

**Ministry of Education and Science of Ukraine
Taras Shevchenko National University of Kyiv
Educational and Scientific Institute of Philology
Department of English Philology and Intercultural Communication**

Master's Thesis

LANGUAGE PROFILE OF RISHI SUNAK

Olha Kudruk,

2nd year student of the Education Program
'English Communication Studies and Translation
and Two Western European Languages'

Field of science: 03 "Humanities"

Specialty: 035 "Philology"

Supervised by:

Anna Karaban, PhD

«Допущено до захисту»

Протокол засідання кафедри англійської філології

та міжкультурної комунікації

Протокол № 8 від 22.04.2024

Зав. кафедри _____ д. філол. н., проф. Алла БЄЛОВА

АНОТАЦІЯ

Актуальність дослідження мовного профілю Ріші Сунака полягає в тому, що воно дозволить нам краще зрозуміти, як політик використовує мову для спілкування з громадськістю, які ефективні мовні засоби він використовує для переконання та мобілізації своєї аудиторії, а також які ідеї та цінності він поділяє та просуває.

Об'єктом дослідження є мовна особистість Ріші Сунака та рамки його мовленнєвої поведінки в дискурсивному просторі політики.

Предметом дослідження є лексичні засоби вербалізації мовної особистості Ріші Сунака.

Мета дослідження - розкрити мовний профіль Ріші Сунака як сучасного мультикультурного британського політика та визначити його ключові цінності, ідеологію та управлінські стратегії в контексті політичної діяльності.

Для досягнення поставленої мети необхідно вирішити такі **основні завдання**:

- розглянути поняття політичного дискурсу та його принципи;
- розкрити поняття мовної особистості та способи її реалізації в політичній культурі;
- порівняти лексичні аспекти індивідуального стилю політика протягом трьох періодів його кар'єри;
- визначити ключові концепти та цінності, якими керується Ріші Сунак на політичній арені, і те, як вони резонують з цінностями Великої Британії як країни.

Основними **методами дослідження** були контент-аналіз, кількісний метод, порівняльний метод, реалізовані за допомогою програми WordSmith Tools.

Робота складається зі вступу, трьох розділів, загальних висновків, списку використаних джерел, списку ілюстративного матеріалу, додатків та анотації українською та англійською мовами.

Використовуючи засоби WordSmith Tools, аналіз промов Ріші Сунака виявив зміни в його стилі спілкування та фокусі протягом різних періодів його

кар'єри. Спочатку наголошуючи на можливостях і сімейних цінностях, він пізніше перейшов до конкретних економічних термінів на посаді канцлера, а на посаді прем'єр-міністра наголошував на нагальності та патріотизмі.

Ключові слова: Ріші Сунак, мовна особистість, політичний дискурс, Wordsmith tools, аналіз промов

ABSTRACT

The **relevance** of the study of Rishi Sunak's language profile is that it will allow us to better understand how the politician uses language to communicate with the public, what effective language means he uses to persuade and mobilise his audience, as well as what ideas and values he embraces and promotes.

The **object** of the study is the linguistic personality of Rishi Sunak and the framework of his speech behaviour in the discursive space of politics.

The **subject** of the study is the lexical means of verbalisation of linguistic personality of Rishi Sunak.

The purpose of the study is to reveal the linguistic profile of Rishi Sunak as a modern multicultural British politician and to identify his key values, ideology and management strategies in the context of political activity.

In order to achieve this goal, **the following main tasks** need to be addressed:

- to consider the concept of political discourse and its principles;
- to reveal the concept of linguistic personality and ways of its realisation in political culture;
- to compare the lexical aspects of Rishi Sunak's individual style during the three periods of his career;
- to identify the key concepts and assets that guide Rishi Sunak in the political arena and how they resonate with the values of the UK as a country.

The main **research methods** used were content analysis, quantitative method, comparative method, implemented with the help of Wordsmith Tools.

The paper consists of an introduction, three chapters, general conclusions, a list of references, a list of illustrative material, an appendix, and an annotation in Ukrainian and English.

Using WordSmith Tools, an analysis of Rishi Sunak's speeches revealed changes in his communication style and focus during different periods of his career. Initially emphasising opportunity and family values, he later shifted to specific economic terms as chancellor, and as prime minister he emphasised urgency and patriotism.

Keywords: Rishi Sunak, linguistic personality, political discourse, Wordsmith tools, speech analysis

CONTENTS

INTRODUCTION.....	8
CHAPTER 1. THEORETICAL FOUNDATIONS OF THE STUDY OF LINGUISTIC PERSONALITY IN POLITICAL DISCOURSE.....	11
1.1.Political discourse as an object of modern linguistic research	11
1.2.Main features of political discourse.....	15
1.3.Functions of political discourse	19
1.4.Corpus-assisted discourse analysis	25
1.5.Basic corpus techniques used in critical discourse analysis	27
1.6.The advantages of applying corpus methods in CADS.....	34
1.7.The disadvantages of applying corpus methods in CADS	36
1.8.The notion of linguistic personality	37
1.9.Rishi Sunak as a political figure	41
CHAPTER 2. METHODOLOGICAL FOUNDATIONS FOR CREATING AND PROCESSING OF THE CORPUS.....	46
2.1.The use of content analysis method.....	46
2.2.Data collection	46
2.3.The use of WordSmith Tools for corpus linguistics.....	51
2.4.Reference corpus description.....	55
CHAPTER 3. LINGUISTIC PECULIARITIES OF RISHI SUNAK'S LANGUAGE PROFILE.....	57
3.1.Analysis with the help of WordSmith Tools	57
3.1.1. Identifying keywords based on the reference corpus.....	57
3.1.2. Identifying speech patterns based on concordances.....	62
3.2.The implementation of Rishi Sunak's linguistic personality at the lexical level.....	64

3.2.1. The beginning of the work in the government (2015 - 2018).....	64
3.2.2. The chancellorship (2019 - 2022)	73
3.2.3. The premiership (2022 – present time)	83
CONCLUSION	95
SUMMARY	99
REFERENCES.....	104
DATA SOURCES.....	110
APPENDIXES	113

INTRODUCTION

Language does not exist outside of culture. It is one of its most important components, a form of thinking, a specific manifestation of human life, which, in turn, is the true essence of language. Therefore, language and culture are inseparable. Researchers note that language is culture, and culture is language. One of the trends in linguistics is the detailed elaboration of the problem of the human factor in speech activity. Considering that humans play a key role in the process of creating and using language, the anthropological approach becomes central in linguistic research.

Active research in language across various fields, especially in politics, has led to the emergence of the concept of "linguistic personality." In 19th-century European linguistics, the issue of linguistic personality arose in the examination of the social nature of language, the relationship between language and speech, and the language of the individual and the collective.

The study was based on 22 speeches by Rishi Sunak collected in a corpus. Three periods of his career were taken as a basis to ensure the objectivity of the study and to trace the development of his personality, namely the beginning of his work in the government (2015- 2018), the chancellorship (2019 - 2022) and the premiership (2022 - present time). The total volume of the analysed material is 32500 words (160000 characters).

The **relevance** of such a study lies in the fact that it will allow us to better understand how Rishi Sunak uses language to communicate with the public, what effective language means he uses to persuade and mobilise his audience, and what ideas and values he embraces and promotes. Such a study could be an important addition to the overall understanding of the political culture and dynamics of the contemporary British political landscape.

The **purpose** of the study is to reveal the linguistic profile of Rishi Sunak as a modern multicultural British politician and to identify his key values, ideology, and management strategies in the context of political activity.

To achieve this goal, the **following main tasks** need to be addressed:

- to consider the concept of political discourse and its principles;
- to reveal the concept of linguistic personality and ways of its realisation in political culture;
- to compare the lexical aspects of Rishi Sunak's individual style during the three periods of his career, namely the beginning of his work in the government, the chancellorship, and the premiership;
- to identify the key concepts and assets that guide Rishi Sunak in the political arena and how they resonate with the values of the UK as a country.

The **object** of the study is the linguistic personality of Rishi Sunak and the framework of his speech behaviour in the discursive space of politics.

The **subject** of the study is the lexical means of verbalisation of communicative goals implemented in Rishi Sunak's speeches.

The study was based on 22 speeches by Rishi Sunak collected in a corpus. Three periods of his career were taken as a basis to ensure the objectivity of the study and to trace the development of his personality. The total volume of the analysed material is 160000 characters.

The main **research method** used in the study was content analysis. Furthermore, the following methods were utilised: a quantitative method (to count the frequency of use of linguistic units in Sunak's speeches), a comparative method (to trace common and distinctive features in speeches for the three periods of his career). These methods were carried through the use of the Wordsmith Tools software. Namely, we used the WordList function (to generate frequency lists for all periods and compare them), the Keywords function (to determine the the words which are most unusually frequent in the given corpus), the Concord function (reveal insights into frequency, context, collocation of language elements).

The **novelty** of the study of Rishi Sunak's language profile lies in several key aspects. Firstly, the analysis of his linguistic style can reveal the uniqueness of the communication approach of a new political actor. Secondly, given his governmental position, the study can reveal how he uses language to represent government policies.

Third, an analysis of his language profile can reveal how he adapts his speech to cultural and social contexts. Finally, the study of linguistic strategies allows us to understand how Rishi Sunak influences public opinion and shapes his political image. Such a study proves to be key to understanding the current political landscape and the role of communication in it.

The **practical significance** of the study lies in the unique lessons in political communication and contributes to the development of communication skills for politicians and public figures. The analysis of the language profile also helps to better understand the cultural and social contexts of a politician's activities, which contributes to building mutual understanding in society.

The structure of the research paper. The paper consists of an introduction, three chapters, general conclusions, a list of references (102 items, including 86 items in foreign languages), a list of illustrative material, an appendix, and an annotation in Ukrainian and English. The total volume of the work is 118 pages, the volume of the main text is 95 pages.

CHAPTER 1. THEORETICAL FOUNDATIONS OF THE STUDY OF LINGUISTIC PERSONALITY IN POLITICAL DISCOURSE

1.1. Political discourse as an object of modern linguistic research

In the twenty-first century, discourse studies occupy a leading place in modern linguistics, and the relevance of such studies arose due to a change in the scientific paradigm: anthropocentric models of research gradually replaced systemic-structural and statistical ones. Modern approaches to the definition of discourse are based on linguistic and cultural studies, in which language, culture and cultural and linguistic personality are central.

Political discourse plays a significant role in shaping the way people think and feel about certain phenomena, events, or individuals, as politicians can use the power of rhetoric and their skills to shape, control, or change public opinion. Therefore, there is considerable linguistic interest in researching this phenomenon.

In general, despite a considerable number of foreign and domestic works on the study of discourse, and political discourse in particular, there is no generally accepted definition of either the general term "discourse" or the specific term "political discourse".

The diversity of approaches to the definition of this concept leads to attempts by linguists to identify common areas of interpretation. The term "discourse" was independently coined by Z. Harris in 1952 [48], defining it as a sequence of spoken or written sentences within a specific context. In this regard, P. Seriot [71] considers several definitions, among which the most widespread in linguistics are: 1) discourse is any specific utterance; 2) discourse (within the framework of pragmatics) is the impact of the utterance on the addressee and its entry into the situation of the utterance, taking into account the subject and addressee, time and place of the utterance; 3) discourse is a system of restrictions imposed on an unlimited number of utterances arising from a certain social or ideological position.

Michel Foucault [39] made use of the findings of Emile Benveniste while delving into the purpose and aims involved in the discourse analysis. Within the French school of discourse analysis, discourse is perceived as either oral or written language,

representing a complex manifestation of linguistic activity and serving as a distinct mode of expression characteristic of particular social-political groups or historical periods.

In his exploration of semiotic theory, Roland Barthes scrutinised various discourses including medicine, history, fashion, customs, myths, advertising, and mass-produced objects [21]. The linguist argued that discourse encompasses social signs carrying socially significant meanings and mythological narratives. Barthes proposed that not only words but also images and objects hold the potential to convey essential semantic messages.

The idea of the social aspect of the discourse was further carried out by Jurgen Habermas [45]. He linked the term discourse to the theory of social action and essentially highlighted the communicative action as an ideal communicative model.

Teun A. van Dijk [76] advocated for discourse analysis in both broad and narrow meanings. As a representative of critical discourse analysis, van Dijk believed that discourse should always be analysed in the context of a given situation, in the existing cultural tradition and the prevailing ideology of the time. According to the linguist, discourse is viewed as a finished product or as an ongoing process, subject to interpretation by its recipients.

F. Batsevich [1] notes that discourse is a type of communicative activity, an interactive phenomenon, a stream of speech that has different forms of expression (oral, written, paralinguistic) that takes place within a certain channel communication, which is managed by participants with the help of strategies and tactics; synthesis of cognitive linguistic and extralinguistic factors that are determined by a certain range of "forms of life" depending on the subject of communication, which results in the formation of different speech genres.

As a result of generalising numerous approaches to the definition of the term discourse, K. Serazhym [13] offers the following definition: discourse is a complex sociolinguistic phenomenon of the modern communicative environment, which, firstly, is directly or indirectly determined by sociolinguistic, political, pragmatic-situational, psychological and other factors; secondly, it has a linguistic and extralinguistic

structure and, thirdly, is characterised by the commonality of the world created by the author in the process of unfolding the discourse and interpreted by its recipient. In the following, I will use K. Serazhym's definition due to its exhaustiveness.

As for political discourse, this concept is as multifaceted as the concept of discourse, which is explained, on the one hand, by numerous approaches to defining this phenomenon (as well as discourse itself), and on the other hand, by the interdisciplinary nature of this term. In order to understand the essence of this phenomenon, we believe it is necessary to first consider the definition of the concept of "politics".

According to Wodak and de Cilia [79], a crucial concern within language and politics research revolves around defining what constitutes politics and political action. At its core, politics comprises the realm of the polity, including the actions and endeavours of professional politicians, formal political institutions, and engaged citizens. Additionally, political practice typically involves the exertion of power struggles and collaborative efforts aimed at advancing the objectives of a society or group [31].

Some scholars have argued that the definition of "political discourse" is closely linked to two understandings of politics - a broad and a narrow, or more specific, one. The broad definition of politics includes the activities of those organisations that belong to civil society and that are not necessarily regulated by the state, but at the same time compete for resources - trade unions, business associations, environmental groups, etc. In the narrowest sense, politics is limited to the activities of institutions such as the government, parliament and parties that play a role in resource allocation. This duality in understanding political discourse, whether broad or narrow, is encapsulated in Paul Bayley's statement:

“There is no such thing as political language, but a wide and diverse set of discourses, or genres, or registers that can be classified as forms of political language. It would be feasible to identify a set of ‘canonical’ forms of political discourse: policy papers, ministerial speeches, government press releases or press conferences, parliamentary discourse, party manifestos (or platforms), electoral speeches, etc. They

are all characterised by the fact that they are spoken or written by (or for) primary political actors – members of the government or the opposition, members of parliament, leaders of political parties, candidates for office” [22].

In broad terms, political discourse refers to the language utilised by mass media or various institutions within social and political communication contexts [22].

The Dutch professor T.A. van Dijk argues that "political discourse is defined by its actors or authors, namely politicians. Context plays a fundamental role in van Dijk's theory of political discourse, which sees it as a mental model through which a participant in the political process perceives the information received [76].

Considering the pragmatic aspect of political discourse, I. Klymenko gives the following definition: political discourse is a complex communicative phenomenon, which aims at the struggle for power through the formation of public opinion, which includes a text as a verbalised result of speech, situational and socio-cultural context, as well as special linguistic means, that meet both the goals and objectives of the discourse [5].

N. Kondratenko, applying the communicative-discursive approach, defines political discourse as a specific manifestation of political communication, which involves the actualisation of a political text in a communicative act of interaction between a political subject (politician, political force, government) and object (audience, electorate, voter) [6].

L. Nagorna emphasises that the political aspect of the concept of "discourse" is most accurately reflected in its definition as a type of linguistic communication, as well as the semantic space in which political meanings are tested. In a broad sense, discourse can be seen as a mechanism for reconciling personal, social and cultural knowledge. In a narrower, applied aspect, it focuses on the analysis of linguistic communication and individual texts, speeches, and interviews [9].

A generalised definition of political discourse is found in the work of K. Serajim: it is a text conditioned by the situation of political communication [13]. This definition correlates with the classical definition proposed by N. Arutyunova, according to which

discourse is "speech immersed in life". It can be clarified that political discourse is "speech immersed in political life".

In the study, we will use Klymenko and Kondratenko's definition of political discourse as a kind of manifestation of the struggle for power through interaction with the public.

It should be noted that political linguistics is formed at the intersection of such different fields of science as linguistics, political science, sociology, psychology, cultural studies, ethnography, etc. Also, for a comprehensive in-depth analysis, linguists also use the achievements of cognitive linguistics, rhetoric, stylistics, text linguistics, linguopragmatics and sociolinguistics.

1.2. Main features of political discourse

Politics and language are closely linked at a fundamental level, as political activity clearly cannot exist without the use of language - its activities are predominantly made up of language [32].

The language of politics is important because it shows us how power is gained and how language is used to influence and persuade audiences. It is only through the language and with the help of the language tied to social and political institutions that commands, threats, questions, proposals and promises can be issued. Verbal communication becomes crucial in political interaction with the audience. Thus, to convince people, they must make them believe that what they say is true.

Political discourse is a phenomenon that society encounters every day. That is why many studies of political discourse deal with the language of professional politicians and political institutions, some of which are discourse analytical [32]. The main pragmatic component of political discourse is the so-called "political performances":

- acts of distrust, demands, unity, solidarity;
- appeals;
- promises;
- slogans.

The language of political discourse is characterised by semantic ambiguity, esotericism, ideological polysemy, complexity of lexical meanings, and the use of emotionally charged vocabulary, which collectively allows the speakers to manipulate the content of their statements. From a cognitive perspective, political discourse is characterised by the use of abstract concepts, the assessment of the truthfulness of statements based on their authorship and alignment with ideological viewpoints. A wide range of texts is utilised in political discourse practice: on one hand, narrative, declarative, and referential texts, and on the other hand, literary, journalistic, and scholarly texts, which serve as the basis for political figures as instruments of influence on the consciousness and experience of the listener.

According to N. Kirvalidze [53], political discourse is studied in terms of its linguistic, cultural and pragmatic aspects. According to this approach, the following special features of political discourse can be identified:

1. Institutional nature and illocutionary function.

A person enters a particular institutional situation not only in a certain social role, but also with a certain purpose; the illocutionary force of political discourse is the struggle for political power in its various manifestations, whether it is coming to power or consolidating it, winning supporters or defeating opponents in various spheres of activity (such as election campaigns, etc.), regulating the distribution of resources, etc.

2. Linguistic diglossia of political discourse, which implies the simultaneous functioning of the propaganda political and general subsystems of language.

On the one hand, political language is a special, functionally determined language, while, on the other hand, it can be interpreted as the political jargon of an ideologically united group of people. Thus, political language should perform functions that seem to be mutually exclusive: it should be understandable to the masses, but at the same time, political discourse should be addressee-oriented, i.e. directed at and understandable only to certain social groups.

3. Informational communicative specificity, which is determined by its illocutionary function and sub-institutional type. Thus, we can distinguish three communicative types of politician's discourse:

- informative (in the linguistic sense of the term), focused on the relevance of the information contained in the speech;
 - conventional and phatic, which is stereotypical and ritualistic and does not aim to provide the addressee with new information (for example, the president's speech at the inauguration ceremony);
 - emotionally expressive, whose illocutionary function is to influence public opinion (e.g., emotional expressiveness of political discourse is considered one of the election strategies of politicians).
4. The polemical and theatrical nature of political discourse.

Political debates often turn into a war of words. Polemicism and theatrical nature of political discourse is especially evident during election campaigns: the victory of a politician representing a particular party is determined not so much by their political platform as by how successful the show was put on by image makers and how much it influences public opinion.

5. A large number of slogans.

Slogans can be seen as a kind of socio-political advertising. A political slogan aims to regulate ideological relations in society by informing people about the ideological and conceptual platforms of political parties and other political actors. The effectiveness of a political slogan depends on the extent to which the ideology declared in the most concise form appeals to the public consciousness.

6. Ambiguity and esotericism.

These characteristics of political discourse can be viewed as communicative and pragmatic strategies of a politician, aimed at disguising and neutralising undesirable information and facts; concealment of the truth and deliberate distraction of the population; avoidance of responsibility through anonymisation and depersonalisation; preservation of own political rating and image; avoiding public conflicts in communication.

7. Metaphorisation and manipulation of symbols.

These features of political discourse create its emotional expressiveness, aimed at enhancing the influence of political discourse on public opinion. Politicians use figurative symbols to manage public consciousness, especially during political and economic crises, when they seek to persuade society to believe and do things they might not otherwise believe or do. There is a direct correlation between the socio-economic and political situation in the country and the degree of metaphorisation in politicians' speeches: the degree of metaphorisation of political discourse can be seen as a linguistic marker or indicator of political and socio-economic crisis in the country.

8. Addressee factor and interactivity.

Political discourse is usually aimed at a dual audience: a specific audience that is physically present at a political event and participates in the discourse, and the so-called implicit audience, which includes the general public and sometimes the entire population of a state. As for the interactive features of the addressee, they are manifested through the audience's reaction, which can be mainly of three types: passive, active expression of solidarity, and oppositional-antagonistic.

9. Special contact with the audience and its dependence on the mass media.

This implies two types of contact with the addressee: direct and mediated through the mass media. None of the other institutional types of discourse depends so much on the mass media as the political one. Among the main functions of the mass media are informing citizens about the activities of institutions, recontextualising and translating institutional language into their own language. Therefore, the mass media can be considered as the main instrument of political influence on public opinion, acting as intermediaries between political subjects and the population. Thus, politics, which since its inception has been predominantly linguistic activity, can relate to many different layers of discourse: dialogue within and between institutions, dialogue between the media and institutions, dialogue between the media and parties, as well as dialogue between the media and citizens.

