NATIONS UNITED THROUGH DISCOURSE: A CORPUS ANALYSIS OF UN GENERAL
ASSEMBLY ADDRESSES

By Nicole Brun-Mercer

A Dissertation

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in Applied Linguistics

Northern Arizona University

May 2018

Approved:

Douglas Biber, Ph.D., Co-chair

Randi Reppen, Ph.D., Co-chair

William Crawford, Ph.D.

Stephen Wright, Ph.D.

ABSTRACT

NATIONS UNITED THROUGH DISCOURSE:
A CORPUS ANALYSIS OF UN GENERAL ASSEMBLY ADDRESSES

NICOLE BRUN-MERCER

The overarching goal of this dissertation is to investigate the linguistic features of United Nations (UN) General Assembly (GA) General Debate addresses delivered in English during 2015. In order to describe UNGA addresses with the greatest accuracy, breadth, and depth, the study explores typical linguistic features across texts as well as variation within the register. Thus, the dissertation also examines the geographic, political, economic, and social profile of the countries represented in the UNGA in order to identify any country-related characteristics associated with systematic linguistic variation. The dissertation uses two corpus-informed methodologies: register analysis and Corpus-Assisted Discourse Studies (CADS), with a secondary goal of discerning the ways in which methodology affects the nature of findings.

The study is based on a corpus of all 92 UNGA addresses delivered in English during the General Debate in 2015. In the examination of linguistic features across UNGA addresses, the UN texts are contrasted with the British National Corpus as well as four comparable registers: prepared speeches, spontaneous speeches, official documents, and conversation. The analysis begins with a description of situational characteristics for the UNGA and the four comparison registers. Then keywords and Multi-Dimensional scores for the five principal dimensions found in Biber (1988) reveal the lexical and grammatical features that characterize the UNGA. The second part of the investigation looks at variation among UNGA addresses, with texts divided into groups based on country-specific characteristics such as region and Gross Domestic Product. Keyword analysis, key feature analysis, and Multi-Dimensional analysis are used to identify

systematic variation of linguistic features based on country group. The final analysis is a CADS study of three frequent and politically salient terms (*terrorism, Security Council,* and *women*) with keyword and collocational analyses to determine how the terms are conceptualized generally in the UNGA and also whether any variation can be identified for country groups.

Results show a remarkable consistency across the UNGA not only in what lexical and grammatical features are preferred and dispreferred but also in how they are used. Lexical features reveal the importance of topic, purpose, and discourse structure. Grammatical features are greatly influenced by written-to-be-spoken production circumstances and purpose. In spite of the large number of features (126) and country groups (31) explored, very few exhibit any systematic patterns of variation. The variation that does occur is primarily associated with level of development and region. The CADS analysis supports the register analysis, showing that UNGA texts tend to conceptualize politically charged terms in similar ways with only occasional variation based on country group. These findings have important implications for the field of political discourse analysis, furthering our understanding of the linguistic features of international organization spoken discourse. In addition, the CADS analysis demonstrates some important methodological considerations, with consequences for the design of future corpus-informed research.

Nicole Brun-Mercer
© 2018

ACKNOWLEDGEMENTS

I also wish to express my appreciation for the friendship and assistance of my colleagues at NAU. In particular, I would like to thank Meixiu Zhang for her conscientious and skillful coding.

Special thanks go to Cheryl Zimmerman, mentor and friend, who took me all the way to New Zealand to show me I was ready for a doctoral program. Your teaching and guidance have been dedicated, selfless, and inspirational, Rally Monkey.

I am also truly grateful for the help, encouragement, and love of my mother. Thank you for believing in me, for coming to the rescue again and again as emergency babysitter, and for keeping us all fed. To both my mom and stepdad, I am deeply appreciative of the moral and logistical support that you provided on a daily basis and that were essential to the completion of this work. I also owe much to Nedjma, whose energy and devotion to Jake and Sophia gave me the time and space I needed to get this done. Your presence brightened our days.

Finally, I thank my children and husband. Jake, for showing me the moon every night. Sophia, for showing me the work it takes to reach it. Jake, for teaching me the pleasure of reading the same text five times in a row. Sophia, for teaching me the pleasure of throwing the book on the floor and dancing. You have both helped me keep this little project in perspective. My most heartfelt gratitude goes to my husband. Fred, you put your life on hold to make this happen. That sacrifice made it impossible to turn back. I thank you for your pragmatism. I thank you for your vision. I thank you for your love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xxiii

DEDICATION
For my grandmothers.

CHAPTER 1: INTRODUCTION

1.1. Introduction

Political discourse has been the subject of extensive inquiry. The language of government officials is powerful in that it can influence public opinion (Becker, 1967) and offers insight into political action, relationships, and ideologies (van Dijk, 1997). Studies on national political discourse abound (e.g., Baker, 2006; Bhatia, 2009; Cap, 2015; Collet, 2009; Halmari, 2005; L'Hôte, 2009; Morley & Bayley, 2009; Otieno, 2015; Partington, Duguid, & Taylor, 2013; Pujante & Morales-Lopez, 2008; Reyes, 2011; Sowinska, 2013); fewer have been carried out at the global level. Yet international organizations offer a unique communicative context. Nations from around the world, often with competing interests, come together to exchange views and ideally, reach agreements promoting peace, environmental protection, and human rights, to name just a few broad aims. Speakers at a venue such as the United Nations (UN) General Assembly (GA) General Debate are high-level diplomats, foreign ministers, and heads of state and government. Discourse is formalized, from lexical features (e.g., *peacekeeping operations*) to discourse moves (e.g., beginning a speech by congratulating the president of the proceedings), creating a distinct register (Donahue & Prosser, 1997). As such, international organization discourse has been a compelling subject of academic inquiry.

Investigations into the spoken and written texts of international organizations have primarily drawn on Critical Discourse Analysis techniques and other qualitative methods (e.g., Donahue & Prosser, 1997; Duchêne, 2008; Krijtenburg & de Volder, 2015; McEntee-Atalianis, 2013; Valentine & Preston, 2002). These studies have provided valuable qualitative data illustrating how certain metaphors, collocations, or even pronouns reflect ideologies and political relationships within an international organization. The few studies on the discourse of

international organizations to have employed quantitative corpus methods have been restricted to written texts and to a small number of target features (e.g., Baker and McEnery, 2005; D'Acquisto, 2017).

This has limited the types of research questions that have been asked. To date, no comprehensive study employing both quantitative and qualitative methods has been carried out on a wide range of lexical and grammatical features in spoken international organization texts. As a result, little is known about what linguistic features distinguish international organization speeches from comparable registers nor how linguistic features vary within the register. In what ways are the linguistic features of international organization speeches different from those in other types of planned speech? In what ways are the linguistic features of individual or groups of texts within the register of international organization speeches different from other texts in the same register? These questions have yet to be explored in a systematic, thorough manner.

This dissertation addresses that gap by undertaking a comprehensive linguistic description of UNGA addresses. This goal is achieved by (1) identifying and explaining the function of particularly frequent lexical and grammatical forms in UNGA General Debate addresses and (2) determining the ways in which those linguistic forms systematically vary within the UNGA.

In addition, this dissertation explores the influence of methodological decisions in the types of findings that result. Section 1.2 presents a brief overview of three methodologies of relevance to this study of international organization discourse: Critical Discourse Analysis (CDA), Corpus-Assisted Discourse Studies (CADS), and register analysis.

1.2. Critical Discourse Analysis, Corpus-Assisted Discourse Studies, and Register Analysis

Much of the research on political speech has followed a CDA approach, qualitatively examining the linguistic features, rhetorical patterns, or argumentation style of one text or a small collection of texts and interpreting findings through a political, social, or cultural lens to explore issues of power and influence (e.g., Bhatia, 2009; Collet, 2009; Pujante & Morales-Lomez, 2008; Sowinska, 2013). Typically with CDA, a text or small collection of texts is read carefully for passages relating to a concept that the researcher has set out to explore (e.g., terrorism in Bhatia, 2009; globalization in Pardo, 2001). Other CDA studies begin with a broader research question, such as investigating how ideological meaning is constructed in a text (e.g., Pujante & Morales-Lopez, 2008). No quantitative data are provided; instead, findings are presented as text excerpts and interpretation is based on the analyst's understanding of a political, social, or cultural phenomenon. The objective of CDA is not to provide generalizable or even replicable findings, but rather to examine issues of power and ideology in one or just a few texts.

Recent advances in corpus linguistics have allowed discourse analysts to explore similar research goals with a large number of texts, incorporating quantitative analysis in addition to the traditional qualitative analyses of CDA. This growing body of research known as CADS (see Baker, 2006; Partington et al., 2013) investigates political, social, or cultural identity, relationships, and power, typically through a detailed analysis of one or two terms selected by the researcher as promising in terms of prospective findings (e.g., *anti-Americanism* in newspapers from the United States, United Kingdom, and Italy in Partington et al., 2013). Because CADS uses computer programs to identify search terms in the texts under analysis, a much larger number of texts can be examined, thereby increasing the generalizability of the

results. Some CADS studies do not report quantitative data such as frequencies, illustrating findings through text excerpts instead, much like CDA. In contrast, a growing number of CADS studies do report frequencies, improving accountability of claims made.

Though CADS methods offer the potential for more generalizable, accountable findings, CADS does not purport to render a comprehensive description of linguistic features in a collection of texts. In contrast, register analysis can do just that. It is important to note that the term "register" has been used to mean different things in different frameworks. In the present study, I use the term to mean a variety of language associated with situations of use and communicative purposes, as defined in Biber and Conrad (2009; see also Biber, 1994). Thus, a "register analysis" is a description of a variety of language that includes the situational features of the register, the linguistic features of the register, and a functional interpretation explaining the relationship between those situational and linguistic features. A register study can identify an extensive range of lexical and grammatical forms that characterize a group of texts (e.g., Biber, 2006, on university language; Forchini, 2012, on movie conversation; Gray, 2015, on research articles) and generally includes both what forms are typical across the register as well as what variation exists within the register. These similarities and differences within the corpus are then interpreted functionally, with the situational characteristics of the register being used to explain the particularly frequent or infrequent use of linguistic structures. For example, in a register study of call-center telephone conversations, Friginal (2009) finds modals are used frequently, often to express ability (e.g., *Yes, I **can** do that for you, no problem, sir*, p. 152). Friginal explains that modals are used frequently in call-center conversations because one of the purposes of the conversation is to provide help, so agents frequently state what they *can* do.

CADS and register analysis both examine large numbers of texts using corpus methods, thereby improving generalizability, but they answer very different research questions. CADS examines just one or two key words in depth, with interpretation of findings based on notions of power and ideology. Register analysis offers a comprehensive account of linguistic features with interpretations that include situational characteristics and communicative purposes. Together, CADS and register analysis can present a fuller description of a register, thereby furthering our understanding of the reasons certain linguistic features are used more or less frequently in a given variety of language.

1.3. Goals of the Dissertation

Section 1.1 has shown the importance of UNGA General Debate addresses and the relative dearth of information on the linguistic features that characterize this register. Section 1.2 has described two promising, complementary corpus-informed methodologies for analyzing linguistic features: CADS and register analysis. The overarching purpose of this dissertation is to use these two methodologies to add greater depth and breadth to the current understanding of international organization discourse. This purpose can be broken down into two principal goals:

1. To identify the distinctive lexical and grammatical characteristics of UNGA addresses when contrasted with comparable registers and explain any differences in linguistic features based on the situational characteristics and discourse functions of UNGA addresses.

2. To distinguish any systematic patterns of lexical and grammatical variation among UNGA addresses and explain differences in linguistic features based on the geographic, social, political, and economic characteristics of the countries represented by the speakers.

5

The first goal of this study is to identify the lexical and grammatical forms that typify UNGA addresses by comparing UNGA addresses to other registers. Differences in the frequencies of linguistic features are explained based on the situational characteristics of the registers. The second goal is to distinguish variation within the UNGA. The objective is to determine whether lexical and grammatical forms in a text differ systematically based on a country's geographic, social, political, and economic characteristics. These first two goals help describe UNGA General Debate addresses as a register and are informative for political scientists, speech writers, and translators, in addition to linguists.

A final methodological consideration that is addressed in this study is whether there is a difference in the nature of the findings made possible with CADS and with register analysis. A CADS investigation examines the use of one or two politically or socially charged words or phrases selected by the researcher as potentially compelling. The objective is not to describe the UNGA as a register but rather shed light on political identities and relationships based on the ways in which key terms are conceptualized. Comparing the types of findings that emerge in CADS with those in a register analysis adds to our understanding of corpus-informed methodology. Data cannot be interpreted without an acknowledgement of the ways in which methods shape findings. By following the steps of both register analysis and CADS using the same data, it is possible to determine how these two approaches differ, to evaluate the strengths and weaknesses of the two methods, and to better appreciate which methodology is best suited for which types of research questions.

1.4. Organization of the Dissertation

This dissertation comprises seven chapters. Chapter 2 lays the foundation for the study by presenting a review of the literature, beginning with spoken discourse more generally and a few

of the methodologies in current practice for the study of spoken discourse, then more detailed

descriptions of research on planned, monologic spoken texts and political speeches in particular.

The chapter ends with specific findings on international organization discourse, presenting the

gap in what is known about planned international organization speeches. I situate the present

study of UNGA General Debate addresses within the current body of research and make the case

that UNGA texts are the most compelling object of study for advancing the field of international

organization discourse. Chapter 3 then describes the corpora and methods used in the analyses of

Chapters 4 and 5. Chapter 4 reports findings on the lexical and grammatical forms that typify

UNGA addresses compared to other registers. Chapter 5 describes variation within the UNGA

based on the geographic, social, political, and economic characteristics of the countries

represented. Chapter 6 presents the methods and results of three case studies using a CADS

approach. For this chapter, I select three words based on the keyword analysis in Chapter 4 and

carry out an in-depth investigation of the ways in which these terms are conceptualized both

across the corpus and by country category. The chapter concludes with an analysis of the

differences in the nature of the findings from the register analysis in Chapters 4 and 5 and from

the CADS in Chapter 6. Finally, in Chapter 7, I summarize the findings from Chapters 4-6 and

discuss implications and suggestions for future research.

CHAPTER 2: REVIEW OF THE LITERATURE

2.1. Introduction

The differences between written and spoken language have been the object of much discussion and research. Today, it is widely recognized that spoken and written discourse differ, but not in consistent ways across all registers (Biber, 1988, 2009; Staples, 2015b): For example, the linguistic features setting apart face-to-face conversation from academic prose are not the same as the features distinguishing university lectures from personal email messages. Thus, spoken discourse is an established domain of inquiry set apart from written discourse, but also an umbrella term covering numerous registers, from telephone conversation to political debate.

Scripted political speeches such as the UNGA addresses in the present study hold a distinctive place within the domain of spoken language. Like face-to-face conversation, the setting for addressor and addressee is shared and the texts are delivered orally. Unlike conversation, the texts are planned, edited, and read out loud. These situational characteristics influence the frequency with which many linguistic features are used. Understanding the situational and linguistic characteristics of scripted political speeches therefore requires a discussion of what is known about spoken discourse generally and also what is known about how scripted political speeches, notably in international organizations, differ from other spoken registers.

The objective of Chapter 2 is therefore to situate international organization planned speeches within the context of spoken discourse. Section 2.2 presents an overview of research on spoken texts. It is by no means an exhaustive review of the literature, but rather an examination of the types of studies that have been carried out using three different theoretical frameworks: Critical Discourse Analysis (CDA), Corpus-Assisted Discourse Studies (CADS), and register

analysis. Studies using these three methodologies were selected because of their potential utility in the analysis of international organization planned speeches. Section 2.3 then presents an overview of some of the research on planned monologic spoken texts and scripted political texts in particular. Section 2.4 narrows the focus to international organization texts, summarizing some of the principal findings from research on international organization discourse and identifying specific gaps in the literature on this important register. Section 2.5 concludes with a synthesis of the chapter.

2.2. Spoken Discourse

Spoken language, be it a telephone conversation or a university lecture, has been examined using diverse methodologies, from conversation analysis to functional grammar. The objective of Section 2.2 is to highlight a few key studies of spoken discourse that have followed three methodologies: CDA, CADS, and register analysis. Far from a comprehensive account of all research on spoken discourse, this section instead focuses on the methods and findings most relevant to the present study on the linguistic features of UNGA addresses.

Section 2.2.1 is a brief description of CDA. Though this methodology is not used in the present study, it is one of the most frequent types of analysis in previous studies on international organization discourse. Understanding the theory and motivation behind CDA also aids in the understanding of the second methodology, presented in Section 2.2.2: CADS. A more recent methodology, the growing number of CADS on political texts has shown CADS to be a promising type of linguistic analysis suitable for investigations of international organization discourse. In Section 2.2.3 I discuss register analysis and Multi-Dimensional analysis as a specialized type of register analysis. In Section 2.2.4 I summarize the principal differences in these three methodologies and suggest that CADS and register analysis are complementary and

that there are benefits to using them together in a register study, particularly for texts from a political, social, or cultural institution.

2.2.1. Critical Discourse Analysis

CDA views language through a social, cultural, and political lens (Woods, 2006). The goal of CDA is to draw connections between language, power, and ideology and reveal the covert ways in which they influence people (Fairclough, 1989). Thus, the language in a text, from grammatical structures and lexical style to rhetorical patterns and argumentation strategies, is seen as reflective of certain assumptions about social identities and relationships and is interpreted qualitatively in light of social, political, institutional, and cultural conditions (van Dijk, 2006). Extensive studies have been carried out on topics such as racism (e.g., van Dijk, 1993), gender (e.g., Lazar, 2005; Wodak, 1997), and capitalism (e.g., Bruner, 2002; Fairclough, 2002). Because the focus of CDA is on power and influence, texts are often collected from political, media, academic, and corporate elites. Spoken texts include political speeches (e.g., Wodak, 1989), television news broadcasts or interviews (see van Dijk, 1988), classroom language (e.g., Kumaravadivelu, 1999), and workplace conversation (e.g., Holmes, 2000). Findings are presented in the form of text excerpts illustrating the ways in which language is used to promote ideologies; quantitative data, such as the frequencies of linguistic features, are relatively rare.

2.2.2. Corpus-Assisted Discourse Studies

CADS uses corpus tools to achieve the same objectives as CDA. Topics and text types are similar to those explored in CDA, again with a view to investigating social, political, and cultural identity, relationships, and influence. Unlike CDA, however, CADS seeks generalizable

patterns, basing analysis on a large corpus rather than just a few texts (see Morley & Bayley, 2009; Partington et al., 2013; Partington, Morley, & Haarman, 2004).

This search for generalizable patterns leads researchers to undertake quantitative analysis in addition to the typical qualitative analysis of CDA. CADS generally begins with frequency data from computer programs. Because this first step of the analysis is automatized, the researcher can sift through much larger collections of texts, which reduces the influence of one idiosyncratic text. The use of corpus methods helps to defend against criticism of cherry-picking occasionally leveled at CDA (e.g., Widdowson, 2000, 2004; see discussion in Baker & Levon, 2015). However, the quantitative analysis in CADS is not entirely objective; like CDA analysts, CADS researchers choose which words will be investigated based on their social, political, or cultural reference. CADS does not analyze all frequently occurring words and many studies do not even report all frequently occurring words. CADS researchers do not claim to present a comprehensive analysis of linguistic features. Rather, quantitative methods are used to uncover patterns and frequencies that are unexpected, even counter-intuitive, forms that would go unnoticed or take much longer to identify in a manual search (Baker, 2006, p. 2).

This quantitative step reveals the extent to which patterns hold true but not whether or why they are meaningful. The patterns that emerge in the quantitative analysis must be qualitatively interpreted in light of political, social, or cultural implications. During this qualitative analysis, like with CDA, the researcher investigates the context surrounding the words and phrases found from the computerized search and interprets the use of these terms based on what is known of the institution from which the texts are drawn (e.g., a political party, a newspaper). Thus, the quantitative step ensures a sound linguistic rationale for the choice of target forms while the qualitative step establishes meaningful connections and patterns.

11

Emphasis on quantitative findings and whether to report frequencies varies, however, from one CADS study to another. In some CADS studies (e.g., Lombardo, 2009, on television reporting of the 2003 Iraq War), frequencies of keywords and collocates are reported. Others, such as the study in Partington et al. (2013) on impoliteness in UK House of Commons debates, the Hutton Inquiry, and BBC interviews with Labour politicians, report no frequency information. Without frequency information, it is difficult for the reader to evaluate the magnitude of the claims made. Findings are assumed to be more generalizable because a large corpus is used, but without evidence that the linguistic associations described in a study are truly patterns, occurring in many of the texts rather than just one or two, generalizability is less certain. The distinction between CDA and CADS is a matter of number of texts used for data analysis, but within CADS, degree of generalizability and emphasis on quantitative reporting vary greatly.

Similarly, there is some variation in the methods used for CADS. The CADS analyst generally selects a linguistic feature for investigation in one of two ways: either based on a priori potential interest or through keyword analysis. In a study such as Kim (2014) looking at US news, the researcher selects a compelling search term (*North Korea*) before beginning the analysis. Collocates are identified through computer programs such as WordSmith (Scott, 2017) and analyzed for patterns. Concordance lines showing the immediate co-text surrounding the search term are also automatically generated and then examined manually for patterns. The patterns may be lexical (e.g., *states/countries like North Korea*) or grammatical (e.g., attributive adjectives such as *rogue*). They may be specific phrases (e.g., *missile threats posed by countries like North Korea*) or broad categories (e.g., countries and cities collocating with *North Korea*). Because the collocates reported are based on frequencies, findings are considered to be more

12

objective. At the same time, the initial choice of research questions and the analysis of patterns are not without some bias. Additional subjectivity is introduced with this type of CADS, where words or phrases are chosen because the researcher anticipates they will produce noteworthy findings.

A more bottom-up type of CADS begins with a keyword search. A keyword analysis determines which terms are statistically more frequent in a given text or corpus when compared to another. Keywords are thus revelatory of what makes the target corpus unusual. In this type of CADS, the researcher picks one or two words that were identified as "key" for further investigation. A keyword analysis can be carried out comparing a target corpus with a very large reference corpus. Alternatively, the target corpus can be divided into two or more sub-corpora and these groups of texts are then contrasted. Many researchers choose to do both types of keyword analyses in order to get a more complete picture of the target register (see Baker, 2006).

