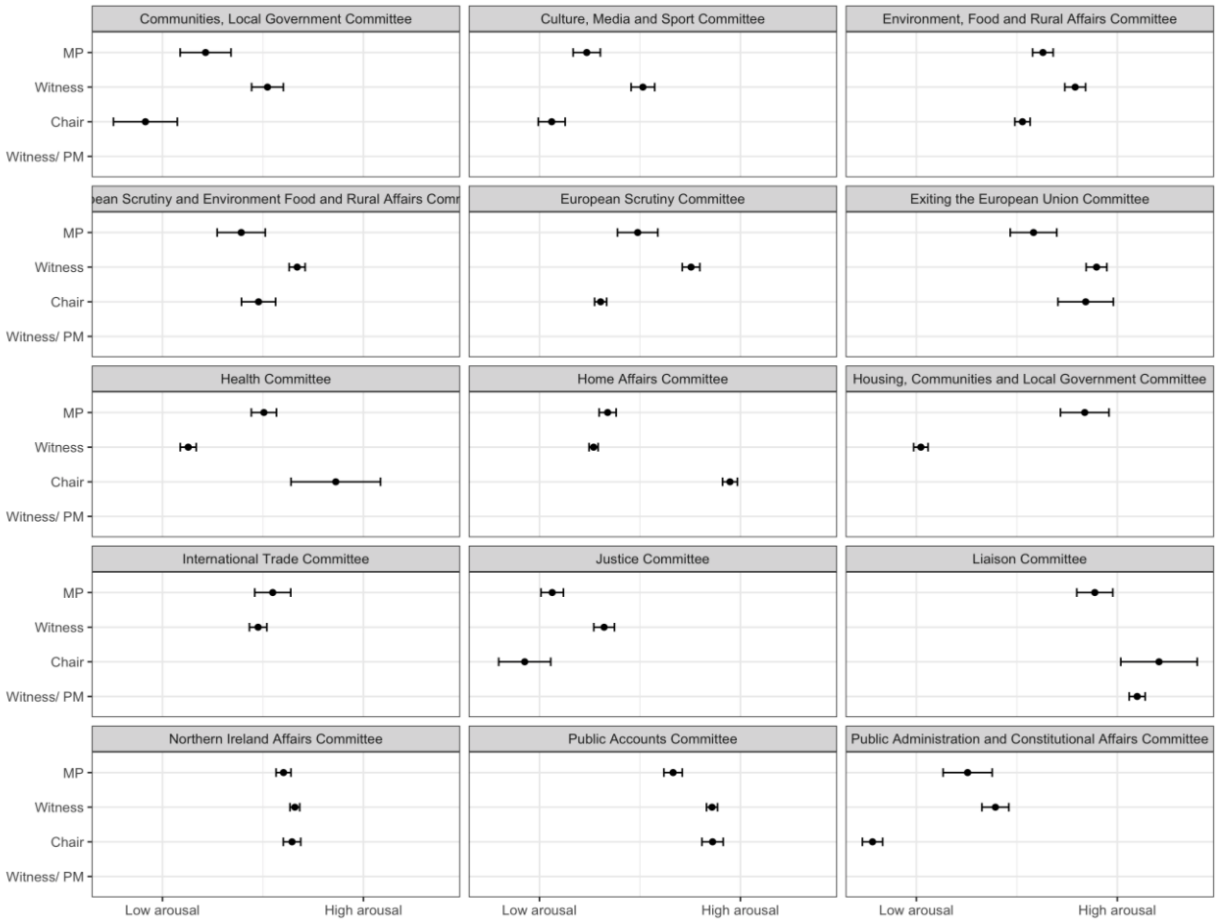


Figure 40: Arousal by role of speaker for each committee

References

Anderson, Rindy C. and Casey A. Klofstad. 2012, December. "Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles". *PLOS ONE* 7 (12): e51216. Publisher: Public Library of Science.

Ansolabehere, Stephen and Shanto Iyengar. 1996, January. *GOING NEGATIVE: How Political Ads Shrink and Polarize the Electorate* (First Edition ed.). New York: Free Press.