In summary, political discourse is intricately linked with language, serving as a tool for power and persuasion. It includes various communication strategies and types, aiming to persuade audiences and manipulate information. Key functions include

disseminating information, setting agendas, and predicting the future. Different factors influence its nature, such as form, speaker, audience, and purpose. Additionally, political discourse relies heavily on mass media to influence public opinion.

1.3. Functions of political discourse

The functional purpose of political discourse is determined by a semiotic understanding of language as a system of signs, which "exists and is used to achieve any goals, otherwise, to perform functions." The functionalist methodology involves studying the structure and functioning of language to identify correspondences between them. Theoretically, functional methodology is based on recognising the interdependence between form and function and considering the influence of language usage on structural characteristics [16].

The functional approach to studying political discourse involves analysing the functions of discourse, which allows for the examination of the language functions of politics in socio-cultural and linguistic contexts. Questions about the functions of language interest many linguists, philosophers, and psychologists to varying degrees, but consensus on this matter has not been reached to date [16].

According to B. Prokopenko [12], political language lies between two poles - a functionally conditioned specialised language and the jargon of a particular group with its characteristic ideology. Therefore, political language must serve contradictory functions, namely: to be accessible for understanding (according to the goals of propaganda) and oriented towards a specific group (for historical and socio-psychological reasons). The latter often contradicts the accessibility of political language.

The main directions in defining language functions in modern linguistics were laid down by K. Bühler and R. Jakobson. K. Bühler [29] focuses on three main components of the communication act (sender, receiver, things and states). The triangular model with three variant relationships provides the basis for identifying the leading functions of language: expression, appeal, and representation.

A similar basis for differentiating language functions is implemented by R. Jakobson [51], who identifies six main functions of language depending on one of the

elements of the communicative act: the addressee, the addresser, the message, the contact, the context, and the code. The referential function of discourse involves reference to objects, events, or situations. The phatic function of language and discourse is also associated with reference. The conative (also known as the appellative or regulatory function) and expressive (also known as the emotive or emotive-affective function) functions are related to the addresser and addressee. The poetic function ensures the realisation of the code and is oriented towards conveying the necessary information to the listener.

The functions identified by R. Jakobson and K. Bühler correspond to the nature of language as a whole. It is important to emphasise that the degree of realisation of these functions in different types of discourse is considered a determining factor for their classification. Thus, political discourse, alongside other discourses (advertising, medical, legal, religious, media), is included in the group of discourses with a predominant conative (regulatory) function [7].

Therefore, based on the intended purpose of functions, we consider the main purpose of the conative function of political discourse to be its application as an active instrument in the political struggle for power (struggle for power acquisition, retention, realisation, stabilisation, and redistribution).

The primary function of political discourse is rightly considered the struggle for power. This is because any form of oral or written expression by politicians is aimed at influencing and positively presenting themselves as a means of achieving or maintaining power. N. Kondratenko [6] identifies three functions of political discourse: informing, influencing, and agitation. L. Nagorna [9] distinguishes informative, orientational, mobilisational, consolidating, and identification functions of political discourse.

According to F.H. Semenchenko, political discourse performs seven main functions: informational, instrumental, prognostic, normative, legitimising, persuasive, and political propaganda [10].

The informational function is implemented through the interaction between political actors and the media. Political discourse always seeks to objectify its

informational elements. The projection of the informational field of discourse onto the audience, mainly through the media, is a necessary condition for acquiring and maintaining power. The process of informing is a necessary condition for maintaining power relations in society [10].

The instrumental function provides mechanisms for the informational reproduction of political discourse in society, with its foundation based on the relationship between power and the media. The essence of these relationships lies in the fact that power seeks to fully control the instruments for reproducing political information in society, which are essentially in the hands of the media. In turn, the media exert significant influence on political discourse, possessing a monopoly on the tools to maintain and alter the balance of political forces in society [10].

The central feature of the normative function is the distribution of resources - the most important element of the dominance-subordination relations in society. Another important aspect of the normative function is the legitimate right to use violence by those in power [10].

The functions of legitimisation and prediction establish an intra-discursive connection between the past, present, and future, ensuring support for certain relationships in society [10].

As noted by V. Pavlutska, among the most important functions of political discourse are persuasion and political propaganda. "The main target of political discourse is to shape in society those opinions and assessments that are necessary for the discourse author. This type of discourse is primarily oriented towards public consciousness, thus being a sufficiently dangerous instrument of influence as it targets all layers of society. The specificity of political discourse lies precisely in implementing policies and achieving political goals, which are predominantly associated with the issue of power" [10].

The contentious nature of political discourse is emphasised by G. Seidel [70], who claims that political discourse must simultaneously fulfil many functions and embody numerous motives. For instance, a party program is intended to persuade, agitate, and propagate (persuasive function); meanwhile, there should be an evident

connection of each specific point of the program with the persuasive ideology of the given party or group, meaning that every demand, every action must be logically linked to the values, traditions, and ideology (argumentative function).

P. Chilton and C. Schaffner [33] identify four types of strategic functions of political discourse: 1) coercion; 2) resistance, opposition, protest; 3) simulation; 4) legitimisation and delegitimation. The more detailed characterisation of these four types of strategic functions is the following one:

1) Coercion function

"Coercion" is part of linguistic acts such as commands or laws that are backed by certain sanctions. Political actors often act coercively through discourse, defining issues, choosing topics in conversation, positioning themselves and others in certain relationships, making assumptions about realities that listeners must, at least temporarily, accept the perception of text and language. Power can also be exercised through control over the use of language by "others" - through various degrees of censorship.

2) Function of resistance, protest, and opposition.

This category of functions of political discourse is the opposite of the coercion function and is implemented in opposition to power relations. Discursive characteristics of this type can be both features of mass media, graffiti (among marginal ethnic groups, posters, etc.), and specific linguistic structures (petitions, requests, slogans, etc.).

3) Function of simulation.

This function is associated with control over information. It combines both the effects of defocusing "undesirable places" through the euphemising of original data and the direct inversion of original data.

4) Function of delegitimation and legitimisation.

Legitimation is closely related to the coercion function, as it implements the mechanism of submission, i.e., legitimacy. Techniques used include appealing to voters' desires, shared ideological principles, projection of a charismatic leader, positive self-presentation, etc. Delegitimation essentially represents opposition:

representing others (foreigners, "internal enemies," institutional opposition, unofficial opposition) carries a negative connotation. Techniques of this category include the use of ideas of distribution.

The renowned researcher D. Graber [44] identifies the following functions of political discourse.

1. Dissemination of information, associated with the creation of virtual reality in the recipients, where one's own empirical practice is eliminated through the use of proposed cognitive schemes, common beliefs. In addition, information dissemination may include implicitly connotative information (positively/negatively colored keywords, primarily based on political orientations and societal values, the use of patriotic symbols and euphemisms).

2. Establishment of the theme of political speech, where the dominant criterion is control over information dissemination. Speakers typically try to avoid including topics on the agenda that could lead to complex and disadvantageous situations when discussed.

3. Projection into the future and past, based on an awareness of past facts and forecasting the future. In political communication, appeal to past experience is often used as an argumentative tactic by politicians.

According to R. Denton and G. Woodward [36], there are contradictions in the functions inherent in political communication. The researchers believe that political communication serves as a cohesive link, often replacing physical violence and enabling changes in society towards order, paving the way for compromises, and making facts and arguments accessible to the wider public. At the same time, it serves as a language of factionalism (division), dividing society into friends and enemies. It can exacerbate differences to an irreversible level or, conversely, reconcile them. The ability of political communication to transform society for the better is optimistic, but widespread abuse of it can lead to despair. Therefore, political rhetoric is multifaceted: it can inform, inspire, reassure, divide, and sow hostility.

Analysis of works dedicated to the study of the nature of political communication allows for the identification of the following functions within the instrumental function of political language [23], [34], [28]:

1. Social control (creating conditions for the unification of behaviour, thoughts, feelings, and desires of a large number of individuals, i.e., manipulating society's consciousness);
2. Legitimation of power (explaining and justifying decisions regarding the distribution of power and societal resources);
3. Reproduction of power (strengthening allegiance to the system, particularly through the ritualistic use of symbols);
4. Orientation (formulating goals and problems, shaping the political reality in the consciousness of society);
5. Social solidarity (integration within society or specific social groups);
6. Social differentiation (alienation of social groups);
7. Agonistic (initiating and resolving social conflicts, expressing disagreement and protest against the actions of authority);
8. Actional (implementing policies through mobilisation or "narcotisation" of the population: mobilisation involves activating and organising supporters, while narcotisation refers to the process of calming and diverting attention, dulling vigilance);
9. Information dissemination (reporting on the state of affairs in politics);
10. Agenda-setting (for example, bringing certain issues to the forefront of societal attention, controlling the dissemination of information);
11. Projection into the past and future (forecasting and reconstructing the past as a method of argumentation).

We consider this classification to be the most complete and representative.

Thus, political discourse occupies a significant portion of public discourse with the aim of presenting facts, opinions, and attitudes of speakers towards a particular topic. However, it is also necessary to note that the consequences of political communication do not occur if the audience does not perceive the information and is

not influenced. For this reason, it is essential to consider the content of the message, historical context, or even the political environment, as well as the position of the audience.

1.4. Corpus-assisted discourse analysis

Corpus-assisted discourse analysis (CADS) is a relatively new methodologically oriented approach to the qualitative analysis of naturally occurring spoken and written language. Unlike traditional forms of discourse analysis based on introspective data, its empirical orientation places CADS at the intersection of discourse analysis and corpus linguistics. Although CADS is by no means the only method for analysing discourse using corpus data [60], what sets CADS apart is its explicit commitment to the integration of qualitative analysis with the insights provided by corpus linguistics and the systematic study of language patterns. This means that despite the highly diverse and at times incommensurable nature of the various forms of discourse analysis and corpus linguistics, CADS must seek to reconcile the differing methodological and theoretical assumptions of these disparate fields. As a result, much of the basic research and development in CADS thus far has been concerned with developing a viable interface between discourse analysis and corpus linguistics.

One of the fundamental questions that arises in the context of using corpora in CADS concerns its size and selection method. The least frequently utilised approach within CADS is conducting research on general language corpora. For example, Hamilton, Svenja, and Brigitte [47] reconstruct the meaning of the word "risk" based on general corpora of British and American English. Maunter [59] focuses his attention on issues with a stronger social connotation, analysing the functioning of the word "elderly" in a selected sub-corpora of the Bank of English. In turn, Orpin [63] compares discursive constructions related to the use of two English words describing dishonest behaviours of public figures: "corruption" and "sleaze." He does this based on the analysis of their occurrences in a sub-corpus of British press within the Bank of English corpus. The common feature of the described studies is the focus on the semantics of the examined words - a detailed analysis of collocations or contexts in which a given word appears provides a very good insight into its meaning and functioning in the

language but leaves a certain dissatisfaction regarding more socially or critically oriented conclusions.

A significantly more popular trend among researchers in CADS is the utilisation of specialised corpora. They can be divided according to two criteria:

1. the variety of language they contain: written versus spoken;
2. the portion of the population being studied: sampled versus complete.

Most commonly, corpora of written texts are subject to investigation, constituting a specific sample of all texts that can be considered relevant to the research issue. However, in many cases, researchers opt for such a detailed formulation of the research problem that allows them to encompass all texts relevant to the issue at hand.

Only a few researchers focus on the analysis of speech: in such cases, transcripts of audio recordings are subject to investigation, either from natural communicative situations [49] or from research conducted using social science methods such as interviews or focus groups [77]. One of the advantages of this type of research is the focus on the most primary and spontaneous form of language, which is speech. It can be assumed that the relatively low utilisation of such corpora in CADS is determined by numerous technical and organisational challenges associated with their creation.

On the other hand, among the examples of research conducted on populations, one can find analyses based on full texts of parliamentary debates concerning a specific topic, for example discourse of same-sex relationships [19], official documents of selected institutions, like military reports on the Kandahar "friendly fire" incident [17], or media texts covering a narrow time frame around an event deemed socially significant by the author [50]. Utilising a corpus encompassing the entire population of texts ensures a high level of accuracy in the presented results. However, on the other hand, it necessitates a focus on narrow issues, the significance of which for the goals set by CADS may raise some doubts.

The application of corpora in discourse research requires the use of specialised computer software primarily due to the size of the analysed research material, as well as the statistical procedures used in various techniques. The most popular software used by researchers includes the free AntConc [18] and the licensed WordSmith Tools [69].

Both allow work on corpora with diacritical marks, calculation of frequency lists, keyword lists, and collocation lists for adjustable settings, as well as exporting results to selected programs. In search of more advanced tools, it is worth noting those available within Wmatrix, developed by a team led by Paul Rayson [65].

1.5. Basic corpus techniques used in critical discourse analysis

Frequency list analysis

A frequency list is defined as "a list of all the words appearing in a corpus along with their frequency of occurrence and their percentage share in the corpus" [20]. It is automatically generated by software used for corpus analysis, such as the AntConc or WordSmith Tool. It also serves as the basis for more complex analytical tools, such as keyword analysis described below. In itself, it can be considered the most basic tool in corpus-assisted discourse analysis (CADS).

Generating a frequency list provides the researcher with information about all the words appearing in the texts under analysis, as well as the frequency of their occurrence. Such data has been utilised, for example, by Chen [30] in research on the impact of political changes in China on the press, for example on the English-spoken newspaper China Daily. This impact is observed through the frequency of positively and negatively charged, as well as neutral, comparative words in a diachronic corpus of newspaper articles. On the other hand, Mohamad et al. [61] used frequency lists created for two sub-corpora of mathematics textbooks to compare gender representation in English-language textbooks published in Qatar and outside its borders. Analysing the frequency of selected pronouns, job titles, or kinship terms enabled the authors to answer questions about the level of sexism in both textbook groups. The findings suggest that while language is not inherently sexist, it can reflect societal biases and attitudes. The analysis revealed linguistic features that conveyed sexism, although gender bias favouring males as dominant figures was not statistically significant. However, the textbooks portrayed females in secondary roles to males, particularly in foreign textbooks.

The potential of the frequency list as a tool supporting the comparison of different discourses was also utilised by Edwards [37] in research on the discourse of

the British National Party (BNP). Comparing the frequency of individual words in BNP manifestos from 2005 and 2010 allowed him to select node-words. These are words or terms that serve as important reference points within a network or graph representing relationships between words or concepts. The node-words play a crucial role in capturing the semantic relationships between concepts, which can be utilised for tasks such as information retrieval, question answering, and knowledge representation. The frequency of the node words in the study significantly changed in the analysed manifestos, and whose contexts of occurrence were then subjected to further qualitative analysis. In the case of frequency list analysis, the software is used to generate them, while the selection of lexical units subjected to further analysis, as well as the interpretation of the obtained frequencies, remain the responsibility of the researcher.

Keyword analysis

Keywords are defined as "points around which ideological battles are fought" [72]. Meanwhile, Mautner [59], citing Williams, points out the validity, and even the necessity, of being interested in keywords, as words should be seen as part of problems. However, these statements do not provide clear guidance on how to determine what is a keyword for a given text, discourse, or problem. Two approaches are encountered here: the first is based on knowledge of the issue being addressed, which the researcher possesses. Based on their significance, the researcher selects keywords, the occurrence and function of which in the text will later be subjected to further analysis [35]. The second way of searching for keywords is more rooted in the tradition of corpus linguistics, according to which a keyword is defined as a word that appears significantly more frequently in one corpus than in another (referred to as a reference corpus) [24]. To calculate such defined keywords, the frequency list of the analysed corpus is compared through the mentioned software with the frequency list of the reference corpus. Then, using statistical measures of keyness, a list of words that occur statistically significantly more frequently in the analysed corpus than in the reference corpus is generated.

In such a procedure for determining keywords, the issue of selecting the reference corpus becomes particularly important, as the result obtained will depend on

its composition. In this context, Bondi and Scott [24] point out that it is rather difficult to speak of a single thematic scope of the text, from which information is obtained by analysing keywords. They consider it more accurate to assume that there are many different thematic scopes of a given text, which can be discovered by comparing it with various reference corpora. Thus, when seeking general guidelines regarding relevant concepts in a given set of texts, researchers usually rely on general corpora for the respective language. However, when a specific context becomes important due to the phenomenon being analysed or the research question posed, reference corpora prepared specifically for the needs of the study are used.

The second of the described approaches provides a focus on expressions characteristic of the discourse under study, rather than, for example, the genre itself: this is utilised by Lukac [58], comparing the blogs he examines with a corpus of blogs on different topics. In this way, he avoids concentrating solely on phenomena characteristic of blogs as an internet genre. Such a strategy also proves very helpful in revealing differences in the discourses of different groups on the same topic. For instance, Subtirelu [74] points out various approaches to non-English speakers among supporters and opponents of the multilingual ballot law. He is able to reveal them through the analysis of keywords obtained by comparing the corpus of statements by supporters of the law with the corpus of statements by its opponents. Baker [20] employs a similar strategy, comparing the voices of supporters and opponents of the fox hunting ban in a parliamentary debate on the issue. This allows him to reveal, among other things, the discursive connection between fox hunting and British identity present in the debate among supporters of hunting, or the tendency toward literal and blunt descriptions of the effects of hunting among its opponents. Meanwhile, using the FLOB corpus of general English as a reference corpus and further analysing the keywords obtained in this way, Weninger [77] identifies the category of facilitated agency, characteristic of American discourse on urban revitalisation, breaking with the classical division of imaging key participants in social situations as passive or active. The examples mentioned above indicate that the mere list of keywords does not answer the question regarding the discourses present in the material under study. Further

analysis is necessary, which involves assigning keywords to thematic groups [41] or detailed analysis of the contexts in which they occur in search of patterns [77]. Therefore, similar to the case of frequency lists, a computer program only generates a list of words, and its further processing belongs to the researcher. However, keywords provide an important indication directing researchers' attention to phenomena characteristic of the discourse they are analysing. They can thus serve as a kind of point of entry into the collected data.

Collocation analysis

Collocation is defined by researchers in corpus linguistics tradition as frequent co-occurrence [72]. For a chosen word, its span is determined, and then with the help of software utilising specific statistical measures, a list of words significantly more frequent within a specified span from the base word is generated. After obtaining such a list, researchers proceed to a more qualitative analysis focused on meaning. Essentially, two different models of procedure are encountered here:

Focus on collocations themselves - words from the obtained list are grouped into thematic clusters, allowing the determination of the functioning of the key concept in discourse by reconstructing the most important domains it is associated with or occurs in the vicinity of. For example, Freake et al. [40] adopt this strategy in their research on the identity of Quebec residents. Collocation analysis proved valuable in the mentioned study by providing a systematic way to identify and categorise the linguistic patterns surrounding keywords related to the discourse on nationhood and belonging in Quebec. Researchers could classify collocates of significant words like "QUEBEC" into attributes of nationhood. It allows them to establish, for instance, that for French-speaking residents of the province, Quebec is primarily associated with the nation understood as a historical community.

Contextual analysis containing the obtained collocations involves generating lists of their occurrences in the text along with the nearest context, and then analysing the obtained fragments in search of patterns. For collocation analysis, the researchers chose the term "Europe" specifically, as opposed to terms like "Union" or "EU," because it offered greater speculative potential surrounding the historical, cultural, and

geopolitical dimensions embedded within its discourse. Such actions lead Forchtner and Kolvraa [38] to establish that crucial for the construction of European identity in the material they studied are commonly shared values, which are placed in the context of a common difficult past.

Another significant issue related to the use of collocation analysis is the selection of words for which these collocations are to be determined. Here, researchers also employ various techniques of Corpus Linguistics utilised in CADS. Subtirelu [74] seeks collocations for previously established keywords. This allows him to focus on the construction of concepts characteristic of the discourse he is studying. On the other hand, Lischinsky [56], due to her interest in the construction of economic crises in the press, decides to use lexicographic sources (dictionaries and FrameNet database) in search of words used to describe the crisis. In her work Mautner [59] determines words that are significant in her opinion for the concept of the entrepreneurial university and focuses her attention on analysing their collocational profiles.

An example of a more theoretically grounded application of collocation can be found in Mulderrig [62], who codes the collocations she examines according to action types proposed by Halliday and Matthiessen [46]. This allows her to identify New Labour's Party discourse subtly enacting hegemony through defining social subjects and structuring power relations, particularly in education policies of the UK.

In summary, collocational profiles generated using linguistic analysis software are used as a basis for identifying general patterns that are the focus of the researcher's interest. They most often serve to obtain more detailed information about the functioning of specific words in the analysed texts. Based on such information, the researcher later identifies discursive structures. Collocation analysis, especially when conducted on a large corpus of general language, also allows for determining how the examined word and the associated concept function socially.

Semantic prosody

One particular use of collocation lists is in the analysis of semantic prosody. This concept was proposed by Louw [57], who defined it as the "consistently recurring aura of meaning that a given word is imbued with by its collocations". The semantic prosody

for a selected word is thus determined by the researcher based on the assessment of the connotation (most often on a positive-negative axis) of its strongest collocations. The concept itself has been subjected to criticism [78] regarding, for example, the possibility of transferring connotations from one lexical unit to another or the type of connotations that should be analysed. Despite this criticism, semantic prosody analysis is used in research within the Corpus Linguistics tradition. For example, Mautner [59] indicates that the word "elderly" has a much stronger negative semantic prosody when used as a noun (an elderly person) than when it appears as an adjective (older). In this case, the semantic prosody was determined based on the collocation profile of a keyword with a positive or negative connotation in the context. Semantic prosody analysis is also utilised by Hamilton, Adolphs, and Nerlich [47] in their research on the meaning of the words "risk" and "to risk" by exploring concordance data around these words, which was divided in five categories. It was found that "risk" most commonly has negative semantic prosody, because of its usage around health issues and property loss, even when it's used in different contexts.