In the first type of keyword analysis, the researcher compares word frequencies in the corpus of interest with the frequencies in a reference corpus. Baker (2006) suggests a large collection such as the British National Corpus (BNC) for the reference corpus because it encompasses a range of situational characteristics and thus serves as an indicator of what can be expected in "normal" language use (p. 43). The CADS researcher then chooses one or several keywords for further analysis, generally from among the top keywords because these represent a greater departure from expected frequencies, from the norm. Some keywords are dismissed from further analysis because they reflect situational characteristics that are not of interest to the researcher. For instance, in his study on parliamentary debates, Baker (2006) finds several proper nouns and *Mr.* on his keyword list because in this register, speakers often refer to other participants involved in the speech act. He omits these words because he is not interested in

studying the relationships between individual parliamentarians but rather the attitude of speakers

toward a concept: hunting. With this rationale, he identifies the keyword *criminal* as promising

and conducts further study on its collocates. Some researchers discard function words like

pronouns, but for the researcher studying interaction or relationships between speakers, the

pronoun *you* can be of utmost importance. The choice of keywords to be studied is therefore

somewhat subjective, though a rationale for the choice is generally offered to support the

decision made. To determine which keywords to pursue, in addition to personal interest,

researchers often examine concordance lines to view the surrounding context (see Baker, 2006;

Partington et al., 2013). When comparing the target corpus to a reference corpus, keywords

reflect the "aboutness" of the texts under investigation overall. For instance, in her study of the

American sitcom *3RFS*, Gregori-Signes (2017) uses the spoken subcomponent of the BNC and

the American English corpus 2006 as reference corpora, finding a number of keywords in *3RFS*

related to women, femininity, and womanhood.

In the second type of keyword analysis, the target corpus is divided in order to explore

differences within the register. The keyword analysis in Bednarek (2010) compares the speech of

six main characters in the television series *Gilmore Girls*. Baker (2005) compares the speech of

British parliamentarians supporting a reform on the age of consent for gay men to that of

parliamentarians opposing the reform. Lombardo's (2009) study examines TV news reporting on

the 2003 Iraq War; her keyword analysis compares texts from CBS to those from the BBC.

These keyword analyses reveal differences within a corpus. Thus, the discourses of two different

political, social, or cultural groups can be contrasted while controlling for a certain number of

situational characteristics: What does one US television station talk about when they talk about

the war in Iraq and how does it compare to one UK television station on the same topic during

14

the same month? Again, the keyword analysis typically is only the beginning of the search. After one or more interesting keywords have been identified, collocates and concordance lines are examined to investigate how the keywords are conceptualized, what patterns can be detected in how the terms are used. For instance, the keyword *big* was identified in Lombardo's CBS texts. This is not particularly meaningful until she examines collocates and discovers that over half of the occurrences of *big* collocate with *picture*. Concordance lines show CBS news anchors use *big picture* as they present the broader context of the war to television viewers. Lombardo argues that CBS news anchors "frame" the news for viewers in a very controlled dissemination of information.

For both types of keyword analyses, the choice of which word or words to investigate is considered to be more objective because it is based on frequency. That objectivity is not absolute, however. The words are selected due to their unusual frequency in a corpus, but not all unusually frequent keywords are discussed. Even with a keyword search, the researcher chooses a word of political, social, or cultural interest. This subjective choice leads to interesting findings and discussions, but not, as CADS researchers are the first to acknowledge (Baker, 2006; Partington et al., 2013), a comprehensive description of the linguistic features that characterize a collection of texts.

2.2.3. Register analysis

2.2.3.1. Overview of register analysis

A more complete investigation of the linguistic features that typify a group of texts can be carried out with register analysis (see Biber & Conrad, 2009; Biber, Conrad, & Reppen, 1998). Both CADS and register analysis draw on corpus methods, but they differ primarily in the types of research questions posed and in the interpretation of linguistic features. CADS

researchers analyze features based on political, social, and cultural identities and influence. Register analysts give a functional explanation of textual features based on broader situational characteristics. Thus, one of the first steps in register analysis is to identify the situational characteristics of a text type. Taxonomies of situational characteristics differ, but Biber and Conrad (2009, p. 40) offer a particularly thorough framework that includes participants, relationships between participants, channel (speech/writing; permanent/transient), production circumstances, setting, communicative purpose, and topic. A register analysis uses situational characteristics such as these to explain why certain linguistic features are more or less frequent in a given communicative context.

For example, in Quaglio's (2009) comparison of the television situational comedy *Friends* to face-to-face conversation, he interprets differences in vague language based on the situational characteristic "audience." In typical face-to-face conversation, addressor and addressee often share a great deal of background knowledge that allows them to be vague (e.g., *and stuff like that*). Though some vague language has been written into the *Friends* script, much less occurs compared to natural face-to-face conversation so that the dialogue can still be followed easily by a general audience. Similarly, Staples (2015a) examines nurse-patient interactions, comparing nurses born and trained in the US to nurses born and trained outside of the US. She interprets differences between the two nurse groups, such as the greater use of questions among the US nurses, based on views of nurse-patient relationships (US nurses tend to have a more patient-centered approach).

The examples of Quaglio (2009) and Staples (2015a) highlight two important characteristics of register studies. First, the frequencies of linguistic features are identified and reported. Second, the frequencies of two registers or sub-registers are compared (e.g., a

16

television show versus face-to-face conversation; US versus international nurses). Frequencies of

linguistic features are meaningful only when compared to another register. In Biber's (2006)

study of university language, nouns are shown to occur in over 300 out of 1,000 words in

textbooks but fewer than 150 in office hours. The importance of nouns in textbooks is manifest

only when compared to office hours. In Friginal's (2009) study of call center language, the high

frequency of polite and respect markers (e.g., *sir, please, ma'am*) is apparent only when a

keyword analysis compares those frequencies to face-to-face conversation. The Longman

Grammar of Spoken and Written English (Biber, Johansson, Leech, Conrad, & Finegan, 1999),

based on a study of four registers (conversation, fiction, news, and academic prose) has an entire

section devoted to the grammar of conversation (pp. 1037-1125) noting, for instance, the high

frequency of repetitions, parenthetical structures, and discourse markers in conversation

compared to other registers.

The types of analyses carried out in a register study vary. Like with CADS, many register

studies use keyword and collocational analyses to investigate lexical features (see Section 2.2.2).

Unlike with CADS, the analyst reports and interprets the full set of keywords or collocates or a

pre-determined number (e.g., the top 100, Biber & Egbert, in press). While CADS uses keywords

to select a term for investigation, with register analysis, the objective of keyword analysis is to

determine what lexical features distinguish the register overall, each keyword serving an

important function that can be explained by situational characteristics. For example, in Friginal's

(2009) study of call center interactions, *number, phone, account, address, code,* and *meter* were

identified as keywords and interpreted as providing information necessary for completing a

transaction, the purpose of the phone call (p. 139). The goal of collocational analysis in register

analysis is much like in CADS: to establish the immediate context of a word and thus how a term

17

is conceptualized or what function it serves. For instance, in Fortanet's (2004) study of *we* in

university lectures, collocates such as *talk* illustrate that the most frequent metadiscourse

function served by *we* is to guide the addressee through the speech event: *we're gonna talk, we'll*

talk about, we talk about.

A number of register studies have also examined categories of lexico-grammatical

features (e.g., conditionals in Staples, 2015a). In order to do so, first the lexico-grammatical

features must be annotated using a program such as the Biber Tagger (see Biber 1988, Biber et

al., 1999). Then, the annotated texts are run through a program like TagCount (see Biber 1988,

Biber et al., 1999) that counts total occurrences of each linguistic feature. The total number of

occurrences of the target structure are counted and frequencies of each lexico-grammatical

feature in the target corpus are compared to a reference corpus. For instance, Quaglio (2009)

compares the frequency of past tense verbs in his corpus of episodes from the television show

Friends to previously published frequencies from face-to-face conversation. With very large or

small differences in frequencies, descriptive statistics suffice to show that two registers are

relatively different or similar. Comparing frequencies is useful with a top-down approach, when

the researcher decides on a limited number of features to be investigated and is looking for both

differences and similarities between the target register and comparison register(s): Quaglio

chooses past tense verbs because he is interested in examining features related to narration, not

because he sees an unusually large or small difference in frequencies.

Reporting frequencies is more prohibitive with a bottom-up approach, where the

researcher investigates a large number of linguistic features to determine which forms distinguish

the target register from the comparison register(s). Rather than report and interpret dozens of

features, which may or may not actually distinguish the register, the analyst needs a means of

detecting which features vary between registers and are therefore worth interpreting. To detect a large effect size in the difference of feature frequencies, an innovative approach coined "key feature analysis" was adopted in a study on web registers (Biber & Egbert, in press) using Cohen's *d*. This method is discussed in more detail in Section 3.7.3. To date, no studies on spoken discourse use "key feature analysis," but the method looks promising.

2.2.3.2. Multi-Dimensional analysis as a specialized type of register analysis

In addition to these investigations of individual lexical and grammatical features, many register studies examine how lexico-grammatical features co-occur using Multi-Dimensional (MD) analysis. MD analysis involves both microscopic and macroscopic analysis. At the microscopic level, it identifies co-occurring linguistic features that serve a similar function (e.g., past tense, perfect aspect, and third person pronouns are frequent in narration). At the macroscopic level, MD analysis identifies the overarching ways in which texts vary (e.g., some texts are more narrative than others).

MD analysis was first introduced to investigate linguistic variation across a wide range of spoken and written registers, from face-to-face conversation to official documents (Biber, 1988). Factor analysis was used to identify the co-occurrence patterns of 67 linguistic features. These co-occurring linguistic features (or *factors*) were interpreted functionally as underlying dimensions of variation in texts. The five principal dimensions identified in the 1988 study were Dimension 1: Involved versus Informational Production, Dimension 2: Narrative versus Non-Narrative Concerns, Dimension 3: Explicit versus Situation-Dependent Reference, Dimension 4: Overt Expression of Persuasion, and Dimension 5: Abstract versus Non-Abstract Information. Table 2.1 presents the primary linguistic features which loaded most heavily on, that is, which were found to be most indicative of each dimension. Negatively loaded features were negatively

correlated with positive features. Thus, for Dimension 1, an "involved" text would have a high frequency of private verbs (e.g., *believe*, *feel*), contractions, and WH- questions (positively loaded), but a relatively low frequency of nouns and prepositions (negatively loaded). The opposite would be true for "informational" texts, at the other end of Dimension 1. See Table 3.12 for the complete list of linguistic features for each dimension.

Table 2.1. Principal linguistic features for Dimensions 1-5 in Biber (1988)

Dimension 1: Involved vs. Informational Production	**Positive**: private verbs, *that* deletion, contractions, present tense verbs, second person pronouns, pro-verb *DO*, analytic negation, demonstrative pronouns, general emphatics, first person pronouns, pronoun *it*, BE as main verb **Negative**: nouns, word length, prepositions, type-token ratio, attributive adjectives
Dimension 2: Narrative vs. Non-Narrative Concerns	**Positive**: past tense verbs, third person pronouns, perfect aspect verbs, public verbs **Negative**: present tense verbs
Dimension 3: Explicit vs. Situation-Dependent Reference	**Positive**: WH relative clauses, pied piping constructions, phrasal coordination, nominalizations **Negative**: Time and place adverbials, adverbs
Dimension 4: Overt Expression of Persuasion	**Positive**: Infinitives, prediction modals, suasive verbs, conditional subordination, necessity modals **(No negative features)**
Dimension 5: Abstract vs. Non-Abstract Information	**Positive**: Conjuncts, passives, past participial clauses and WHIZ deletions **Negative**: Type-token ratio

The dimensions found in Biber (1988) have been used in numerous subsequent studies in order to compare a new register to previously explored registers (e.g., Quaglio, 2009, on the television sitcom *Friends*; Sardinha & Pinto, 2017, on television language). These studies compare dimension scores for their target register to one or more registers in Biber (1988). For instance, in his study comparing the television sitcom *Friends* to face-to-face conversation in Biber (1988), Quaglio finds a nearly identical mean score for Dimension 1, involved versus

informational production, and concludes that the language of *Friends* is, on average, as involved as face-to-face conversation. In contrast, the standard deviation for Dimension 1 is much smaller in *Friends* than in face-to-face conversation. He attributes this difference to the wider range of speakers and situations in Biber (1988); in *Friends,* there is little to no variation in terms of age, dialect, and situations. By comparing the dimension scores of *Friends* and face-to-face conversation, Quaglio was able to show that *Friends* shares the core linguistic features characterizing conversation.

Because MD analysis looks at the frequencies of co-occurring linguistic features, they must be tagged and counted as described in Section 2.2.3.1. The TagCount program (see Biber 1988, Biber et al., 1999), in addition to counting lexico-grammatical features, also calculates dimension scores for each individual text (in the case of Quaglio, 2009, for each episode of *Friends*). The mean and standard deviation for the entire corpus is then calculated and compared to the means and standard deviations in Biber (1988) or other subsequent studies reporting data along the same dimensions.

Similarities and differences between registers are explained based on situational characteristics (e.g., the smaller range in speakers and situations in *Friends* results in a lower standard deviation when compared to Biber's face-to-face conversation). A functional interpretation is then proposed, illustrating the relationship between the linguistic features that typify a dimension and the situational characteristics of a register. Quaglio (2009) shows through textual examples how first and second person pronouns, contractions, present tense, private verbs, hedges, and discourse particles are used in *Friends* for the same reasons they are used in face-to-face conversation: shared communicative context, real-time production, and interactivity.

2.2.4. Summary of the three methodologies

CDA, CADS, and register analysis address different types of research questions and each offers its own advantages. CDA allows the researcher to investigate qualitatively a small number of texts in depth. Findings are not generalizable, however, and critical discourse analysts are sometimes accused of "cherry-picking." CADS uses both quantitative corpus methods and qualitative analysis to examine a large number of texts. The results are therefore more generalizable than for CDA. The primary objective of CADS is to investigate political, social, or cultural ideologies, identities, and relationships through the use of one or just a few terms selected by the researcher. The focus of this approach is to explore the use of specific words. In contrast, register analysis, which also uses quantitative corpus methods followed by qualitative analysis, does not set out to examine the use of just a few pre-selected words. Rather, the goal of register analysis is to describe as thoroughly as possible the lexical and grammatical features that distinguish a register. Typically, a researcher will carry out either CADS or register analysis. However, the two methodologies are complementary. The examination of politically or socially charged words from CADS can add greater depth to a register analysis and offers a potentially sizable contribution in an analysis of texts from a political, social, or cultural institution.

2.3. Planned Monologic Spoken Texts

Certain linguistic features emerge as characteristic of spoken texts overall, but as many of the studies described in Section 2.2.3 have shown, a great deal of variation occurs across speech registers due to differing situational characteristics. Understanding this variation is instrumental when situating the object of this study, UNGA addresses, within the larger domain of spoken discourse. UNGA addresses are planned, monologic, scripted texts and are expected to differ from spontaneous dialogic speech in their frequencies of certain linguistic features. The objective

of Section 2.3 is to tease apart these elements, first by examining planned, monologic speech that is not scripted to see the ways in which it differs from spontaneous, dialogic speech and then by adding the scripted component to see how scripted speech differs from unscripted.

Planned monologic speech has three distinctive characteristics compared to spontaneous dialogic speech: planning, interactivity, and communicative purpose. Section 2.3.1 is a discussion of these three situational characteristics and of the linguistic features that result from a text being planned and monologic, with a review of some key findings on these types of spoken texts. An additional consideration within the domain of planned monologic spoken texts is whether the text is scripted. Scripted texts are typically distinguished from non-scripted and sometimes referred to as a "hybrid" register because they originate in writing but are performed in speech. In order to account for this additional distinction within planned monologic texts, in Section 2.3.2 I consider how the unique written-to-be-spoken production circumstances can be expected to affect the choice of linguistic features, with examples from television news broadcasts and sermons. Section 2.3.3 then narrows the focus to scripted political speeches, showing how the features of scripted language emerge in political texts and pointing out certain key distinctions in the language of political speeches when compared to other scripted texts.

2.3.1. Select research on planned, unscripted monologic speech

Within the domain of spoken discourse, planned monologic texts can be distinguished from unplanned, dialogic speech in terms of production circumstances (planning) and interactivity (monologic versus dialogic). Section 2.3.1 begins with a discussion of the relationship between these two situational characteristics and the linguistic features that have been identified as more or less prevalent in planned monologic texts, with findings from studies of non-scripted texts to illustrate the association between situational context and language use. In

addition, Section 2.3.1 considers the key situational characteristic of communicative purpose.

Planned monologic texts are typically highly informative; they are planned and monologic

precisely because the speaker, generally an expert in a field, has something important to say on a

subject, has information to convey. This communicative purpose differs markedly from many

other types of spoken text (e.g., impromptu small talk with a neighbor), again resulting in

differences in the frequencies of a number of linguistic features.

The importance of planning as a situational characteristic of spoken discourse is widely

recognized. Cutting (2011) classifies spoken texts as unplanned (e.g., spontaneous conversation),

semi-planned (e.g., marriage proposal), semi-scripted (e.g., university lecture), and scripted (e.g.,

formal political speech), though she notes that these are not distinct categories but rather exist on

a cline. The same can be said about the linguistic features that have been found to be associated

with planning: Rather than a complete absence or presence of these features, they may be more

or less prevalent in a text depending on *how* planned it is. Typically, complex noun phrases and

complete sentences are possible with planning and revision. In contrast, real-time unplanned

speech such as most face-to-face conversation includes more vague nouns, hedges, ellipsis,

repairs, and structurally incomplete sentences (Biber & Conrad, 2009, p. 68; see also Cutting,

2011; Staples, 2015b).

One would expect a planned, but unscripted register such as classroom teaching to fall

somewhere between the two extremes of a spontaneous, unplanned text and a prepared, revised

text. Indeed, in Biber's (2006) study of university language, the frequencies of many linguistic

features for classroom teaching lie between less planned speech (e.g., office hours) and written

texts (e.g., textbooks). For instance, the mean score for hedges in office hours is 3.6 per 1,000

words, in teaching 2.5 per 1,000, and in written course management only 0.4 per 1,000 (see

Biber, et al., 2004, p. 312). Camiciottoli's (2007) study of business studies lectures with American, British, and non-native-speaking lecturers shows frequent reduced forms, vagueness, idioms, and syntactic informality such as ellipsis. Camiciottoli also finds numerous discourse dysfluencies such as pause fillers, false starts, and repetitions, though their relative frequency is difficult to evaluate without direct comparison to data from dialogic discourse. It is initially surprising that the planned, relatively monologic discourse of classroom teaching would include so many features typically associated with spontaneous speech, but the pressures of real-time production are factors for both classroom teaching and entirely unplanned speech: The addressor is thinking and speaking at the same time, resulting in a less "polished" text.

Interactivity between addressor and addressee is also a key situational characteristic that distinguishes monologic from dialogic speech. Definitions of interaction vary, some including applause (Halmari, 2005) and others comprehension checks produced by the speaker (Camiciottoli, 2007). In the present, interaction will be defined as an exchange of speech produced by two or more participants to build shared discourse, based on the theoretical framework of Biber and Conrad (2009). The rationale for excluding non-verbal interaction is to account for the difference between interaction and involvement, where involvement is defined as engagement, as opposed to detachment (Chafe, 1982; see also Biber, 1988). The distinction between interaction and involvement is critical in terms of communicative function, as will be illustrated in the following discussion.

Some linguistic features such as questions and first and second person pronouns are characteristic of both interactive and involved speech, but they are used differently for interactive purposes than for involvement. When questions and first and second person pronouns are used for interaction, they help co-construct a text, building shared discourse as participants in a

conversation exchange information. This communicative function is illustrated in Text 2.1, a

dialogue between a customer and a clerk at a university bookstore, from Biber (2006, p. 188).

Text 2.1. Dialogue showing interaction (Biber, 2006, p. 188, questions and first and second person pronouns in bold)

> Customer: Can **I** ask **you** something**?**
> Clerk: Yeah.
> Customer: **We**'re at the previews and of course my book is back there with my husband. Do **you** have coupons**?**
> Clerk: No **we** don't have them here. **You** guys only get them. Yeah.
> Customer: OK.
> Clerk: Did **you** want to come back**?** Cos **I** can hold onto your stuff.
> Customer: Could **you** hold all this stuff? […]

In Text 2.1, personal pronouns and questions reflect interaction between the customer and

the clerk. Questions are asked by one participant and answered by the other. There is an

exchange of information in which both speakers convey topical information in order to arrive at

a solution. Only one instance of the second person also illustrates involvement, with expression

of desire: *Did you want to come back?*

In contrast, in a true monologue, these features are not interactive. They are involved:

involving both the addressor and the addressee, but without necessarily resulting in any verbal

exchange between participants. This is illustrated in Text 2.2, an excerpt from a lecture during a

graduate-level humanities class, from Biber (2006, p. 191).

Text 2.2. Monologue showing involvement (Biber, 2006, p. 191, questions and first and second person pronouns in bold)

> Instructor: […] every established genre, every approach, every way that everything that's accepted for all of those things that seems like there's, there are things that are rejected, not done, not looked at, and as **you** were just pointing out, some of the ways that particular systems get promoted are not exactly ethical or correct or at the very least not everybody has access to being able to promote certain kinds of discourses and systems, so **I** guess, **I** don't know whether this is really a question or just a comment though **I**

26

think **we**'re, think a lot of **us** feel very sort of trapped between buying into this is how it is and since it's this way, let'**s** do the best **we** can with it and sort of saying, who's got the agency here**? you** know, how is the system beneficial for people**?** who is it hurting**?**

In Text 2.2, there is no exchange of productive speech between two or more participants. Personal pronouns and questions reflect involvement rather than interaction. Many first person pronouns are used in expressions of stance (*I guess, I think*), showing engagement of self in the proposition; second person pronouns are used to refer to members of the audience (*as you were just pointing out*) and in discourse markers (*you know*), both of which involve the listeners, recognizing them as participants in the communicative event. Questions are rhetorical, not meant to be answered by the audience.

Texts 2.1 and 2.2 exemplify the importance of interpreting the function of linguistic features, which are not necessarily used for the same purpose in different texts or even within the same text. They also explain why "interactive" features occur in a monologue: These "interactive" features are also involved features.