Archer, Dawn. 2017, December. "Mapping Hansard Impression Management Strategies through Time and Space". *Studia Neophilologica* 89 (sup1): 5–20. Publisher: Routledge _eprint: https://doi.org/10.1080/00393274.2017.1370981.

Archer, Dawn. 2018, July. "Negotiating difference in political contexts: An exploration of Hansard". *Language Sciences* 68 : 22–41.

Aristotle. 2001, September. *The Basic Works of Aristotle* (Revised ed. edition ed.). New York: Modern Library.

Arnold, Taylor. 2017. "A Tidy Data Model for Natural Language Processing using cleanNLP". *R Journal* 9 (2). arXiv:1703.09570 [cs, stat].

Ashcroft, Lord. 2016, June. "How the United Kingdom voted on Thursday... and why".

Bakker, Bert N., Gijs Schumacher, and Matthijs Rooduijn. 2021, February. "Hot Politics? Affective Responses to Political Rhetoric". *American Political Science Review* 115 (1): 150–164. Publisher: Cambridge University Press.

Barry, Ellen. 2019, January. "John Bercow, Shouting for 'Order' Amid Chaos, Is Brexit's Surprise Star and Villain". *The New York Times*.

Beauchamp, Nick. 2011. "Using Text to Scale Legislatures with Uninformative Voting".

Beauchamp, Zack. 2016, June. "Brexit was fueled by irrational xenophobia, not real economic grievances".

Benton, Meghan and Meg Russell. 2013, October. "Assessing the Impact of Parliamentary Oversight Committees: The Select Committees in the British House of Commons". *Parliamentary Affairs* 66 (4): 772–797.

Bezooijen, Renée van. 1984, December. *Characteristics and Recognizability of Vocal Expressions of Emotion:*. Foris Publications.

Bodnaruk, Andriy, Tim Loughran, and Bill McDonald. 2015, August. "Using 10-K Text to Gauge Financial Constraints". *Journal of Financial and Quantitative Analysis* 50 (4): 623–646. Publisher: Cambridge University Press.

Boersma, Paul and David Weenink. . "Praat: doing Phonetics by Computer".

Britain Stronger In Europe. 2016. "Get the Facts". `http://www.strongerin.co.uk/get_the_facts`. Archived at https://archive-it.org/collections/7397.

Brown, Michael E., Owen R. Coté Jr, Sean M. Lynn-Jones, and Director of Studies Center for Science and International Affairs Steven E. Miller (Eds.)2000, July. *Rational Choice and Security Studies: Stephen Walt and His Critics*. Cambridge, Mass: MIT Press.

Burden, Barry C. and Joseph Neal Rice Sanberg. 2003. "Budget Rhetoric in Presidential Campaigns from 1952 to 2000". *Political Behavior* 25 (2): 97–118. Publisher: Springer.

Cap, Piotr. 2016, November. *The Language of Fear: Communicating Threat in Public Discourse* (1st ed. 2017 edition ed.). London: Palgrave Macmillan.

Cochrane, Christopher, Ludovic Rheault, Jean-François Godbout, Tanya Whyte, Michael W.-C. Wong, and Sophie Borwein. 2022, January. "The Automatic Analysis of Emotion in Political Speech Based on Transcripts". *Political Communication* 39 (1): 98–121. Publisher: Routledge _eprint: https://doi.org/10.1080/10584609.2021.1952497.

Colantone, Italo and Piero Stanig. 2018, May. “Global Competition and Brexit”. *American Political Science Review* 112 (2): 201–218. Publisher: Cambridge University Press.

Commission, European. 2015, July. “Standard Eurobarometer 83 - Spring 2015 - Factsheets”. Technical report.

Commission, European. 2024, May. “CAP at a glance - European Commission”.

Council, European. . “Timeline - The EU-UK withdrawal agreement”.

Demianyk. 2020, September. “End The Shouting, Jeering And Heckling In Parliament, Say Most People”. *HuffPost UK*. Section: Politics.

Dietrich, Bryce J., Ryan D. Enos, and Maya Sen. 2019, April. “Emotional Arousal Predicts Voting on the U.S. Supreme Court”. *Political Analysis* 27 (2): 237–243. Publisher: Cambridge University Press.