Semantic preference

Another way of more complex utilisation of collocation lists within Corpus Linguistics tradition is the analysis of semantic preference. It is defined as the tendency of a specific lexical unit to frequently collocate with a series of units belonging to one semantic field [68]. Similar to semantic prosody, preference is determined based on a list of collocations. Yet, in this case, the researcher assesses not the connotation of individual collocations but their belonging to specific thematic groups (semantic fields). The semantic preference for words like "defensive," "collective," and "limited" with "jihad" suggests a specific conceptualisation of jihad as localised and bounded, rather than expansive or global in nature.

Among studies in which the analysis of semantic preference leads to the identification of social constructions of selected phenomena, we can mention the aforementioned work by Hamilton et al. [47], where preference analysis is used to identify dominant discourses related to risk. This leads, for example, to the observation of a tendency to evaluate the perception of the term "risk" in the corpus under study.

Similarly, Salama [68] establishes, through the analysis of collocations, that the representation of Wahhabism in one of the sources he examined is associated with threat, conspiracy, and the concept of a police state. Thus, focusing on thematic groups to which collocations of words describing phenomena studied by the researcher belong leads to revealing ways of representation characteristic of the analysed discourse.

One weakness of this analysis may be considered the reliance on the classification into specific semantic fields solely based on the researcher's subjective view (there is neither a single list of semantic fields used by researchers nor an algorithm allowing for assigning words to such fields in a manner characterised by a high level of intersubjectivity) [52].

Concordance analysis

While all the techniques described above generally support qualitative analysis or provide the researcher with clues on where to start such an analysis or what it might focus on, on the other hand, concordance analysis can be considered the closest to classically understood quantitative discourse analysis. Concordance itself is defined as "a list of all occurrences of a search term in a corpus, presented along with the context in which the term appears" [20]. The length of the context is determined by the researcher and measured in terms of the number of words or characters (depending on the software used), and the concordance list is generated according to selected parameters by the software used.

One of the advantages of concordance analysis is the ability to examine the context of occurrence of significant words in relation to the research question even in very large text corpora. The ability to automatically generate a list of all occurrences of a selected word along with its immediate context significantly shortens therapy (there is certainty that all occurrences of the given word have been analysed) analysis process and increases its accuracy and repeatability (different researchers for the same corpus will always obtain the same concordance list). It allows one to gain a deeper understanding of how the word is used in different contexts. Searching for patterns among the obtained concordances leads researchers to identify dominant discourses and ways of discursive representation of the analysed phenomena. However, a drawback is the focus on the word/words themselves rather than the issues that may be discussed in the text without using the searched word (e.g., using pronouns) [52].

1.6. The advantages of applying corpus methods in CADS

The use of corpus methods in research within the field of Corpus-Assisted Discourse Analysis (CADS) entails significant changes in this research domain.

Increasing the number of analysed data enhances the accuracy of the obtained results. Studies based on tens or hundreds of thousands of words from newspapers allow for a more comprehensive answer to questions concerning, for example, immigrant representation than analysing a few articles. Furthermore, the ability to work with a relatively large corpus of data gives researchers the opportunity for a more

balanced selection and enables the application of complex algorithms for text selection. This reduces the risk of focusing on analysing texts chosen solely by intuition. Moreover, one of the criticisms raised against CADS is precisely the lack of representativeness of the analysed texts [72] or the reliance on personal biases in their selection [25]. The use of large datasets partially addresses this criticism.

Furthermore, increasing the level of transparency of the applied procedures enhances the possibility of replicating conducted research. It is also significant in the context of comparative research: collocation profiles or keywords calculated using specific methods for specific corpora in one language or time period can be compared with data obtained in a similar manner for another language or time period. Apart from that, the transparency of methods leading to the obtained results can increase their credibility level and thus make them more convincing. Similar to the case of analysed data, the lack of clear methods is one of the problems pointed out by critics of CADS [25]. The application of selected corpus techniques provides a partial solution to this problem.

Using quantity-based pattern lists of keywords, concordances, or collocations allows researchers to focus on the most frequently occurring words or concepts in texts. This provides a starting point for further analysis, grounded in the data itself rather than the researcher's intuition or knowledge. Recipients of research results thus gain clarity regarding the criteria for selecting words that have undergone further analysis, facilitating understanding and interpretation of the presented results. This can also lead to a reduction in the level of bias in the obtained results – the researcher's attention is directed towards issues most frequently appearing in the texts under investigation, rather than those that intuitively seem most significant [25].

The capability obtained through software and corpus techniques to observe and interpret quantitative patterns in the analysed texts can facilitate the uncovering of hidden and non-obvious trends present in the material under analysis. Collocations of selected words or the repetition of certain modes of representation observed through concordance analyses allow attention to be focused on regularities present in a relatively large dataset. Shifting the emphasis from what is repeatable in the

researcher's opinion to what is repeatable in terms of frequency of occurrences can also be a step towards increasing the intersubjectivity of the conducted analyses [25].

1.7. The disadvantages of applying corpus methods in CADs

The assumption of analysing large corpora of texts can modify the scope of research topics undertaken. It directs the researcher's attention to texts that are easily accessible in electronic form and that allow for the compilation of relatively large datasets. This tendency is evident, for example, in conducting a significant portion of the analyses discussed on newspaper texts [54], [41] or transcripts of parliamentary debates [19], [74]. In this way, questions about discursive constructions present in less accessible utterances, such as influential priests' sermons, speeches by local community leaders, or lectures by teachers in schools and universities, are omitted. Concentrating on large electronic text resources can thus limit the research field in two ways: on the one hand, it limits the sources used to seek answers to research questions. On the other hand, it also limits the questions posed to those for which answers can only be obtained through the analysis of existing electronically available texts.

In turn, focusing on quantitative patterns prompts the question of whether the most frequent implies the most important. There is a lack of convincing evidence that frequency of occurrence in discourse is the strongest factor influencing the societal and individual representations of a chosen phenomenon. For instance, it can be speculated that certain utterances, either due to their authors or the nature of the utterance itself (e.g., particular expressiveness), may have a stronger influence on the discursive constructions present among recipients.

Also worth noting is the problem of multiple variables that influence the results obtained using corpus techniques. Keyword lists depend on the accuracy of selecting both the main corpus and the reference corpus, as well as on the adjustment of the statistical measure used. Additionally, the obtained collocational profiles depend on the coefficients used for their calculation, and their reliability as a source of information regarding discursive constructions also depends on the accuracy of selecting the words for which they are calculated. A minor error or inaccuracy appearing at any stage of

using corpus techniques can lead to skewed results that may be difficult to detect both by the researcher and the recipients of their research [52].

Another problem, also related to the corpus techniques themselves, is the presence of numerous assumptions in them, often implicit and frequently overlooked in interpretations of results. For example, the use of keyword techniques leads to a focus on what is different and overlooks what is similar. The analysis focuses on words that differ most in frequency in the studied corpora, while omitting those that occur very frequently in both the main and reference corpora. Concordance analysis for specific words is associated with the reconstruction of how a chosen word functions in discourse, rather than the concept it represents, which may also be expressed using other words. Additionally, it leads to the assumption that the immediate context of a given word is the most significant and decisive factor in the research results. The choice of collocations as a tool for examining discourse assumes that only the frequency of co-occurrence is relevant, rather than other characteristics of word combinations. Each of these assumptions can be questioned and each influences the scope and manner of interpreting the obtained results [52].

Most of the problems mentioned above can be addressed through careful and precise corpus construction, as well as by considering possible doubts in the development and interpretation of results. However, their omission can lead to a dangerous situation: the results obtained using corpus methods create the impression of being very reliable due to the amount of data on which they are based, as well as the limitation of the role of the researcher-interpreter in favour of using statistical measures.

1.8. The notion of linguistic personality

The term "linguistic personality" is the core and system-forming philological concept. It is considered in three directions:

- Linguodidactic (i.e., the regularity of language learning);
- Literary (study of the peculiarities of language in fiction);

- Psycholinguistic (the relationship between language and thought, the characteristics of perception, understanding, and the influence of speech).

Purely linguistic research in this field is dedicated to the development of models of linguistic personality [4], analysis of the concept, studying the peculiarities of linguistic personality as a bearer of national language and culture, and so on.

In Ukrainian linguistics, scholars such as S. Yermolenko, V. Melnychayko, M. Pentyliuk, and L. Matsko have devoted a series of works to the study of the essence and problem of linguistic personality.

M. Pentyliuk [11] emphasises that linguistic personality manifests itself in speech. That is, the personality itself is formed under the influence of various psychological processes. The researcher also notes that linguistic personality can be considered from the perspective of national-cultural specificity, taking into account ethno-psychological and ethno-linguistic factors. This, in turn, develops the national component of personality.

This means that based on the analysis of a person's speech activity who belongs to a particular nation, characteristic features of the entire nation can be determined. These features manifest themselves in the way of thinking, self-awareness, character, mentality, and motivation.

Also interesting is the approach of L. Matsko [8] - a linguist who argues for the formation of linguopersonology as a science that studies linguistic personality, factors of its formation, and types.

The researcher believes that in speech, the concept of "linguistic personality" is embodied through background knowledge, linguistic mode, and behaviour. Linguistic personality relates to a certain time-space continuum. This is reflected in the relevance of ideas, used concepts, embodied meanings, images, stereotypes, and linguistic behaviour.

According to L. Matsko [8], the components of linguistic personality are:

- Linguistic ability;
- Linguistic consciousness;

- Linguistic communicative activity.

The researcher emphasises that linguistic personality is a specific individual who expresses themselves in and through language.

According to A. Zahnitko [4], linguistic personality is analysed in terms of abilities and characteristics, as well as the individual's understanding of communication products, which differ in the following:

- Degree of linguistic-structural complexity;
- Depth, consistency, and accuracy of reflecting reality;
- Communicative orientation;
- Aesthetic load;
- Psychological effect.

Considering linguistic personality from this perspective, A. Zahnitko [4] proposes the following definition: it is a combination of a person of linguistic competence, a desire for creative self-expression, and the free realisation of speech activity.

Some scholars are also inclined to believe that the integral nature of the concept of linguistic personality is the cause of the synthesis of modern linguistics, interdisciplinary integration of knowledge, and the expansion and increase of the research object. This also applies to the concept of discourse in which linguistic personality is situated.

L. Slavova [15] emphasises the special interest in examining linguistic personality in political discourse, as it has a significant impact on public opinion. According to her, political language is an inexhaustible source of information about it. Language itself is part of the character that defines the politician as a personality and is filled with symbols.

Linguistic personality in political discourse is a phenomenon through which politicians create their unique image and influence society. In political discourse, the temperament of the personality is reflected, the ability to perform certain types of

activities (including communicative), dominant feelings and motives of activity, individual psychological characteristics, and so on.

By demonstrating their linguistic personality, politicians create a special image that sets them apart from others and provides a powerful basis for research in the fields of sociology, psychology, linguistics, philology, and accordingly, translation studies.

Some linguists also equate linguistic personality with a politician's individual linguistic style. Therefore, when studying the peculiarities of political rhetoric, speech, and the challenges of translation, one cannot ignore the concept of idiolect.

The term "idiolect" is not used often; a more well-known and straightforward term is *idiostyle* or *linguistic portrait*. In addition to certain external, psychological, ethical characteristics, and signs, every player in the political arena of any state has their own linguistic portrait.

V. Voloshuk [2] explains idiolect as "a system of meaningful and formal linguistic characteristics typical of the works/speeches of a particular author. It demonstrates the uniqueness of the author's way of linguistic expression in their works."

S. Yermolenko [3] asserts that the systematicity of an individual's linguistic style is based on the connection between language and thought, on the formation of a linguistic worldview that combines the general and the individual, the general and the singular. It is in the individual linguistic worldview that the author's, personal, embodied in an individual writing style, is combined with the common and typical. Most often, it is the individual style of the speaker that deserves careful attention, as it reflects not only the individual characteristics of the speaker's rhetorical style but also the cultural and national-specific characteristics of linguistic personality.

On the current world political stage, there are many prominent, original personalities whose individual traits are reflected in their political speeches. Therefore, the term "idiolect of a political figure" has emerged in political linguistics.

Thus, linguistic personality is the subject of study in various fields of science. Because of this, there is no clear definition of the concept, and therefore there are different approaches to analysing personalities.

The linguistic personality in politics is also subject to debate. In this context, it represents a combination of various attributes of politicians that give rise to neologisms and unique expressions.

Public speaking encourages the oratorical creativity of political leaders, as it is always relevant and sometimes unpredictable. Political speeches are often highly journalistic, as they contain statements of certain ideas, views, directions, ideological positions, inevitably involving the alienation of other perspectives, even if they are not explicitly mentioned.

Therefore, the linguistic personality of a politician forms the basis of their political rhetoric. It must constantly evolve and improve. Expressing ideological positions, especially to a large audience, always requires rhetoric, which can be expressed in various ways: successful, "striking" paraphrases, comparisons and contrasts, epithets and metaphors, moral-ethical antitheses, firm refutations, repetition of certain ideological themes, hyperboles, emphasising particular words and phrases suitable for any situation, occasional neologisms, rhythmic syntax, and expressive intonation. These are the techniques that politicians use in their speeches. And it is these techniques that create their linguistic personality, which is shaped by the character, image, style, and communication peculiarities of politicians.

1.9. Rishi Sunak as a political figure

Rishi Sunak was born in 1980 in Southampton, Hampshire, the eldest of three siblings. He is of Indian and Afro-European heritage. Sunak attended the prestigious public school Winchester College, where he was head boy. This way Rishi Sunak is the middle-aged representative belonging to the upper class ethnic minority. He subsequently read Philosophy, Politics, and Economics at Oxford University; his expenses being paid partly by a student loan. Sunak then attended Stanford University in California where he obtained an MBA. After two years at Goldman Sachs, Sunak became director of investment firm Catamaran Ventures, owned by his father-in-law, Indian businessman N. R. Narayana Murthy. Sunak's mother Usha was also involved in a pharmaceuticals business. Sunak's entry into politics was influenced by a meeting with the then-Conservative leader William Hague at an International Studies reception

during his time at Oxford University. This prompted an interest in global politics, leading him to join the Conservative Party in the early 2000s.

Chancellor of the Exchequer

On February 13, 2020, Sajid Javid stepped down from his role as Chancellor after a proposed takeover of the Treasury led by Dominic Cummings and supported by Downing Street. Rishi Sunak was selected as his replacement.

This transition occurred alongside the emergence of the COVID-19 pandemic, which has since been a defining aspect of Sunak's tenure at the Treasury. The challenges posed by the pandemic, coupled with Sunak's subsequent rise in prominence, have significantly boosted his public and media presence.

Sunak's inaugural budget occurred on March 11, 2020, incorporating an additional £30 billion in spending, of which £12 billion was earmarked for mitigating the economic repercussions of the COVID-19 pandemic.

On March 17, Sunak unveiled a £330 billion emergency 'furlough' initiative to aid businesses and workers impacted by the pandemic. These furlough disbursements have undergone multiple extensions and reintroductions during Britain's three lockdown phases.

Furthermore, the government's 'Eat Out to Help Out' program, which subsidised food and non-alcoholic beverages at participating eateries by 50% up to £10 per person, provided another platform for Sunak to enhance his visibility. The meticulously orchestrated marketing campaign prominently featured Sunak. Overall, the scheme subsidised £849 million worth of meals.

In 2022, Rishi Sunak's standing as Chancellor experienced a substantial decline. Amidst a cost-of-living crisis, he faced backlash regarding his wife's tax matters and the revelation of his recent relinquishment of his American 'green card' visa.

Additionally, in April 2022, Mr. Sunak received a fixed penalty notice for attending a lockdown gathering in the Cabinet Office with the prime minister. Consequently, he stepped down from his government position in July 2022.

In summary, Rishi's impact as Chancellor of Exchequer was shown by his economic response to the pandemic which includes debt management, support of

businesses and individuals, and tax policies. Due to the difficulties during this period he was not particularly popular among self-employed people. Sunak's speech during this period is likely to feature a wealth of economic vocabulary and discussion related to the COVID-19 pandemic [64].

Election of the head of the Conservative Party

On July 8, 2022, Sunak announced his candidacy for the leadership elections of the Conservative Party, aiming to replace Johnson. Sunak launched his campaign with a video posted on social media, stating that he would "restore trust, rebuild the economy, and unite the country." He declared his values as "patriotism, fairness, and hard work." The domain readyforrishi.com was initially registered with GoDaddy on December 23, 2021, and ready4rishi.com on July 6, 2022, just two days after Sunak resigned from the position of Chancellor. The first domain serves as a redirect to the second one. Conservative politicians who supported Johnson criticised Sunak for "leading accusations of ousting the prime minister," and Johnson's key ally, Jacob Rees-Mogg, labelled him "the highest tax chancellor."

Sunak and Foreign Secretary Liz Truss became the final two candidates in the contest on July 20, who were put forward for the ultimate vote to elect the leader. However, among the Members of Parliament, he garnered the highest number of votes, with Sunak receiving 137 votes compared to Truss's 113 in the final round. In the vote among all party members, Truss secured 57.4% of the votes, making her the new leader of the Conservative Party.

Sunak's campaign promises included tax cuts only when inflation is under control, the cancellation of the 5% VAT rate on household energy for one year, the introduction of a temporary fine of £10 for patients who miss appointments with general practitioners, and limitations on the number of refugees.

During the campaign in July 2022, a clip from the 2001 BBC documentary "Middle Classes: Their Rise and Sprawl" surfaced, in which he remarked: "I have friends who are aristocrats, I have friends from the upper class, I have friends who, you know, the working class, but... well, not working class." Sunak commented on the clip, saying, "We all talk nonsense when we are younger." In a video appearing in August

2022, where Sunak addressed an audience in Tunbridge Wells, Kent, he stated that he had changed funding formulas that "shove" money into "deprived urban areas," "to make sure that such areas receive the funding they deserve." Sunak responded that he wanted to "raise the level everywhere," not just assist "very large urban cities." After Liz Truss's victory in the 2022 Conservative Party leadership elections, Sunak faded into the background.

After Liz Truss's resignation on October 20, 2022, Sunak was considered a potential contender to participate in an accelerated leadership contest alongside Penny Mordaunt and Boris Johnson, who had stepped down as Prime Minister before Truss's appointment. On October 22, it became known that Sunak had the necessary number of supporters—100 members of the House of Commons—to participate in the vote on October 24. Tobias Ellwood wrote on Twitter that he "had the honour of being the hundredth Conservative MP to support #Ready4Rishi." The total number of MPs publicly declaring their support exceeded 100 by the second half of October 22. On October 23, Sunak announced his candidacy for the elections.

Sunak was supported by several members of the Cabinet and prominent party members, such as Lord Frost, Kwasi Kwarteng, Nadhim Zahawi, Sajid Javid, Jeremy Hunt, Tom Tugendhat, Suella Braverman, Michael Gove, and Dominic Raab [73].

Prime Minister of the United Kingdom

On October 25, the day after becoming the leader of the Conservative Party, Sunak was appointed as the Prime Minister of the United Kingdom by King Charles III. Sunak's appointment also marked him as the youngest Prime Minister since Robert Banks Jenkinson, 2nd Earl of Liverpool.

He started his time as Prime Minister by choosing new members for his cabinet. His initial pledges centred on restoring economic stability following Truss's drastic tax cut policies, which eroded investor confidence in the UK.

He acknowledged the necessity of making difficult decisions, especially concerning taxation and debt reduction. He pointed out: "The government I lead will not leave the next generation, your children and grandchildren, with a debt to settle that we were too weak to pay ourselves" [67].

Sunak has implemented several alterations and reversals of previous policies, deviating from some of the commitments he made during the summer competition. However, he remains dedicated to upholding promises outlined in the 2019 Conservative manifesto, notably the agenda of levelling up to address economic disparities across the UK.

Among his initial actions were the reinstatement of the fracking ban as pledged in the manifesto and the implementation of tax hikes across various sectors. His inaugural budget statement was due for November 17, 2022.

In summary, Rishi Sunak is a prominent political figure in the United Kingdom, known for his tenure as Chancellor of the Exchequer and subsequent rise to the position of Prime Minister. His political journey began with his entry into politics, where he quickly ascended through various roles within the Conservative Party. Sunak's tenure as Chancellor was marked by significant economic challenges, including the COVID-19 pandemic, during which he implemented extensive economic support measures. His unique path in career should provide useful insight into British political discourse in his speeches. It would be interesting to discover peculiarities the politician employed to talk about a variety of issues, such as budget allocation or COVID-19 response strategy [73].

CHAPTER 2. METHODOLOGICAL FOUNDATIONS FOR CREATING AND PROCESSING OF THE CORPUS

2.1. The use of content analysis method

Content analysis is usually used in the linguistic domain in order to identify specific words, expressions or concepts. These are usually found together with the qualitative data for the analysis. With the help of content analysis one can examine the occurrence, significance, and connections of these particular expressions. For instance, one could analyse the use of the word *unemployment* in political campaigns. This way the connection with other words such as inequality, economy, laziness could be traced. This would further allow for uncovering the relationships and meanings behind these words to understand the core message of this or that campaign.

In our study, we will use content analysis to trace the transformation of Rishi Sunak's language strategies from the beginning of his career to the position of Prime Minister at the lexical, grammatical and stylistic levels, to analyse the context in which they are used, the connotations they carry, the purpose and effect they are aimed at, and the extent to which they influence the image of the exemplary British politician.

The analysis of each period will be divided into blocks, where the above levels will be considered. After analysing all the periods, the main trends in the use of certain language units will be highlighted and a comparative analysis will be carried out.