Because these features are used for both interaction and involvement, they can be expected to occur more frequently in dialogic texts, where they serve both functions, than monologic texts, where they serve just the one (involvement). Indeed, the research has shown this to be the case. For instance, in Biber (2006), university teaching tends to fall between more spontaneous speech and written texts for first person pronouns. The mean score for first person pronouns in office hours is 58.2 per 1,000 words, in teaching 37.5, and in textbooks 6.6. The mean for second person pronouns in office hours is 42.9 per 1,000 words, in teaching 30.0, and in textbooks 2.7. In her study of medical conference presentations, Webber (2005) also finds high frequencies of first and second person pronouns. Camiciottoli (2007) reports frequent first person pronouns and questions in her business studies lectures as well.

The use of involved features in monologic speech can be attributed to two primary functions: expression of stance and signaling discourse structure for signposting. Both of these functions arise from the communicative purpose (conveying information) and production circumstances (oral, real-time delivery). Unlike informative written registers (e.g., textbooks, newspapers, user manuals), a speech involves face-to-face contact. This results in more personal expression of stance. Rather than use impersonal expressions (e.g., *evidently, it appears*), speakers such as university lecturers and conference presenters have been shown to use personal expression of stance and as a result the first person singular (e.g., *I guess*, Biber, 2006, p. 191; *I don't think this is true*, Webber, 2005, p. 164).

In addition, speakers recognize and involve the audience by referring to the addressees' own background information (e.g., *you all know that*, Biber, 2006, p. 191, *as we all know*, Webber, 2005, p. 165) and to the addressees' information processing and the exchange of information occurring between participants (e.g., *I hope you see the multidisciplinary approach we're taking to this problem*, Webber, 2005, p. 163).

The second function served by first and second person pronouns as well as questions is to signal key information and guide listeners through the discourse structure of the text. This is crucial because the information is delivered orally. Most informative texts are written, making them easier to process. Real-time delivery makes a spoken informative text more challenging to follow because the listener cannot pause to reflect or return to a previous utterance if something is unclear (see Biber & Conrad, 2009, pp. 43-44 for a discussion). Speakers can aid the listener by signposting: cuing particularly important information and by signaling discourse organization (*I'm going to illustrate*, Biber, 2006, p. 118; *We'll talk about, let's move on to, as I said before*, Camiciottoli, 2007, p. 85; *Now I will turn to*, Webber, 2005, p. 173).

Questions, too, can serve this purpose of signposting. Rhetorical questions are often used to signal a particularly important idea (e.g., *who's got the agency here?* in Text 2.2). Questions that are immediately answered by the speaker alert and prepare the listener for the information that is to follow (*Why is that? Well because…,* Camiciottoli, 2007, p. 107). In addition, they break up what might otherwise be an overly dense text. In many question-answer pairs, the question does not introduce any new information; it simply sets up the answer itself: *So what do we do? We legislate the minimum wage* (Camiciottoli, 2007, p. 107). These question-answer pairs demonstrate a key difference between "interaction" (dialogic production of speech) and "involvement" (engagement). With dialogic interaction, the question would be asked by one participant and the answer would be provided by the other. With monologic involvement, both question and answer are uttered by the lecturer, without actual language production from the audience.

Other involved features that signpost include expressions of evaluation and comprehension checks. Expressions of evaluation signal important information: *The key thing is; what is crucial is* (Camiciottoli, 2007, p. 97). Comprehension checks also tend to signal that the preceding information was important. The audience rarely responds to checks such as *right?* and *okay?* (Camiciottoli, 2007, p. 108) except, perhaps, with a silent nod of the head. The speaker hopes, however, that listeners are reviewing and taking note of what they just heard. Comprehension checks, however brief, provide a moment of pause, again to break up any dense information.

Thus, planned monologic speech such as classroom lectures and conference presentations can be distinguished from unplanned conversation in relation to production circumstances (planning but real-time delivery), interaction (monologic versus dialogic), as well as purpose

(conveying information). Because lectures are delivered in real time, much like spontaneous conversation, hedges and other vague language as well as discourse dysfluencies such as false starts may occur. However, because lectures are planned (to varying degrees), there are typically fewer of these features when compared to unplanned conversation. The discussion and examples distinguishing interaction and involvement show that true monologues (without dialogic episodes) are not interactive, but they are involved. Personal pronouns and questions are not used to co-construct shared discourse but rather to mark engagement with a text. In addition to involving both addressor and addressee with expression of stance, these features are also used to signal important information or the discourse structure of the text in order to reduce the difficulty of real-time processing.

2.3.2. Select research on scripted monologic speech

Scripted monologues share many of the characteristics of planned, monologic, unscripted texts, but also features that are commonly associated with writing (Fortanet, 2005). Because they are written to be spoken and heard, scripted texts are sometimes considered a hybrid register, falling between written and oral discourse (Archer, Aijmer, & Wichmann, 2012, p. 14).

Scripted texts generally involve more planning, making them easier to produce for the addressor but not necessarily easier to grasp for the listener. If read word for word, speakers are less affected by real-time delivery because they are not attempting to think about what they are going to say while they are speaking. In contrast, real-time delivery remains problematic for the listener, who must follow the pace of the speaker without pausing to reflect on the content of the speech. As a result, the features associated with real-time constraints for speakers (e.g., discourse dysfluencies) are typically not present in scripted texts; the features associated with real-time constraints for listeners are, however.

Real-time delivery is generally recognized as an important situational characteristic distinguishing, for instance, television news from print (Montgomery, 2007), with ensuing differences in linguistic features. Television news makes frequent use of many of the involved features noted in Section 2.3.1 on unscripted speech: first and second person pronouns and first person plural imperatives such as *let's* (Ferrarotti, 2009; Lombardo, 2009).

As with the classroom lectures and conference presentations in Section 2.3.1, in scripted texts such as news broadcasts, these features are typically used to acknowledge the audience's intelligence and understanding (*as you can see, you may remember,* Ferrarotti, 2009, p. 80) and to help guide the listener with discourse organization markers (*let's listen to, we take you,* Ferrarotti, pp. 75; *this evening we'll bring you reports,* Ferrarotti, p. 83).

Interestingly, related findings have been reported in a very different scripted register: sermons. Kuczok (2014) notes certain "characteristic formulas" in John Henry Newman's *Parochial and Plain Sermons* that are highly involved, with frequent use of first and second person pronouns (e.g., *I wish to speak to you, my brethren*; *let us pray to God to enlighten us;* pp. 80-81). These expressions can serve discourse organizing functions such as introducing a new topic and calling to prayer. In a study of African American sermons, Wharry (2003) found sermonic expressions such as *amen, hallelujah,* and *praise God* serve as textual boundary markers, signaling a change in text type, a topic or subtopic boundary, or topic continuity. Though specific to the domain of religious texts, they remain relevant to the present discussion of scripted speeches in that they demonstrate the speaker's acknowledgment of the need to guide listeners in the structure of the speech.

2.3.3. Scripted political texts

Scripted political texts are distinct from other scripted monologues primarily in terms of purpose and topic. The purpose of persuasion in particular helps to explain differences in how linguistic features are used in political speeches compared to other scripted monologues. For instance, like other scripted texts, political speeches are notable for their high frequency of involved features like first and second person pronouns and rhetorical questions. However, these features do not appear to be used to organize discourse. Rather, most analysts have interpreted them as serving a persuasive function (e.g., Cap, 2015; Halmari, 2005; Otieno, 2015). Of particular relevance to the present study are four persuasive devices: expression of personal stance, polarization and implied solidarity between addressor and addressee, the use of questions to elicit emotional response, and legitimation through appeals to logical reasoning.

Expression of personal stance is not the most common persuasive device used in political speeches, but it merits some discussion nonetheless. When *I* is used in political speeches, it is often coupled with persuasive/modality verbs such as *I ask you to strengthen our democracy* in a speech delivered by President Clinton (Halmari, 2005, p. 119) and is thus a marker of overt persuasion. The first person singular is also frequently used with mental verbs such as *believe* (e.g., *I opposed the war in Iraq precisely because I believe that we must exercise restraint in the use of military force, and always consider the long-term consequences of our actions* in an Obama speech, Reyes, 2011, p. 799). With mental verbs, *I* is used to express personal opinion, a more indirect persuasive technique.

Overall, however, the first person plural is much more frequent than the first person singular in scripted political texts (e.g., Halmari, 2005, on Reagan and Clinton State of the Union addresses; Miller & Turci, 2006, on Martin Luther King, Jr. speeches; Otieno, 2015, on Obama

speeches). The more frequent *we* in political discourse often implies solidarity and shared goals between speaker and audience, as exemplified in a State of the Union address delivered by Reagan: *we have four basic objectives tonight* (Halmari, 2005, p. 119). This is in stark contrast to many other registers, where *I* is more frequent than *we* (e.g., conversation, fiction, news, see Biber et al., 1999, p. 334). In other registers, it is less important for speakers to establish common views and objectives with the audience, but it is key in the art of political persuasion.

The use of *we* has important implications in the interpretation of relationships between speaker and listener. In Miller and Turci's (2006) analysis of political speeches and religious sermons delivered by Martin Luther King, Jr., they note that King uses *we* more than twice as frequently in political speeches than in sermons. Miller and Turci suggest that this is due to the audience. In his political speeches, King is addressing a large, diverse audience with potentially different views and goals; in his sermons, his audience typically shares his views and goals. Thus, Miller and Turci posit that inclusive *we* is used in political speeches to help create a sense of shared goals and values between speaker and listener (see also Mulderrig, 2012 on forcing membership through the use of *we*). *We* is not necessary in sermons because solidarity is presumed, both addressor and addressee acknowledge shared membership in a group. This use of inclusive *we* is an important point for the analysis of pronoun use and involvement. The use of first person pronouns is often interpreted as reflective of very involved language. This does not necessarily mean that the relationship between addressor and addressee is closer than in a text with fewer occurrences of first person pronouns. Rather, it may indicate simply the *desire* of the speaker to *imply* a close relationship.

Moreover, the inclusive *we* is also used frequently in political speeches in contrast to *they*, where the third person plural refers to the "other," the threat. This polarization between *us*

and *them* has been shown in countless studies of scripted political texts. Four studies of different speeches delivered by President George W. Bush (Bhatia, 2009; Cap, 2015; Collet, 2009; Sowinska, 2013) show how Bush polarizes the *civilized world* (*us*) and *terrorists* (*them*). Similarly, Pujante and Morales-Lomez (2008) show how Spanish Prime Minister Azar uses *we* to refer to the international community and *they* to terrorists in one of his speeches before Parliament justifying the war in Iraq. Van Dijk (2006) analyzes a fragment of one of UK Prime Minister Blair's speeches before the House of Commons on UK involvement in Iraq, also finding a contrast between "democratic *Us* and undemocratic *Them*" (p. 168) in the terms *our* and *others*. In Reyes (2011), *us* versus *them* polarization was found in speeches delivered by US Presidents Bush and Obama (e.g., Obama's *We know that Al-Qaeda and other extremists seek nuclear weapons, and we have every reason to believe that they would use them*, p. 795). In Oddo (2011), it was found in addresses delivered by US Presidents Franklin D. Roosevelt and George W. Bush. Polarization reinforces the solidarity implied with the inclusive *we* by juxtaposing the opinions and goals presumed to be shared between addressor and addressee with the opposing beliefs and objectives of a common enemy (*they*).

Questions are another involved feature that has been attributed to persuasion in scripted political speech (Halmari, 2005; Otieno, 2015). Both unanswered rhetorical questions and questions immediately answered by the speaker are interpreted as persuasive rhetorical devices. Rhetorical questions are often phrased to elicit a strong, sometimes emotional response in the audience as with this excerpt from an Obama speech: *So tonight let, let us ask ourselves – if our children should live to see the next century, if my daughters should be so lucky to live as long as Ann Nixon Cooper, what change will they see? What progress will we have made?* (Otieno, 2015, p. 90). Question-answer pairs suggest the answer was reached by mutual consent with the

audience, as illustrated in an excerpt from a Clinton State of the Union address: *Do you believe we can become one nation? The answer cannot be to dwell on our differences, but to build on our shared values* (Halmari, 2005, p. 117).

These involved features, first and second person pronouns and questions, have been identified as frequently occurring in other types of scripted texts as well as in non-scripted, planned monologues. The difference readily apparent in the examples from political speech is that these features are used to persuade.

Other features have also been noted as serving a persuasive function by implying that a proposition is inevitable. They frame opinion as necessity and turn subjectivity into objective fact. These features include modals, factive verbs, and logical connectors.

The use of modals (*must, should*) frame a personal opinion as a factual requirement. This is illustrated in an excerpt from a Clinton State of the Union address: *We have to address the real reason for the explosion in campaign costs* (Halmari, 2005, p. 125), in a speech from King: *We must forever conduct our struggle on the high plane of dignity and discipline* (Miller & Turci, 2006, p. 418), and in a speech from President Bush: *We learned a lesson: the dangers of our time must be confronted actively and forcefully, before we see them again in our skies and our cities* (Cap, 2015, p. 551). As illustrated in these three examples, modals of necessity are used frequently with the inclusive *we* and the passive, both of which reduce the force of imposition, *we* because the speaker implies a shared goal between speaker and audience, and the passive because the agent is left unnamed. The high frequency of *we* collocating with a modal or semi-modal is indicative once again of the speaker implying shared objectives or responsibilities. In State of the Union addresses delivered by Reagan and Clinton, 113 out of 243 occurrences of *we*

collocate with a modal or semi-modal (Halmari, p. 129), as in Reagan's *we must protect that peace and deter war* (Halmari, p. 127).

Another way to express the factual certainty of a proposition is through the use of factive verbs (e.g., *know, realize, recognize, see, understand*). When Clinton states *everyone knows elections have become too expensive* (Halmari, 2005, p. 119), the force of the proposition increases with the factive verb *knows*. When Bush says *Yet, it's important for our fellow citizens to understand that failure in Iraq would be a disaster for our future* (Reyes, 2011, p. 794), the factive verb *understand* imposes certainty. When King says *Many of our white brothers [...] have come to realize that their destiny is tied up with our destiny and their freedom is inextricably bound to our freedom*, the factive verb *realize* signals inevitability.

Finally, logical connectors (e.g., *if, when,* and *because* clauses) imply indisputable causal relationships. Examples from George W. Bush include *as a result, we must look at security in a new way* (Cap, 2015, p. 553), *the cause we serve is right, because it is the cause of all mankind* (Sowinska, 2013, p. 802), and *because democracies respect their own people and their neighbors, the advance of freedom will lead to peace* (Sowinska, p. 805). Logical connectors can also be used in thesis-antithesis patterns. For instance, Spanish Prime Minister Aznar states *retreat is simply not an option in the face of terror. If we retreat, then all our efforts to date will have been in vain* (Pujante & Morales-Lopez, 2008).

Overall, the literature on scripted political speeches shows that these texts are highly persuasive. Some personal stance is expressed through *I* and either persuasive verbs or mental verbs, but primarily politicians prefer the use of *we* to imply solidarity and shared goals with the audience. Similarly, rhetorical questions and questions posed and then answered by the speaker suggest addressor and addressee agree on the response. Finally, politicians frame opinions as

36

logical inevitability and fact through the use of modals of necessity, factive verbs, and logical connectors.

The features analyzed in these studies of scripted political texts do not serve the same function as the involved features noted in studies of planned monologic speech (classroom lectures, conference presentations) or even other types of scripted discourse (television news broadcasts, sermons). Most notably, speakers use few to no linguistic devices to help the listener process a highly informative text in real time.

Many of the studies reviewed in Section 2.3.3 follow a CDA or CADS approach. They do not attempt to describe in detail a comprehensive set of linguistic features characterizing scripted political discourse. They do, however, reveal a widespread interest in certain compelling features and their functions. Some of these features have also been noted in international organization discourse as well and are discussed further in Section 2.4.

2.4. International Organization Discourse

Scripted speeches delivered at international organizations are similar scripted national political discourse in their use of some features, but they are also distinctive in their topic, purpose, and rhetorical conventions. Section 2.4 reviews the relevant literature on international organization texts, both written (Section 2.4.1) and spoken (Section 2.4.2), and Section 2.4.3 summarizes findings specifically on linguistic features. Section 2.4.4 argues that of all the different types of international organization texts, the addresses delivered during the UN General Assembly General Debate represent a highly compelling register for analysis.

2.4.1. International organization written discourse

A number of studies have examined the written texts of international organizations, which are typically more readily accessible to the general public. Few speeches or oral interventions during meetings remain on public record in any form other than a summary record, which is written by a UN precis writer and which reflects the content but not necessarily the language forms of the original spoken text. In spite of the differences in linguistic features inherent in these written texts, some of the results of these studies are pertinent to the present study on UNGA addresses. This section gives a brief overview of the different types of research carried out on international organization written documents and details specific findings of relevance to the present linguistic analysis of scripted texts.

Some of the research on international organization written documents has focused on text organization and the text creation process. For example, the objective of Duchêne's (2008) study of UN discourse on linguistic minorities is to determine how documents on the rights of linguistic minorities have emerged. He examines transcripts, summary records, annual reports, working papers, and resolutions on linguistic minorities, focusing on how information is modified as discussion in working groups is synthesized in summary reports and then finalized in resolutions and treaties. Duchêne's primary focus is on the content that is retained or omitted rather than the specific linguistic structures used or eliminated. However, of relevance to the present study is his observation that in the process of creating a summary record from a spoken intervention, precis writers omit emotional or persuasive language for the sake of added objectivity. This finding reveals first that objectivity is a valued attribute in a UN text. In addition, it indicates that spoken texts do feature some emotional and persuasive language.

Valentine and Preston (2002) also study the process of writing a summary record at two UN agencies, the International Seabed Authority and the Caribbean Environment Programme. Like Duchêne (2008), they investigate the qualities of the summary record, finding the importance of discernment, discretion, and diplomacy. Their investigation centers on additions or omissions to the content of a spoken intervention as well as style and tone. Two key findings from Valentine and Preston are pertinent to the present study: (1) the use of anonymous subjects (e.g., *one speaker*, *some argue*, and the passive voice) allowing for greater neutrality and obscured agency and (2) the frequency of expressions of agreement (e.g., *commend, support, hail, applaud, issues of common concern, consensus*), implying shared values and objectives.

Other studies on written texts have looked primarily at content and lexical forms at a broad level. Krijtenburg and de Volder (2015) explore how the term *peace* is contextualized in four key UN documents: the UN Charter, *An Agenda for Peace (Agenda 1992), Supplement to an Agenda for Peace (Supplement 1995),* and the *World Summit Outcome (Outcome 2005)*. They create cultural scripts (see Wierzbicka, 1997), showing both similarities and differences between the UN notion of *peace* and *peace* in the Giryama culture of Kenya. Their focus is on themes rather than specific lexical features, but their finding that *peace* is linked to other terms like *democracy* and *development* in these UN documents reveals an interesting lexical association of potential import to the present study.

A much more detailed lexical analysis is undertaken in Baker and McEnery (2005), investigating the terms *refugee* and *asylum seeker* in United Nations High Commissioner for Refugees (UNHCR) documents and in newspapers. Their CADS presents frequency data, collocations, and metaphors associated with the terms *refugee* and *asylum seeker*. Collocates are found through computer programs and analyzed for patterns, both lexical (e.g., *flood of refugees*)

and grammatical (e.g., pre-modifying quantifiers before *refugees*). They examine specific phrases (e.g., *rejected asylum seekers*) and broad categories (e.g., terms related to the honesty of asylum seekers such as *genuine, illegal, bogus*) to find themes in associations made with these two terms (e.g., *refugees* in metaphors of movement, tragedy, and crime). The study underscores some of the advantages of undertaking both quantitative and qualitative analysis. The corpus methods employed by Baker and McEnery allow for the examination of a much greater number of texts, thereby increasing generalizability. In addition, the metaphors suggested by Baker and McEnery are solidly supported with patterns of association based on the most frequent collocates.

Finally, D'Acquisto (2017) examines the language of Security Council and General Assembly Resolutions on Palestine, looking specifically at sentence structure (patterns of coordination and subordination), legal style, and the use of modals and semi-modals. Of particular relevance to the present study is her finding that modals are often in the passive voice or in other impersonal constructions in order to mitigate directive force.

These studies on written documents do not set out to undertake a complete lexical or grammatical analysis of their texts, but they do shed some light on the types of features prevalent in international organization discourse.

2.4.2. International organization spoken discourse

Studies of international organization speeches have reported findings similar to both written international organization documents (see Section 2.4.1) and to other political speeches (see Section 2.3.3). Like with the research on written international organization texts, none has set out to produce a comprehensive analysis of lexical and grammatical features, but some findings are salient to the present study.

Some studies of international organization speeches have focused on text organization but have identified certain lexical and grammatical patterns. Donahue and Prosser (1997), for instance, explore discourse and rhetorical patterns in their corpus of UN General Assembly General Debate addresses from 1945 to 1995. They find UNGA addresses follow a relatively strict format: congratulating or at least greeting the current president of the proceedings, asserting the importance of the UN and its objectives, and then commenting on regional and world issues. They also present a small case study examining the use of *we* and *I* in United Nations addresses. Donahue and Prosser note how the first person singular is used more frequently among speakers who might have felt the need to assert their "footing" or credentials, such as African leaders from newly decolonized countries. This opposition between *we* and *I* shows a contrast between solidarity and individuality.

Other studies of international organization speeches have examined content, with some analysis of lexical features but at a broad topical level. For instance, Baturo, Dasandi, and Mikhaylov (2017) compiled UNGA General Debate addresses from 1970-2014 and suggest researchers use text scaling methods to identify national preferences on policy issues based on the topics raised in these texts. They illustrate how this can be done using correspondence analysis to show a positive relationship between expressing concern over security and terrorism in a UNGA address and signing a nonsurrender agreement. Baturo et al. argue that UNGA General Debate addresses are the best texts for analyzing country positions because speakers are relatively free to express their country's views on a wide range of issues and often use the UNGA as a forum to influence international perceptions of their country. The topics raised in UNGA addresses offer insight into issues of national concern and can therefore be used to predict a country's position on policy.

Other studies examining content and, in a very general sense, lexicon have focused on a very different purpose: identifying how speakers construct identity of their organization and position political relationships. In a critical metaphor analysis, McEntee-Atalianis (2013) compares the speeches of two consecutive Secretary Generals at the International Maritime Organization (IMO). This qualitative analysis investigates lexical features only through their metaphorical use, but some of her excerpts illustrate features noted in other studies of international organization discourse. For instance, in a text excerpt from Secretary General O'Neil about the process of creating a new convention, he expresses implicit disapproval of the diminished enthusiasm of those responsible for developing a convention over the course of the project (see Text 2.3).