Dietrich, Bryce J., Matthew Hayes, and Diana Z. O’Brien. 2019, November. “Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech”. *American Political Science Review* 113 (4): 941–962. Publisher: Cambridge University Press.

Dietrich, Bryce J., D. Schultz, and Tracey Jaquith. 2018. “This Floor Speech Will Be Televised : Understanding the Factors that Influence When Floor Speeches Appear on Cable Television”.

Ekman, Paul, Maureen O’Sullivan, Wallace V. Friesen, and Klaus R. Scherer. 1991, June. “Invited article: Face, voice, and body in detecting deceit”. *Journal of Nonverbal Behavior* 15 (2): 125–135.

Erisen, Cengiz and José D. Villalobos. 2014, October. “Exploring the invocation of emotion in presidential speeches”. *Contemporary Politics* 20 (4): 469–488.

Filipovic, Jill. 2020, September. “Why Trump is too emotional to be president (opinion)”.

Fitzgerald, Ian, Ron Beadle, and Kevin Rowan. 2022, February. "Trade unions and the 2016 UK European Union Referendum". *Economic and Industrial Democracy* 43 (1): 388–409. Publisher: SAGE Publications Ltd.

Foxlee, Neil. 2018, June. "Pivots and Levers: Political Rhetoric around Capitalism in Britain from the 1970s to the Present". *Contributions to the History of Concepts* 13 (1): 75–99. Publisher: Berghahn Journals Section: Contributions to the History of Concepts.

Friedman, Jeffrey. 1996, September. *Rational Choice Controversy: Economic Models of Politics Reconsidered*. New Haven, CT: Yale University Press.

Gelatt, Julia Gelatt Julia. 2005, October. "Schengen and the Free Movement of People Across Europe".

Goldthorpe, John H. 1998. "The Quantitative Analysis of Large-Scale Data-Sets and Rational Action Theory: For a Sociological Alliance 1". In *Rational Choice Theory And Large-Scale Data Analysis*. Routledge. Num Pages: 23.

Gray, J. A. 1985, January. "The Neuropsychology of Anxiety". *Issues in Mental Health Nursing* 7 (1-4): 201–228.

Grimmer, Justin. 2010, January. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases". *Political Analysis* 18 (1): 1–35. Publisher: Cambridge University Press.

Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts". *Political Analysis* 21 (3): 267–297.

Haag, Stefan. 2017, June. "The historical evolution of EU-UK relations".

Helm, Toby. 2016, September. "Brexit camp abandons £350m-a-week NHS funding pledge". *The Observer*.

Huddy, Leonie, Stanley Feldman, and Erin Cassese. 2007, September. "On the Distinct Political Effects of Anxiety and Anger". In W. R. Neuman, G. E. Marcus, M. MacKuen, and A. N. Crigler (Eds.), *The Affect Effect: Dynamics of Emotion in Political Thinking and Behavior*. Chicago, IL: University of Chicago Press.

Huddy, Leonie, Stanley Feldman, and Christopher Weber. 2007. "The Political Consequences of Perceived Threat and Felt Insecurity". *The Annals of the American Academy of Political and Social Science* 614 : 131–153. Publisher: [Sage Publications, Inc., American Academy of Political and Social Science].

Huddy, Leonie, Lilliana Mason, and Lene Aarøe. 2015, February. "Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity". *American Political Science Review* 109 (1): 1–17. Publisher: Cambridge University Press.

Hughes, Ceri. 2019, June. "It's the EU immigrants stupid! UKIP's core-issue and populist rhetoric on the road to Brexit". *European Journal of Communication* 34 (3): 248–266. Publisher: SAGE Publications Ltd.

Jagers, Jan and Stefaan Walgrave. 2007. "Populism as political communication style: An empirical study of political parties' discourse in Belgium". *European Journal of Political Research* 46 (3): 319–345. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-6765.2006.00690.x.

Kellstedt, Paul M. 2000. "Media Framing and the Dynamics of Racial Policy Preferences". *American Journal of Political Science* 44 (2): 245–260. Publisher: [Midwest Political Science Association, Wiley].

Kim, Yoon, Yi-I Chiu, Kentaro Hanaki, Darshan Hegde, and Slav Petrov. 2014, June. "Temporal Analysis of Language through Neural Language Models". In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, Baltimore, MD, USA, pp. 61–65. Association for Computational Linguistics.