2.2. Data collection

Multiple sources were found for assembly of the corpus. The first one is Rishi Sunak's own YouTube channel. There is a large collection of politicians' speeches and other videos regarding various topics such as local policies, economy, responses to the current events etc. It would take a lot of time to listen to and document every necessary video in text format, but YouTube has a video-to-text option, which significantly simplifies the task.

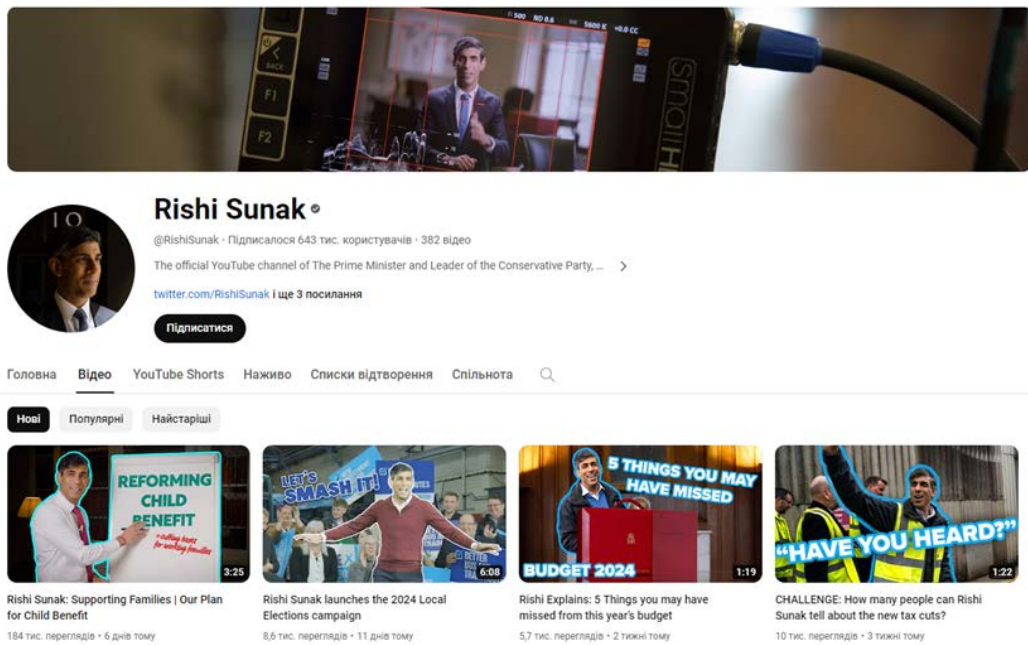


Figure 2.1. Rishi Sunak’s YouTube channel

Additionally, Rishi Sunak has his own web portal [66]. The website is designed to serve multiple purposes, such as providing information about Sunak's political platform, detailing his achievements and initiatives, offering news and updates on his activities, as well as a section for speeches and media appearances. Its landing page contains multiple sections: a small paragraph about the minister, a web form for communication, and multiple links to the topics of interest.



Figure 2.2. Rishi Sunak’s official website

An additional resource for accessing speeches by Rishi Sunak following his appointment as Prime Minister is the UK Government official website [43]. It is equipped with search functionality that includes multiple filters, providing aid in location of relevant articles. This portal represents a valuable tool for the compilation of the last part of our corpus.

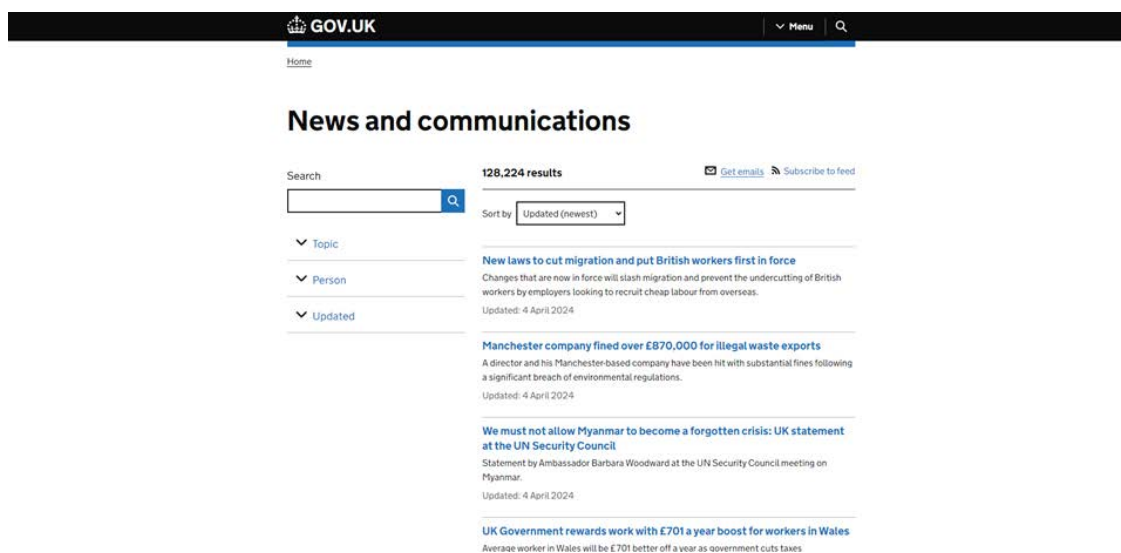


Figure 2.3. UK government official website

Occasionally, other sources were also checked, for example official Facebook and Twitter pages.

Manually searching through hundreds of pages is not only time-consuming but also prone to errors, making it an inefficient method for gathering information. Recognising these challenges, it was decided to implement a web scraping method to do this task automatically.

Web scraping is the automated process of extracting data from websites, web services, and web applications. This approach eliminates the need for manual data collection from individual websites, which is often a lengthy and inefficient task. The process uses automated scripts or programs that access a web page's structure, analyse the data present, and extract specific elements required for further analysis. Typically, web scrapers are short scripts written in languages like Ruby, Python, or PHP.

One of two technologies is almost always used by the web browser to display a webpage: the Hypertext Transfer Protocol (HTTP) protocol allows it to connect to the website and request resources, like documents, images, or videos. The language used to create websites is called HyperText Markup Language (HTML) [55].

HyperText Markup Language represents the foundational language employed in the creation and design of web pages on the internet. Introduced by Tim Berners-Lee in 1991 at the European Organisation for Nuclear Research (CERN), HTML was initially conceived as a straightforward markup language [80]. Over the years, it has undergone significant evolution, progressing through various versions from HTML 2.0 to HTML5, the latter being the most recent iteration as of 2024.

Usually, web scrapers attempt to read the HTML code of the page to extract necessary information. However, data that is manually extractable from a webpage may not always be accessible for automated scraping. Several factors can impede the process of automatic extraction, including:

1. HTML that is poorly organised or inconsistently constructed, sometimes as a result of using office or graphic editing software that is largely concerned with the page's appearance. Such coding techniques are commonly found on out-of-date government websites.
2. User authentication systems intended to prevent automated access. These could include points of entry that need a subscription or CAPTCHA verification systems.
3. The use of session-based processes to track user behaviour by utilising browser cookies.
4. The steps system administrators take to limit access to particular datasets.
5. Changing the way data is structured as a result of a website redesign. The most insidious thing is that such a redesign can be invisible to a human reader, but completely confusing to a crawler program. For example, for some reason, the site developers have started to present in a list what they used to present in a stream of named blocks. From the point of view of semantic ethics, they did the

best thing, but the crawler created for the previous design will not read anything from the page [55].

In the present research, three web scrapers were developed. The initial scraper is designed to aggregate articles from Rishi Sunak's official website. This process is defined in three distinct steps:

1. Collecting all hyperlinks to articles located within the media section of the website.
2. Individually accessing each hyperlink to extract the article's title, publication date, and content.
3. Archiving the acquired data on a computer in the form of text files.
4. Notably, certain pages might not present a direct transcript of speeches but rather offer a hyperlink to an external source, commonly YouTube. These hyperlinks are to be saved as well.

The second scraper will extract Rishi Sunak's speech transcripts from YouTube. It will use links extracted by the first scraper to access video pages and use previously mentioned website's video-to-text option to extract the transcript itself. It will perform in a similar manner:

1. Navigate to the YouTube video page using the previously saved links.
2. Programmatically activate the "Show text version" feature available on the page.
3. Extract the transcript and archive it on a computer as a text document.

The third scraper will work almost in the same way as the first one, however, it will work on articles from the UK Government website. All scrapers will be written using Python programming language.

Using previously mentioned web scrapers, a total of 74 speeches was extracted from which three groups were manually selected, which represent three prominent periods in Rishi Sunak's political career:

Table 1. Speeches from the first group (the beginning of the work in the government)

Date of publication	Title
11 June 2015	My maiden speech in Westminster
26 February 2016	Why I will vote for Britain to leave the EU
2 November 2016	Defending the rural economy
9 March 2017	Westminster Hall debate on the Coast to Coast Walk
19 April 2017	Supporting measures to help small businesses grow
6 November 2017	Transport for the North
27 November 2017	Budget debate

Table 2. Speeches from the second group (as chancellor)

2 October 2019	Party Conference Speech: big investment in our public services
29 July 2020	Summer Statement: Plan for Jobs
11 March 2020	Budget Speech 2020
3 March 2021	Full Speech: Budget 2021

Table 3. Speeches from the third group (as PM)

25 October 2022	Rishi Sunak's first speech as Prime Minister
17 April 2023	PM speech on improving attainment in mathematics
6 May 2023	Prime Minister's words to mark the Coronation of King Charles III and Queen Camilla
22 May 2023	PM statement to the House of Commons on G7 summit
5 June 2023	PM statement on illegal migration delivery update
12 June 2023	PM London Tech Week speech
30 June 2023	PM speech launching the NHS Long Term Workforce Plan
12 July 2023	PM statement to leaders at NATO Summit in Lithuania
13 July 2023	PM statement on public sector pay review
8 November 2023	PM statement on protests following meeting with Metropolitan Police Commissioner
15 January 2024	PM statement on defending the UK and its allies

2.3. The use of WordSmith Tools for corpus linguistics

There are several tools available for analysing large corpora, including AntConc, Sketch Engine, and WordSmith Tools, among others. AntConc provides basic corpus analysis features such as concordancing (context extraction for chosen words or phrases), keyword analysis (search of descriptive words about given text), and collocation analysis (searching for words which frequently appear together). It's a

lightweight tool suitable for quick text analysis tasks. Sketch Engine is a commercial tool which offers more advanced features including extensive corpus querying capabilities, detailed linguistic analysis tools, and support for multiple languages and specialised corpora. WordSmith Tools focuses on concordancing, word frequency analysis, and collocation analysis. It also generates wordlists and finds keywords.

Most of these programs offer similar functionalities, such as creating word frequency lists, generating concordances, and identifying keywords. WordSmith Tools was selected because it has a trial version freely available which provides all the necessary functionalities, along with multiple additional features like clustering, tokenisation, bulk analysis, and many others. Furthermore, this tool was chosen due to previous experience with the program.

WordSmith Tools was developed by Mike Scott in 1996 and new versions were being released up until 2021. This software is designed for Windows systems only. Start menu presents three main functions for corpus analysis: *Concord*, *WordList*, and *KeyWords* [69].

Concord is used to create concordances of a word or phrase. It finds context for the given input and sorts all results alphabetically. It is possible to query a single word or a phrase, also wildmarks are supported, giving the possibility to search for different forms of the same word. For instance, this is a concordance of word “spleen”:

N	Concordance	Set	Tag	Sent	Sent	Para	Para	H...	H...	Sect	S
				.#	.#	.#	.#	.#	.#	.#	
1	the technique of linking up the donor's spleen accomplished? ¶A thin			15	47	5	47	n/a	n/a	n/a	0
2	active functioning drops. ¶The donor's spleen braces up the organism,			23	12	6	...	n/a	n/a	n/a	0
3	¶In addition, it is akin to human spleen, having tissues that are			13	33	4	...	n/a	n/a	n/a	0
4	shield appeared to be futile. ¶Then surgeons tried to use the spleen as a natural filter. ¶Pumped through the live tissue, the blood remains as intact as if it were driven through the organism's own spleen. ¶The effect was marvellous. ¶The new method was suggested by Professor Anatoly Tsy-pin, head of department of the Research Institute of Transplantology and Artificial Organs. ¶Pig spleen			9	...	3	...	n/a	n/a	n/a	0
5	and Artificial Organs. ¶Pig spleen reacts most actively to any			12	4	4	...	n/a	n/a	n/a	0
6	tube connects the artery of the pig's spleen kept in a thermostat with the			16	53	6	53	n/a	n/a	n/a	0
7	futile. ¶Then surgeons tried to use the spleen as a natural filter. ¶Pumped			8	31	3	...	n/a	n/a	n/a	0
8	by pumping it through the spleen of a pig. ¶As is known, the			3	...	1	...	n/a	n/a	n/a	0
9	her blood was purified through the spleen of a pig, the girl was sitting in			1	83	0	...	n/a	n/a	n/a	0
10	the spleen of a pig. ¶As is known, the spleen plays a very important role in			4	17	2	17	n/a	n/a	n/a	0
11	discovered. ¶Attempts to transplant the spleen with a view to fortify the			7	27	3	27	n/a	n/a	n/a	0

Figure 2.4. Concord tool in WordSmith Tools

The Clusters feature in Concord allows searching for a specific cluster of words. For example, trying to locate verbs in the passive voice. The equivalent used in texts of this style and the transformation do not distort the semantic meaning of the word.

This is a very useful feature which can significantly help in analysis of speech patterns and writing styles, show important context about certain words and phrases, and reveal insights into frequency, context, collocation of language elements.

WordList is the second (but equally important) tool which generates frequency lists from a given corpus. Wordlists can be used in variety of ways, including but not limited to:

- just to examine the kind of words that are used;
- to spot word clusters that are frequently used;
- to obtain a concordance of one or more words in your list;
- to compare the frequency of a word in various text files or genres;
- to compare the frequencies of words that are cognate or have translation equivalents in various languages.

The Frequency display might look like this:

N	Word	Freq.	%	Texts	% Dispersion	Lemmas	Set
1	THE	172,107	6.08	480	100.00	0.99	
2	OF	80,383	2.84	480	100.00	0.99	
3	TO	76,462	2.70	480	100.00	0.99	
4	AND	69,677	2.46	480	100.00	0.99	
5	A	63,517	2.24	480	100.00	0.99	
6	IN	53,980	1.91	480	100.00	0.99	
7	THAT	37,605	1.33	479	99.79	0.97	
8	#	33,326	1.18	430	89.58	0.93	

frequency | alphabetical | statistics | filenames | notes

64,042 entries | Row 1 | 0% | T S | THE

Figure 2.5. WordList tool in WordSmith Tools

The top 8 categories in a wordlist derived from 480 texts are displayed in the section. For instance, the word “of” appeared 80,383 times in the 480 texts; the frequency column indicates how frequently each word occurred, and the percentage column indicates that the frequency accounts for 2.84% of the running words in those texts. Word "of" appears in 480 texts, or 100% of the texts used for the wordlist, according to the Texts column. With a dispersion value of .99, it is evident that the word is distributed rather evenly across the 480 texts in the corpus. The frequency list can be saved in a file to be accessed later, which is needed in the *KeyWords* tool.

The *KeyWords* tool compares two pre-existing wordlists, which must have been created using the WordList tool. One of these is assumed to be a large wordlist which will act as a reference file. The other is the wordlist based on one text which is to analyse. Keywords are not the most frequent words but the words which are most unusually frequent in the given corpus. Keywords usually give a reasonably good clue to what the text is about. Although frequently used, the term "keyword" lacks a definition in the field of linguistics. This software uses frequency patterns to mechanically identify significant words (A human being, though, might select a sentence as a keyword.).

A word is considered to be "key" if:

1. It appears in the text at least as frequently as the user has indicated that it should;
2. If the frequency of the item in the text is compared to its frequency in a reference corpus, it can be determined that the statistical likelihood, as calculated by a suitable process, is less than or equal to a p value that the user specifies.
3. Moreover, the keyness strength needs to be at least equal to the user-specified minimum log ratio.

N	Key word	Freq.	%	ext	RC. Freq.	Rc. %	BIC
1	ROMEO	115	0.5%	1	115	0.01%	491.18
2	TYBALT	47	0.2%	1	47		192.68
3	JULIET	40	0.2%	1	45		155.60
4	NURSE	39	0.2%	1	83	0.01%	115.35
5	PARIS	27	0.1%	1	81		61.42
6	MONTAGUE	21	0.1%	1	41		58.55
7	MERCUTIO	14	0.1%	1	14		47.81
8	THURSDAY	14	0.1%	1	17		44.12
9	COUNTY	16	0.1%	1	28		43.97
10	CAPULET	13	0.1%	1	13		43.42
11	FRIAR	22	0.1%	1	80		40.94
12	ROMEO'S	12	0.0%	1	12		39.03
13	TYBALT'S	10	0.0%	1	10		30.26
14	DEAD	49	0.2%	1	528	0.06%	29.79
15	MANTUA	11	0.0%	1	18		27.05
16	CELL	12	0.0%	1	29		23.72
17	VERONA	10	0.0%	1	19		21.14
18	WILT	30	0.1%	1	299	0.04%	16.01
19	LAURENCE	7	0.0%	1	8		15.82

Figure 2.6. KeyWords tool in WordSmith Tools

2.4. Reference corpus description

As a reference corpus for keyword generation, British National Corpus (BNC) was chosen [26]. BNC is a large collection of various spoken and written samples of specifically British English starting from the 20th century. It is balanced, meaning it includes various discourse styles such as spontaneous speech or academic writing. Its size exceeds 100 million words, which makes the entire corpus too large for current purposes, so only a fraction of such corpus will be used. Also, even though the corpus is freely available for academic research online, however it's not possible to download it to analyse with custom tools. Fortunately, a downloadable sample of British National Corpus is available with the name BNC Baby Edition [27].

This archive contains approximately 4 million words from text samples divided into 4 categories: academic prose, unscripted conversation, imaginative prose, and newspapers and periodicals. Each text sample is no longer than 40000 words. The

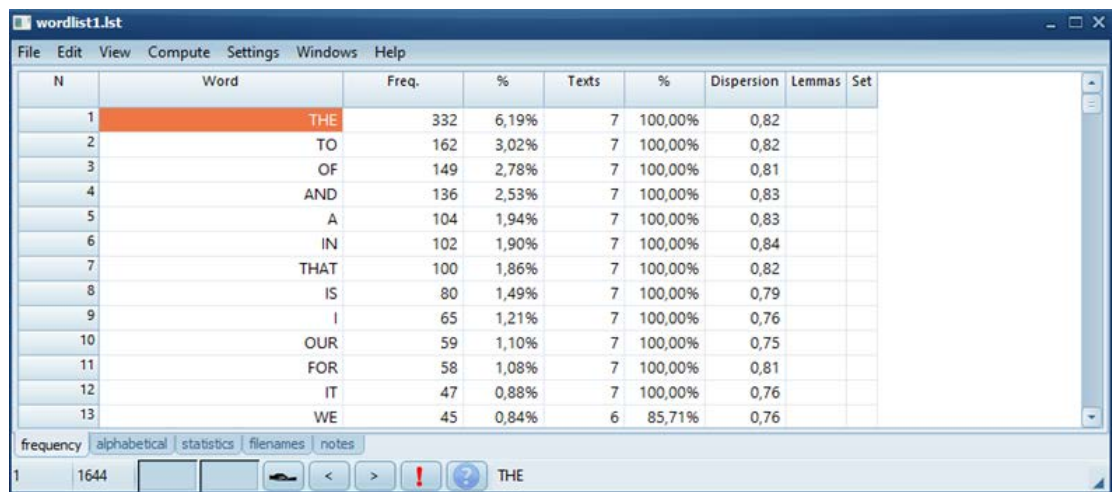
corpus sample is completely annotated with the speaker's ages, sex, text author and publication date etc. All data files are stored in XML format, which is recognisable by the custom software Xaira Tools which is also provided in BNC Baby Edition [27]. However, Xaira is not needed for the current purposes, and since WordSmith Tools only accepts raw text for frequency list generation (which is the main purpose of reference corpus), this XML data must be converted to just text. To extract raw text from these files an additional Python script will be used. The script will just take only necessary data and compose a text file.

CHAPTER 3. LINGUISTIC PECULIARITIES OF RISHI SUNAK'S LANGUAGE PROFILE

3.1. Analysis with the help of WordSmith Tools

3.1.1. Identifying keywords based on the reference corpus

The vital step in keyword generation is making a wordlist to serve as an input data for the tool. Overall, three wordlists were generated, one for each period of Rishi Sunak's political career. The next figure presents a part of the first wordlist:



The screenshot shows a window titled 'wordlist1.lst' with a menu bar (File, Edit, View, Compute, Settings, Windows, Help) and a table of wordlist data. The table has columns for Rank (N), Word, Frequency (Freq.), Percentage (%), Number of Texts, Percentage of Texts (%), Dispersion, Lemmas, and Set. The word 'THE' is highlighted in orange in the first row.

N	Word	Freq.	%	Texts	%	Dispersion	Lemmas	Set
1	THE	332	6,19%	7	100,00%	0,82		
2	TO	162	3,02%	7	100,00%	0,82		
3	OF	149	2,78%	7	100,00%	0,81		
4	AND	136	2,53%	7	100,00%	0,83		
5	A	104	1,94%	7	100,00%	0,83		
6	IN	102	1,90%	7	100,00%	0,84		
7	THAT	100	1,86%	7	100,00%	0,82		
8	IS	80	1,49%	7	100,00%	0,79		
9	I	65	1,21%	7	100,00%	0,76		
10	OUR	59	1,10%	7	100,00%	0,75		
11	FOR	58	1,08%	7	100,00%	0,81		
12	IT	47	0,88%	7	100,00%	0,76		
13	WE	45	0,84%	6	85,71%	0,76		

Figure 3.1. Wordlist of the 1st period (2015-2018)

The wordlists themselves provide very little information regarding the nature and style of the text since most of the words at the top are very commonly used in English. There weren't any descriptive words until the 39th one, "Britain," which was used 15 times (0.28% of all words) in 5 speeches out of 7 given.