Text 2.3. Excerpt from Secretary General O'Neil (McEntee-Atalianis, 2013, p. 329)

> When the journey begins and a new convention is being developed, everyone is happy to crowd on board. But many seem to disembark before the next leg of the journey – that is ratification of the convention – and still more have left the boat by the time we get around to the chore of fully implementing the convention and all its requirements. And when it comes to making sure ships maintain their quality throughout their working lifetimes our ferry is almost empty.

Rather than directly criticize IMO members, he uses anonymous subjects (*everyone, many, still more*). He also uses the passive (*a new convention is being developed*) to focus on the goal rather than the human agents. Finally, he avoids agency by describing a final state (*our ferry is almost empty*) rather than explicitly name the agents who have jumped ship.

Few studies of international organization speeches have specifically set out to explore linguistic features. However, one qualitative analysis of two speeches delivered before the United Nations is noteworthy in its investigation of a wide range of lexical and grammatical features. Najarzadegan, Dabaghi, & Eslami-Rasekh (2017) compare two addresses, one delivered

by US President Obama and the other by Iranian President Rouhani at the General Assembly

General Debate. Their findings illustrate many of the features noted in other studies of scripted

political texts, from rhetorical questions to polarization. Like with other political speeches, these

two UNGA addresses express stance through the use of *I* with mental verbs (e.g., Obama's *I*

believe it is in the security interest of the United States, p. 770). Najarzadegan et al. also note the

use of factive verbs such as *acknowledge* to show the certainty of a proposition. For instance,

Obama states: *There are going to be moments where the international community will need to*

acknowledge that the multilateral use of military force may be required to prevent the very worst

from occurring (p. 772).

Some of their most interesting findings include nominalizations and passives to obscure

agency. When Rouhani declares that *unjust sanctions, as manifestation of structural violence,*

are intrinsically inhumane and against peace (p. 769), he places *sanctions* in the subject position

so that the agent who imposed the sanctions is left unnamed. Passives are also used to avoid

placing direct blame on the accused. In *coercive economic and military policies and practices*

geared to the maintenance and preservation of old superiorities and dominations have been

pursued in a conceptual mindset that negates peace, security, human dignity, and exalted human

ideals (p. 769), Rouhani uses the passive voice rather than explicitly designate the parties who

have pursued these coercive economic and military policies and practices. Obama also uses the

passive to avoid explicitly naming an agent in *to succeed, conciliatory words will have to be*

matched by actions that are transparent and verifiable (p. 771).

The two speeches in Najarzadegan et al. (2017) offer numerous illustrations of lexical

and grammatical features that have been noted in other studies on political speech. Though their

study is limited to just two UNGA addresses, their findings are most relevant to the present

study, offering a glimpse of the types of frequent lexical and grammatical features to be expected in a scripted international organization text.

2.4.3. Summary of findings on linguistic features in international organization discourse

Few studies on international organization discourse, written or spoken, have focused specifically on the use of linguistic features, but the literature has nonetheless revealed some interesting findings. Studies on international organization texts have found an emphasis on joint objectives and cooperation. Features that accomplish this include frequent use of the first person plural as well as lexicon related to support and consensus (e.g., *agreement, endorse, without objection*). Studies have also shown a preference for neutrality and the tendency to obscure agency through anonymous subjects, nominalizations, and the passive voice. The passive voice is used in particular to mitigate the directive force of modals such as *must*. Finally, opinion is expressed through the use of mental verbs such as *believe* but also presented as fact through the use of factive verbs such as *acknowledge*. The high frequency of these features reveals some commonalities with other political speeches, but much remains to be explored.

Research on the language of international organizations has revealed a compelling area of study, but the focus of most international organization discourse analyses, as with most of the studies on scripted political texts, has not been on linguistic features. Moreover, to date, no study has undertaken a comprehensive register analysis with a detailed description of linguistic features characterizing international organization discourse and a functional interpretation to explain particularly high and low frequencies. Register studies have greatly informed our understanding of other discourses such as university language (Biber, 2006). Similarly, a register study of international organization discourse could help fill a substantial gap in the current literature.

2.4.4. Making a case for the study of UNGA addresses

The review of the literature on international organization discourse in Sections 2.4.1-
2.4.3 has revealed a sizeable gap in terms of research questions posed, with no study to date
undertaking a full register analysis to describe the lexical and grammatical features of this
important text type. Section 2.4 has also shown that the language of international organizations is
vast and comprises several sub-registers, arguably registers in their own right (e.g., summary
records, the UN Charter, General Debate addresses), each with distinctive situational
characteristics. An investigation of speeches would certainly result in an analysis of the use of
first person pronouns, as the literature on political speeches has shown this to be a compelling
area of study; in contrast, first person pronouns are unlikely to emerge in a study of summary
reports, which are intended as factual accounts of the proceedings of a meeting without any
involvement from the precis writer. Controlling for situational characteristics as much as
possible and focusing on one precise register within the larger domain of international
organization discourse is essential in order to home in on the linguistic features that are truly
typical of one text type. In Section 2.4.4, I present my rationale for selecting UNGA addresses as
my target register.

Addresses delivered at the United Nations General Assembly General Debate are
compelling for two reasons: they are spoken texts and they are delivered by high-ranking
politicians. Most studies on international organization discourse have focused on written
documents because fewer spoken texts are publicly available. Generally only addresses delivered
by high-ranking diplomats (e.g., Secretary Generals, heads of state or government) are kept in
the public archives. Unlike summary records, such speeches, texts written to be spoken, reflect

the actual language use of politicians. A summary record reports the content of an intervention, but only the speech itself reveals the linguistic forms and rhetorical devices.

UNGA addresses are some of the few speeches made available to the public because of their importance in both international and national politics. These texts are delivered by presidents, prime ministers, foreign ministers, and other high-ranking representatives from every UN member state. The General Debate is a forum for presenting national positions and policies and making recommendations for future UN action. The General Debate is also an opportunity for member states to solidify their relationships with other countries and to improve their role within the organization itself. These addresses are thus of critical importance. Becker's "hierarchy of credibility" (1967) argues that statements made by powerful people are more influential. The opinions of the highest ranking individuals in a group are considered to be more credible as the powerful are assumed to have full knowledge of a given situation. The more powerful or influential the language, the more effort must inevitably be made to avoid misstatement and misinterpretation. Just as eloquent words can have impressive consequences in fostering diplomacy, so too can a misstep have disastrous ramifications. The powerful nature of international organization discourse makes it a compelling subject of study. Addresses delivered by high ranking officials at a venue such as the UNGA General Debate are particularly appealing because they are afforded more credibility. One of the driving questions behind research on diplomacy and international organization discourse is how powerful political players use language to exert influence, and UNGA General Debate texts are well positioned to address that question.

The importance placed on UNGA addresses by the UN organization itself is reflected in the records kept and made readily available to the public (United Nations General Assembly,

n.d.). Each statement is accompanied by an English-language summary, a video, and audio files of the original and of translations in all the official UN languages. This is in stark contrast with other UN speeches, which are either not available to the public or available only in the form of a summary report prepared by a precis writer.

UNGA addresses represent a key type of text in international organization discourse. The register is narrow enough to find commonalities across texts, with a sufficiently restricted set of situational characteristics, but also broad enough to investigate variation, given that speakers potentially come from every country in the world.

2.5. Conclusion

Chapter 2 has reviewed some of the most relevant literature on spoken discourse and scripted political speeches in particular. It began with a description of three productive methodologies used to analyze spoken discourse today: CDA, CADS, and register analysis. In this overview of these three types of studies on spoken language, I recognized the importance of two types of interpretation of linguistic features: (a) interpretation based on political, social, or cultural identity and relationships (CDA and CADS) and (b) interpretation based on situational characteristics such as production circumstances and purpose (register analysis). I showed how techniques such as keyword analysis, collocational analysis, key feature analysis, and Multi-Dimensional analysis have been used in studies on the linguistic features of diverse types of spoken texts, from television situational comedies to parliamentary debate, from classroom lectures to news broadcasts.

In the last sections of Chapter 2, I presented an overview of findings on the linguistic features from planned monologues (classroom lectures, conference presentations) and from scripted texts (television broadcasts, sermons). I then gave a more detailed account of findings

from scripted political texts and from written and scripted international organization discourse. This review of the literature has revealed two fundamental gaps in the literature:

1. No systematic investigation of linguistic features has been carried out of scripted political texts. In fact, very little has been reported about the linguistic differences between "hybrid" scripted texts and other text types.

2. No comprehensive linguistic analysis has been undertaken on international organization discourse, and more generally, little attention has been paid to the linguistic features of scripted international organization texts and how these features differ from other registers such as prepared speeches, spontaneous speeches, and written texts.

This study addresses that gap by identifying the distinctive lexical and grammatical characteristics of UNGA addresses when contrasted with other spoken and written registers, providing a functional interpretation of any differences in linguistic features based on the situational characteristics and discourse functions of UNGA addresses. This study also distinguishes systematic patterns of lexical and grammatical variation within UNGA addresses, accounting for any differences in linguistic features based on the geographic, social, political, and economic characteristics of the countries being represented by the speakers. To accomplish these goals, the study uses two of the methodologies described in Chapter 2: register analysis (Chapters 4 and 5) and CADS (Chapter 6). Register analysis is the primary methodology adopted because it allows for a comprehensive description of the linguistic features that characterize UNGA addresses. CADS is used in addition to register analysis in order to explore expression of political ideology, identity, and relationships through the use of a few pre-selected terms. As

discussed in Section 2.2.4, register analysis and CADS are potentially complementary methodologies. Adopting both can present a fuller picture of the linguistic features of the UNGA register.

Before reporting my findings in Chapters 4 through 6, I describe in Chapter 3 my corpora and the procedure I followed to carry out the study.

CHAPTER 3: CORPUS AND DATA ANALYSIS

3.1. Introduction

The goal of this chapter is to describe the methods used to collect, clean, code, annotate, and analyze the UNGA data. In Section 3.2, I present an overview of the analyses undertaken in this study. In Section 3.3, I explain my motivation for collecting speeches from the UNGA General Debate and then describe the contents of the UNGA corpus, with word count and country name for each file. Section 3.3 also offers a brief description of five comparison corpora. Section 3.4 details how the text files were converted and cleaned and Section 3.5 how texts were annotated or "tagged" for part of speech. Section 3.6 reports the taxonomy for text coding, and Section 3.7 outlines the data analysis procedure, including keyword analysis, Multi-Dimensional analysis, and key feature analysis.

3.2. Overview of the Analyses

This study uses quantitative and qualitative analyses in order to investigate the linguistic features of UNGA addresses. My first research goal is to describe the linguistic features of the UNGA as a whole compared to other spoken and written registers. In order to identify distinctive lexical features, I carried out a keyword analysis using the British National Corpus (BNC) as a reference corpus. Keywords for the UNGA were identified using WordSmith 7.0 (Scott, 2017) and categorized into semantic domains. By examining concordance lines and the co-text around each keyword, I was able to see how the keywords were being used. To clarify some particularly frequent or ambiguous keywords (e.g., *situation*), I did a collocational analysis to determine whether any words co-occurred with the keyword with significantly high frequency. In order to interpret the function these keywords served, I considered which situational characteristics of the UNGA explained their particularly frequent use. To investigate the grammatical features of the

UNGA, I used Multi-Dimensional (MD) scores along five dimensions, which I had obtained

using the Biber Tagger and TagCount (see Biber, 1988; Biber et al., 1999). I compared MD

scores for the UNGA with scores for four comparison registers: prepared speeches, spontaneous

speeches, conversation, and official documents. I interpreted higher or lower dimension scores

(representing higher or lower frequencies of co-occurring linguistic features) by analyzing the

relationship between the situational characteristics of the UNGA and these linguistic features,

with particular attention to how situational characteristics differed between the UNGA and the

four comparison registers. That is, I explained the relative importance of one function and

unimportance of another function based on the communicative context of the UNGA.

My second research goal is to identify variation within the UNGA. To achieve this goal, I

first needed to identify how situational characteristics varied within the UNGA and I determined

that variation occurred based on geographic, political, economic, and social factors of the

countries represented by the texts. I grouped each text (each country) into categories (e.g., which

region the country was located in). For each analysis relating to the second research goal, the

target corpus was the collection of texts in one country group (e.g., Africa) and the reference

corpus was the remaining texts, those not in the country group (e.g., Americas, Asia, Europe, and

Oceania). To investigate lexical features, like with my first research goal, I carried out keyword

analysis. To investigate grammatical features, I carried out key feature analysis and MD analysis.

For key feature analysis, I calculated the mean (M) and standard deviation (SD) of 126 lexico-

grammatical features for each target corpus (each country group) and for each reference corpus

(the texts not in the country group). I calculated Cohen's d to determine whether there was a

large effect size for the difference between the M of each target corpus (country group) and the

M of its reference corpus. Where there was a large effect size, I interpreted the frequency of the

51

feature based on the situational characteristics of the country category. The MD analyses by

country group were similar to the analysis for the UNGA as a whole, described above, with the

exception that not all results were discussed (though they were all reported). Like with the key

feature analysis, I analyzed only dimension scores for a country category when there was a

notable difference in MD scores. Table 3.1 summarizes this overview of the analyses and

indicates the chapter in which the results can be found. More detailed descriptions of the steps

for each of these analyses can be found in Section 3.7.

Table 3.1. Overview of the Analyses

Research Goal	Analysis	Programs Used	Target Corpus	Reference Corpus	Chapter
Identify the lexical features of the UNGA	Keywords	WordSmith 7.0 (Scott, 2017)	UNGA	British National Corpus	4
Identify the grammatical features of the UNGA	Multi-Dimensional	Biber Tagger, TagCount (Biber, 1988; Biber et al., 1999)	UNGA	1. Prepared speeches 2. Spontaneous speeches 3. Conversation 4. Official documents (Biber, 1988)	4
Identify patterns of variation in lexical features of the UNGA	Keywords	WordSmith 7.0 (Scott, 2017)	UNGA texts representing a country category	Remaining UNGA texts not in the country category	5
Identify patterns of variation in grammatical features of the UNGA	Key features	Biber Tagger, TagCount (Biber, 1988; Biber et al., 1999)	UNGA texts representing a country category	Remaining UNGA texts not in the country category	5
Identify patterns of variation in grammatical features of the UNGA	Multi-Dimensional	Biber Tagger, TagCount (Biber, 1988; Biber et al., 1999)	UNGA texts representing a country category	Remaining UNGA texts not in the country category	5

3.3. Data Collection

Section 3.3 describes the data collection procedure. In Section 3.3.1, I present my motivation for collecting a corpus of addresses delivered in English during the General Debate of the 70th Session of the UNGA in fall 2015. In Section 3.3.2, I describe the corpora used: the UNGA corpus, including countries represented and word counts for each text (Section 3.3.2.1) and the five comparison corpora (Section 3.3.2.2).

3.3.1. Corpus motivation

I chose to examine UNGA General Debate addresses because the primary objective of this study is to describe the linguistic features of scripted international organization spoken discourse. The UNGA is arguably the most powerful and certainly the most representative international body. Speakers at the General Debate are high-ranking politicians such as presidents and prime ministers. Written texts, audio files, and videos of the addresses are made available to the public, so speakers respect a high level of decorum. In addition, the sample is most representative of truly global discourse because all 193 UN member states and "observers" such as the Holy See and Palestine are invited to speak during the General Debate, and it is rare for a country to decline. Finally, unlike UN specialized agencies and international institutions outside of the UN family, all speech files from the UNGA General Debate can be downloaded at no cost or special request, so it is easy to collect the texts.

I collected speeches from the fall 2015 UNGA General Debate. I selected only the texts that had been originally written and delivered in English. UNGA addresses are translated and published in all six of the official UN languages (viz., Arabic, Chinese, English, French, Russian, and Spanish), but I wanted to avoid the moderating variable of translation.

3.3.2. Corpus description

Section 3.3.2 offers a detailed description of the contents of the UNGA corpus (Section 3.3.2.1) as well as a brief sketch of five comparison corpora (Section 3.3.2.2).

3.3.2.1. UNGA corpus description

The UNGA corpus was built from all English-language addresses delivered during the General Debate of the 70th session of the UNGA (28 September – 3 October 2015). Each UNGA speech is available on line as a written pdf file, an audio file, and a video file. I listened to each audio file to ensure that I had every speech delivered in English and only speeches delivered in English. The result was a collection of 92 speeches, amounting to 174,994 words from all five regions of the world (see Table 3.2 for word count by country).

It deserves mention that the 92 countries in the UNGA corpus represent just under half of the total number of countries that spoke during the UNGA General Debate in 2015. Not included in this corpus are the addresses from 101 other countries (including Palestine, not a UN member state but with observer status). Certain regions are particularly under-represented. Based on the geographical sub-regions designated by the UN Statistics Division (2016), the UNGA corpus includes no countries from Northern Africa, Middle Africa, or Central Asia. Only four of the 27 countries from Northern, Central, and South America are represented in the UNGA. In addition, seven powerful countries, either permanent members of the Security Council or members of the G7 industrialized countries, are not represented: Canada, China, France, Germany, Italy, Japan, and Russia. Though I have called the corpus "UNGA," it comprises only texts that were delivered in English, again for the very important reason that including the other texts would have introduced a potentially critical moderating variable: translation.

54

Table 3.2. UNGA Word Count by Country

Region	Country	Tokens
Africa		**38,554**
	Botswana	2942
	Eritrea	821
	Ethiopia	1496
	Ghana	2635
	Kenya	2764
	Liberia	1008
	Malawi	1558
	Mauritius	2689
	Namibia	2666
	Nigeria	1755
	Rwanda	548
	Seychelles	1529
	Sierra Leone	2453
	Somalia	1776
	South Africa	2061
	South Sudan	1161
	Swaziland	2008
	Tanzania	1931
	Uganda	958
	Zambia	2090
	Zimbabwe	1705
Americas		**28,667**
	Antigua	2232
	Bahamas	1790
	Barbados	2237
	Belize	1999
	Grenada	1547
	Guyana	1551
	Jamaica	2532
	Saint Kitts and Nevis	1919
	Saint Lucia	2420
	Saint Vincent and the Grenadines	2108
	Suriname	1306
	Trinidad	2379
	US	4647
Asia		**36,379**
	Afghanistan	1814
	Azerbaijan	1588

Region	Country	Tokens
Asia (continued)		
	Bhutan	1724
	Brunei Darussalam	1405
	Cambodia	716
	Cyprus	2022
	Georgia	2151
	Indonesia	1196
	Israel	3706
	Jordan	1346
	Malaysia	2221
	Maldives	1552
	Mongolia	2829
	Myanmar	2119
	Nepal	2129
	Pakistan	1757
	Philippines	2083
	Turkey	2520
	Viet Nam	1501
Europe		**45,701**
	Albania	1014
	Austria	1406
	Bulgaria	1817
	Croatia	3235
	Denmark	1485
	Estonia	1827
	EU	997
	Finland	1085
	Greece	1659
	Holy See	2827
	Iceland	1836
	Ireland	2549
	Latvia	1482
	Liechtenstein	1506
	Lithuania	767
	Malta	2370
	Moldova	2105
	Netherlands	1839
	Norway	1436
	Romania	1962
	San Marino	1013
	Slovakia	1785

Region	Country	Tokens
Europe (continued)		
	Slovenia	1452
	Sweden	1425
	UK	1708
	Ukraine	3114
Oceania		**25,643**
	Australia	1863
	Fiji	2415
	Kiribati	2299
	Marshall Islands	1701
	Micronesia	1680
	Nauru	1616
	New Zealand	1828
	Papua New Guinea	1599
	Samoa	2492
	Solomon Islands	2834
	Tonga	1642
	Tuvalu	2153
	Vanuatu	1521
Total		**174,994**

3.3.2.2. *Comparison corpora description*

My first research goal is to describe the UNGA as a register in contrast to other spoken and written registers. To this end, I needed to find comparison corpora that would serve as an appropriate and revelatory contrast, to show most clearly what makes the UNGA distinctive. Comparison corpora can be general or specialized, depending on the research goals. A very large reference corpus is used to identify unusually frequent or infrequent linguistic features in the target register (the UNGA). In contrast, specialized corpora are used to tease out nuances in similar registers.

One reference corpus and four specialized, smaller corpora were used in the present study. I selected the British National Corpus (BNC) as a reference corpus (see Baker, 2006) for the lexical (keyword) analysis. The BNC is often chosen as a reference corpus because of its size

and representativeness; the 100-million-word corpus is composed of approximately 90% written and 10% spoken texts and spans a wide range of registers. Comparing the UNGA to the BNC reveals the features distinguishing the UNGA from written and spoken British English as a whole.

I also chose four comparison registers for the MD analysis of grammatical features: prepared speeches, spontaneous speeches, conversation, and official documents. These collections of texts are described in Biber (1988) and are drawn from two large corpora: the Lancaster-Oslo-Bergen Corpus of British English (LOB, see Johansson, Leech, & Goodluck, 1978) and the London-Lund Corpus of Spoken English (Svartvik & Quirk, 1980). Prepared speeches include sermons, university lectures, court cases, political speeches, and popular lectures; spontaneous speeches are court cases, dinner speeches, radio essays, and speeches in the British House of Commons; face-to-face conversation includes interactions between intimates as well as "distants;" official documents comprise government documents, foundation reports, industry reports, and college catalogues. Table 3.3 summarizes these four comparison corpora and details their text and word counts along with counts for the UNGA. Word counts are calculated based on descriptions of the LOB Corpus, whereby each text is approximately 2,000 words and of the London-Lund Corpus, whereby each text is approximately 5,000 words.

Table 3.3. Corpora Used in the Grammatical (Multi-Dimensional) Analysis

Corpus	Text Description	Text Count	Word Count
UNGA	Addresses delivered at the General Assembly General Debate of the 70th Session	92	174,994
Prepared speeches (Biber, 1988)	Sermons, university lectures, court cases, political speeches, popular lectures	14	70,000
Spontaneous speeches (Biber, 1988)	Court cases, dinner speeches, radio essays, speeches in the House of Commons	16	80,000
Conversation (Biber, 1988)	Interactions between intimates and "distants"	44	220,000
Official documents (Biber, 1988)	Government documents, foundation reports, industry reports, college catalogues	14	28,000

I selected these four comparison registers based on their situational characteristics. Prepared speeches most resemble the UNGA register; spontaneous speeches would allow me to identify which features could be attributed to advance preparation of a text; conversation would serve as a sharp contrast because it is the most informal, unplanned spoken register; official documents would also allow for a contrast but at the other end of the spectrum as the most formal, least involved written text type. Section 4.3.2 presents the situational characteristics of these four registers as support for their use as comparison corpora.