Klofstad, Casey A. 2016. “Candidate Voice Pitch Influences Election Outcomes”. *Political Psychology* 37 (5): 725–738. Publisher: [International Society of Political Psychology, Wiley].

Knox, Dean and Christopher Lucas. 2021, May. “A Dynamic Model of Speech for the Social Sciences”. *American Political Science Review* 115 (2): 649–666. Publisher: Cambridge University Press.

Kruger, Haidee, Bertus van Rooy, and Adam Smith. 2019, September. “Register Change in the British and Australian Hansard (1901-2015)”. *Journal of English Linguistics* 47 (3): 183–220. Publisher: SAGE Publications Inc.

Kubala, M. 2011, October. “Select Committees in the House of Commons and the Media”. *Parliamentary Affairs* 64 (4): 694–713.

Labour Leave. 2016. “Labour’s Vote-to-LEAVE Campaign (campaign brochure)”. `http://www.labourleave.org/`. Archived at https://archive-it.org/collections/7397.

Laver, Michael and John Garry. 2000. “Estimating Policy Positions from Political Texts”. *American Journal of Political Science* 44 (3): 619–634. Publisher: [Midwest Political Science Association, Wiley].

Liu, Bing. 2015. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge: Cambridge University Press.

Loughran, Tim and Bill Mcdonald. 2011. “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks”. *The Journal of Finance* 66 (1): 35–65. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2010.01625.x.

Loughran, Tim and Bill Mcdonald. 2016. “Textual Analysis in Accounting and Finance: A Survey”. *Journal of Accounting Research* 54 (4): 1187–1230. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-679X.12123.

MacDonald, Paul K. 2003, November. “Useful Fiction or Miracle Maker: The Competing Epistemological Foundations of Rational Choice Theory”. *American Political Science Review* 97 (4): 551–565.

Mair, Christian. 2019, August. “Brexitiness: The Ebbs and Flows of British Eurosceptic Rhetoric since 1945”. *Open Library of Humanities* 5 (1). Number: 1 Publisher: Open Library of Humanities.

Makortoff, Kalyeena. 2015, July. “Europe’s migrant crisis: Thousands storming UK border”. Section: Europe News.

Marcus, George E. 1988, September. “The Structure of Emotional Response: 1984 Presidential Candidates”. *American Political Science Review* 82 (3): 737–761. Publisher: Cambridge University Press.

Marcus, George E. 2000. *Affective intelligence and political judgment / George E. Marcus, W. Russell Neuman, Michael MacKuen.* Chicago: University of Chicago Press.

Marcus, George E. 2023. “Evaluating the status of theories of emotion in political science and psychology”. *Frontiers in Political Science* 4 .

Martin, Garret. 2022, February. “Order, order! A guide to ‘partygate’ and the UK’s rambunctious Parliament”. *The Conversation*.

Mauss, Iris B. and Michael D. Robinson. 2010. “Measures of emotion: A review”. In J. d. Houwer and D. Hermans (Eds.), *Cognition and emotion: reviews of current research and theories*. Hove, East Sussex ; New York, NY: Psychology Press.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. “Distributed Representations of Words and Phrases and their Compositionality”. In *Advances in Neural Information Processing Systems*, Volume 26. Curran Associates, Inc.

Mohammad, Saif M and Peter D Turney. 2010, June. “Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon”. In *Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, LA, pp. 9.

Mohammad, Saif M. and Peter D. Turney. 2013. “Crowdsourcing a Word-Emotion Association Lexicon”. *Computational Intelligence* 29 (3): 436–465,. arXiv: 1308.6297.

Mollin, Sandra. 2007, November. “The Hansard hazard: gauging the accuracy of British parliamentary transcripts”. *Corpora* 2 (2): 187–210. Publisher: Edinburgh University Press.

News, BBC. 1975, April. “1975: Labour votes to leave the EEC”.

News, BBC. 2016a, June. “Brexit: David Cameron to quit after UK votes to leave EU”. *BBC News*.

News, BBC. 2016b, June. “Immigration focus is turning point in EU campaign, says Farage”. *BBC News*.