N	Word	Freq.	%	Texts	%	Dispersion	Lemmas	Set
33	HE	19	0,35%	4	57,14%	0,46		
34	THAN	19	0,35%	6	85,71%	0,62		
35	AT	18	0,34%	7	100,00%	0,71		
36	BY	18	0,34%	7	100,00%	0,77		
37	MORE	17	0,32%	6	85,71%	0,70		
38	WAS	16	0,30%	4	57,14%	0,52		
39	BRITAIN	15	0,28%	5	71,43%	0,71		
40	ITS	15	0,28%	7	100,00%	0,75		
41	LIKE	15	0,28%	7	100,00%	0,79		
42	ONLY	15	0,28%	6	85,71%	0,68		
43	SO	15	0,28%	6	85,71%	0,71		
44	THEY	15	0,28%	7	100,00%	0,75		

Figure 3.2. The word “Britain” in the wordlist

The other two wordlists look very similar to the first one and were created in the same way. The next step would be feeding the acquired data into the KeyWords tool and checking what important words will emerge. The British National Corpus was used as a reference corpus. WordSmith Tools only accepts reference corpora in the WordList format, so the software generated a wordlist from the corpus as well. After that, it was possible to calculate keywords of the first group of speeches:

Table 1. Keywords of the 1st period exported from WordSmith Tools

N	Keyword	Frequency	%	Texts	RC Frequency	RC. Freq. %
1	OUR	59	1,0995	7	2565	0,0654
2	FUTURE	14	0,2609	7	557	0,0142
3	GROUSE	13	0,2423	1	8	0,0002
4	NORTH	13	0,2423	3	656	0,0167
5	EU	11	0,2050	4	4	0,0001
6	SHOOTING	10	0,1864	1	80	0,0020
7	DEBT	10	0,1864	2	114	0,0029
8	RURAL	9	0,1677	3	94	0,0024
9	NORTHERN	9	0,1677	1	309	0,0079
10	EIS	7	0,1305	1	0	0,0000
11	LEON	7	0,1305	1	13	0,0003
12	HARRIER	6	0,1118	2	3	0,0001
13	MOORS	6	0,1118	2	13	0,0003
14	BAN	6	0,1118	1	94	0,0024

15	POWERHOUSE	5	0,0932	1	4	0,0001
16	PREDECESSOR	5	0,0932	2	30	0,0008
17	NATION'S	5	0,0932	4	37	0,0009
18	SMES	4	0,0745	1	0	0,0000
19	HMRC	3	0,0559	1	0	0,0000
20	STREETER	3	0,0559	1	0	0,0000

The extracted keywords provide much more insight into Rishi Sunak's speeches. Immediately it can be seen that the politician addressed a variety of issues like ecological matters, for instance, by using the word "grouse" (a species of bird) 13 times, "shooting" 10 times (related to the extermination of animals), and that's only in one speech. Additionally, he actively discussed economics and finances in the local government, focusing audiences' attention on critical matters by using the word "debt" 10 times in 2 speeches, "SMEs", which is an abbreviation for "small and medium-sized enterprises" – at least 4 times in a single speech. The first word "our" is dominating the table with 59 entries making up for 1.1% of the whole text. It means that Rishi Sunak was trying to create a sense of inclusivity and unity among the audience. The frequent usage of such word was employed to persuade the audience by sharing responsibility and concern for problems in the country.

To make an even better visual representation of keywords in the given corpus, a word cloud can be formed. Our word clouds for the first and subsequent periods are located in appendixes.

In this picture, the more frequent words appear larger and bolder, while the less frequent words appear smaller. Word clouds are often used to quickly visualise the key themes, topics, or keywords in a body of text, making it easy to identify patterns and trends.

The second keyword tables and word clouds are made in a similar manner:

Table 2. Keywords of the 2nd period exported from WordSmith Tools

N	Keyword	Frequency	%	Texts	RC Frequency	RC. Freq. %
1	OUR	249	1,4122	4	2565	0,0654
2	WILL	249	1,4122	4	8468	0,2159

3	PEOPLE	102	0,5785	4	4033	0,1028
4	NEW	95	0,5388	4	3750	0,0956
5	JOBS	94	0,5331	4	368	0,0094
6	BUSINESSES	84	0,4764	3	129	0,0033
7	MILLION	71	0,4027	3	673	0,0172
8	SPEAKER	65	0,3686	3	274	0,0070
9	BILLION	60	0,3403	4	103	0,0026
10	BUSINESS	52	0,2949	4	1099	0,0280
11	FUTURE	37	0,2098	3	557	0,0142
12	CORONAVIRUS	31	0,1758	3	0	0,0000
13	FISCAL	30	0,1701	2	9	0,0002
14	RESPONSE	28	0,1588	3	399	0,0102
15	PROTECT	24	0,1361	3	147	0,0037
16	OBR	20	0,1134	2	0	0,0000
17	INVEST	18	0,1021	3	45	0,0011
18	PROSPERITY	17	0,0964	3	35	0,0009
19	SECTORS	16	0,0907	3	42	0,0011
20	FURLOUGH	11	0,0624	2	0	0,0000

In the second period, the word “our” is still at the top of the keyword list, appearing 249 times, which corresponds to 1.41% of the whole text. However, the shift in the main topics is noticeable since Rishi Sunak was Chancellor of the Exchequer at that period, and one of his main responsibilities was managing the budget for the upcoming year. Of course, his speeches greatly covered this topic, and it is obvious in the keyword table and word cloud. For instance, the word “businesses” is used 84 times, “jobs” and “billion” (most likely referring to the quantity of money) 60 times. Also, the word “will” is also used very often – 249 times, interestingly, which is the same as “our.” This is explained by more talks about the future and plans rather than the past or present. The word “future” itself is used 37 times, which further confirms this. Next is the keyword table and word cloud of the third group:

Table 3. Keywords of the 3rd period exported from WordSmith Tools

N	Keyword	Frequency	%	Texts	RC Frequency	RC. Freq. %
1	WILL	83	0,8676	11	8 468	0,2159
2	MORE	63	0,6585	9	7 748	0,1975
3	MATHS	37	0,3867	1	60	0,0015
4	COUNTRY	27	0,2822	9	768	0,0196

5	NHS	26	0,2718	4	81	0,0021
6	WORLD	25	0,2613	8	1 644	0,0419
7	AI	21	0,2195	3	15	0,0004
8	PAY	21	0,2195	5	876	0,0223
9	SECURITY	17	0,1777	4	286	0,0073
10	PLAN	16	0,1672	5	407	0,0104
11	NATO	14	0,1463	3	5	0,0001
12	UKRAINIAN	11	0,1150	3	3	0,0001
13	SUMMIT	11	0,1150	5	39	0,0010
14	GLOBAL	11	0,1150	6	132	0,0034
15	HOUTHIS	9	0,0941	1	0	0,0000
16	STRIKES	8	0,0836	2	44	0,0011
17	DELIVER	8	0,0836	3	49	0,0012
18	TECH	7	0,0732	1	15	0,0004
19	CANNOT	5	0,0523	5	0	0,0000
20	UKRAINE'S	5	0,0523	2	0	0,0000

During the third period, there is a variety of words in terms of the number of repetitions, namely the word will (83 times), more (63 times), we (50 times), maths (37 times), AI (artificial intelligence) (21 times), etc. The prime minister is responsible for leading the government and making strategic decisions on the economy, social development, legislation, foreign policy, and many other aspects of public administration. To be successful in this role, one needs to understand a wide range of topics and be prepared to work with experts and consultants from various fields. That is why Rishi Sunak actively raises not only political issues but also science, artificial intelligence, and education aspects, which are all part of the process of technological change. This indicates that the politician addresses the most relevant issues for his country so that it can continue to serve as a voice for development.

In the quantitative analysis, we took pronouns into account precisely because of the frequency of their usage. At the beginning of his career, Rishi Sunak used the pronoun “I” 65 times, the pronoun “we” 45 times, the pronoun “you” 11 times, and the possessive pronoun “our” 59 times. During his time in the government, he used the pronoun “I” 170 times, the pronoun “we” 278 times, the pronoun “you” 30 times, and the possessive pronoun “our” 249 times. During Sunak's premiership, the pronoun “I” was used 111 times, the pronoun “we” 141 times, the pronoun “you” 38 times, and the

possessive pronoun “our” 156 times. In terms of percentage, we can see that the usage of the pronoun “I” has increased. This can be explained by the fact that taking on the new positions of Chancellor of the Exchequer and Prime Minister, Rishi Sunak makes more statements about his personal opinion, outlining certain phenomena. It can be assumed that at the beginning of his career, the politician did not have many powers and therefore he was more active in sharing the opinion of the top leadership. A similar tendency is evident in the use of the pronoun we. In fact, this is an interesting case of the use of this pronoun because the UK is a country with an individualistic culture. That is, they do not often use the word “we” for identification and identity as such [75]. However, in Rishi Sunak's speeches, this pronoun is dominant. This is due to the fact that although Sunak grew up in the UK, his family is from Punjab, a state in India that, in general, belongs to a collectivist culture. Collectivism is characterised by the use of the pronoun we, which is a symbol of unity. It is evident that Sunak's family has passed on these values to the future politician, who incorporates them in his speeches. As for the pronoun you, its usage has not changed much over the three periods, which can be explained by the fact that it is more important for him to address the community than individual members.

3.1.2. Identifying speech patterns based on concordances

Having analysed the keywords, the last step is using Concord tool in WordSmith Tools to make a concordance analysis of specific words and phrases. Multiple ideas were considered for this, but not all results were particularly useful. In some cases, however, the tool provided an interesting insight into Rishi Sunak’s speech style.

For this analysis, two words were selected, and concordances were generated with the help of the WordSmith Tools. These tables shows that the common feature that unites all the speeches is the use of action verbs. As mentioned earlier, action verbs convey a sense of activity and determination, making the message more direct and impactful. They can help politicians to clearly communicate their intentions and goals to the public. Here are some examples of these words:

way we fund science in this country. I can confirm that we will invest at least £800m in a
an investment-led recovery. So today, I can announce the Super Deduction. For the next
a scientific superpower is something we can achieve. I don't think that's hubristic or
growing parts of the global economy. We can use economic innovations like free ports to

be completely unaffected. And third, we will introduce a taper above £50,000 so that only
if Russia ever invades Ukraine again, we will provide swift and sustained assistance,
at we were too weak to pay ourselves. I will unite our country, not with words, but
try, not with words, but with action. I will work day in and day out to deliver for

It was also interesting to analyse the conjunction “and” at the beginning of the sentence using concord analysis. The politician not only logically structures his thoughts, but also introduces a more informal style. It is worth mentioning that the politician does not use more formal structures very often, such as moreover, furthermore, in addition, apart from that etc. This could indicate that the text is intended for a casual audience or that the speaker is adopting a less formal tone:

with a new payment of £2,000 per apprentice . And we will introduce a brand new bonus for
kickstarters in their new jobs this autumn . And I urge every employer, big or small,
- Google DeepMind, OpenAI and Anthropic . And I'm pleased to announce they've committed to
could need to be off work at any one time . And business supply chains are being disrupted
position that “Ukraine’s future is in NATO .” And we've taken a big step this week towards
in all four nations of the United Kingdom . And I'm taking further specific steps with

Based on these results, it can be concluded that Rishi Sunak's linguistic personality exemplifies informal features. Yet, he presents the information consistently rather than chaotically. His phrases often include action verbs to provide guarantees of his further steps to improve the quality of life of British citizens.

3.2. The implementation of Rishi Sunak's linguistic personality at the lexical level

3.2.1. The beginning of the work in the government (2015 - 2018)

Modern society is constantly faced with such phenomena as influence, stimulation, manipulation. The problem of studying the factors that influence decision-making and choice is of particular importance. Every day, people hear speeches of other people that have the potential to change their life attitudes, adjust the system of values. The choice of vocabulary, specific grammatical structures, stylistic devices - all of this plays a key role in shaping the social and political processes.

As a person of Punjabi origin, Rishi Sunak has a multifaceted linguistic profile. Through qualitative analysis, we have identified which tendencies in the use of linguistic structures dominate in the speeches of this particular politician and how they represent the political life of the UK as such.

At the lexical level of Rishi Sunak's speeches, there are specific lexical items that are divided into two groups: general politician vocabulary which includes philosophical, sociological, cultural, religious terms widely used in the media, as well as words related to the functioning of various unions, associations, NGOs, etc., as well as lexical items used to refer to processes and events in the field of education and cultural life of the country; special political vocabulary including job titles, state bodies, political parties and movements, terms of international politics, legal and economic terms, as well as electoral and military vocabulary.

Philosophical Terms

Term	Frequency	Relative Frequency
OPPORTUNITY	5	0,093%
CONTROL	4	0,075%
BEAUTY	3	0,056%
CONTRIBUTION	3	0,056%
DEBATE	3	0,056%
GENERATION	3	0,056%
SPIRIT	3	0,056%
SUCCESS	3	0,056%
AMBITION	2	0,037%

ASSURANCE	2	0,037%
TRUTH	2	0,037%
VALUES	2	0,037%
FREEDOM	1	0,019%
Total	36	0,671%

Among philosophical terms, the word *opportunity* (5 times) is the most common. This indicates that Rishi Sunak is starting to promote his ideas at the government level, emphasising the amount of work that can be done.

Sociological Terms

Term	Frequency	Relative Frequency
FAMILY	4	0,075%
GENERATION	3	0,056%
POPULATION	3	0,056%
COMMUNITY	2	0,037%
IMMIGRATION	2	0,037%
MEMBER	2	0,037%
UNION	2	0,037%
Total	18	0,336%

Cultural Terms

Term	Frequency	Relative Frequency
HERITAGE	4	0,075%
HISTORY	3	0,056%
VALUES	2	0,037%
Total	9	0,168%

In terms of sociological and cultural words, Rishi Sunak is guided by the values of family and cultural heritage. These values are a solid foundation for nation building.

Unions, Associations, and NGOs

Term	Frequency	Relative Frequency
EU	11	0,205%
EIS	7	0,130%
NHS	1	0,019%
Total	19	0,354%

Mentioning the names of unions and associations, the EU appears 11 times in Rishi Sunak's speeches. At this stage, the EU has become central to discussions on Britain's political strategy and Brexit.

Education Terms

Term	Frequency	Relative Frequency
SCHOOLS	4	0,075%
EDUCATION	2	0,037%
UNIVERSITIES	2	0,037%
Total	8	0,149%

Educational terms are rarely used due to Sunak's relatively early career and focus on broader issues.

Job Titles

Term	Frequency	Relative Frequency
SPEAKER	8	0,149%
CHAIRMAN	6	0,112%
SECRETARY	4	0,075%
MINISTER	2	0,037%
CHANCELLOR	1	0,019%
Total	21	0,391%

We see that Rishi Sunak often addresses the *speaker*. This is a way of expressing formality and politeness to officials.

State Bodies

Term	Frequency	Relative Frequency
POLICE	12	0,224%
GOVERNMENT	12	0,224%
OFFICERS	5	0,093%
STATE	3	0,056%
HMRC	3	0,056%
METROPOLITAN	3	0,056%
Total	38	0,708%

State terms include the frequent use of the words *police* and *government*. They are evidence that Rishi Sunak refers to authoritative bodies when addressing a particular topic.

Political Parties and Movements

No explicit mentions of political parties and movements were identified in the current period.

Terms of International Politics

Term	Frequency	Relative Frequency
TRADE	5	0,093%
POLICY	4	0,075%
NATION	3	0,056%
FOREIGN	2	0,037%
RELATIONS	1	0,019%
Total	15	0,280%

At the international level, the politician refers most often to *trade* as a means of "connecting" with the world and spreading the UK's influence.

Legal Terms

Term	Frequency	Relative Frequency
ANONYMITY	4	0,075%
REGULATIONS	3	0,056%
INVESTIGATION	2	0,037%
JUSTICE	2	0,037%
CRIMINALITY	1	0,019%
DEFENSE	1	0,019%
EVIDENCE	1	0,019%
LAW	1	0,019%
LEGAL	1	0,019%
RIGHTS	1	0,019%
Total	17	0,317%

We can also notice that Rishi Sunak speaks on legal topics, namely *anonymity* of expression.

Economic Terms

Term	Frequency	Relative Frequency
ECONOMY	11	0,205%
DEBT	10	0,186%
GROWTH	8	0,149%
INVESTMENT	7	0,130%
GDP	7	0,130%
FINANCE	6	0,112%
TRADE	5	0,093%
BUDGET	4	0,075%
CAPITAL	4	0,075%
MARKET	4	0,075%
INTEREST	4	0,075%
TAX	3	0,056%
FISCAL	3	0,056%
BUSINESS	2	0,037%
FINANCES	2	0,037%
INCOME	1	0,019%
EXPORT	1	0,019%
CURRENCY	1	0,019%
Total	83	1,547%

Electoral Vocabulary

Term	Frequency	Relative Frequency
VOTE	2	0,037%
VOTERS	1	0,019%
CAMPAIGN	1	0,019%
Total	4	0,075%

Military Vocabulary

Term	Frequency	Relative Frequency
RAF LEEMING	1	0,019%
ARMY	1	0,019%
ARMED FORCES	1	0,019%
CATARACT GARRISON	1	0,019%
GREEN HOWARDS	1	0,019%
Total	5	0,094%

In contrast to the military and electoral vocabulary, Rishi Sunak focuses on economic issues, specifically the words *economy* and *debt*, which became critical for the adoption of the general budget of the country.

Nouns with Positive Connotation

Term	Frequency	Relative Frequency
GROWTH	8	0,149%
OPPORTUNITY	5	0,093%
BEAUTY	3	0,056%
LAUGHTER	3	0,056%
SUCCESS	3	0,056%
RESPECT	3	0,056%
TRUTH	2	0,037%
FRIEND(s)	2	0,037%
WARMTH	1	0,019%
PEACE	1	0,019%
RELIEF	1	0,019%
Total	32	0,596%

During the first period, the politician used nouns with positive connotations 32 times, the most frequent of which are *growth* and *opportunity*. These words refer to local businesses, which, in Sunak's opinion, have great potential to grow to the national level. Rishi Sunak logically turns to local regions to understand the internal structure of economic issues. This will allow him to further develop the strategy at the national level.

Nouns with Negative Connotation

Term	Frequency	Relative Frequency
DEBT	10	0,186%
SHOOTING	10	0,186%
BAN	6	0,112%
CANCER	2	0,037%
FAILINGS	2	0,037%
VICTIMS	2	0,037%
DELAY	1	0,019%
FEAR	1	0,019%
DISASTER	1	0,019%
STRUGGLE	1	0,019%
RISK	1	0,019%
ACCUSATION	1	0,019%
ACCUSER	1	0,019%
CRIMINALITY	1	0,019%
DESTRUCTION	1	0,019%
Total	41	0,764%

Nouns with negative connotations are used 42 times, with the focus on the words *debt* and *shooting*. It is interesting that words with negative connotations prevail in this case, as politicians deliberately avoid negative connotations. However, in the existing context, Rishi Sunak wants to draw attention to issues that really affect the welfare of citizens. This fact is a bright example of not being afraid to raise serious topics.

Adjectives with Positive Connotation

Term	Frequency	Relative Frequency
EXCELLENT	6	0,112%
BETTER	6	0,112%
PROUD	5	0,093%
HONOURABLE	4	0,075%
STRONG	4	0,075%
BRIGHT	4	0,075%
OUTSTANDING	2	0,037%
FIT	2	0,037%
REMARKABLE	2	0,037%
WELCOMING	2	0,037%
FINEST	2	0,037%
AMBITIOUS	2	0,037%
PROSPEROUS	2	0,037%
HEALTHY	1	0,019%
GOOD	1	0,019%
FORTUNATE	1	0,019%
COMPASSIONATE	1	0,019%
Total	47	0,876%

An interesting trend is the use of adjectives with strong positive connotations. We can see that the word *excellent* prevails. Instead, the word *good* is used only once. In fact, strong adjectives appeal more to people's emotions. Therefore, it is important for politicians to connect with the public in this way, which is what Rishi Sunak is doing. In this way, he expresses a firm position even when appealing to emotions.

Adjectives with Negative Connotation

Term	Frequency	Relative Frequency
HARD	3	0,056%
POORER	2	0,037%
DIFFICULT	2	0,037%

DANGEROUS	1	0,019%
DEVASTATING	1	0,019%
DISABLED	1	0,019%
INADEQUATE	1	0,019%
MISLEADING	1	0,019%
ODD	1	0,019%
SAD	1	0,019%
SEVERE	1	0,019%
OUTDATED	1	0,019%
TOUGH	1	0,019%
WRONG	1	0,019%
UNREGULATED	1	0,019%
SHAMEFUL	1	0,019%
UNJUSTIFIED	1	0,019%
GUILTY	1	0,019%
Total	22	0,410%

In contrast to adjectives with positive connotations, adjectives with negative connotations are used only 22 times, of which the word *hard* is used three times. Politicians deliberately avoid such words in order not to spread pessimism. The example of Rishi Sunak shows that the politician does not see any issues that cannot be solved. Therefore, the use of such adjectives is not frequent enough.