3.4. Corpus Cleaning

The comparison corpora needed no further preparation, but the UNGA corpus had to be converted to plain text and annotated for part of speech (POS) for the analysis of grammatical features and computation of dimension scores. Concordance programs used in the analyses also require texts to be in plain text format. All pdf files were converted into plain text files using Zilla. The text files were cleaned in order to ensure there were no spelling errors due to reformatting (e.g., *development* becoming *deveBopment*) and to delete all words that were not

spoken (e.g., cover page with name of country and speaker, page numbers). However, any lexical

or grammatical errors in the original pdf files were not corrected (e.g., *crush* rather than *crash* in

Malaysian MH17 plane crush, Ukraine).

3.5. Grammatical Annotation

UNGA texts were then annotated with part of speech and other lexico-grammatical

information. Like numerous other corpus-based register studies (e.g., Biber, 2006; Quaglio,

2009), I used the Biber Tagger (see Biber, 1988; Biber et al., 1999), a program that identifies 145

grammatical features such as proper nouns, first person pronouns, predicative attitudinal

adjectives, suasive verbs, modals of prediction, *that* deletion, WH- clauses, , and phrasal

connector coordinating conjunctions. After tagging, Biber's TagCount was run to get frequencies

of each feature normed to 1,000 as well as total word count, type-token ratio, and average word

length for each text. The TagCount program also calculates factor scores for five dimensions

found to distinguish text varieties (see discussion in Section 3.7.2).

Some of the features identified using the Biber Tagger are extremely detailed and not

particularly relevant to UNGA texts, so they were not retained for analysis. For instance,

intransitive activity phrasal verbs (e.g., *come on, sit down*) did not appear in 84 of the 92 UNGA

texts and appeared only once or twice in the remaining texts so this feature was not analyzed.

126 features were ultimately retained for analysis. (See Appendix A for a list of these features).

3.6. Country Categorization

My second research goal is to identify any systematic linguistic variation within the

UNGA. In order to do this, I first had to establish where there was variation in the situational

characteristics of the UNGA so that I could group and compare texts by those situational

characteristics. I determined that the variation that occurred stemmed from the country

represented by the text (see Section 4.3). Thus, the most natural way to group texts was based on

country characteristics. The next step was to identify which country characteristics might be

meaningful for the description of linguistic variation across types of countries. Section 3.6.1

reports the procedure for categorizing texts based on country characteristics for comparison

within the UNGA corpus. Section 3.6.2 then details the results of the coding.

3.6.1. Category selection and operationalization

In order to explore variation within the UNGA, countries were categorized based on

geographic, political, economic, and social factors. In order to determine categories, I examined

documents from international organizations (the UN; the WTO), a national organization (the

Central Intelligence Agency), and a nonprofit organization (Pew Research Center) as well as

academic literature on international relations (McCarthy, 1998; Sullivan, 1990; Vital, 1967) and

political discourse (Donahue & Prosser, 1997) to investigate how countries are commonly

classified. I also re-read the UNGA addresses themselves to determine whether any documented

country categories appeared particularly relevant.

This review of the literature resulted in a list of country features comprising region, area,

population, combined economic and social development (Human Development Index, HDI),

economic power (Gross Domestic Product, GDP), Least Developed Country (LDC) status, Small

Island Developing State (SIDS) status, Landlocked Developing Country (LLDC) status, military

spending, and religious non-affiliation. These country categories were operationalized, whenever

possible, based on UN classifications. Where no UN classification was found, data from other

international organizations or government agencies (e.g., the World Bank) were used. Each

country category is described here in turn, with a rationale for selecting it and a description of

how it was operationalized. Table 3.4 summarizes country category operationalization.

Table 3.4. Country Category Operationalization

Country category	Operationalization	Levels
Region	UN Statistics Division	(1) Africa (2) Americas (3) Asia (4) Europe (5) Oceania
Area	Central Intelligence Agency's *World Factbook*	(1) Small (0-7,999 km²) (2) Medium (8,000-79,999 km²) (3) Large (80,000-379,999 km²) (4) Very large (380,000-9,000,000 km²)
Population	Central Intelligence Agency's *World Factbook*	(1) Small (0-499,999) (2) Medium (500,000-4,999,999) (3) Large (5,000,000-19,999,999) (4) Very large (20,000,000 +)
Human Development Index (social and economic development)	UN Development Program Based on indicators related to life expectancy, years of schooling, and per capita Gross National Income	(1) Low (2) Medium (3) High (4) Very high
Gross Domestic Product (economic power)	World Bank	(1) Low ($0-$10,000) (2) Medium ($10,001-$20,000) (3) High ($20,001-$30,000) (4) Very high ($30,001 +)
Least Developed Country	UN Statistics Division	(1) Yes (2) No
Small Island Developing State	UN Statistics Division	(1) Yes (2) No
Landlocked Developing Country	UN Statistics Division	(1) Yes (2) No
Military Spending	World Bank	(1) Low (0-1.9% GDP) (2) High (2% + GDP)
Religious Non-affiliation	Pew Research Center	(1) Low (less than 1%) (2) High (1% or more unaffiliated)

Countries are typically categorized by region in international organizations (e.g., UN), national organizations (e.g., CIA), and in academic texts and research (e.g., Donahue & Prosser, 1997). Region was operationalized according to the classification used by the UN Statistics Division, with five macro geographical regions: Africa, the Americas, Asia, Europe, and Oceania.

Size is also a common country classification (e.g., *World Factbook*), with particularly frequent reference to small state status (see World Bank, UN Statistics Division). As a result, area and population were added to the list of country categories. For both area and population, countries were coded as small, medium, large, or very large based on figures from the CIA's *World Factbook.* The levels, small through very large, were operationalized to create close to equal numbers of countries in each group while respecting natural breaks in the data. For instance, no country fell between 71,740 km² and 83,871 km² so 80,000 km² was a natural cut-off.

Economic might is an important factor distinguishing countries (see UNDP, World Bank). Increasingly, however, organizations such as the UN Development Program (UNDP) recognize the importance of social development in addition to economic power. Therefore, both economic power and combined social and economic development were selected as country categories in this study. Social and economic development was operationalized based on the UNDP Human Development Index (HDI), which is calculated using indicators related to life expectancy, years of schooling, and per capita Gross National Income. The UNDP classifies countries as low, medium, high, and very high HDI (see United Nations Development Programme, n.d.), and these levels were retained for this study.

Both Gross Domestic Product (GDP) and Gross National Income (GNI) are common indicators of economic power (see World Bank). For the purposes of this study, total per capita GDP was used for economic might because GNI was already a variable in HDI. GDP data came from the World Bank (2016a), and countries were categorized as low, medium, high, or very high, with levels determined based on logical breaks in the data.

In addition to these two indicators of development, the UN Statistics Division classifies three types of countries based on potential for growth: Least Developed Countries (LDCs), Small Island Developing States (SIDSs), and Landlocked Developing Countries (LLDCs). According to the United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States (UN-OHRLLS, 2018), LDCs are low income countries faced with serious obstacles to sustainable development; SIDS are island nations with similar sustainable development challenges (e.g., few resources, small domestic market, high costs for energy and transportation); LLDCs are low income countries facing trade and development constraints due to their remoteness. I initially hesitated to add these three developmental groupings, skeptical of their importance. Only after re-reading the texts did it become evident that these categories were fundamental to how many states viewed their position within the international community. Numerous speakers referred to their country's status (see Text 3.1) or to the unique concerns of their group (see Text 3.2).

Text 3.1. Excerpt from Barbados

> Barbados, a Small Island Developing State, has been, and remains a staunch supporter of the United Nations, committed to the principles of multilateralism and the sovereign equality of States, and to the inextricable link between peace and security on the one hand and economic and social development on the other.

Text 3.2. Excerpt from Botswana

> Botswana, like other LLDCs will require substantial augmentation of its capacity to implement any one of these areas.

Military spending potentially impacts a nation's economy (Faini, Annez, & Taylor, 1984) and more generally can affect international relations (Sullivan, 1990). Military expenditure is therefore an important country indicator, often cited as percentage of GDP. However, no indicators appear to determine the percentage of GDP that constitutes "low" or "high" military spending. I therefore examined spending figures from the World Bank (2016b) and noted that the majority of UNGA countries spend between 1-1.5% GDP on defense, with a mode of 1.1%. Dividing the countries in half was illogical because the cut-off would be at 1.3%, with 10 countries in the "low" category spending 1.2-1.3% and eight countries in the "high" category spending 1.4-1.5%. I therefore created somewhat uneven groups but based on a natural break in the data at 1.9%.

The final category was religious non-affiliation. This group was added after an examination of the texts revealed numerous textual references to God and therefore the possibility that religion was a variable (see Text 3.3).

Text 3.3. Excerpt from Malaysia

> If God promptly rebuked the Prophet Muhammad, how much more will we, the community of Muslim world leaders especially, stand to be rebuked by our Creator if we "frown and turn our faces away" from our fellow-Muslim poor and marginalized, now fleeing Syria in massive numbers – causing social and economic stresses in Europe?

References to *God, Prophet Muhammad, Muslim*, and *Creator*, make this passage from Malaysia particularly religious. Initial attempts to categorize countries by religious majority (e.g., Christian, Muslim) based on data from the Pew Research Center (2012) revealed potential

problems with such classification. Papua New Guinea, for instance, has a population that is over

99% Christian and less than 1% unaffiliated. Nigeria has a population that is 49.3% Christian,

48.8% Muslim, and 0.4% unaffiliated. Sweden's population is 67.2% Christian, 4.6% Muslim,

and 27% unaffiliated. These three countries have Christian majorities, but their religious

composition differs greatly. As a result, religion was operationalized instead as degree of

religious non-affiliation based on data published by the Pew Research Center. A natural break in

the data appeared at 1% religious non-affiliation, with most countries in the UNGA under the 1%

cut-off, indicating near full religious affiliation of their populations. Countries were coded as low

(less than 1% non-affiliation) and high (1% or more).

 This list of country categories is not meant as all-inclusive. Countries can be categorized

in numerous other ways, but these classifications were either irrelevant given the countries

represented in the UNGA or difficult to systematically operationalize. For instance, classifying

countries based on Security Council permanent membership would have resulted in a target

corpus of two (UK and US), so results would not have been very meaningful. Classifying

countries based on English language status can be contentious. Some countries, such as the

United States and Australia, have no official language at the federal level. Others claim English

as the sole official language though most of the population has a different mother tongue (e.g.,

Zambia, Nigeria). In some "English-speaking" countries, less than 1% of the population speaks

the language (e.g., Mauritius, Solomon Islands). Finally, some countries with more than one

official language may not be represented in the UNGA by a speaker who is actually bilingual

(e.g., Ireland, New Zealand).

 Finally, categorizing addresses based not on country categories but rather on

communicative purpose of the text proved challenging during a pilot study and resulted in very

low interrater reliability. Diplomatic rhetoric sometimes requires veiled references, and most criticism, for instance, was not explicit. For example, *Indeed, where successfully implemented, the MDGs have had a positive impact on the lives of millions of people. But the results have been mixed and uneven across countries* [Kenya]. Was this criticism? Even after extension discussion, in many instances raters were unable to reach agreement on communicative purpose.

3.6.2. Country categorization results

Because country-specific characteristics were based on official statistics rather than subjective interpretation, I, alone, coded each country. Results of the country categorization are reported in Tables 3.5 and 3.6. In Table 3.5, data are presented by category, with total number of countries in each group. In Table 3.6, data are presented by individual country.

Table 3.5. Country Classification by Category

Country category	Levels	Number of countries
Region	(1) Africa	21
	(2) Americas	13
	(3) Asia	19
	(4) Europe	26
	(5) Oceania	13
Area	(1) Small (0-7,999 km²)	22
	(2) Medium (8,000-79,999 km²)	24
	(3) Large (80,000-379,999 km²)	23
	(4) Very large (380,000-9,000,000 km²)	22
	(99) Missing	1
Population	(1) Small (0-499,999)	24
	(2) Medium (500,000-4,999,999)	24
	(3) Large (5,000,000-19,999,999)	22
	(4) Very large (20,000,000 +)	22
Human Development Index (social and economic development)	(1) Low	19
	(2) Medium	15
	(3) High	27
	(4) Very high	25
	(99) Missing	6
Gross Domestic Product (economic power)	(1) Low ($0-$10,000)	57
	(2) Medium ($10,001-$20,000)	12
	(3) High ($20,001-$30,000)	4
	(4) Very high ($30,001 +)	14
	(99) Missing	5
Least Developed Country	(1) Yes	19
	(2) No	73
Small Island Developing State	(1) Yes	26
	(2) No	66
Landlocked Developing Country	(1) Yes	15
	(2) No	77
Military Spending	(1) Low (0-1.9% GDP)	47
	(2) High (2% + GDP)	18
	(99) Missing	27
Religious Non-affiliation	(1) Low (less than 1%)	29
	(2) High (1% or more unaffiliated)	60
	(99) Missing	3

As shown in Table 3.5, for region, 21 countries were in Africa, 13 in the Americas, 19 in Asia, 26 in Europe, and 13 in Oceania. For geographical size, 22 countries were coded as small (0-6,000 km²), 24 countries medium (9,000-71,000 km²), 23 countries large (83,000-338,000 km²), 22 countries very large (390,000-9,000,000 km²), and one missing. For population, 24 countries were coded as small (0-436,000), 24 medium (585,000-5,000,000), 22 large (5,200,000-19,000,000), and 22 very large (21,000,000 +). For social and economic development, 19 were categorized as low, 15 medium, 27 high, and 25 very high; six were not categorized by the UNDP and were therefore coded for the purposes of this study as missing. For GDP, 57 countries were categorized as low ($0-$10,000), 12 medium ($10,001-$20,000), four countries were high ($20,001-$30,000), 14 very high ($30,001 +), and five missing. For LDC status, 19 were coded yes and 73 no. For SIDS status, 26 were coded yes and 66 no. For LLDC status, 15 were coded yes and 77 no. For military expenditure, 47 countries were coded as low (0-1.9% GDP), 18 high (2% + GDP), and 27 missing. For religious non-affiliation, 29 were coded low (less than 1% unaffiliated), 60 were high (1% or more), and three missing.

Table 3.6 shows how each country was grouped into the categories. It should be noted that the economic indicators HDI and GDP are by no means synonymous. They produce the same classification for only 31 (about one-third) of the countries (e.g., low for Afghanistan, high for the Bahamas, very high for Australia).

Table 3.6. Country Classification by Individual Country

Country	Region	Area	Pop	HDI	GDP	Mil	LDC	LLDC	SIDS	Rel
Afghanistan	3	4	4	1	1	1	1	1	2	1
Albania	4	2	2	3	1	1	2	2	2	2
Antigua & Barbuda	2	1	1	3	2	99	2	2	1	2
Australia	5	4	4	4	4	2	2	2	2	2
Austria	4	3	3	4	4	1	2	2	2	2
Azerbaijan	3	3	3	3	1	2	2	1	2	1
Bahamas	2	2	1	3	3	99	2	2	1	2
Barbados	2	1	1	3	2	99	2	2	1	2
Belize	2	2	1	3	1	1	2	2	1	2
Bhutan	3	2	2	2	1	99	2	1	2	1
Botswana	1	4	2	2	1	2	2	1	2	2
Brunei Darussalam	3	1	1	4	3	2	2	2	2	1
Bulgaria	4	3	3	3	1	1	2	2	2	2
Cambodia	3	3	3	2	1	99	1	2	2	1
Croatia	4	2	2	4	2	1	2	2	2	2
Cyprus	3	2	2	4	3	1	2	2	2	2
Denmark	4	2	3	4	4	1	2	2	2	2
Eritrea	1	3	3	1	99	99	1	2	2	1
Estonia	4	2	2	4	2	2	2	2	2	2
Ethiopia	1	4	4	1	1	1	1	1	2	1
EU	4	99	4	4	4	1	2	2	2	99
Fiji	5	2	2	3	1	1	2	2	1	1
Finland	4	3	3	4	4	1	2	2	2	2
Georgia	3	2	2	3	1	2	2	2	2	1
Ghana	1	3	4	2	1	1	2	2	2	2
Greece	4	3	3	4	2	2	2	2	2	2
Grenada	2	1	1	3	1	99	2	2	1	2
Guyana	2	3	2	2	1	1	2	2	1	2
Holy See	4	1	1	99	99	99	2	2	2	1
Iceland	4	3	1	4	4	99	2	2	2	2
Indonesia	3	4	4	2	1	1	2	2	2	1
Ireland	4	2	2	4	4	1	2	2	2	2
Israel	3	2	3	4	4	2	2	2	2	2
Jamaica	2	2	2	3	1	1	2	2	1	2
Jordan	3	3	3	3	1	2	2	2	2	1
Kenya	1	4	4	1	1	1	2	2	2	2
Kiribati	5	1	1	2	1	99	1	2	1	1
Latvia	4	2	2	4	2	1	2	2	2	2

Country	Region	Area	Pop	HDI	GDP	Mil	LDC	LLDC	SIDS	Rel
Liberia	1	3	2	1	1	1	2	2	2	2
Liechtenstein	4	1	1	4	99	99	2	2	2	2
Lithuania	4	2	2	4	2	1	2	2	2	2
Malawi	1	3	3	1	1	1	1	1	2	2
Malaysia	3	3	4	3	1	1	2	2	2	1
Maldives	3	1	1	3	1	99	2	2	1	1
Malta	4	1	1	4	3	1	2	2	2	2
Marshall Islands	5	1	1	99	1	99	2	2	1	2
Mauritius	1	1	2	3	1	99	2	2	1	1
Micronesia	5	1	1	2	1	99	2	2	1	99
Moldova	4	2	2	2	1	1	2	1	2	2
Mongolia	3	4	2	3	1	1	2	1	2	2
Myanmar	3	4	4	1	1	2	1	2	2	99
Namibia	1	4	2	2	1	2	2	2	2	2
Nauru	5	1	1	99	1	99	2	2	1	2
Nepal	3	3	4	1	1	1	1	1	2	1
Netherlands	4	2	3	4	4	1	2	2	2	2
New Zealand	5	3	2	4	4	1	2	2	2	2
Nigeria	1	4	4	1	1	1	2	2	2	1
Norway	4	3	3	4	4	1	2	2	2	2
Pakistan	3	4	4	1	1	2	2	2	2	1
Papua New Guinea	5	4	3	1	99	99	2	2	1	1
Philippines	3	3	4	2	1	1	2	2	2	1
Romania	4	3	4	3	1	1	2	2	2	1
Rwanda	1	2	3	1	1	1	1	1	2	2
Saint Kitts & Nevis	2	1	1	3	2	99	2	2	1	2
Saint Lucia	2	1	1	3	1	99	2	2	1	2
Saint Vincent	2	1	1	3	1	99	2	2	1	2
Samoa	5	1	1	3	1	99	2	2	1	2
San Marino	4	1	1	99	99	99	2	2	2	2
Seychelles	1	1	1	3	2	1	2	2	1	2
Sierra Leone	1	2	3	1	1	1	1	2	2	1
Slovakia	4	2	3	4	2	1	2	2	2	2
Slovenia	4	2	2	4	2	1	2	2	2	2
Solomon Islands	5	2	2	1	1	99	1	2	1	1
Somalia	1	4	3	99	1	1	1	2	2	1
South Africa	1	4	4	2	1	1	2	2	2	2
South Sudan	1	4	3	1	1	2	1	1	2	1
Suriname	2	3	2	3	1	99	2	2	1	2
Swaziland	1	2	2	1	1	1	2	1	2	2

Country	Region	Area	Pop	HDI	GDP	Mil	LDC	LLDC	SIDS	Rel
Sweden	4	4	3	4	4	1	2	2	2	2
Tanzania	1	4	4	1	1	1	1	2	2	2
Tonga	5	1	1	3	1	99	2	2	1	1
Trinidad & Tobago	2	1	2	3	2	1	2	2	1	2
Turkey	3	4	4	3	1	2	2	2	2	2
Tuvalu	5	1	1	99	1	99	1	2	1	2
Uganda	1	3	4	1	1	1	1	1	2	1
UK	4	3	4	4	4	1	2	2	2	2
Ukraine	4	4	4	3	1	2	2	2	2	2
US	2	4	4	4	4	2	2	2	2	2
Vanuatu	5	2	1	2	1	99	1	2	1	2
Viet Nam	3	3	4	2	1	2	2	2	2	2
Zambia	1	4	3	2	1	1	1	1	2	1
Zimbabwe	1	4	3	1	1	2	2	1	2	2

Note. Region: 1 = Africa, 2 = Americas, 3 = Asia, 4 = Europe, 5 = Oceania. Area: 1 = small, 2 = medium, 3 = large, 4 = very large. Pop (Population): 1 = small, 2 = medium, 3 = large, 4 = very large. GDP (Gross Domestic Product): 1 = low, 2 = medium, 3 = high, 4 = very high. HDI (Human Development Index): 1 = low, 2 = medium, 3 = high, 4 = very high. Mil (military spending): 1 = low, 2 = high. LDC (Least Developed Country): 1 = yes, 2 = no. LLDC (Landlocked Developing Country): 1 = yes, 2 = no. SIDS (Small Island Developing State): 1 = yes, 2 = no. Rel (religious non-affiliation): 1 = low, 2 = high. 99 designates missing data.

The categorization of UNGA countries serves as a reminder that the UNGA represents just under half of the countries of the world (see discussion in Section 3.3.2.1). On the whole, there is a relatively equal distribution of states, but the country categorization shows, for instance, a much greater representation of countries from Oceania (13 in the corpus out of a total of 25, or 52%) than the Americas (13 out of 55, or 24%). Likewise, the UNGA includes a greater proportion of countries with higher social and economic development (HDI). Of the 106 total countries classified by the UNDP as high and very high HDI, 52 are in the UNGA (49%). Of the 82 countries classified as low and medium HDI, only 34 are in the UNGA (41%). The difference is not flagrant, but should be noted, nonetheless.