Nielsen, Finn Årup. 2011, May. “A new ANEW: Evaluation of a word list for sentiment analysis in microblogs”. In *'Making Sense of Microposts': Big things come in small packages*. arXiv: 1103.2903.

Osnabrügge, Moritz, Sara B. Hobolt, and Toni Rodon. 2021. “Playing to the Gallery: Emotive Rhetoric in Parliaments”. *American Political Science Review*: 1–15. Publisher: Cambridge University Press.

Owren, Michael J. and Jo-Anne Bachorowski. 2007. “Measuring emotion-related vocal acoustics”. In *Handbook of emotion elicitation and assessment*, Series in affective science, pp. 239–266. New York, NY, US: Oxford University Press.

Parliament, European. 2022, April. “The Treaty of Lisbon”.

Reid, David. 2019, September. "Anger and tears as the UK Parliament's mood boils over". *CNBC*.

Rheault, Ludovic. 2016. "Expressions of Anxiety in Political Texts". In *Proceedings of the First Workshop on NLP and Computational Social Science*, Austin, Texas, pp. 92–101. Association for Computational Linguistics.

Rheault, Ludovic, Kaspar Beelen, Christopher Cochrane, and Graeme Hirst. 2016, December. "Measuring Emotion in Parliamentary Debates with Automated Textual Analysis". *PLOS ONE* 11 (12): e0168843. Publisher: Public Library of Science.

Rheault, Ludovic and Christopher Cochrane. 2020, January. "Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora". *Political Analysis* 28 (1): 112–133. Publisher: Cambridge University Press.

Schroeder, Juliana and Nicholas Epley. 2016, August. "Mistaking Minds and Machines: How Speech Affects Dehumanization and Anthropomorphism". *Journal of Experimental Psychology: General* 145 .

Slembrouck, Stef. 1992, May. "The parliamentary Hansard 'verbatim' report: the written construction of spoken discourse". *Language and Literature* 1 (2): 101–119. Publisher: SAGE Publications Ltd.

Smelser, Neil J. 1992, October. "The Rational Choice Perspective: A Theoretical Assessment". *Rationality and Society* 4 (4): 381–410. Publisher: SAGE Publications Ltd.

Sprunt, Barbara. 2020, July. "'I Could Not Allow That To Stand': Ocasio-Cortez Rebukes Republican For Vulgar Insult". *NPR*.

Sundby, Alex. 2016, June. "Brexit result shocks some voters who wanted U.K. to leave European Union - CBS News".

Tarr, Alexander, June Hwang, and Kosuke Imai. 2023, October. "Automated Coding of Political Campaign Advertisement Videos: An Empirical Validation Study". *Political Analysis* 31 (4): 554–574.

Tasch, Barbara. 2016, March. "This map shows how much the refugee crisis is dividing Europe".

Tolvanen, Juha, James Tremewan, and Alexander K. Wagner. 2021, October. "Ambiguous Platforms and Correlated Preferences: Experimental Evidence". *American Political Science Review*: 1–17. Publisher: Cambridge University Press.

Treasury, HM. 2015, December. "European Union Finances 2015: statement on the 2015 EU Budget and measures to counter fraud and financial mismanagement". Technical Report Cm 9167.

UK Parliament. . "Select Committees".

UK Parliament. 2019, June. "Alternative arrangements to the backstop examined - Committees - UK Parliament".

Vasilopoulou, Sofia and Markus Wagner. 2017, September. "Fear, anger and enthusiasm about the European Union: Effects of emotional reactions on public preferences towards European integration". *European Union Politics* 18 (3): 382–405. Publisher: SAGE Publications.

Vote Leave. 2016. "Why should we Vote Leave on 23 June? (campaign brochure)". `http://www.voteleavetakecontrol.org/`. Archived at https://archive-it.org/collections/7397.

Young, Lori and Stuart Soroka. 2012, April. "Affective News: The Automated Coding of Sentiment in Political Texts". *Political Communication* 29 (2): 205–231. Publisher: Routledge _eprint: https://doi.org/10.1080/10584609.2012.671234.

Zahniser, David, Dakota Smith, and Emily Alpert Reyes. 2020, April. "Garcetti: L.A. is 'under attack' and will need to furlough thousands of city workers".