Verbs with Positive Connotation

Term	Frequency	Relative Frequency
SUPPORT	4	0,075%
CREATE	3	0,056%
HELP	3	0,056%
IMPROVE	2	0,037%
LEARN	2	0,037%
SAVE	2	0,037%
THRIVE	2	0,037%
WELCOME	2	0,037%
AMUSED	1	0,019%
APPRECIATE	1	0,019%
CELEBRATE	1	0,019%
ACHIEVE	1	0,019%
ATTRACT	1	0,019%
LOVE	1	0,019%
SUCCEED	1	0,019%
UPGRADE	1	0,019%
Total	28	0,522%

As for verbs with positive connotations, they are used 28 times. All these words are action verbs aimed at development, such as *support*, *create*, *help*. In this way, Rishi Sunak emphasises the call to action, to unite, to a common goal.

Verbs with Negative Connotation

Term	Frequency	Relative Frequency
LOST	3	0,056%
FAILED	2	0,037%
FORGET	2	0,037%
STRUGGLE	1	0,019%
ABANDONED	1	0,019%
BANNING	1	0,019%
DIED	1	0,019%
LIE	1	0,019%
Total	12	0,224%

As for verbs with negative connotations, they project a similar situation with adjectives, i.e. they are not quite frequent in use (12 times). The fact is that they inherently sound harsh, and that is why politicians also try to avoid such words. So, again, for Rishi Sunak, it is important to show action, not mistakes, growth, not failed attempts.

An analysis of Rishi Sunak's early career speeches helped to reveal the tendencies in the usage of general political and special political vocabulary. Philosophical terms such as "opportunity" are frequently repeated, illustrating his focus on unlocking the UK's growth and development potential. Sociologically and culturally, Sunak emphasises family and heritage values, portraying them as fundamental to nation-building efforts. The reference to alliances and associations, particularly the EU, underlines the centrality of the politician's discussions of UK political strategy and Brexit.

Frequent references to official positions help Rishi Sunak to give his speeches a more formal look. The use of economic terms and references to government agencies suggests that at the beginning of his career, the politician was actively involved in these

discussions and was not afraid to express his opinion, although at that time his concerns did not include these particular issues. A balance was found between positive and negative nouns in his speeches, while other politicians would have avoided using terms with negative connotations. This again shows how Sunak actively referred to current issues in Britain. His use of words such as *excellent* instead of the regular *good* demonstrates his broader vocabulary and adds formality to his speeches.

Based on the vocabulary used by Rishi Sunak during the first period, it is clear that at the beginning of his career it was important for the politician to show his commitment to the country and lay the foundations of trust by demonstrating the capabilities of the UK, appealing to the core values of the country and thereby showing that he also shared them.

3.2.2. The chancellorship (2019 - 2022)

Philosophical Terms

Term	Frequency	Relative Frequency
TIME	35	0,199%
OPPORTUNITY	12	0,068%
GENERATION	6	0,034%
TRUTH	5	0,028%
NATURE	5	0,028%
POWER	5	0,028%
SUCCESS	4	0,023%
AMBITION	3	0,017%
FREEDOM	3	0,017%
DEBATE	2	0,011%
SPIRIT	2	0,011%
VALUES	2	0,011%
CONTROL	1	0,006%
CONTRIBUTION	1	0,006%
Total	86	0,488%

In this period, Rishi Sunak started using the term *time* remarkably often. Most likely, he tried to emphasise urgency in ongoing political matters. Also, the *opportunity* is at the top of philosophical terms used by the politician, projecting the same tendency as in the previous period. Since these two terms complement each other, Sunak used both fairly frequently to convey his ideas to the public. Compared to the previous

period, philosophical terms were used slightly less frequently (0,488% now, 0,671% before).

Sociological Terms

Term	Frequency	Relative Frequency
MEMBER	9	0,051%
GENERATION	6	0,034%
FAMILY	3	0,017%
COMMUNITY	3	0,017%
SOCIETY	3	0,017%
POPULATION	2	0,011%
IMMIGRATION	2	0,011%
UNION	2	0,011%
Total	30	0,170%

Cultural Terms

Term	Frequency	Relative Frequency
HISTORY	4	0,023%
VALUES	2	0,011%
CUSTOMS	1	0,006%
Total	7	0,040%

Both sociological and cultural terms are used less frequently compared to the first period (0,170% from 0,336% and 0,040% from 0,168%, respectively). Together with lesser usage of philosophical terms, we can conclude that Rishi Sunak has moved away from persuading the public with abstract concepts and makes more use of precise facts.

Unions, Associations, and NGOs

No mentions of unions, associations, or NGOs were identified in the given period.

Education Terms

Term	Frequency	Relative Frequency
SCHOOLS	7	0,040%
EDUCATION	6	0,034%

TRAINING	6	0,034%
SCHOOL	2	0,011%
COLLEGE	2	0,011%
ACADEMIC	2	0,011%
TEACHERS	1	0,006%
UNIVERSITIES	1	0,006%
LEARNING	1	0,006%
Total	28	0,159%

Education terms were used consistently rarely in the first and second period of Sunak's career (0,159% now, 0,149% before). *Schools* and *education* were mentioned a few times during budget talks, when the politician tried to draw attention to the allocation of funds for this area.

Job Titles

Term	Frequency	Relative Frequency
SPEAKER	65	0,369%
SECRETARY	14	0,079%
MINISTER	5	0,028%
CHANCELLOR	3	0,017%
Total	87	0,493%

During the second period, Rishi Sunak most often appealed to or mentioned the *speaker* 65 times. While also mentioning other roles like *secretary*, it's apparent that most of his speeches were addressed towards the speaker itself, while in the first period, the distribution of roles Rishi spoke to was fairer. Also, Sunak started mentioning the job titles themselves more often than in previous period as a way to appeal to the authority (0,493% now, 0,391% before).

State Bodies

Term	Frequency	Relative Frequency
GOVERNMENT	33	0,187%
BANK	15	0,085%
NHS	14	0,079%
TREASURY	5	0,028%
OFFICE	4	0,023%
HMRC	4	0,023%
PARLIAMENT	4	0,023%

Total | 79 0,448%

Among the mentions of state bodies, the *government* appeared the most often – 33 times – likely as an implicit way of referring to his colleagues. Other terms like *bank*, *NHS* (National Health Service) appeared 15 and 14 times, respectively, in discussions related to the budget.

Political Parties and Movements

Term	Frequency	Relative Frequency
CONSERVATIVE(S)	11	0,062%
LABOUR(^S)	6	0,034%
Total	17	0,096%

In Rishi Sunak’s budget speeches, mentions of the Conservative and Labour parties helped the politician appeal either to his colleagues (Conservatives) or to his opposition, depending on the argument. This technique wasn’t employed in the first period.

Terms of International Politics

Term	Frequency	Relative Frequency
POLICY	19	0,108%
NATION	11	0,062%
TRADE	6	0,034%
UNIVERSAL	6	0,034%
TRANSITION	5	0,028%
Total	47	0,267%

Legal Terms

Term	Frequency	Relative Frequency
EVIDENCE	3	0,017%
JUSTICE	1	0,006%
Total	4	0,023%

It is evident that Rishi Sunak used legal vocabulary much more rarely than in the first period (0,0317% to 0,023%). However, the topic of international politics was consistent between both periods. The most frequent terms are *policy* and *nation*, with

19 and 11 appearances, respectively. This suggests a focus on issues related to governance, national identity, and strategic direction affecting Britain and other countries.

Economic Terms

Term	Frequency	Relative Frequency
JOBS	94	0,533%
MILLION	71	0,403%
BILLION	60	0,340%
TAX	59	0,335%
ECONOMY	56	0,318%
BUDGET	53	0,301%
BUSINESS	52	0,295%
INVESTMENT	36	0,204%
PLAN	36	0,204%
PAY	36	0,204%
FUNDING	32	0,181%
FISCAL	30	0,170%
CUT	29	0,164%
RATE	23	0,130%
GROWTH	22	0,125%
COST	20	0,113%
FINANCES	13	0,074%
WAGE	12	0,068%
TAXES	10	0,057%
INCOME	10	0,057%
CAPITAL	9	0,051%
WAGES	9	0,051%
GDP	8	0,045%
DEBT	7	0,040%
INTEREST	7	0,040%
TRADE	6	0,034%
INFLATION	5	0,028%
FINANCE	3	0,017%
MARKET	3	0,017%
EXPORT	3	0,017%
Total	814	4,617%

During this period, economic terms are the most prominent, with a relative frequency of 4,167% (1,547% in the first period), which means that on average, among 200 words randomly chosen from the corpus, a term related to the economy is almost certain to appear. Very high usage of the words *jobs*, *business*, *investment*, *tax*, and

wage is obviously meant to convince the government to boost the stagnating economy during the COVID-19 crisis. Words *million* and *billion* are very common in the context of large-scale economic figures, such as budget allocations, investments, or economic output since discussions of this kind require accuracy and certainty.

Electoral Vocabulary

Term	Frequency	Relative Frequency
PROGRAMME	3	0,017%
ELECTION	3	0,017%
VOTE	1	0,006%
VOTING	1	0,006%
Total	8	0,045%

Rishi Sunak doesn't use terms related to elections very often, even less than in the previous period (0,045% now, 0,075% before). It could suggest that he is comfortable with his current position and his government.

Military Vocabulary

Term	Frequency	Relative Frequency
DEFENCES	3	0,017%
ARMED FORCES	3	0,017%
Total	6	0,034%

The situation is similar for military vocabulary (0,034% now, 0,075% before). The politician uses military terms, usually in an economic context, during his argumentation for funding entities related to the UK armed forces.

Nouns with Positive Connotation

Term	Frequency	Relative Frequency
SUPPORT	40	0,226%
GROWTH	22	0,125%
PROSPERITY	17	0,096%
COMMITMENT	15	0,085%
RELIEF	15	0,085%
OPPORTUNITY	12	0,068%
HOSPITALITY	11	0,062%

PRODUCTIVITY	10	0,057%
SECURITY	9	0,051%
HEALTH	9	0,051%
RECOVERY	8	0,045%
DEVELOPMENT	7	0,040%
GUARANTEE	7	0,040%
CONFIDENCE	6	0,034%
PROMISE	6	0,034%
SUCCESS	4	0,023%
OPPORTUNITIES	2	0,011%
Total	200	1,133%

Nouns with Negative Connotation

Term	Frequency	Relative Frequency
CORONAVIRUS	31	0,176%
CRISIS	15	0,085%
DEBT	7	0,040%
UNEMPLOYMENT	7	0,040%
DAMAGE	4	0,023%
POLLUTION	4	0,023%
RISK	4	0,023%
PROBLEM	4	0,023%
Total	76	0,431%

In Sunak's speeches, the usage of nouns with positive connotation is preferred, even more so in the second period (1,133% now, 0,596% before). The top 1 positive noun is *support*, while top 1 negative is *coronavirus*. These terms are closely related, since the pandemic led to the economic crisis, and small/medium businesses needed support from the government in forms of tax cuts, credits, and other means just to be afloat. It's obvious that the politician would consistently bring up this problem in his budget speeches. Other words with positive connotation used are *growth*, *prosperity*, *commitment*, and *relief*. These are closely tied to the previous argument according to which if other politicians agree and cooperate with the Chancellor, all these terms will come to reality. However, it's important to note that *growth* can sometimes have negative connotations, such as when referring to an increase in the number of sick people, even though the word itself is generally positive. Other fairly used negative terms include *crisis*, *debt*, and *unemployment* as immediate consequences of the COVID-19 pandemic.

Adjectives with Positive Connotation

Term	Frequency	Relative Frequency
HONOURABLE	13	0,074%
BETTER	12	0,068%
GENEROUS	10	0,057%
FAIR	9	0,051%
YOUNG	8	0,045%
PROUD	6	0,034%
GOOD	6	0,034%
INNOVATIVE	6	0,034%
FREE	6	0,034%
STRONG	5	0,028%
COMPETITIVE	5	0,028%
CONFIDENT	5	0,028%
INCREDIBLE	4	0,023%
PLEASED	4	0,023%
PRODUCTIVE	3	0,017%
EXCELLENT	2	0,011%
AMBITIOUS	2	0,011%
CREATIVE	2	0,011%
EXCEPTIONAL	2	0,011%
PROSPEROUS	1	0,006%
HEALTHY	1	0,006%
Total	112	0,635%

The politician continues to consistently use adjectives with strongly positive connotations, with a slight decrease from the previous period (0,635% now, 0,876% before). Interestingly, the most commonly used adjective is *honourable*, with 13 appearances over the corpus. Upon further inspection of speeches, it was found that Sunak added the word frequently when mentioning his fellow colleagues, often referring to them as his “honourable friends”. It gives the public the impression that Rishi Sunak is a friendly and polite person. Words *better* and *generous* also make up a fair amount of positively connotated adjectives, mostly referring to economic reforms and tax cuts. In the second period, the word *good* is used more often (6 times) than its strong counterpart *excellent* (2 times). It means that Rishi considered the word *good* more fitting in his budget speeches.

Adjectives with Negative Connotation

Term	Frequency	Relative Frequency
VULNERABLE	9	0,051%
SICK	8	0,045%
WORRIED	6	0,034%
HARD	6	0,034%
UNEMPLOYED	4	0,023%
IRRESPONSIBLE	3	0,017%
UNABLE	3	0,017%
WORST	3	0,017%
AGGRESSIVE	1	0,006%
Total	43	0,244%

As for the adjectives with a negative connotation, the most frequently used include *vulnerable*, *sick*, and *worried*. All these refer to people suffering during the COVID-19 crisis. However, the usage of negative adjectives decreased compared to the previous period (0,244% now, 0,410% before), which shows that Sunak tried to keep his arguments on a more positive note to persuade his audience.

Verbs with Positive Connotation

Term	Frequency	Relative Frequency
SUPPORT	50	0,283%
HELP	24	0,136%
PROTECT	24	0,136%
INCREASE	24	0,136%
CREATE	17	0,096%
PROMISED	17	0,096%
PROVIDE	16	0,091%
CREATING	14	0,079%
SUPPORTING	13	0,074%
GROW	11	0,062%
DELIVER	9	0,051%
IMPROVE	7	0,040%
PROTECTING	7	0,040%
PROMISE	6	0,034%
LEADING	6	0,034%
ACHIEVE	5	0,028%
LEAD	4	0,023%
SAVE	3	0,017%
WELCOME	3	0,017%
SUCCEED	3	0,017%
LEARN	1	0,006%

ATTRACT	1	0,006%
Total	265	1,502%

Verbs with a positive connotation are widespread through the corpus, and their usage has risen significantly since the previous period (1,502% now, 0,522% before). The most frequent term is *support* at 50 occurrences, and the next three are a tie: *help*, *protect*, and *increase* at 24 occurrences. It should be noted that *support* was also the most frequent noun with a positive connotation, but nouns and verbs were counted separately. These verbs are most likely related to business support during the pandemic. Rishi Sunak appears to use a wide range of vocabulary as of positive verbs in his speeches.

Verbs with Negative Connotation

Term	Frequency	Relative Frequency
STRUGGLING	3	0,017%
LOST	2	0,011%
STRUGGLE	1	0,006%
LOSE	1	0,006%
CEASE	1	0,006%
Total	8	0,045%

Compared to the first period, the usage of verbs with a negative connotation is almost non-existent (0,045% now, 0,764% before). With the increased number of positive terms, the contrast is quite high. It is evident that for the political figure it was even more important to showcase strength in his speeches in the second period.

The comparative analysis revealed trends in the use of lexical items in the second period, which in some cases differ significantly. During his second period as Chancellor of the Exchequer, Sunak frequently emphasises terms such as "time" and "opportunity". These underline his efforts to convey a sense of urgency and potential for action in addressing current political issues.

An interesting fact is that Rishi Sunak takes a more precise approach, with less use of philosophical, sociological and cultural terms. Instead, Sunak seems to rely more on concrete facts and figures to convey his ideas, especially in the economic context.

Economic terms dominate his speeches, showing efforts to overcome the challenges posed by the COVID-19 crisis and to stimulate economic recovery. In addition, Sunak's use of positive nouns and adjectives, combined with a decrease in the use of negative terms, indicates that he is maintaining a positive tone and instilling confidence in his policies and initiatives.

It is worth noting that Sunak actively mentions state bodies, positions, and political parties. In this way, he actively appeals to various stakeholders and positions himself in the political landscape.

Based on the selected vocabulary during the second period, it is clear that the politician no longer appeals to all the values he emphasised in the first period. The position of the Chancellor required the politician to speak in the language of facts (often economic data), and thus to change the style of communication as such. In this way, we can see how the politician's flexibility with the new responsibilities in his life is manifested.

3.2.3. The premiership (2022 – present time)

Philosophical Terms

Term	Frequency	Relative Frequency
TIME	12	0,125%
FREEDOM	9	0,094%
PRINCIPLES	7	0,073%
OPPORTUNITY	5	0,052%
SPIRIT	4	0,042%
PURPOSE	4	0,042%
GENERATION	3	0,031%
VALUES	3	0,031%
CONTROL	2	0,021%
NATURE	2	0,021%
POWER	2	0,021%
CONTRIBUTION	1	0,010%
Total	54	0,564%

In the third period, the usage of philosophical terms stays consistently average. *Time* is again one of the most frequent occurrences (12 times), which, as it was mentioned, conveys a sense of urgency. Also, *freedom* appeared on the list with 9

mentions, which is likely related to talking about Ukrainians during the Russian invasion. Generally, philosophical terms employed by Sunak are positive by nature.

Sociological Terms

Term	Frequency	Relative Frequency
GENERATION	3	0,031%
FAMILY	2	0,021%
COMMUNITY	2	0,021%
UNION	2	0,021%
POPULATION	1	0,010%
IMMIGRATION	1	0,010%
MEMBER	1	0,010%
Total	12	0,125%

Cultural Terms

Term	Frequency	Relative Frequency
HISTORY	4	0,042%
VALUES	3	0,031%
CULTURE	3	0,031%
LANGUAGE	2	0,021%
Total	12	0,125%

Sociological and cultural terms usage remained low throughout all periods. Compared to the second period, sociology-related words were used slightly less frequently (0,125% now, 0,170% before), and cultural terms appeared a bit more often (0,125% now, 0,040% before). *History*, *values*, and *generation* being at the top of local frequency list suggest that the politician tried to raise some sense of patriotism in his audience.

Unions, Associations, and NGOs

Term	Frequency	Relative Frequency
NATO	14	0,146%
EU	1	0,010%
CPTPP	1	0,010%
Total	16	0,167%

As for mentions of organisations, NATO is the most prominent in the context of military discussions. Generally, mentions of unions, associations, and NGOs are not very frequent among the corpus (0,354% in the first period, 0% in the second period, 0,167% in the third period). It is evident that Rishi Sunak usually addresses individual people instead of organisations in his speeches.

Education Terms

Term	Frequency	Relative Frequency
TRAINING	11	0,115%
EDUCATION	6	0,063%
TEACHERS	5	0,052%
SCHOOLS	3	0,031%
UNIVERSITIES	1	0,010%
LEARNING	1	0,010%
Total	27	0,282%

Education term usage stays low over three periods corpus (0,149% in the first period, 0,159% in the second period, 0,282% in the third period). In the current period, the small increase is linked to the word *training*, which in the given context is not related to education, but rather to the training of Ukrainian and British troops.

Job Titles

Term	Frequency	Relative Frequency
DOCTOR(S)	10	0,103%
NURSE(S)	10	0,103%
SPEAKER	8	0,084%
MINISTER	6	0,063%
CHANCELLOR	2	0,021%
SECRETARY	1	0,010%
Total	37	0,387%

As Prime Minister, Rishi Sunak now addresses a separate primary audience – the general public of the UK – instead of his fellow colleagues. So, it is consequent that mentions of *speaker*, *minister*, *chancellor*, and *secretary* became much rarer.

State Bodies

Term	Frequency	Relative Frequency
NHS	26	0,272%
GOVERNMENT	24	0,251%
COUNCIL	6	0,063%
POLICE	4	0,042%
METROPOLITAN	2	0,021%
BANK	2	0,021%
PARLIAMENT	1	0,010%
Total	65	0,679%

In terms of state bodies, the most frequent usage of these terms is still *NHS* with 26 occurrences and *government* with 24. While mentions of *NHS* are mostly isolated in a single speech, the *government* appears consistently throughout all speeches in the period. Overall, occurrences of state bodies in speeches of Rishi Sunak are consistent and have average frequency (0,708% in the first period, 0,448% in the second period, 0,679% in the third period).

Political Parties and Movements

No explicit mentions of political parties and movements were identified in the current period.

Terms of International Politics

Term	Frequency	Relative Frequency
SECURITY	17	0,178%
WAR	11	0,115%
POLICY	8	0,084%
AID	5	0,052%
TRADE	4	0,042%
PARTNERSHIP	4	0,042%
ALLIANCE	4	0,042%
AGREEMENT	3	0,031%
NATION	2	0,021%
FOREIGN	2	0,021%
DIPLOMACY	2	0,021%
PEACE	2	0,021%
SANCTIONS	1	0,010%
NEGOTIATIONS	1	0,010%
CONFLICT	1	0,010%

Obviously, national *security* and *war* mentions have been at the top of the frequency list since the Russian invasion. What is also notable is that Rishi Sunak is not afraid to use the word *war*, while other politicians would try to avoid being so direct. It shows his solid stance in the ongoing world conflict. Words such as *aid*, *partnership*, *alliance*, and *agreement* are also related to these events, even though these were not used as often. Compared to the first and second periods, terms of international politics started occurring much more frequently in the third period (0,280% in the first period, 0,267% in the second period, and 0,700% in the third period).

Legal Terms

Term	Frequency	Relative Frequency
ILLEGAL	8	0,084%
LAW	5	0,052%
EVIDENCE	3	0,031%
CRIME	2	0,021%
CASE	2	0,021%
JUDGE	1	0,010%
LEGAL	1	0,010%
SANCTIONS	1	0,010%
Total	23	0,240%

In the current period, the politician returned to his usage of legal vocabulary, even though it was mostly related to Russian war crimes in Ukraine, with terms like *illegal*, *evidence*, and *crime*. It explains the jump from 0,023% in the second period to 0,240% in the third period.