3.7. Data Analysis

Section 3.7 describes the three different types of analyses carried out in the present study: keyword analysis, Multi-Dimensional (MD) analysis, and key feature analysis. These three analyses are commonly used in CADS, register studies, or both, as noted in Section 2.2. Keyword analysis (Section 3.7.1) was used to investigate lexical features for the description of both the UNGA as a whole (the first research goal) and patterns of variation within the UNGA (the second research goal). MD analysis (Section 3.7.2) was used to examine grammatical features, again for both commonalities across the UNGA and differences within the UNGA. The objective of using MD analysis in this study was to identify co-occurrence patterns of grammatical features rather than look at every feature individually. Key feature analysis (Section 3.7.3) is a much more detailed investigation of each individual grammatical feature. Key feature analysis was not deemed necessary for the first research goal. Describing the UNGA as a whole when compared to other spoken and written registers was expected to produce some very broad, basic distinctions in text type and thus in combinations of grammatical features, as such distinctions have been found in other MD analyses comparing distinct registers or sub-registers (e.g., Biber, 2006; Friginal, 2009). In contrast, more minute differences in individual grammatical features were expected for the second research goal looking at variation within the UNGA, again based on the literature (e.g., Quaglio, 2009; Staples, 2015a). To detect this variation of individual features when comparing groups of countries in the UNGA, I used the more detailed key feature analysis in addition to MD analysis.

3.7.1. Keyword analysis

Keyword analysis compares the words in two texts or groups of texts to determine which terms are statistically more frequent in a target text or corpus (the object of study) when

compared to a reference text or corpus. The objective is to determine what a text or corpus is primarily about, so the reference corpus should be chosen with a view to eliciting distinctive characteristics from the target corpus.

As described in Section 2.2.2, there are two types of keyword analysis. With the first type, the target corpus is compared to a large corpus such as the BNC. This identifies the lexical features that are distinctive in the target corpus as a whole. This is therefore the type of keyword analysis that I carried out to address the first research goal of describing the UNGA overall. With the second type of keyword analysis, the target corpus is divided into different parts and each part is compared to the rest of the corpus in order to find variation within the collection of texts. This type of keyword analysis was carried out to explore the second research goal. For example, in the keyword analysis for Africa, all the texts from Africa made up the target corpus; the remaining texts (the texts from the Americas, Asia, Europe, and Oceania) were the reference corpus. In the keyword analysis for the Americas, all the texts from the Americas became the new target corpus, and the texts from Africa, Asia, Europe, and Oceania became the reference corpus.

After keywords are identified, they must be interpreted. As described in Section 2.2.2, CADS typically does not discuss all keywords identified. Instead, the researcher selects one keyword (or occasionally several related keywords) and interprets its use based on political, social, or cultural identities and relationships. As indicated in Section 2.2.3, keyword analysis is also carried out in some register studies. Typically, however, in register studies all keywords or a designated number such as the top 100 (e.g., Biber & Egbert, in press) are reported and analyzed. Interpretation is based on the situational characteristics of the register represented by the corpus.

74

In this study, I interpreted up to the top 100 keywords based on the situational characteristics of the UNGA, which are described in Section 4.3.

The steps involved in keyword analysis can be summarized as follows, with optional steps indicated in italics:

Step 1: Create a word list for the target corpus

Step 2: Create a word list for the reference corpus

Step 3: Identify keywords (up to top 100)

Step 3a. Identify all keywords by running a computerized program comparing the two word lists from Steps 1 and 2

Step 3b. Identify meaningful keywords by eliminating all keywords that do not meet a priori distribution criteria

Step 4: Put all keywords (up to top 100) into semantic categories

Step 5: Determine meaning and use of keywords

Step 5a: Examine concordance lines where keywords were identified

Step 5b: Examine broader co-text surrounding keywords

Step 5c: Identify collocates of keywords by running a computerized program

Step 6: Interpret keywords by considering the relationship between keywords and situational characteristics

As illustration of how these steps are implemented, I will take an example from Chapter 5 where I analyzed the keywords for European countries. My objective in this analysis of European keywords was to identify any distinct lexical forms for European texts when compared to the rest of the UNGA, as part of my second research goal looking at variation within the

register. My target corpus in this example was all 26 European texts. My reference corpus was the remaining 66 texts from Africa, the Americas, Asia, and Oceania.

For Steps 1 and 2, create a word list for the target and reference corpora, I used the WordList function in WordSmith Tools 7.0 (Scott, 2017), which simply requires uploading the texts into the program in text format. For Step 3a, identify all keywords, I used the KeyWords function, again in WordSmith Tools 7.0, uploading the two word lists I had created in the first steps. This resulted in the keyword list shown in Table 3.7.

Table 3.7. Keywords in European Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
Ukraine	57	0.12	17	10		103.07	4.02	0.00
European	59	0.12	17	23	0.02	75.44	2.87	0.00
Ireland	30	0.06	1	1		72.57	6.42	0.00
Europe	68	0.14	22	39	0.03	66.27	2.31	0.00
Romania	19	0.04	1	0		51.17	137.72	0.00
Croatia	19	0.04	1	0		51.17	137.72	0.00
Iceland	17	0.04	1	0		45.79	137.56	0.00
Moldova	17	0.04	4	0		45.79	137.56	0.00
Netherlands	16	0.03	1	0		43.09	137.47	0.00
crimes	45	0.10	15	27	0.02	42.20	2.25	0.00
Syria	62	0.13	20	54	0.04	39.26	1.71	0.00
Crimea	17	0.04	7	1		38.66	5.60	0.00
Russian	20	0.04	4	5		31.86	3.61	0.00
Malta	14	0.03	1	1		30.96	5.32	0.00
Greece	14	0.03	2	1		30.96	5.32	0.00
responsibility	56	0.12	18	55	0.04	30.09	1.53	0.00
situation	36	0.08	18	27	0.02	27.18	1.92	0.00
Marino	10	0.02	1	0		26.93	136.79	0.00
Latvia	10	0.02	1	0		26.93	136.79	0.00
refugees	46	0.10	17	43	0.03	26.52	1.61	0.00
Ukrainian	12	0.03	5	1		25.87	5.09	0.00
migration	41	0.09	18	36	0.03	25.69	1.70	0.00
Donbas	9	0.02	1	0		24.24	136.64	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, all texts except Europe: Africa, Americas, Asia, Oceania).

The keyword list includes the keyword, the number of times the keyword appears in the target corpus ("frequency"), the percentage of the text that this frequency represents ("%"), the number of target texts it appears in ("texts"), the number of times the keyword appears in the reference corpus ("R.C. frequency"), and the percentage of the text that this frequency represents ("R.C. %"). The keyword list also provides the results of two statistical analyses, log likelihood and log ratio. The log likelihood test measures the statistical significance of the keyness, while the log ratio indicates the size of the keyness. Finally the *p* value is given, a standard indicator of statistical significance. For a term to appear on the keyword list, it must pass both the log likelihood and the log ratio tests. (For more information on statistical tests used in keyword analysis, see Scott, n.d.).

Step 3b, identify meaningful keywords by eliminating all keywords that do not meet distribution criteria, is an optional step. Whether or not a distribution cut-off is set (e.g., only keywords appearing in 25% of the texts will be analyzed) has important implications for the interpretation of results. With a small data set, if there is no distribution cut-off, a risk is run of identifying a keyword that is not truly key for the corpus but merely for one text or group of texts (see Baker, 2004, for a discussion). For instance, as shown in Table 3.7, *Ireland* appears 30 times in the European target corpus, but all 30 occurrences are in one text (the Irish text). Thus, its keyness is not reflective of the *aboutness* of European texts but only the *aboutness* of the Irish text. The fewer the texts in a corpus, the more likely it is that a keyword will be identified even though it is not representative of the entire corpus but rather just one or two texts in the corpus.

For my second research goal of examining differences within the UNGA based on country categories, I determined that it was crucial to have a distribution cut-off. Some of my target corpora for country categories were very small, in the most extreme case as few as four

texts (high GDP, see Table 3.5). Moreover, my country categories were exploratory. No previous studies had identified systematic patterns of linguistic variation in international organization texts based on geographic, political, economic, and social factors. Above all, I did not want to conclude that a particular group of countries used a particular word or words with unusual frequency based on texts from just two or three countries. For my analyses by country category, I set my distribution cut-off to 50%. That is, at least one keyword had to appear in 50% of the texts for me to continue keyword analysis with that country group.

An additional criterion for keyword analyses of regional country groups was that the keyword could not be a place name in the region under investigation (e.g., *Caribbean* for American texts, *Africa* for African texts). Nearly all speakers referred multiple times to their country or to places in their region. Thus, a regional place name did not help distinguish unusual lexical patterns for any country group, which was the objective of the keyword analysis.

If no keywords (outside of regional place names) appeared in 50% or more of the texts, I deemed that country category lexically unproductive, without clear patterns in lexical features, and did not proceed with Steps 4-6. Returning to the European texts example and Table 3.7, nine of the 23 keywords appeared in at least 13 out of the total 26 European texts (50%). Of these, three were regional place names (*Ukraine, European, Europe*) and thus not indicative of any unusual lexical pattern for European texts. This left six keywords that were identified as "meaningful" because they appeared in at least 50% of the texts and were not place names. Only one keyword meeting these criteria was required; I therefore considered European texts to have clear patterns in lexical features and continued with the keyword analysis.

Though I set a distribution cut-off for analyses of the smaller sub-corpora for country categories, I did not need to do so for my keyword analysis of the UNGA as a whole. With a

78

large data set and with a target corpus that is relatively distinct from the reference corpus, a distribution cut-off is not always critical. A large number of keywords can be expected because of the size and distinctness of the target corpus. Given that keywords represent the "aboutness" of a collection of texts, many of the keywords are semantically related: they are about the same thing or things. Patterns can be discerned based on these semantically related terms. Every keyword will not appear in every text but a sufficient number of semantically related words should appear across a sufficient number of texts. While I did not set a distribution cut-off, I did examine the number of texts in which my keywords for the UNGA appeared. I noted that in the top 100 keywords, 75% of them appeared in at least half the texts (46 texts total). Only seven appeared in less than 25% of the texts (23 texts). Moreover, all the keywords were related to a few distinct categories, so even if one word appeared in only 20 texts, it nonetheless represented a semantic concept that was prevalent in the UNGA overall. Given my distribution criteria, the keyword analysis for the UNGA would have easily "passed": The UNGA had over 75 keywords (not just the requisite one keyword) each occurring in at least half the texts. Comparing the number of keywords in the UNGA as a whole to the number of keywords in the European texts highlights the difference in the expected results of keyword analysis when the target corpus is large and distinct from the reference corpus (e.g., the UNGA compared to the BNC) and when the target corpus is small and similar to the reference corpus (e.g., European UNGA texts compared to other UNGA texts). Though top keywords can be *expected* to be distributed across a majority of texts when the target corpus is large and distinct, this cannot be *assumed*. Reporting keyword frequencies and distribution ensures greater accountability. In this study, a keyword list like the one shown in Table 3.7 was included for the top 100 keywords for the UNGA (see Table 4.3).

Step 4 is to identify two or more keywords that relate to the same topic or semantic domain and put them together in one category. (As a reminder, this step was carried out only for productive country categories with at least one keyword meeting the 50% distribution criteria.) For this step, all keywords for the country group were used; distribution criteria no longer needed to be as stringent for each individual word. If seven different words, each from seven different countries, all relate to violence, the semantic domain is viable. This is particularly important when there are synonyms (e.g., *aggression, attack, violence*) or words in the same family (e.g., *violate, violated, violation*) on the keyword list. Each individual term might not meet the distribution cut-off, but when analyzed as a semantically related group, the words reveal a distinct pattern. Therefore, when classifying keywords into semantic categories, I analyzed all keywords regardless of distribution. In the interest of transparency, however, I italicized all keywords that appeared in 50% or more of the texts to indicate words that met the distribution criteria (see Table 3.8).

Table 3.8. Keyword Categories for European Texts

Places in Europe	Places outside of Europe	Refugees	Abstract nouns
Ukraine	*Syria*	*refugees*	*crimes*
European		*migration*	*responsibility*
Europe			*situation*
Ireland			
Romania			
Greece			
Croatia			
Iceland			
Netherlands			
Romania			
Latvia			
Iceland			
Malta			
Marino			
Ukrainian			
Crimea			
Russian			
Donbas			
Moldova			

Note. Keywords in at least 50% of texts are italicized.

Classifying keywords helps lay the groundwork for the interpretation of keyword meaning and use. For instance, rather than interpret *refugees* and *migration* as distinct concepts, the analyst recognizes their similar meaning. The use of *refugees* and *migration* will be different, but topically, it is logical to consider *refugees* and *migration* together in the analysis. Step 4 categorizes such semantically transparent keywords.

However, the use and sometimes even meaning of many words is not always clear. *Responsibility* and *situation* in particular could be used in many different contexts to mean different things (e.g., *in a comfortable situation, in the Situation Room*). To clarify a keyword (Step 5), the analyst can examine concordance lines where the keyword has been identified (Step

5a), examine the broader co-text surrounding the keywords (Step 5b), and identify collocates of keywords by running a computerized program (Step 5c). Not all of these steps are necessary, but they oftentimes prove to be useful.

Table 3.9 shows the interest of looking at concordance lines (Step 5a) for understanding the use of a keyword. All of the 10 occurrences of *Latvia* are in the Latvian text, and nine of them indicate the position of the country relative to the work of the UN or the international community. When a country name was used in no other text but from the country itself, I called this use of country name "self-reference." This occurred very frequently in the UNGA. In the European texts alone, self-reference occurred in eight texts: in addition to Latvia, with Ireland, Romania, Croatia, Iceland, the Netherlands, Malta, and San Marino. Because self-reference occurred in almost every country category, it quickly became apparent that self-reference was not a distinguishing lexical characteristic for any country group. It was common across nearly all UNGA texts.

Table 3.9. Concordance Lines for *Latvia* in European Texts

Concordance Line	Text
Republic, Mali and other places. **Latvia** provides a voluntary contribution to UN	Latvia
capabilities of prevention of conflicts. **Latvia** welcomes the ongoing UN review of	Latvia
greatly strengthened in the past 10 years. **Latvia** strongly supports the work of the	Latvia
that this review process, where **Latvia** plays a leading role, is successful and	Latvia
costly or even unfeasible. **Latvia** supports the reduction of the greenhouse gas	Latvia
negotiations on this important issue. **Latvia** supports expansion of the Security	Latvia
Raimonds Vÿjonis, President of **Latvia**, at the 70th session of the United Nations	Latvia
UN in situations of concern is crucial. **Latvia** supports the proposal to voluntarily	Latvia
Iraq and the wider region. **Latvia** welcomes the recent nuclear agreement with Iran.	Latvia
rest of the international community, **Latvia** will continue to support Ukraine's	Latvia

Note. Concordance lines have been abridged slightly to the left and right in the interest of space.

Concordance lines do not always give a clear picture of how or why a term is being used, however. When concordance lines are not clear, examining the broader co-text around the keyword, expanding the concordance line (Step 5b), is necessary. For instance, the co-text in which *Ukraine* appears shows a concern among European countries about the Russian occupation of Ukraine. This is not obvious in all the concordance lines themselves (see Table 3.10), but expanding the concordance lines reveals the association between Ukraine and the occupation (see Text 3.4 for an example). The concordance line has been underlined in Text 3.4.

Table 3.10. First 10 Concordance Lines for *Ukraine* in European Texts

Concordance Line	Text
solution that provides for a free and stable **Ukraine**, which enjoys strong ties with	Austria
attempts for further destabilization of **Ukraine**. The full implementation of the	Bulgaria
integrity, for us, Crimea is Ukraine and **Ukraine** is Europe. My country condemns	Bulgaria
territorial integrity. For us, Crimea is **Ukraine** and Ukraine is Europe. My country	Bulgaria
as well as to assisting and supporting **Ukraine** on its reformist path towards its	Croatia
On the European continent, in the **Ukraine**, the implementation and the viability	Croatia
more complex crises and threats. In **Ukraine**, we have seen completely	Denmark
on our continent by force, like in **Ukraine**, in violation of the Charter of the	EU
put an end to the aggression against **Ukraine**. Compared to a year ago, the armed	Estonia
disappointing that in cases of Syria and **Ukraine** the Council debates have brought	Estonia

Note. Concordance lines have been abridged slightly to the left and right in the interest of space.

Text 3.4. Expanded Concordance Line from Austria

> Ladies and Gentlemen, There is another crisis area we shouldn't forget about. Last year, like most of you I referred to the Russian-Ukrainian conflict as probably the most serious challenge to peace and security in Europe. One year later, we still cannot see a positive development. The conflict is ongoing, the number of deaths has even increased and the socio-economic consequences are huge. There is only one way ahead: A peaceful settlement based on dialogue and negotiation. We must find a <u>solution that provides for a free and stable Ukraine, which enjoys strong ties with</u> both the European Union and the Russian Federation.

For some particularly frequent and ambiguous keywords, a collocational analysis was useful (Step 5c) because it revealed patterns that I did not necessarily see in my qualitative examination of the concordance lines. Collocational analysis identifies the words that co-occur within five words on either side of a search term, thus eliciting patterns in how a word is used based on the other terms immediately preceding or following it.

To take an example from the European texts, the keyword *situation* (see Table 3.7) is ambiguous. A text is not *about* the same thing if the speaker is describing a *work, financial, political,* or *legal situation,* and the semantic prosody (the positive or negative evaluation) differs depending on whether the *situation* is deemed to be *absurd, ideal, alarming,* or *hypothetical.* In order to clarify the use of *situation* in the European texts, I carried out collocational analysis using the *Concord* function in WordSmith Tools 7.0. The tab *collocates* at the bottom of the concordance line screen issues a list of collocates as shown in Table 3.11, with the number of texts the collocation occurs in, the total number of times the collocation occurs, and where the collocate occurs, to the left or to the right of the search term. Table 3.11 is abbreviated, showing up to the center placement, with the five places to the right of the search term deleted in the interest of space.

Table 3.11. Abbreviated Table of Collocates for *Situation* in European Texts

Word set	Texts	Total	Total left	Total right	L5	L4	L3	L2	L1	Centre
the	15	38	29	9	2	3	8	2	14	0
situation	18	36	0	0	0	0	0	0	0	36
in	13	21	5	16	0	1	1	3	0	0
to	7	11	6	5	1	1	2	2	0	0
a	7	10	7	3	1	0	0	1	5	0
of	6	9	5	4	0	2	0	3	0	0
is	7	8	3	5	1	1	1	0	0	0
and	6	7	3	4	0	0	1	2	0	0
security	4	4	3	1	0	0	0	0	3	0
Ukraine	3	4	0	4	0	0	0	0	0	0
Middle	3	3	0	3						
refugee	3	3	3	0					3	
which	3	3	0	3						
with	2	3	1	2			1			
this	3	3	3	0				2	1	
where	3	3	0	3						
East	3	3	0	3						
by	2	2	0	2						
unfortunately	2	2	2	0	1			1		
complex	2	2	1	1				1		
especially	2	2	0	2						
conflict	2	2	0	2						
while	2	2	2	0			1	1		
an	2	2	0	2						
years	1	2	0	2						
we	2	2	1	1	1					
been	2	2	2	0		1	1			
as	1	2	2	0		1		1		
have	2	2	1	1	1					
must	2	2	1	1		1				
has	2	2	1	1		1				
I	2	2	0	2						
humanitarian	2	2	2	0					2	
human	2	2	2	0				2		
region	2	2	0	2						
rights	2	2	2	0					2	
for	2	2	0	2						
or	1	2	1	1	1					
present	2	2	2	0				2		
global	2	2	1	1				1		

Table 3.11 illustrates the need to filter information in the collocate list. Top collocates are very often function words simply due to the high frequency of function words generally. It is not surprising and not particularly meaningful to note that *the situation* occurs 14 times. An analysis of function word collocates is useful for answering questions concerning the grammatical behavior of a word. However, if the objective of the analysis is to find common semantic associations with the search term, these grammatical words are often ignored and only content words are analyzed. Only three content words are in the top 10 collocates of *situation*: the search term itself (*situation*) along with *security* and *Ukraine.* However, in the next 30 collocates, certain patterns emerge. *Situation* collocates with geographic areas (*Ukraine, Middle East*), with terms related to conflict (*security, conflict*), and with a few commonly cited problems addressed by the UN (*refugee, humanitarian, human rights*). Further analysis of *situation* using concordance lines and co-text (as described in Steps 5a and 5b), showed that all these terms were, in fact, related to conflict: Ukraine and the Middle East were designated as areas of conflict; *refugees*, *humanitarian,* and *human rights* referred to problems resulting from conflict (e.g., *the degrading humanitarian situation in several conflict zones* in the text from Bulgaria). Identifying the collocates of keywords such as *situation* and *responsibility* made it easier to find semantic patterns and thus to interpret the use of these words.

The final step in the keyword analysis is to interpret the keywords based on situational characteristics. For the interpretation, rather than providing an in-depth analysis of every keyword identified, I interpreted groups of keywords that appeared to be in the same semantic category, serving the same function. Some of these semantic categories had been identified in Step 4 (e.g., place names), but generally interpretation required more thorough analysis of situational characteristics. For example, in 2015 when the UNGA addresses were delivered,

Europe was facing a political, social, and economic crisis due to the large influx of refugees from Syria. The concern expressed about *refugees, migration,* and *Syria* could be explained, at least in part, by this situational characteristic. For more on the interpretation of the results from European texts, see Section 5.3.3.

Section 3.7.1 has described the steps of the keyword analysis and illustrated those steps with an example from Chapter 5 using the European texts. The results of all the keyword analyses that produced meaningful keywords can be found in Chapters 4 and 5. Chapter 4 investigates the first research question, so Section 4.4 describes the results of the keyword analysis comparing the UNGA to the BNC, to identify the distinctive lexical characteristics of UNGA addresses. Chapter 5 investigates the second research question, so Section 5.3 describes the results of the keyword analysis for all country categories that produced "meaningful" keywords (i.e., country categories with at least one keyword that appeared in 50% or more of the texts).

3.7.2. Multi-Dimensional analysis

Multi-Dimensional (MD) analysis was used to identify co-occurrence patterns of grammatical features for both the UNGA overall (research goal 1) and for country groups (research goal 2). For my description of grammatical features in the UNGA as a whole, I analyzed MD scores for the first five dimensions found in Biber (1988), comparing MD scores for the UNGA to MD scores for four comparison registers: prepared speeches, spontaneous speeches, conversation, and official documents. For my description of grammatical features by country group, I compared the MD scores for each country group (e.g., low GDP) to the remaining texts (medium, high, and very high GDP). Close MD scores between the target corpus (low GDP) and the reference corpus (remaining texts) indicated similarities in co-occurring

grammatical features. A sizeable difference in scores revealed a systematic pattern of variation

for the country group in terms of lexico-grammatical forms.

MD analysis was first introduced to explore patterns of linguistic variation in spoken and

written texts, examining such diverse registers as romantic fiction, mathematics academic prose,

and sports broadcasts (Biber, 1988). Factor analysis was used to identify *factors* (co-occurring

linguistic features), which were then interpreted functionally as underlying dimensions of

variation in texts. Table 3.12 presents a summary of the factorial structure for the five principal

factors in Biber (1988). That is, it shows the co-occurring linguistic features that were identified

for each of the five factors and their factor loadings. The larger the factor loading, the more

representative a feature is for that factor. Positively loaded features co-occur with other

positively loaded features but tend not to co-occur with negatively loaded features.

Table 3.12. Factorial Structure for the Five Principal Factors (from Biber, 1988, pp. 89-90)

FACTOR 1

private verbs	.96
that deletion	.91
contractions	.90
present tense verbs	.86
2nd person pronouns	.86
do as pro-verb	.82
analytic negation	.78
demonstrative pronouns	.76
general emphatics	.74
1st person pronouns	.74
pronoun *it*	.71
be as main verb	.71
causative subordination	.66
discourse particles	.66
indefinite pronouns	.62
general hedges	.58
amplifiers	.56
sentence relatives	.55
WH questions	.52
possibility modals	.50
non-phrasal coordination	.48
WH clauses	.47
final prepositions	.43
(adverbs	.42)
(conditional subordination	.32)
nouns	-.80
word length	-.58
prepositions	-.54
type-token ratio	-.54
attributive adjectives	-.47
(place adverbials	-.42)
(agentless passives	-.39)
(past participial WHIZ deletions	-.38)
present participial WHIZ deletions	-.32)

FACTOR 2

past tense verbs	.90
3rd person pronouns	.73
perfect aspect verbs	.48
public verbs	.43
synthetic negation	.40
present participial clauses	.39
(present tense verbs	-.47)
(attributive adjectives	-.41)
(past participial WHIZ deletions	-.34)
(word length	-.31)

FACTOR 3

WH relative clauses on object positions	.63
pied piping constructions	.61
WH relative clauses on subject positions	.45
phrasal coordination	.36
nominalizations	.36
time adverbs	-.60
place adverbs	-.49
adverbs	-.46

FACTOR 4

infinitives	.76
prediction modals	.54
suasive verbs	.49
conditional subordination	.47
necessity modals	.46
split auxiliaries	.44
(possibility modals	.37)

FACTOR 5

conjuncts	.48
agentless passives	.43
past participial clauses	.42
by passives	.41
past participial WHIZ deletions	.40
other adverbial subordinators	.39
(predicative adjectives	.31)
(type-token ratio	-.31

Note. Features in parentheses were not used in the computation of factor scores.

These five factors were then interpreted functionally as dimensions of textual variation: Dimension 1: Involved versus Informational Production distinguishes very involved registers such as face-to-face and telephone conversation from very informational registers such as academic prose and official documents (with a number of other registers such as fiction and broadcasts falling in between). Dimension 2: Narrative versus Non-Narrative Concerns distinguishes highly narrative texts (e.g., fiction) from non-narrative texts (e.g., official documents). Dimension 3: Explicit versus Situation-Dependent Reference distinguishes texts with explicit and elaborated endophoric reference (e.g., official documents, academic prose) from texts with situation-dependent reference (e.g., sports broadcasts). Dimension 4: Overt Expression of Persuasion distinguishes texts with high overt persuasion (e.g., editorials) from texts with low overt persuasion (e.g., broadcasts). Finally, Dimension 5: Abstract versus Non-Abstract Information distinguishes abstract texts (e.g., academic prose) from non-abstract texts (e.g., face-to-face conversation).

As described in Section 2.2.3.2, numerous studies have explored new registers by analyzing dimension scores for these first five dimensions found in Biber (1988). In such studies, the dimension scores of a new target register (e.g., the television sitcom *Friends* in Quaglio, 2009) are compared to the dimension scores from previously explored registers (e.g., face-to-face conversation in Biber, 1988). The steps involved for comparing dimension scores are outlined below. Steps in italics are optional.

Step 1: Convert files to text format

Step 2: Annotate files for part of speech (POS) using a tagger such as the Biber Tagger (see Biber, 1988; Biber et al., 1999)

Step 3: Run the TagCount program, also developed by Biber, which counts the lexico-grammatical features (the "tags" from Step 2) and calculates dimension scores for each individual text

Step 4a: Calculate the mean (M) and standard deviation (SD) for the target corpus from the scores of individual texts produced in Step 3

Step 4b: Calculate the M and SD for the reference corpus

Step 5: Compare M and SD for the target corpus and reference corpus and report findings

Step 6: Analyze findings based on situational characteristics

Step 1 was described in Section 3.4 and Steps 2 and 3 in Section 3.5. For Step 4a, I calculated the M and SD using Excel. The M and SD for each target feature and each dimension score in the UNGA is reported in Appendix A. I did not need to carry out Step 4b for the first research goal, comparing the entire UNGA to the four comparison registers because the M and SD of the comparison registers can be found in Biber (1988). For the second research goal, I divided the UNGA into country categories and calculated the M and SD for each country category and its reference corpus using Excel.

For Step 5, comparing the M and SD for the target corpus to several other registers shows which other registers the target register most resembles. For instance, Quaglio (2009) shows the Dimension 1 score for *Friends* in relation to telephone conversation, face-to-face conversation, personal letters, public conversations, prepared speeches, fiction, editorials, prose, and official documents (p. 66). The *Friends* score is closest to face-to-face conversation, showing that the texts in *Friends* and in face-to-face conversation are extremely similar in terms of involved language (the characteristic of high Dimension 1, as discussed above). I, too, reported M and SD for the UNGA and the four comparison corpora as part of my first research question. The scores

91

were distinct enough in the four comparison corpora to show clear differences and occasionally similarities. With only four comparisons for each dimension (the UNGA with each of the four reference corpora), it was possible to analyze and discuss all scores.

In contrast, for the second research question, there were 10 country categories and a total of 31 levels, meaning 31 target corpora and 31 reference corpora to compare. Because these texts were all UNGA addresses, many of the country categories shared similar dimension scores (e.g., for Dimension 2, the average for Africa was – 3.05 and the average for the remaining texts was – 3.08). It was prohibitive and not particularly meaningful to analyze all of these dimension scores. Therefore, I reported the M and SD of scores in Appendices B-K, but discussed only the scores that showed large differences compared to the other texts because my second research goal was to find systematic patterns of variation (rather than similarities).

Finally, Step 6 is to interpret findings based on situational characteristics. I considered the meaning of each dimension and what functions it served. To provide an example with simplified interpretation (but for a full detailed analysis, see Section 4.5.1), Dimension 1 represents involved versus informational production; one function of involved language is to focus on interpersonal or affective content and one function of informational language is to inform. The importance of these functions for each register can be explained in terms of the situational characteristics of the registers. For example, face-to-face conversation shows higher Dimension 1 scores, indicating more involved language because typically in face-to-face conversation there is a focus on revealing and inquiring about personal feelings, attitudes, and desires. In contrast, official documents have low Dimension 1 scores, more informational production. The purpose of an official document is to convey specific information from an impersonal source rather than reveal feelings and attitudes. All dimension scores were

interpreted for research goal 1 because scores were sufficiently distinct. However, as indicated above in the discussion for Step 5, for the second research goal examining variation within the UNGA, only dimension scores that were sufficiently distinct were interpreted.

3.7.3. Key feature analysis

In addition to MD analysis, key feature analysis was used to investigate grammatical features when comparing groups of countries for the second research question. Key feature analysis is a means of detecting "key" grammatical features, meaning features that are used with a statistically higher or lower frequency in the target corpus than in the reference corpus. Typically, the notion of "keyness" is associated with keywords, that is, with word types (e.g., *situation*, see Section 3.7.1). However, conceptually, there is no reason that grammatical features (e.g., abstract nouns) cannot be considered "key." Biber and Egbert (in press) propose a previously unexplored means of detecting "keyness" for lexico-grammatical features. They point out that grammatical features are generally much more frequent than a single word type and therefore their relative frequency cannot be calculated in the same way. Instead, Biber and Egbert measure keyness for grammatical features using the formula for Cohen's *d*:

$$d = \frac{M_1 - M_2}{SD_{pooled}}$$

where:

$$SD_{pooled} = \sqrt{\frac{SD_1{}^2 + SD_2{}^2}{2}}$$

Typically, Cohen's *d* calculates an effect size, but logically it represents the difference between two means expressed in standard deviation units to ensure comparability across scores. That is precisely what is being analyzed with the keyness of feature categories. The mean

frequency of one linguistic feature in the target corpus (M_1) is being compared to the mean

frequency of that feature in the reference corpus (M_2), as shown in the example below (see "Step

5b"). A feature is considered "key" if Cohen's *d* reveals a large effect size for that feature.

The steps for carrying out key feature analysis can be summarized as follows:

Step 1: Convert files to text format

Step 2: Annotate files for part of speech (POS) using a tagger such as the Biber Tagger (see Biber, 1988; Biber et al., 1999)

Step 3: Run the TagCount program, also developed by Biber, which counts the lexico-grammatical features (the "tags" from Step 2) and calculates dimension scores for each individual text

Step 4a: Calculate the mean (M) and standard deviation (SD) for 126 lexico-grammatical features for each target corpus

Step 4b: Calculate the M and SD for each grammatical feature for each reference corpus

Step 5: Compare M and SD for each target corpus and its respective reference corpus and report findings

Step 5a: Report M and SD

Step 5b: Calculate Cohen's *d* and report large effect size (*d* > ±.80)

Step 6: Analyze findings based on situational characteristics

Step 1 was described in Section 3.4 and Steps 2 and 3 in Section 3.5. Calculating M and

SD (Step 4) was described in Section 3.7.2. For Step 5, I reported all M and SD for all country

groups in Appendices B-K. As with MD scores (described in Section 3.7.2), I did not analyze all

126 grammatical features for 31 categories, which would have resulted in 3,906 analyses. Like

with MD scores, this would not have been productive because many country categories had

similar frequencies for these grammatical features (e.g., the smallest area countries had M =

124.83 prepositions per 1,000 words and the reference corpus of medium, large, and very large

area countries had M = 124.90 prepositions per 1,000 words). Instead, I calculated Cohen's *d*

using the formula above and analyzed only the features with a large effect size ($d > \pm.80$).

      To illustrate how this was done, I will take the example of second person pronouns for

African texts from Chapter 5. When comparing the African group to its reference corpus (all

UNGA texts except Africa), second person pronouns appeared on average 1.9 times more

frequently, with $M_1 = 4.7$ and $M_2 = 2.5$ (see Table 3.13).

Table 3.13. Second person pronoun frequency for Africa

Feature	Africa		Reference corpus	
	Mean	**SD**	**Mean**	**SD**
2nd person pronoun	4.71	2.67	2.51	2.17

Note. SD = standard deviation. Reference corpus = all UNGA texts except Africa. Frequencies
normed per 1,000 words.

      The pooled standard deviation (SD) for second person pronouns was calculated using the

formula above, where $\sqrt{(2.67^2 + 2.17^2)/2} = 2.43$. Cohen's *d* was then computed: (4.71 −

2.51) / 2.43 = 0.91. Interpretation of Cohen's *d* values differs, but generally $d > \pm.80$ is

considered to be large, $d > \pm.50$ medium, and $d > \pm.20$ small (Cohen, 1977). Thus, there was a

large effect size for second person pronouns in African texts. Like for M and SD, I entered the

formulas for pooled SD and for Cohen's *d* into Excel for calculation.

      For the features that had a large effect size, I offered a functional interpretation to explain

the difference between the target country group and the remainder of the corpus (Step 6). As

with the keyword analysis (Section 3.7.1) and the MD analysis (Section 3.7.2), I used the situational characteristics of the country groups in my analysis.

Key feature analysis was extremely useful for my bottom-up approach to examining patterns of variation within the UNGA. Unlike many other studies where analysts know a priori which grammatical features they are going to investigate (e.g., involved features in Quaglio, 2009; stance features in Staples, 2015a), the present study is much more exploratory. Few linguistic analyses have been carried out for international organization spoken discourse or scripted political speeches more generally, so I wanted to examine a wide range of grammatical features and let the data tell me which features deserved further investigation. I calculated the M and SD for 126 features and it would have been prohibitive to analyze all 126 features for 31 country groups. Cohen's d provided a systematic way of determining whether there was a large effect size in the difference in frequencies for any given feature.

It was particularly important to examine individual grammatical features for the investigation of country groups within the UNGA because the texts within the UNGA are relatively similar. Situational characteristics vary in only minute ways. I therefore expected to see only minute differences in grammatical features as well. Though MD analysis would capture some patterns of variation in co-occurring features, it might not be enough to see the finer distinctions across country groups. In contrast, key feature analysis was not as critical for the analysis of the UNGA as a whole because, when compared to the other four registers, I expected much broader lexico-grammatical differences that would be detected in the MD analysis. As a result, I carried out key feature analysis only for the second research goal, to identify patterns of variation within the UNGA.

3.7.4. Summary of analyses

    Section 3.7 has described the three different analyses carried out in the present study. Each analysis served a different purpose and produced different types of results. Keyword analysis identified the lexical features that were distinctive of the UNGA as a whole as well as the lexical features distinguishing country groups within the UNGA. MD analysis identified variation in co-occurrence patterns of grammatical features, both for the entire UNGA and for groups of countries within the UNGA. Finally, key feature analysis allowed for a more in-depth account of variation in individual grammatical features for groups of countries. The results of these analyses can be found in Chapter 4 (the description of the UNGA as a whole) and in Chapter 5 (the description of variation within the UNGA).

CHAPTER 4: GENERAL REGISTER DESCRIPTION OF UNITED NATIONS GENERAL ASSEMBLY ADDRESSES

4.1. Introduction

The purpose of this chapter is to describe UNGA addresses as a register in contrast to other spoken and written registers. Section 4.2 presents an overview of UNGA texts with two addresses to serve as illustration. Section 4.3 examines the situational characteristics of the UNGA and compares them to the situational characteristics of four other registers: prepared speeches, spontaneous speeches, official documents, and conversation. Understanding the situational characteristics of a register is essential for the functional interpretation of linguistic features because the communicative context helps to explain why certain linguistic features are particularly frequent or infrequent. Section 4.4 presents the lexical features of the UNGA through a comparison to the British National Corpus (BNC), a 100-million-word reference corpus. In Section 4.5, Multi-Dimensional scores for the UNGA are compared to scores from the four other registers (prepared speeches, spontaneous speeches, official documents, and conversation). The grammatical features that typify each dimension are analyzed and interpreted. These results are synthesized in Section 4.6 and a functional explanation of the situational context of the UNGA helps to elucidate the use of linguistic features characterizing this register.

4.2. Overview and Examples of the UNGA Address

The General Assembly is the primary policymaking body of the UN. Each annual session opens with a General Debate in which high-ranking diplomats such as presidents and foreign ministers deliver a formal address. The General Debate is a forum for outlining national opinion and action, recommending future steps at the international level, and building and maintaining sound diplomatic relationships with fellow member states (Donahue & Prosser, 1997). These

addresses are thus of paramount importance with far-reaching consequences (Baturo et al., 2017). The texts also tend to follow similar patterns in terms of topics raised and communicative purpose.

Section 4.2 presents two UNGA addresses with notes on topic and purpose in the margin. The objective of this section is to introduce UNGA addresses and provide a macroscopic view to better situate the more detailed analyses that follow in Chapters 4 and 5, showing overall the structure of the speech as well as the topics that are raised and the goals speakers are trying to achieve. The notes on topic and communicative purpose are intended to illustrate what is meant by "self-reference" or "request joint efforts," terms which will be used in the analyses in later sections.

These notes are qualitative and subjective. As explained in Section 3.6.1, a pilot study was carried out to determine whether it was possible to code UNGA texts for communicative purpose, but interrater reliability was very low and even after extensive discussion, agreement could not always be reached. As a result, no quantitative findings for communicative purpose are being presented in this study. Coders did agree, however, that certain communicative goals were extremely frequent overall in the UNGA addresses. These communicative goals included praising, criticizing, defending/justifying, and persuading (requesting). These same goals were kept for the present description and reference to them will be made in the functional interpretation of UNGA texts.

The first address is from San Marino, a small European country (see Text 4.1).

99

Text 4.1. Address from San Marino

Text	Topic	Purpose
Mr. President, Mr. Secretary General, Excellencies, Ladies and Gentlemen,	Greeting to GA President, UN Secretary General, Assembly	Praise UN officials and UN
I would like to congratulate you on your election as President of the 70th Session of the General Assembly.		
I would like to commend the work of H.E. Sam Kutesa, President of the 69th GA.	Work of UN elected officials	
I also wish to express sincere thanks to the Secretary General, H.E. Ban Ki-moon, for his tireless efforts in leading the United Nations in the past years.		
The result of such efforts culminated with very successful "UN Summit for the adoption of the post-2015 development agenda", and General Debate as well as the historic visit of the Holy Father, Pope Francis, to the General Assembly.	Development (Agenda 2030) Pope	
Moreover, San Marino appreciates the constant and impartial attention paid by the Secretary General to all the UN member States.		
Mr President,	Vocative marking new topic	
After 70 years, the United Nations not only has passed the test of time, but it continues to represent for all our Countries a shining beacon of hope.	Situate UN past, present, and future	Persuasion: Request joint efforts
I believe that in the future we have to be even more keen and inspired to empower and improve the role of this great Institution.		
In fact, the first and last line of defence for San Marino, a small State without any army, with limited territory and resources, is the UN.	Self-reference and national circumstances (area, military)	
Even the strongest and largest Countries cannot deal alone with the complexity and the multitude of today problems. They need the UN, too.	Importance of UN	

With the help of the UN system, we, member States, have built, over the years, the foundations for peace by preventing and ending numerous disputes; we were able to free millions of people from extreme poverty, we made important progress in advancing human rights, democracy, promoting justice and international law.	Past achievements of UN: peace and security, poverty and development, human rights and democracy	Inform Praise UN
Mr. President,		
Unfortunately, despite all our efforts and successes, the international community continues to face the same challenges: wars, poverty, inequalities, discrimination, violence against women and children and many more.	Vocative: new topic	Inform
In addition, today we face new global challenges, like climate change, economic and financial crises, unrelenting unemployment, food uncertainty, international migration and terrorism.	Challenges facing UN: peace and security, poverty and development, human rights and democracy, climate change, economic crisis, migration, terrorism	
These new threats are interdependent and together with the UN we should be able to seize and defeat them.		Persuasion: Request joint efforts
Around the world, our people look up to the UN for leadership.		
That's why the reform of the United Nations becomes fundamental and must remain at the core of our collective efforts during the 70th GA.		
We must make this Organization more effective, while preserving its integrity and enhancing its status.	UN reform	
Mr. President,	Vocative: new topic	
Every day we are witnessing the tragic migration from Africa and Asia towards Europe. These desperate people leave their Countries and their families behind to flee from conflicts, violence, and persecution. Thousands of them already died in the Mediterranean Sea.	Migration	Inform
We are pleased that a few days ago the High-level meeting on "Strengthening cooperation on migration and refugee		Praise UN

movements in the perspective of the new development agenda" was organized.		
Mr. President,	Vocative: new topic	
San Marino cannot be silent about what is happening in Iraq and in Syria, where an ongoing inhuman ethnic and religious cleansing is carried out with unprecedented ferocity.	Iraq Syria	Criticize events in other region
We hope that, under the able stewardship of Mr. Staffan De Mismra, a diplomatic solution can be achieved in the near future.		Persuasion: Request joint and personal efforts
San Marino, a strong supporter of the NPT, believes that one of the most important tasks that we have to accomplish is the total elimination of nuclear weapons, thousands of which are ready to be deployed at any time. We cannot afford to wait any longer.	Nuclear non-proliferation	
Mr. President,	Vocative: new topic	Praise UN, Request joint efforts
San Marino strongly believes in revitalizing and empowering even more the General Assembly, the forum where solutions to today's challenges can be found and consensus on issues of common interest must be reached.	UNGA	
We commend the delicate and challenging work accomplished over the years for peace and security by the Security Council. However, it is time for a reform. We believe that a reform of the Council shall include an enlargement in the category of non-permanent members and a more balanced geographical distribution. We are convinced that this reform, to be legitimated, must be achieved with the widest possible consensus.	Security Council reform	
Mr. President,	Vocative: new topic	
We welcome the post 2015 Development Agenda adopted few days ago. It represents a unique opportunity to prepare a brighter future for our citizens.	Development (Agenda 2030)	Praise UN
The 15-years goals at the core of the Agenda are ambitious: they chart our efforts in pursuit of poverty eradication, people		Inform

empowerment, protection of human dignity, shared prosperity, decent job for all, and protection of the planet. They will promote cooperation on sustainable development and guide public and private action,		Predict
But it is important that each Country, after adopting its national Agenda, will take responsibility towards its own citizens and the international community for its implementation.		Persuasion: Request national efforts
San Marino commends that the GA and the ECOSOC will assess the implementation progress. Therefore, it is critical to produce a comprehensive and effective mechanism for review and accountability.		Praise and request joint efforts
Mr. President,	Vocative: new topic	
The Republic of San Marino has always paid special attention to the most vulnerable groups, such as women, children, the elderly and the disabled.	Women: discrimination and violence	Connect speaker's country to issue / Praise speaker's country
Today, women are still the victims of discrimination and violence in many parts of the world, including in the most developed countries. Trafficking in women is far from being solved. Women living in conflict and post-conflict situations are often subject to sexual violence, torture and summary executions.		Inform
Children, the most vulnerable group, are subject to violence, abuse and exploitation as well. A huge number of children all over the world do not yet have access to basic services, health care and education.	Children: violence and human rights	
We have the duty to protect and help our children.		
Moreover, the international community must strive for the full inclusion and participation of the elderly and of people with disabilities in the economic, social and cultural life of our communities.	Elderly and people with disabilities	Request joint efforts
Mr President,	Vocative: new topic	Connect speaker's

103

In conclusion, freedom and peace have always represented the core of the Republic of San Marino in its centuries-old history. San Marino is a small Country, but strongly committed in favour of the implementation of the universal values of the United Nations and of the principles enshrined in its Charter. We are honoured to offer our contribution to the UN Community.	Freedom, peace Self-reference and national circumstances: size	country to UN Promise support Praise speaker's country

Text 4.1 from San Marino begins, like most UNGA addresses, with a greeting to the UNGA President, the UN Secretary General, and the Assembly at large. H.E. Bodini, the permanent representative of San Marino to the UN, then congratulates and thanks UN elected officials. After these introductory remarks, he discusses the importance of the UN, reflecting on past achievements as well as current challenges. He enumerates a few key UN issues: peace and security, poverty and development, human rights and democracy, climate change, economic crisis, migration, and terrorism. It is in his discussion of the key role played by the UN in tackling these problems that he calls for UN reform. He briefly touches on migration from Africa and Asia to Europe, on violence in Iraq and Syria, and on nuclear non-proliferation before returning to UN and in particular Security Council reform. H.E. Bodini speaks at greater length on development and the UN Agenda 2030 for Sustainable Development (see Section 4.3.1.6) and on the need to protect the rights of women, children, the elderly, and people with disabilities. In this last passage on these vulnerable populations, H.E. Bodini refers to his country, indicating this issue is of particular national importance. Finally, he concludes with a broad statement on San Marino's commitment to UN principles such as freedom and peace. Throughout the address, he uses the vocative *Mr. President* to mark the beginning of a new topic.