Economic Terms

Term	Frequency	Relative Frequency
PAY	21	0,220%
PLAN	16	0,167%
ECONOMY	9	0,094%
MILLION	8	0,084%
BILLION	8	0,084%

INVESTMENT	6	0,063%
CUT	5	0,052%
TRADE	4	0,042%
COST	4	0,042%
DEBT	3	0,031%
TAX	3	0,031%
JOBS	3	0,031%
MONEY	3	0,031%
INDUSTRY	3	0,031%
GROWTH	2	0,021%
CAPITAL	2	0,021%
TAXES	2	0,021%
BUSINESS	2	0,021%
INFLATION	2	0,021%
MARKET	1	0,010%
Total	107	1,118%

Economic vocabulary continues to be prevalent in Sunak's speeches, even though not as in previous periods (4,617% in the second period, 1,118% in the third period). Similarly to the state bodies part, the word *pay* was used mostly in a single speech regarding public sector pay review, while *plan* was consistently used throughout the corpora. The economy, in general, has become less significant since Rishi Sunak assumed the role of Prime Minister, as he now addresses a wider range of topics.

Electoral Vocabulary

No election-related terms were identified in the current period.

Military Vocabulary

Term	Frequency	Relative Frequency
SECURITY	17	0,178%
NATO	14	0,146%
DEFENCE	13	0,136%
WAR	11	0,115%
MILITARY	8	0,084%
STRIKES	8	0,084%
ATTACKS	7	0,073%
RANGE	7	0,073%
AIR	6	0,063%

AID	5	0,052%
TANKS	4	0,042%
TROOPS	4	0,042%
VESSELS	4	0,042%
WEAPONS	4	0,042%
COMBAT	4	0,042%
MISSION	3	0,031%
FORCES	3	0,031%
NAVY	3	0,031%
OPERATION	3	0,031%
AMMUNITION	2	0,021%
AUKUS	2	0,021%
Total	132	1,380%

In the third period military terms became noticeably prevalent due to events with Russia (0,094% in the first period, 0,034% in the second period, 1,380% in the third period). Sunak used a multitude of words related to the military, such as (national) *security, NATO, defence, and war*. Also, he mentioned several warfare assets such as *tanks, air (forces), troops, weapons, navy, etc.*, either regarding military aid or suggesting that each element of British armed forces should be strengthened. Moreover, the usage of words *strikes, attacks, combat* are meant to focus public opinion regarding Ukraine and invoke empathy to people living there.

Nouns with Positive Connotation

Term	Frequency	Relative Frequency
SUPPORT	12	0,125%
FREEDOM	9	0,094%
HOME	9	0,094%
SAFETY	7	0,073%
HOPE	6	0,063%
INNOVATION	6	0,063%
PROGRESS	6	0,063%
AID	5	0,052%
OPPORTUNITY	5	0,052%
OPPORTUNITIES	5	0,052%
TRIBUTE	5	0,052%
LEADERSHIP	5	0,052%
TRUST	4	0,042%
CONFIDENCE	3	0,031%
GENEROSITY	2	0,021%
PROSPERITY	2	0,021%

RESPECT	2	0,021%
Total	93	0,972%

As for nouns with a positive connotation, they are still very prevalent in the corpus (0,596% in the first period, 1,133% in the second period, 0,972% in the third period). However, this time the purpose of such terms is closer to invoking patriotism and urgency in the ongoing crisis, rather than trying to improve the general mood in speeches.

Nouns with Negative Connotation

Term	Frequency	Relative Frequency
WAR	11	0,115%
ATTACKS	7	0,073%
RISK	5	0,052%
RISKS	5	0,052%
SACRIFICE	4	0,042%
DEBT	3	0,031%
PROBLEM	3	0,031%
CRIME	2	0,021%
SHORTAGE	2	0,021%
DISRUPTION	2	0,021%
CRISIS	1	0,010%
DESTRUCTION	1	0,010%
DEATH	1	0,010%
Total	47	0,491%

Regarding nouns with a negative connotation, the term *war* was used 11 times and attacks 7 times, which are related to Ukraine as well. Some other negative nouns were encountered such as *risk*, *debt*, *problem*, *crisis*, which are linked to other issues the Prime Minister addressed at the time of the migration crisis and post-coronavirus economy restoration. Generally, he doesn't use such nouns unless he must convey the problem to the public. Compared to previous periods, this period has less than average number of nouns with a negative connotation (0,764% in the first period, 0,431% in the second period, 0,491% in the third period).

Adjectives with Positive Connotation

Term	Frequency	Relative Frequency
NEW	24	0,251%
BEST	11	0,115%
CLEAR	10	0,105%
FAIR	10	0,105%
BETTER	10	0,105%
PROUD	8	0,084%
RIGHT	7	0,073%
GREAT	5	0,052%
INCREDIBLE	5	0,052%
MODERN	5	0,052%
SAFE	5	0,052%
SURE	5	0,052%
EXTRAORDINARY	5	0,052%
GOOD	4	0,042%
ABLE	4	0,042%
CREATIVE	4	0,042%
POWERFUL	4	0,042%
FREE	4	0,042%
CONFIDENT	3	0,031%
GRATEFUL	3	0,031%
STRONG	2	0,021%
BRILLIANT	1	0,010%
COURAGEOUS	1	0,010%
HONEST	1	0,010%
NOBLE	1	0,010%
FORTUNATE	1	0,010%
Total	143	1,495%

Rishi Sunak seems to consistently use positively connotated adjectives throughout his speeches. However, such adjectives appear especially often in the last period (0,876% in the first period, 0,635% in the second period, 1,495% in the third period). This implies that he is trying to create a better perception of the changes he made during his premiership. Words *new*, *best*, *fair*, *better*, *modern* all convey such feeling to whatever the politician described in the speech.

Adjectives with Negative Connotation

Term	Frequency	Relative Frequency
ILLEGAL	8	0,084%
DIFFICULT	5	0,052%
WRONG	5	0,052%
HARD	4	0,042%

CRIMINAL	3	0,031%
DANGEROUS	3	0,031%
BAD	2	0,021%
LETHAL	2	0,021%
POOR	2	0,021%
WORSE	2	0,021%
DISRESPECTFUL	1	0,010%
DISRUPTIVE	1	0,010%
HORRIFIC	1	0,010%
IMPOSSIBLE	1	0,010%
TERRIBLE	1	0,010%
TRAGIC	1	0,010%
UNACCEPTABLE	1	0,010%
UNFAIR	1	0,010%
DEVASTATING	1	0,010%
Total	45	0,470%

Adjectives with a negative connotation are used the most frequently in the third period. Yet, the difference between the last and first two periods is not big (0,410% in the first period, 0,244% in the second period, 0,470% in the third period). The term *illegal* appears 8 times in the speech concerning immigration problem, *difficult* and *wrong* – 5 times both sprinkled over multiple speeches. Overall, Rishi Sunak seems to avoid overusing adjectives with a negative connotation, unless he wants to point out some specific issue (as evident in the case with illegal migration).

Verbs with Positive Connotation

Term	Frequency	Relative Frequency
SUPPORT	12	0,125%
HELP	10	0,105%
PROVIDE	10	0,105%
LEAD	7	0,073%
GROW	7	0,073%
ACHIEVE	5	0,052%
INCREASE	5	0,052%
CREATE	4	0,042%
IMPROVE	4	0,042%
KEEP	4	0,042%
PROTECT	4	0,042%
TRUST	4	0,042%
CELEBRATE	3	0,031%
ACT	3	0,031%

COMMIT	2	0,021%
INNOVATE	2	0,021%
LEARN	2	0,021%
STRENGTHEN	2	0,021%
UNITE	2	0,021%
VALUE	2	0,021%
WIN	2	0,021%
WISH	2	0,021%
Total	98	1,024%

The top 3 verbs with a positive connotation are *support*, *help*, and *provide*. It's worth noting that these words are related partly to the aid for Ukraine and other unrelated topics, rather than support for businesses like in the second period. Using various positive verbs, Rishi Sunak displayed a lot of empathy in his latest speeches and made himself known as a strong and compassionate leader. Compared with the previous two periods, this time, the usage of such verbs is right in the middle (0,522% in the first period, 1,502% in the second period, 1,024% in the third period).

Verbs with Negative Connotation

Term	Frequency	Relative Frequency
SEIZE	2	0,021%
AVOID	1	0,010%
DISAGREE	1	0,010%
ELIMINATED	1	0,010%
ABDUCTED	1	0,010%
Total	6	0,063%

The usage of verbs with the negative connotation remained low, similar to the second period. The motive is likely the same – to evoke a sense of strength and positivity in the politician's speeches.

In the third period, Rishi Sunak's speeches show some consistent patterns and changes in his language use. He still uses philosophical terms at an average rate, with "time" and "freedom" being common, likely to show urgency and talk about current events like the Russian invasion of Ukraine. These terms generally have positive meanings. Sociological and cultural terms are not used much, though there's a slight increase in cultural terms compared to before. Sunak often mentions patriotism-related words like "history" and "values." When talking about organisations, Sunak mainly

focuses on individuals rather than groups. He mentions educational terms very rarely, except when talking about training troops. As Prime Minister, the politician now speaks more to the general public than to his colleagues, so he mentions job titles less often. He still talks a lot about state bodies like the NHS and the government, especially during times of national security concerns. Military terms become more common due to events like the conflict in Ukraine. The prime minister emphasises the need to strengthen the UK's armed forces. Positive nouns are used to invoke patriotism and urgency, while negative ones are mentioned less often, usually to highlight specific issues. Rishi Sunak continues to use positive adjectives, especially in his latest speeches, to create a positive image of his leadership. Negative adjectives are used carefully, usually to point out specific issues. Furthermore, positive verbs like "support" and "help" are common, showing empathy and strength in his speeches. In conclusion, Sunak's language reflects his efforts to address current events, show leadership, and inspire confidence in his actions.

Based on the words of the third period, it can be understood that with the post of Prime Minister, Rishi Sunak's attention is now focused on international issues that require him to take a clear position. The fact that he addresses individuals, not just political representatives, confirms that he has his own opinion and authority among the community.

CONCLUSION

Language, as a constantly evolving form of communication, has gained immense importance, particularly in our globalised world where borders between nations are fading. It serves as a crucial tool for bridging cultural divides and facilitating understanding across diverse societies.

The political sphere remains crucial in the 21st century due to its profound impact on shaping societies, economies, and global affairs. In an era characterised by rapid technological advancements, interconnectedness, and complex challenges such as climate change and global health crises, political decisions hold significant sway over numerous aspects of life. This fact further leads us to the notion of political discourse which involves analysing how politicians communicate their ideas and persuade others.

We arrived at the conclusion that political discourse is an interdisciplinary field, influenced by linguistic, cultural, and sociopolitical factors. We analysed various definitions and perspectives on political discourse, ranging from its role in shaping public opinion to its function as a tool for power negotiation and communication in political contexts. Additionally, we explored the broad and narrow understandings of political discourse, considering its manifestations in both formal political institutions and civil society organisations. Through this detailed analysis, our study aimed to provide insights into the complexities of political communication and its implications for contemporary society.

Political discourse is intricately linked with language, serving as a tool for power and persuasion. It involves specific strategies like appeals, promises, and slogans to sway opinions. It's characterised by ambiguity, emotional language, and abstract concepts, enabling effective manipulation of information. It targets both explicit and implicit audiences, with mass media playing a vital role in its dissemination and shaping public opinion.

We also delved deeper into the functions of political discourse which aim to inform, influence, and agitate the intended audience, legitimising power structures while delegitimising opposition. Through coercion, resistance, protest, and simulation,

it controls perceptions and shapes agendas. It disseminates information, projects into the past and future, and fosters social control, solidarity, and differentiation. Political discourse also mobilises populations, implements policies, and projects past experiences into future arguments.

A linguistic personality is a multifaceted phenomenon, a combination of various skills and abilities of language communication, a set of linguistic abilities and intentions to perform certain linguistic actions that can be systematised both by types of speech activity and according to the levels of the language system. The study highlighted its manifestation in speech, projecting national-cultural specificity.

The analysis of the linguistic personality of Rishi Sunak at the lexical level, a young and knowledgeable politician, was interesting for several reasons. Analysing his language and communication style improved understanding of the way he interacted with the public and political colleagues. The study of Sunak's linguistic personality also helped to reveal his values and views on politics, as well as their expression through language.

The analysis conducted using WordSmith Tools allowed for the accurate empirical data regarding the usage of lexical means by Rishi Sunak.

Initially, the process involved generating wordlists for different periods of Sunak's career, which helped identify commonly used English words and their frequency in his speeches. Additionally, the KeyWords tool was employed to analyse the data and identify significant words, with reference to the British National Corpus.

In the list of keywords from the three periods, we found an increased frequency of use of the pronouns "I" and "we" in the later periods. This can be explained by the politician's career progression over several years, which gave him more opportunities to express his personal opinion. The connection between the consistently high use of the pronoun "our" and Rishi Sunak's Indian origin, which is a collectivist culture, was established.

In addition, the main key themes in each period of the politician's speeches were identified. At the beginning of his work in the government, various topics were mentioned, such as the problem of hunting local animals or the development of the

Northern economy, which was manifested in the keywords *shooting, grouse, debt, northern*. During his tenure as Chancellor of the Exchequer, the focus shifted to Britain's economic development during the pandemic, which was evident in the dominance of the following vocabulary: *jobs, businesses, million, coronavirus, invest*. In the third period, as Prime Minister, the range of issues mentioned in the politician's speeches expanded again. These comprised the global scope such as security of the UK's allies or the development of digital technologies in the country, exemplified in the words *world, security, NATO, tech, AI, Maths*.

Rishi Sunak's early speeches showed that he often talked about "opportunity" and emphasised family and heritage values to build the nation. He mentioned alliances like the EU, focusing on UK strategy and Brexit. Sunak's tone was polite, especially with officials, and he talked a lot about "police" and "government." He talked about trade internationally, aiming for economic diplomacy and UK influence. His speeches balanced positive and negative ideas, showing hope while recognising challenges. During the first period Sunak's speeches were hopeful and realistic, calling for action to develop the country and earning trust by sharing its values.

Concerning the second period, the politician emphasised terms like "time" and "opportunity" to convey urgency and potential for action in addressing political issues. Sunak adopted a more precise approach, relying less on philosophical terms and more on concrete facts, especially in economic contexts, amidst the challenges posed by the COVID-19 crisis. His speeches were dominated by economic terms, reflecting efforts to stimulate recovery. A positive tone was maintained through the use of words with a positive connotation, indicating confidence in his policies. Sunak actively mentioned state bodies and political parties to appeal to various stakeholders and position himself politically. This adaptation of language and messaging demonstrated Sunak's strength, competence, and optimism to the public, reflecting his flexibility in adjusting his communication style to his role as Chancellor.

In his third period, Rishi Sunak's speeches maintained philosophical terms like "time" and "freedom" to address current events, such as the Russian invasion of Ukraine. Cultural references, particularly patriotism-related words, increased slightly.

The important feature became the focus on individuals over groups when discussing organisations. The fact that he consistently addressed individuals, not just political representatives, confirmed that he had his own opinion and authority among the community. Positive language prevailed, emphasising patriotism and urgency, while negative terms were almost not used.

SUMMARY

Мова не існує поза культурою. Вона є однією з найважливіших її складових, формою мислення, специфічним проявом людського життя, яке, в свою чергу, є справжньою сутністю мови. Активні дослідження мови в різних сферах, особливо в політиці, призвели до появи поняття "мовна особистість". У європейській лінгвістиці 19 століття питання мовної особистості постало в контексті вивчення соціальної природи мови, взаємозв'язку між мовою і мовленням, мовою індивіда і колективу.

Вивчення мовного профілю Ріші Сунака є важливим у контексті розуміння його політичної стратегії та методів комунікації. Ріші Сунак відомий як впливовий політик і провідний економіст, тому аналіз його мовного профілю може розкрити ключові аспекти його політичного підходу і стилю керівництва.

Актуальність такого дослідження полягає в тому, що воно дозволить нам краще зрозуміти, як Ріші Сунак використовує мову для спілкування з громадськістю, які ефективні мовні засоби він використовує для переконання та мобілізації своєї аудиторії, а також які ідеї та цінності він приймає та просуває. Таке дослідження може стати важливим доповненням до загального розуміння політичної культури та динаміки сучасного британського політичного ландшафту.

Мета дослідження - розкрити мовний профіль Ріші Сунака як сучасного мультикультурного британського політика та визначити його ключові цінності, ідеологію та управлінські стратегії в контексті політичної діяльності.

Для досягнення цієї мети було необхідно вирішити такі **основні завдання**:

- розглянути поняття політичного дискурсу та його принципи
- розкрити поняття мовної особистості та способи її реалізації в політичній культурі
- порівняти лексичні аспекти індивідуального стилю Ріші Сунака протягом трьох періодів його кар'єри, а саме початку роботи в уряді, канцлерства та прем'єрства

- визначити ключові концепти та цінності, якими керується Ріші Сунак на політичній арені, і те, як вони резонують із цінностями Великої Британії як країни.

Об'єктом дослідження є мовна особистість Ріші Сунака та рамки його мовленнєвої поведінки в дискурсивному просторі політики.

Предметом дослідження є лексичні засоби вербалізації мовної особистості Ріші Сунака.

Матеріалом дослідження слугували 22 промови Ріші Сунака, зібрані в корпус. Для забезпечення об'єктивності дослідження та простеження розвитку його особистості за основу було взято три періоди його кар'єри: початок роботи в уряді (2015 – 2018 роки), канцлерство (2019 – 2022 роки) та прем'єрство (2022 – по теперішній час). Загальний обсяг проаналізованого матеріалу становить 32500 слів (160000 знаків).

Основним **методом дослідження** був контент-аналіз. Крім того, були використані кількісний метод (для підрахунку частоти вживання мовних одиниць у промовах Сунака), порівняльний метод (для простеження спільних і відмінних рис у промовах за три періоди його кар'єри). Ці методи були реалізовані за допомогою програмного забезпечення Wordsmith Tools. Зокрема, ми використовували функцію WordList (для створення частотних списків за всі періоди та їх порівняння), функцію Keywords (для визначення слів, які є найбільш незвично частотними в даному корпусі), функцію Concord (для виявлення зв'язків між частотою, контекстом, сполучуваністю мовних елементів).

Новизна дослідження мовного профілю Ріші Сунака полягає в кількох ключових аспектах. По-перше, аналіз його мовного стилю може виявити унікальність комунікативного підходу нового політичного актора. По-друге, враховуючи його урядову посаду, дослідження може виявити, як він використовує мову для репрезентації урядової політики. По-третє, аналіз його мовного профілю може виявити, як він адаптує своє мовлення до культурного та

соціального контексту. Нарешті, вивчення мовних стратегій дозволяє зрозуміти, як Ріші Сунак впливає на громадську думку і формує свій політичний імідж. Таке дослідження виявляється ключовим для розуміння сучасного політичного ландшафту та ролі комунікації в ньому.

Практичне значення дослідження полягає в тому, що воно дає унікальні уроки політичної комунікації та сприяє розвитку комунікативних навичок політиків і громадських діячів. Аналіз мовного профілю також допомагає краще зрозуміти культурний і соціальний контексти діяльності політика, що сприяє побудові взаєморозуміння в суспільстві.

Структура дослідження. Робота складається зі вступу, трьох розділів, загальних висновків, списку використаних джерел (102 позиції, з них 86 позицій іноземними мовами), списку джерел даних, додатку та анотації українською та англійською мовами. Загальний обсяг роботи становить 118 сторінок, обсяг основного тексту - 95 сторінок.

Проведений аналіз дозволяє зробити такі висновки.

Ранні промови Ріші Сунака показали, що він часто говорив про "можливості" і підкреслював цінності сім'ї та культурної спадщини для розбудови нації. Він згадував альянси, такі як ЄС, зосереджуючись на стратегії Великої Британії та її виходу з ЄС. Ріші Сунак зосереджував свою увагу на міжнародній торгівлі, прагнучи до економічної дипломатії та впливу Великої Британії. Його промови показали збалансований підхід до лексики з позитивною та негативною конотацією, демонструючи надію, але водночас визнаючи виклики. У перший період (2015-2018 роки) промови Сунака були сповнені надії та реалізму, закликаючи до дій для розвитку країни та завоювання довіри, поділяючи її цінності.

Що стосується другого періоду (2019 – 2022 роки), то політик наголошував на таких термінах, як "час" і "можливість", щоб передати нагальність і потенціал для дій у вирішенні політичних питань. Сунак застосував більш точний підхід, покладаючись менше на філософські терміни, а більше на конкретні факти, особливо в економічному контексті, на тлі викликів, спричинених кризою

COVID-19. У його промовах переважали економічні терміни, що відображало зусилля зі стимулювання відновлення. Позитивний тон підтримувався завдяки використанню слів з позитивною конотацією, що свідчить про впевненість у його політиці. Сунак активно згадував державні органи та політичні партії, щоб апелювати до різних зацікавлених сторін та політично позиціонувати себе. Така адаптація мови та меседжів демонструвала громадськості силу, компетентність та оптимізм Сунака, відображаючи його гнучкість у пристосуванні свого стилю спілкування до ролі канцлера.

У третій період (2022 – теперішній час) у промовах Ріші Сунака збереглися філософські терміни, такі як "час" і "свобода", які стосуються поточних подій, таких як російське вторгнення в Україну. Дещо збільшилася кількість культурних посилань, зокрема, слів, пов'язаних з патріотизмом. Важливою особливістю став фокус на індивідуумах, а не на групах при обговоренні організацій. Той факт, що він постійно звертався до окремих осіб, а не лише до політичних представників, підтверджує, що він має власну думку та авторитет серед громади. Переважала позитивна лексика, що підкреслювала патріотизм і нагальність, тоді як негативні терміни майже не використовувалися.

Аналіз, проведений за допомогою WordSmith Tools, надав цінну інформацію про політичні промови Ріші Сунака, висвітлюючи різні аспекти його мовних одиниць і стилю спілкування.

Спочатку процес включав створення списків слів для різних періодів кар'єри Сунака, що допомогло визначити загальноповживані англійські слова та їхню частоту в його промовах. Крім того, для аналізу даних і виявлення значущих слів було використано інструмент KeyWords, який посилається на Британський національний корпус.

Виділені ключові слова дали уявлення про теми, які порушував Сунак, починаючи від екологічних питань і закінчуючи економікою та фінансами. Крім того, методи візуалізації, такі як хмари тегів (хмари слів, або зважений список, представлений візуально), дали чітке уявлення про основні теми в різні періоди його кар'єри.

Аналіз займенників виявив зміни в стилі спілкування Сунака, зокрема збільшення використання особистих займенників, таких як "я" і "ми", що свідчить про перехід до більш особистих думок і колективного лідерства. Аналіз відповідності виявив такі закономірності в його мові, як часте використання дієслів дії для передачі рішучості та стратегічне використання сполучників, таких як "і", для логічного структурування.

Насамкінець, аналіз за допомогою WordSmith Tools дав всебічне розуміння мовленнєвих форм Сунака, ключових слів і комунікаційних стратегій, надаючи цінну інформацію про етапи його політичної кар'єри і стиль лідерства.

REFERENCES

1. Бацевич Ф.С. Словник термінів міжкультурної комунікації. К.: Довіра, 2007. 205 с.
2. Волошук, В. І. (2008). Індивідуальний авторський стиль, ідіолект, ідіостиль: питання термінології. Наукові праці, 79, 5–9.
3. Єрмоленко, С. Я. (2010). Формування української мовної особистості. Українознавство, 1, 120–123.
4. Загнітко, А. (2017). Linguopersonology Theory: Monograph. Vinnytsia: Nilan Ltd.
5. Клименко І.В. Теоретичні засади лінгвістичного аналізу політичного дискурсу / І.В. Клименко // Лінгвістичні студії: зб. наук. праць. – Донецьк: ДонНУ, 2009. – Вип. 19. – С. 182–186.
6. Кондратенко Н. В. Український політичний дискурс: Текстуалізація реальності: Монографія. Одеса: Чорномор'я, 2007. 156 с.
7. Кунцевич В.Е. Психологічні характеристики та складові політичного дискурсу / В.Е. Кунцевич // . – 2001. – № 5 (10). – С. 30-72.
8. Мацько, Л. І. (2014). Мовна особистість Тараса Шевченка як чинник формування національної ідентичності. Наукові записки Національного університету «Острозька академія», 50, 8–15.
9. Нагорна Л. Б. Політична мова і мовна політика: діапазон можливостей політичної лінгвістики / НАН України; Інститут політичних і етнонаціональних досліджень. К.: Світогляд, 2005. 316 с.
10. Павлуцька, В. О. (2008) Політичний дискурс: особливості та функції. ВІСНИК Житомирського державного університету імені Івана Франка (39). pp. 218-221.
11. Пентиліук, М. (2014). Мовна особистість майбутнього вчителя-словесник в контексті професійної підготовки. Збірник наукових праць, 2, 290–297.
12. Прокопенко Б. В. Предвиборчий американський дискурс: головні особливості та семантично-структурні та прагматично-комунікаційні характеристики: / Б. В. Прокопенко. – Харків, 2008. – 72 с.

13. Серажим К.С. Дискурс як соціолінгвальне явище: методологія, архітектоніка, варіативність: (на матеріалах сучасної газетної публіцистики) / К.С. Серажим. – К.: Київ. нац. ун-т імені Тараса Шевченка, 2002. – 392 с.
14. Славова Л. Л. Мовна особистість у процесі ведення українського та американського політичного дискурсу – типологічне й історичне мовознавство / Л. Л. Славова. – К., 2008. – 89 с.
15. Славова, Л. Л. (2018). Віртуальний образ мовної особистості. Мовні і концептуальні картини світу, 2, 240–245.
16. Фоменков П. Д. Дискурсивний аналіз сучасного явища дискурсу США, 1999. – 59 с.
17. Albakry, M. (2004). U.S. “Friendly Fire” Bombing of Canadian Troops: Analysis of the Investigative Reports. *Critical Inquiry in Language Studies*, 1(3), 163–178.
18. AntLab Software. (n.d.). AntLab Software. Retrieved from www.antlab.sci.waseda.ac.jp/software.html.
19. Bachmann, I. (2011). Civil partnership – “gay marriage in all but name”: a corpus-driven analysis of discourses of same-sex relationships in the UK Parliament. *Corpora*, 6(1), 77–105.
20. Baker, P. (2006). *Using corpora in discourse analysis*. London, New York: Continuum.
21. Barthes, R. (1988). *The Semiotic Challenge* (R. Howard, Trans.). Hill and Wang.
22. Bayley, P. (2005). *Analysing Language and Politics*. *Mediazioni: Online Journal of Interdisciplinary Studies of Language and Cultures*. DOI: 10.1473/media15.
23. Bergsdorf, W., & Corcoran, P. (1978). *Politik und Sprache*. Political science, 186 p.
24. Bondi, M., & Scott, M. (2010). *Keyness in texts*. Amsterdam, Philadelphia: John Benjamins Publishing.

25. Breeze, R. (2011). Critical Discourse Analysis and Its Critics. *Pragmatics*, 21(4), 493–525.
26. British National Corpus. Retrieved from <http://www.natcorp.ox.ac.uk/corpus/index.xml>.
27. British National Corpus (Baby Edition). Retrieved from <https://lids.ling-phil.ox.ac.uk/lids/xmlui/handle/20.500.14106/2553>.
28. Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
29. Bühler, K. (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Stuttgart: Gustav Fischer Verlag.
30. Chen, L. (2012). Reporting news in China: Evaluation as an indicator of change in the China Daily. *China Information*, 26(3), 303–329.
31. Chilton, P. (2004). *Analysing political discourse: Theory and practice*. London: Routledge.
32. Chilton, P. (2005). Missing links in mainstream CDA: Modules, blends, and the critical instinct. In R. Wodak & P. Chilton (Eds.), *A New Agenda in (Critical) Discourse Analysis: Theory, methodology and interdisciplinarity* (pp. 19–51). Amsterdam: John Benjamins.
33. Chilton, P., & Schaffner, Ch. (Eds.). (2002). *Politics as Text and Talk: Analytic Approaches to Political Discourse* (4th ed.). *Discourse Approaches to Politics, Society and Culture*. John Benjamins.
34. Corcoran, P. (1979). *Political language and rhetoric*. University of Queensland Press, St. Lucia, Queensland, 216 p.
35. Degano, C. (2007). Dissociation and Presupposition in Discourse: A Corpus Study. *Argumentation*, 21(4), 361–378.
36. Denton, R. E. Jr., & Woodward, G. C. (1985). *Political communication in America*. New York: Praeger.
37. Edwards, G. O. (2012). A comparative discourse analysis of the construction of ‘in-groups’ in the 2005 and 2010 manifestos of the British National Party. *Discourse & Society*, 23(3), 245–258.

38. Forchtner, B., & Kolvraa, C. (2012). Narrating a 'new Europe': From 'bitter past' to self-righteousness? *Discourse & Society*, 23(4), 377–400.
39. Foucault, M. (1972). *Archaeology of Knowledge and the Discourse on Language* (A. M. Sheridan Smith, Trans.). Pantheon Books.
40. Freake, R., Gentil, G., & Sheyholislami, J. (2010). A bilingual corpus-assisted discourse study of the construction of nationhood and belonging in Quebec. *Discourse & Society*, 22(1), 21–47.
41. Gabrielatos, C., & Baker, P. (2008). Fleeing, Sneaking, Flooding: A Corpus Analysis of Discursive Constructions of Refugees and Asylum Seekers in the UK Press, 1996-2005. *Journal of English Linguistics*, 36(1), 5–38.
42. GeeksforGeeks. HTML Introduction. Retrieved from <https://www.geeksforgeeks.org/html-introduction/>.
43. Gov.uk. Retrieved from <https://www.gov.uk/>.
44. Graber, D. (1981). Political Languages. *Handbook of Political Communication*, 14, 1–30.
45. Habermas, J. (1998). *Between Facts and Norms. Contribution to a Discourse Theory of Law and Democracy* / Tr. by W. Rehg. Cambridge, MA: The MIT Press.
46. Halliday, M., & Matthiessen, C. (2004). *An introduction to functional grammar*. London, New York: Arnold.
47. Hamilton, C., Adolphs, S., & Nerlich, B. (2007). The meanings of 'risk': a view from corpus linguistics. *Discourse & Society*, 18(2), 163–181.
48. Harris, Z. (1981). Discourse Analysis. *Papers on Syntax. Synthese Language Library*, 14, 1–30.
49. Herbel-Eisenmann, B., & Wagner, D. (2010). Appraising lexical bundles in mathematics classroom discourse: obligation and choice. *Educ Stud Math*, 75(1), 43–63.
50. Hidalgo Tenorio Encarnacion (2011) Critical Discourse Analysis, An overview. "Nordic Journal of English Studies", vol. 10, no. 1, s. 184–210

51. Jakobson, R. (1960). Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in language* (pp. 350–377). Cambridge, MA: MIT Press.
52. Kamasa, V. (2013a). Naming “In Vitro Fertilisation”: Critical Discourse Analysis of the Polish Catholic Church’s Official Documents. *Procedia – Social and Behavioural Sciences*, 95, 154–159.
53. Kirvalidze, N., & Samnidze, N. (2016). Political Discourse as a Subject of Interdisciplinary Studies. *Journal of Teaching and Education*, 5, 161–170.
54. Koller, V. (2004). Businesswomen and war metaphors: ‘Possessive, jealous and pugnacious’? *Journal of Sociolinguistics*, 8(1), 3–22.
55. Lindenberg, F. (2012). *The Data Journalism Handbook*. Retrieved from <https://texty.org.ua/archive-books/40161/zhurnalistyka-danykh-posibnyk-40161/#42409>.
56. Lischinsky, A. (2011). In times of crisis: a corpus approach to the construction of the global financial crisis in annual reports. *Critical Discourse Studies*, 8(3), 153–168.
57. Louw, B. (1993). Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies. In M. Baker (Ed.), *Text and Technology* (pp. 157–176). Amsterdam: John Benjamins.
58. Lukac, M. (2011). Down to the bone: A corpus-based critical discourse analysis of pro-eating disorder blogs. *Jezikoslovlje*, 12.2, 187–209.
59. Mautner, G. (2005). The Entrepreneurial University. *Critical Discourse Studies*, 2(2), 95–120.
60. McEnery, T., & Wilson, A. (2001). *Corpus linguistics: An introduction*. Edinburgh: Edinburgh University Press.
61. Mohamad Subakir, M. Y., et al. (2012). Linguistic Sexism In Qatari Primary Mathematics Textbooks. *GEMA Online Journal of Language Studies*, 12(1), 53–68.
62. Mulderrig, J. (2011). Manufacturing Consent: A corpus-based critical discourse analysis of New Labour’s educational governance. *Educational Philosophy and Theory*, 43(6), 562–578.

63. Orpin, D. (2005). Corpus Linguistics and Critical Discourse Analysis: Examining the ideology of sleaze. *International Journal of Corpus Linguistics*, 10(1), 37–61.
64. Politics.co.uk. (n.d.). Rishi Sunak. Retrieved from <https://www.politics.co.uk/reference/rishi-sunak/>.
65. Rayson, P. (2009). Wmatrix: a web-based corpus processing environment. Computing Department, Lancaster University. Retrieved from <http://ucrel.lancs.ac.uk/wmatrix>.
66. Rishi Sunak. Retrieved from <https://www.rishisunak.com/>.
67. Rishi Sunak. YouTube. Retrieved from <https://www.youtube.com/@RishiSunak>.
68. Salama, A. H. Y. (2011). Ideological collocation and the recontextualisation of Wahhabi-Saudi Islam post-9/11: A synergy of corpus linguistics and critical discourse analysis. *Discourse & Society*, 22(3), 315–342.
69. Scott, M. (2024). Lexically.net. Retrieved from <https://lexically.net/downloads/version9/HTML/index.html>.
70. Seidel, G. (1985). Political Discourse Analysis. In T. A. van Dijk (Ed.), *Handbook of Discourse Analysis* (Vol. 4, pp. 43–60). London: Academic Press.
71. Sériot, P. (2012). *Structure and the Whole: East, West and Non-Darwinian Biology in the Origins of Structural Linguistics* (A. Jacobs-Colas, Trans.). Mouton de Gruyter.
72. Stubbs, M. (1997). Whorf's Children: Critical comments on Critical Discourse Analysis (CDA). In A. Ryan & A. Wray (Eds.), *Evolving models of language* (pp. 100–116). Clevedon: British Association for Applied Linguistics.
73. Study Smarter. Rishi Sunak. Retrieved from <https://www.studysmarter.co.uk/explanations/politics/uk-government/rishi-sunak/>.

74. Subtirelu, N. C. (2013). 'English... it's part of our blood': Ideologies of language and nation in United States Congressional discourse. *Journal of Sociolinguistics*, 17(1), 37–65.
75. Toppan Digital. Understanding the 6 dimensions of UK culture. Retrieved from <https://toppandigital.com/us/blog-usa/understanding-6-dimensions-uk-culture/>.
76. van Dijk, T. A. (1997). What is political discourse analysis? Retrieved from <https://pdfs.semanticscholar.org/680a/8ba68935487bcb63aecc68b62be8fb216de0.pdf?ga=2.165676030.1768819635.1578304873-484743753.1578304873>
77. Weninger, C. (2010). The lexico-grammar of partnerships: corpus patterns of facilitated agency. *Text & Talk – An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 30(5), 591–613.
78. Whitsitt, S. (2005). A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics*, 10(3), 283–305.
79. Wodak, R. (Ed.). (1989). *Language, power, and ideology: Studies in political discourse*. Amsterdam: John Benjamins.
80. W3C. HTML Standard. Retrieved from <https://www.w3.org/MarkUP/draft-ietf-iiir-html-01.txt>.

DATA SOURCES

81. Sunak, R. (2015, June 11). My maiden speech in Westminster. [YouTube Video]. Retrieved from <https://www.youtube.com/watch?v=JoQUqlH4eRg&t=30s>
82. Sunak, R. (2016, November 2). Defending the rural economy. [Facebook Video]. Retrieved from <https://www.facebook.com/rishisunak/videos/defending-the-rural-economy/1806207469648938/>
83. Sunak, R. (2016, February 26). Why I will vote for Britain to leave the EU. Retrieved from <https://www.rishisunak.com/news/why-i-will-vote-britain-leave-eu-0>

84. Sunak, R. (2017, November 27). Budget debate. [YouTube Video]. Retrieved from <https://www.youtube.com/watch?v=iT-RD-FeqZA>
85. Sunak, R. (2017, March 9). Westminster Hall debate on the Coast to Coast Walk. [YouTube Video]. Retrieved from <https://www.youtube.com/watch?v=AKqjYlkbIsI>
86. Sunak, R. (2017, November 6). Transport for the North. [YouTube Video]. Retrieved from <https://www.youtube.com/watch?v=3TDdEkDfKm4>
87. Sunak, R. (2017, April 19). Supporting measures to help small businesses grow. [YouTube Video]. Retrieved from <https://www.youtube.com/watch?v=cPkU3u19JxI>
88. Sunak, R. (2019, October 2). Party Conference Speech: big investment in our public services. Retrieved from <https://www.rishisunak.com/news/party-conference-speech-big-investment-our-public-services>
89. Sunak, R. (2020, July 29). Summer Statement: Plan for Jobs. [YouTube Video]. Retrieved from <https://www.youtube.com/watch?v=Z3TUdYxr95E>
90. Sunak, R. (2020, March 11). Budget Speech 2020. Retrieved from <https://www.gov.uk/government/speeches/budget-speech-2020>
91. Sunak, R. (2021, March 3). Full Speech: Budget 2021. [YouTube Video]. Retrieved from <https://www.youtube.com/watch?v=pzGCXznuza4>
92. Sunak, R. (2022, October 25). Rishi Sunak's first speech as Prime Minister. Retrieved from <https://www.gov.uk/government/speeches/prime-minister-rishi-sunaks-statement-25-october-2022>
93. Sunak, R. (2023, April 17). PM speech on improving attainment in mathematics. Retrieved from <https://www.gov.uk/government/speeches/pm-speech-on-improving-attainment-in-mathematics-17-april-2023>
94. Sunak, R. (2023, May 6). Prime Minister's words to mark the Coronation of King Charles III and Queen Camilla. Retrieved from <https://www.gov.uk/government/news/prime-ministers-words-to-mark-the-coronation-of-king-charles-iii-and-queen-camilla>

95. Sunak, R. (2023, May 22). PM statement to the House of Commons on G7 summit. Retrieved from <https://www.gov.uk/government/speeches/pm-statement-to-the-house-of-commons-on-g7-summit-22-may-2023>
96. Sunak, R. (2023, June 5). PM statement on illegal migration delivery update. Retrieved from <https://www.gov.uk/government/speeches/pm-statement-on-illegal-migration-delivery-update-5-june-2023>
97. Sunak, R. (2023, June 12). PM London Tech Week speech. Retrieved from <https://www.gov.uk/government/speeches/pm-london-tech-week-speech-12-june-2023>
98. Sunak, R. (2023, June 30). PM speech launching the NHS Long Term Workforce Plan. Retrieved from <https://www.gov.uk/government/speeches/pm-speech-launching-the-nhs-long-term-workforce-plan-30-june-2023>
99. Sunak, R. (2023, July 12). PM statement to leaders at NATO Summit in Lithuania. Retrieved from <https://www.gov.uk/government/speeches/pm-statement-to-leaders-at-nato-summit-in-lithuania-12-july-2023>
100. Sunak, R. (2023, July 13). PM statement on public sector pay review. Retrieved from <https://www.gov.uk/government/speeches/pm-statement-on-public-sector-pay-review-13-july-2023>
101. Sunak, R. (2023, November 8). PM statement on protests following meeting with Metropolitan Police Commissioner. Retrieved from <https://www.gov.uk/government/news/pm-statement-on-protests-following-meeting-with-metropolitan-police-commissioner-8-november-2023>
102. Sunak, R. (2024, January 15). PM statement on defending the UK and its allies. Retrieved from <https://www.gov.uk/government/speeches/pm-statement-on-defending-the-uk-and-its-allies-15-january-2024>

APPENDIXES

Appendix A. Word cloud of the 1st period



Appendix D. Python script for XML parsing of the reference corpus

```
import xml.etree.ElementTree as ET
import os

def list_files(directory):
    entries = os.listdir(directory)
    files = [os.path.join(directory, entry) for entry in entries if
os.path.isfile(os.path.join(directory, entry)) and entry.endswith('.xml')]

    return files

def extract_values_from_xml(xml_file_path):
    tree = ET.parse(xml_file_path)
    root = tree.getroot()

    lines = []

    s_tags = root.findall('.//s')

    for s in s_tags:
        w_tags = s.findall('.//w')
        line = ' '.join([elem.text for elem in w_tags if elem.text])

        line = line.replace("`", "\\`")
        line = line.replace(" \\", "\\")
        line = line.replace(" n\\", "n\\")

        lines.append(line)

    return lines

def write_to_text_file(output_file_path, lines):
    with open(output_file_path, 'w', encoding='utf-8') as f:
        for line in lines:
            f.write(line + '\n')

def do_folder(xml_folder):
    files = list_files(xml_folder)

    for xml_file_path in files:
        base_name = xml_file_path.rsplit('.', 1)[0]
        output_file_path = f'{base_name}.txt'
        values = extract_values_from_xml(xml_file_path)
        write_to_text_file(output_file_path, values)
        print(f'Done {output_file_path}')

if __name__ == '__main__':
```

```
do_folder('Texts\\aca')
do_folder('Texts\\dem')
do_folder('Texts\\fic')
do_folder('Texts\\news')
```

Appendix E. Python script for scraping speeches at gov.uk

```
import requests
from bs4 import BeautifulSoup
import os
import time

directory = './\output'

def clean_filename(filename):
    forbidden_chars = ['\\', '/', ':', '*', '?', '"', '<', '>', '|']
    for char in forbidden_chars:
        filename = filename.replace(char, '')
    return filename

def scrape_and_save_content(url):
    response = requests.get(url)

    if response.status_code == 200:
        soup = BeautifulSoup(response.content, 'html.parser')
        gov_speak_divs = soup.find_all('div', class_='gov_speak')
        h1_tag = soup.find('h1').get_text(strip=True)

        filename = clean_filename(h1_tag) + '.txt'
        filepath = os.path.join(directory, filename)

        with open(filepath, 'w', encoding='utf-8') as f:
            for div in gov_speak_divs:
                p_tags = div.find_all('p')
                for p in p_tags:
                    f.write(p.get_text() + '\n')

        print(f'Contents from {url} saved to {filename}')
    else:
        print(f'Failed to fetch the URL {url}. Status code:', response.status_code)

urls = ['https://www.gov.uk/government/speeches/pm-london-tech-week-speech-12-june-2023',
        'https://www.gov.uk/government/speeches/pm-statement-on-public-sector-pay-review-13-july-2023',
        'https://www.gov.uk/government/speeches/pm-statement-on-defending-the-uk-and-its-allies-15-january-2024']

if not os.path.exists(directory):
    os.makedirs(directory)

for url in urls:
    scrape_and_save_content(url)
    time.sleep(1) # to prevent too many requests at once
```