The overall purpose of the address is to highlight the UN issues that are of primary importance to San Marino; a related purpose is to persuade other UN member states to devote

their efforts to addressing these challenges. As a result, the text has both informative passages, explaining why an issue is critical, and persuasive, calling on the international community to take action. Persuasion is sometimes explicit (e.g., *the international community must strive*), but more frequently it is less direct, marked instead through the expression of opinion (e.g., *it is important*) Passages of praise and criticism are also related to this overall purpose of highlighting key UN issues and goals: The speaker praises the accomplishment of UN objectives and expresses displeasure over situations (in Iraq and Syria) that run counter to the achievement of UN objectives. Though expression of praise and appreciation is overt, criticism is veiled, with indirect disapproval in the phrase *San Marino cannot be silent about….* Finally, the speaker promises the support of his country in addressing these issues and achieving these goals.

It is important to note that the speaker uses the first person singular only at the beginning of the address in his opening congratulations and thanks and in one expression of personal stance: *I believe*. In the remainder of the speech, he speaks from the perspective of his country: *San Marino believes, San Marino commends, we welcome, we are pleased, we are convinced.* Also noteworthy is how he chooses to describe San Marino, mentioning that it has no army and referring twice to its small size, with additional emphasis on size by speaking of its *limited territory and resources*. With this description of San Marino, H.E. Bodini presents his country as dependent on the UN for maintaining peace and security.

The second address is from Viet Nam, a large Asian country with a very large population. Viet Nam has low economic development (based on GDP) and medium social/economic development (based on HDI) and is in the high military spending group.

Text 4.2. Address from Viet Nam

Text	Topic	Purpose
Mr. President, On behalf of the Vietnamese Delegation, I would like to warmly congratulate you, Mr. Lykketoft, upon your election as the President of this momentous Seventieth Session of the General Assembly. I believe that, with your wisdom and experience, you will guide our Session to a very successful outcome. My delegation applauds the important contributions made by Mr. Sam Kutesa, President of the Sixty-ninth Session of the General Assembly, and by Mr. Ban Ki-moon, Secretary-General of the United Nations, during the past year.	Greeting to GA President Work of UN elected officials	Praise UN officials
Mr. President, Born out of the ashes of the Second World War, during the last seven decades, the United Nations has grown to embrace 193 States, to be the most representative global organization, and to become the true center for the coordination of global efforts to tackle common challenges.	Vocative marking new topic Situate UN past, present, and future	Praise UN
As a beacon of hope, the United Nations has indeed helped make the world a better place by playing an increasingly important role in maintaining peace and security, in protecting and promoting human rights and in advancing development and progress. Guided by its own Charter, the United Nations has worked to uphold the principles of respect for sovereignty and territorial integrity of states, non-interference and non-threat or use of force, to provide guidance for peaceful and amicable relations among nations, underlining in particular the importance of mutual respect, good faith and peaceful settlement of disputes.	Past achievements of UN: peace and security, human rights, development, territorial integrity	Inform
It is in such spirit that we welcome the recent progress made in addressing the Iranian nuclear issue, which is a first step towards a long-term solution that could bring peace, security and stability to the larger region.	Iran nuclear deal	Praise UN

We welcome in particular the normalization of relations between the United States and Cuba, and call on the United States to lift all sanctions and embargoes currently imposed against the brotherly people of Cuba.	US-Cuba relations	Praise other UN countries Persuasion: Request national efforts
Mr. President,	Vocative marking new topic	
We can all look back with pride on the last 70 years of accomplishments of the United Nations, but must also recognize what the Organization has been unable to do and what challenges lie ahead.	Challenges facing UN: peace and security, terrorism, territorial integrity,	Praise and criticize UN
Wars and conflicts were not uncommon in the past decades, and still rage on today. Terrorism, violent extremism, religious and ethnic intolerance pose serious threats to regional and international peace and security. Territorial and maritime disputes endanger peace and stability in many regions.		Inform
Threats to the very survival of mankind remain. The nuclear weapons arsenal, though reduced, is still capable of destroying our home planet many times over.	nuclear weapons,	
Climate change is threatening the livelihoods of billions of people and the existence of many countries.	climate change,	
Poverty remains one of the greatest challenges to sustainable development. Violence, discrimination, social injustice and humanitarian crises plague the enjoyment by all of their human rights.	poverty, human rights	
Immense challenges require equally immense efforts, and the United Nations must continue to play a central role in those endeavours.		Persuasion: Request joint efforts
Viet Nam welcomes the adoption of the Agenda 2030 for Sustainable Development, setting the vision and creating the framework for all nations to strive for zero hunger and no poverty, for more sustainable patterns of production and consumption, for better preservation of our common planet against the threats of climate change and over-exploitation. The full and effective implementation of the Agenda, with a global partnership in its core, will address deeply rooted causes of our	Development: Agenda 2030	Prediction

present problems and transform this world towards a more sustainable future. To transform our world, the UN needs to transform itself. The work of the General Assembly needs to be revitalized to make it more focused, efficient and relevant. The Security Council needs to be more representative, democratic, transparent and effective. The UN development system needs to be strengthened. Institutional reforms must go along side with innovations in substance, while ensuring a balance among various aspects of the UN work.	UN reform (GA and Security Council)	Persuasion: Request joint efforts
Our experience with the implementation of the MDGs have shown that peace and stability are the prerequisite for sustainable development. In the words of the 2030 Agenda itself, and I quote: "There can be no sustainable development without peace and no peace without sustainable development" - end of quote.	Development: Agenda 2030	Prediction
It is therefore incumbent upon each and every state and the United Nations to do their utmost to ensure peace and security at national, regional and international levels. Each state needs to act responsibly in accordance with the basic principles of the UN Charter and international law. Regional and international organizations, with the UN in the lead, also need to act accordingly, and assist relevant states in finding amicable solutions to their differences and disputes by peaceful means as provided for in the UN Charter. Only by so doing can we hope for a future free from violence, threat or use of force, intimidation, coercion and inequities, thus creating an enabling environment for our common development.	Connect peace and security to development	Persuasion: Request joint / national, regional, international efforts
At the same time, in order to succeed, sustainable development must truly be of the people, by the people and for the people. In other words, the people must be at the center of this Agenda, as the owner, driver and beneficiary of all development efforts.	Development: Agenda 2030	
And no agenda can succeed without a strong global partnership for sustainable development. Viet Nam has always believed that domestic resources play the decisive role in the development of any nation. But we also believe that no individual country can do it alone. We call on developed countries to take the lead in assisting developing countries in	Assistance to developing countries	

the implementation of the 2030 Agenda especially in financing, technology transfer and human resource development. We also support and take part in South - South cooperation activities, which we believe are an effective channel for experience sharing, technical support and capacity building.		
Mr. President,	Vocative: new topic	
Like the United Nations, this year Viet Nam celebrates 70th anniversary of its foundation. Right after regaining independence from almost a century of colonialism, Viet Nam immediately associated itself with the principles and values of the Charter of the United Nations. President Ho Chi Minh, the founder of modern Viet Nam, wrote on several occasions to world leaders requesting the recognition of the new Viet Nam and its admission to the United Nations.	Self-reference and national circumstances (history)	Inform
And since becoming a member of the United Nation in 1977, Viet Nam has always done its best to uphold the principles and purposes of the UN Charter and contribute constructively and responsibly to the work of the Organization.		Connect speaker's country to UN
Viet Nam always puts the people at the center of all development strategies. Comprehensive human development encompasses physical and mental well-being, including human rights. Viet Nam has introduced laws and policies to ensure that all human rights of its people, including economic, social, civil and political rights, are respected, protected and promoted. Viet Nam is today party to seven out of nine core international human rights treaties, with the most recent being the Convention on the Rights of Persons with Disabilities.	Self-reference and development / human rights	Praise speaker's country
Viet Nam is strongly committed to the successful implementation of the Agenda 2030 for Sustainable Development.		
At the national level, we will mobilize all available resources, engage all relevant government agencies and other stakeholders, and involve the people and their communities in this major undertaking. Viet Nam pledges its support for the conclusion of a meaningful agreement at the upcoming COP-21 in Paris, and on its part has recently submitted its Intended Nationally Determined Contributions (INDC).	Climate Change	Promise support

At the regional level, we join efforts with our ASEAN brothers and sisters to work towards the ASEAN community. As an integral part of the dynamically growing Asia-Pacific region, ASEAN plays a crucial role in regional development. But for our region to thrive, peace and stability must come first.	Self-reference and national circumstances (region)	Inform Praise regional organization
It is therefore of vital importance that peace and stability are maintained, maritime safety and security and freedom of navigation and overflight in the South China Sea are secured. Viet Nam is working tirelessly with other ASEAN members and all parties concerned to that end, including by calling upon all parties to refrain from the threat or use of force and settle all disputes by peaceful means in accordance with international law, including the 1982 UN Convention on the Law of the Sea. We are also engaging with partners to ensure the full and effective implementation of the Declaration of Conduct of Parties in the South China Sea, and to work for the early conclusion of a Code of Conduct.	Region and peace South China Sea: territorial integrity	Persuasion: Request joint efforts Praise speaker's country / Promise support
At the international level, Viet Nam is prepared to further our active contributions to world peace and security and the well being of all. That is why we step up our participation in UN peace-keeping operations and has put forward candidatures for the Economic and Social Council for 2016-2018, UNESCO Executive Board for 2015-2019 and the Security Council for 2020-2021. We look forward to receiving your continued support.	Peace and security Candidacy (speaker's country) for UN offices	Request support
Mr. President,	Vocative: new topic	
To free our people from fear, from want, and leave no one behind, let us all join actions on this path towards a better and more sustainable future. Let the UN Charter be our source of inspiration and the 2030 Agenda be our guide in building a world of peace, security and prosperity for our people and succeeding generations.	Future Development: Agenda 2030	Persuasion: Request joint action Thank
I thank you, Mr. President.		

The address from Viet Nam also begins with greetings to the President of the Assembly, and with congratulations and thanks to elected officials for their efforts. Like with the address from San Marino, the next passage reflects on past UN achievements, with mention of the general importance of the UN for peace and security, human rights, development, and territorial integrity. Then Vietnamese Vice Minister of Foreign Affairs Nga speaks about two specific recent achievements: the Iran nuclear deal and normalization of relations between the US and Cuba. Next, she addresses the challenges that lie ahead, again mentioning peace and security and territorial integrity as well as terrorism, nuclear weapons, climate change, poverty, and human rights. H.E. Nga speaks at greater length on development and in particular the UN Agenda 2030 for Sustainable Development, with a passage on reform of the General Assembly and Security Council sandwiched between longer sections on development and the Agenda. In this passage on development, H.E. Nga connects peace and security to development and speaks about the importance of aiding developing countries in the implementation of these goals. The next section of the address connects Viet Nam to the UN, with some discussion of the foundation of Viet Nam and admission of the country into the UN. The speaker then returns to the topic of development and human rights, but this time in reference to Viet Nam, with specific mention of national efforts for and commitment to development. Next, H.E. Nga discusses the environment, again referring to Viet Nam's support for UN action on climate change (see more on the Paris Conference and Agreement in Section 4.3.1.6). The last issue to be raised is regional development, peace, and security, with reference to the Association of Southeast Asian Nations (ASEAN) and the South China Sea dispute. This is the third time H.E. Nga speaks about territorial integrity/disputes and the second time she refers to maritime security, but this is the first time they are linked (not entirely explicitly) to the South China Sea. Viet Nam is one of

111

several countries making claims to territory in the South China Sea, but the speaker does not mention either this fact or the tension and even violence that have resulted from the dispute. The difference of opinion between Viet Nam and China in particular is left unsaid. She then goes on to express Viet Nam's commitment to peace and security, announcing her country's candidacy for three UN seats and asking for the support of other UN member states for their candidatures. She concludes on a positive note about a *better and more sustainable future*, again referring to the Agenda 2030.

The general and specific purposes of the address from Viet Nam are similar to those in the address from San Marino. The primary goals are to raise issues of critical importance to Viet Nam and to convince the audience of their importance. In order to achieve these goals, the speaker conveys information and makes predictions such as *The full and effective implementation of the Agenda, with a global partnership in its core, will address deeply rooted causes of our present problems and transform this world towards a more sustainable future*. There is praise for past achievements and calls for joint efforts. There is no overt criticism. Rather than state what is wrong, the speaker focuses on what needs to be done, for example *The Security Council needs to be more representative, democratic, transparent and effective*.

Like the text from San Marino, the address from Viet Nam has only a few instances of the first person singular and they are in the introductory remarks congratulating and thanking UN officials. In the remainder of the text, H.E. Nga speaks on behalf of Viet Nam: *Viet Nam has always believed, Viet Nam is strongly committed to, Viet Nam pledges its support, we welcome, we support.*

Overall, these two texts share many similar features. The discourse structure follows the same pattern: greet UN officials and Assembly, congratulate and thank UN officials, note past

UN accomplishments, note future UN challenges, focus on specific areas of concern, relate

specific areas of concern to the speaker's country including what has been done and/or support

for future endeavors, conclude. General topics are similar (peace and security, human rights,

development, climate change), though specific concerns differ (violence against women and

children, South China Sea dispute). The primary communicative purposes are to emphasize UN

issues and convince other UN member states of the need to take action on these issues. In order

to do so, there are passages in which the speakers inform, request action, praise, and occasionally

(very indirectly) criticize. Finally, the speakers position themselves not as individuals expressing

personal stance but rather as representatives of their countries expressing the attitudes and

desires of their people and governments.

It is worth noting that some texts deviate from these general patterns. Occasionally, a

speaker expresses overt criticism or personal stance using the first person singular. Such

exceptions will be pointed out in the analyses in Chapters 4, 5, and 6. However, it would be

misleading to present them here because the vast majority of the addresses adhere by and large to

this format.

Texts 4.1 and 4.2 have provided a macroscopic view of UNGA addresses. Two

situational characteristics (topic and purpose) have been illustrated explicitly; others (participants

and setting) have been touched on only indirectly. Section 4.3 will now examine in greater depth

the full set of situational characteristics for the UNGA.

4.3. Situational Characteristics

Register analysts use taxonomies of situational characteristics to systematically

investigate the communicative context (e.g., Gray, 2015; Staples, 2015a). The taxonomy

presented by Biber and Conrad (2009, p. 40) is particularly thorough, comprising participants,

relationships among participants, channel of communication, production circumstances, setting, communicative purpose, and topic. This framework was used to describe the situational features for the UNGA (Section 4.3.1) and to compare those features in the UNGA to four other registers: prepared speeches, spontaneous speeches, official documents, and conversation (Section 4.3.2).

4.3.1. Situational characteristics of the UNGA

Like with many other scripted political texts, the UNGA address is written by a speechwriter, not by the speaker who actually delivers the text orally. However, unlike many other political texts, the address is intended to represent the entire country. Not only is any mention of the speechwriter wholly absent from the address, but numerous speakers efface their own role. As shown in Texts 4.1 and 4.2, speakers tend to use the first person singular primarily in the opening remarks where they congratulate and thank UN officials. After this introduction, they use their country name (e.g., *Viet Nam welcomes*) or the first person plural (e.g., *we believe*) because they are speaking on behalf of their countries. In fact, many speakers make their role as representative explicit at the very outset of the speech. Just a few examples include: *I am here representing Jordan*; *I am honoured to address the General Assembly today on behalf of Ireland; I am pleased to bring our usual warm greetings from the Government and people of Grenada, on whose behalf I am privileged to address this renowned institution and distinguished gathering; It gives me great pleasure on behalf of the Islamic Republic of Afghanistan to express my sincere congratulations on your well-deserved election as President.* In the case of UNGA addresses, speakers (like speechwriters) do not necessarily express their own personal views; they express the views of their countries.

The situational characteristics necessarily reflect this unique communicative context. As shown in Table 4.1, the UNGA address is the text written by the speechwriter, the text spoken by the politician or diplomat, and the text symbolically representing the country.

Table 4.1. Situational Characteristics of the UNGA

Feature	UNGA Description		
	Written Text	**Spoken Text**	**Symbolic**
IA. Participants: Addressor	Speechwriter	High-ranking politician	Country
IB. Participants: Addressees	Politicians, diplomats, and general public	Over 100 high-ranking politicians	Other countries
IIA. Relations: Interactiveness	No interaction		
IIB. Relations: Social roles	Differing status and power	Same or similar level of status and power	In theory, same status and power; in practice, some imbalance in power
IIC. Relations: Personal	Strangers	Fellow high-ranking politicians	Some favored economic and political relations
IID. Relations: Shared knowledge	Specialist knowledge not expected of general public	Shared knowledge expected for all references to past events; some present events potentially shared for the first time	
IIIA. Channel: Mode	Written to be read	Written to be spoken	Written to be read/spoken
IIIB. Channel: Specific	Permanent: recorded, transcribed, translated, published		
IV. Production circumstances	Scripted, monologic		
VA. Setting: Time	Not shared time	Shared time, 28 September – 3 October 2015	
VB. Setting: Place	Not shared place	Public setting: UN New York	
VIA. Purpose: General	Report, persuade, maintain harmonious relations		
VIB. Purpose: Specific	Summarize events, describe past/current/future national policy, explain expectations for future international policy		
VIC. Purpose: Factuality	Factual, some opinion		
VID. Purpose: Stance	Epistemic, attitudinal, and style		
VIIA. Topic: General	International and national events and policies		
VIIB. Topic: Specific	Peace and security, environmental protection, UN reform, migration, human rights, development		
VIIC. Topic: Status	Importance and relevance varies	Highest importance and relevance	

4.3.1.1. Participants

The participants involved in UNGA addresses include the addressor and addressee. For the written text, the addressor is the speechwriter and the addressee is potentially the world at large. Speechwriters can expect the addresses to be read by politicians and diplomats both present and not present at the General Assembly, as well as the general public. For the oral text, the addressor is a high-ranking politician or diplomat. Speakers include heads of state (i.e., 31 presidents, three kings, one chief executive, and one sultan), heads of government (i.e., 26 prime ministers), or ministers of foreign affairs (24). A few deputy officials spoke (i.e., three vice presidents, one deputy prime minister, one vice minister of foreign affairs). Finally, one country's permanent representative delivered an address. The addressee for the oral text is the General Assembly, made up of high-ranking politicians and diplomats. However, speechwriters and speakers represent their country, not themselves. Therefore, addressors and addressees are, at least symbolically, the UN member states. In fact, who wrote the speech and the rank of the politician who delivered the speech are likely to be less influential in the features of the text than the characteristics of the country represented by the text. If the addressor represents a Small Island Developing State with limited resources and low economic and social development, his concerns and the way he expresses those concerns will probably be different from a large military power that is highly developed and enjoys valuable natural and human resources.

4.3.1.2. Relationships among participants

The category relationships among participants includes interactiveness, social roles, personal relationships, and shared knowledge. The first element is the most straightforward: there is no interaction, either for the written or spoken text. For social roles and personal relationships, again the positions of the speechwriter and the speaker are of limited importance.

Rather, the roles and relationships among the countries are likely to be more influential. Is the country speaking being criticized because it has invaded a neighbor? Is the country speaking the most influential economic power in its region? Is the country speaking recognized as an independent state by the UN but not by all other countries? These factors can affect the content and form of an address. Finally, for shared knowledge, the speech is written with the General Assembly audience in mind. Addressees are expected to have the same background knowledge as the speechwriter and speaker, particularly for past events. Potentially, some information on national events and policy could be divulged for the first time during the UNGA, but this is highly unlikely. In contrast, the general public, free to read or listen to the texts after they have been delivered, may not share this background knowledge, but no accommodations are made to facilitate their understanding. For instance, in Text 4.2, H.E. Nga speaks about territorial integrity in the context of the South China Sea without explaining any background to the maritime dispute. Though speechwriters do not provide background knowledge for a potentially uninformed audience, it is possible that the language used is different from the language in speeches not made available to the public.

4.3.1.3. Channel of communication and production circumstances

The channel of communication is written to be spoken. Texts are drafted and edited with great care by professional speechwriters. They are recorded, transcribed, translated, and published for the permanent archives, so words are not chosen lightly. The production circumstances allow for careful crafting with dense information-packing and speechwriters might be particularly inclined to be concise because there is a suggested 15-minute time limit for UNGA addresses. The written texts of these monologic speeches will not show any hesitations, repetitions, self-repairs, interruptions, or non-rhetorical questions, though the oral delivery

118

occasionally results in minimal discourse dysfluencies. In contrast, because the texts are written

to be spoken before an audience, the speechwriter could very well incorporate lexico-

grammatical features that facilitate comprehension in the spoken mode (e.g., first and second

person pronouns and questions for signposting, as discussed in Section 2.3.2). As a "hybrid"

register, UNGA addresses might have qualities characteristic of both written and spoken texts,

but this remains an empirical question, particularly given the relative dearth of information on

the degree to which and the ways in which channel of communication and production

circumstances affect linguistic features in scripted political texts.

4.3.1.4. Setting

The setting for the UNGA corpus is highly constrained. All addresses were delivered

between 28 September and 3 October 2015 before the General Assembly at the United Nations

in New York. In addition to impacting topic, as discussed below, the shared setting between

participants affects type of reference, allowing for situation-dependent reference such as *at this

podium* and *this year*. Though these texts may be read decades after delivery, they are

nonetheless written to be spoken at one specific venue, and therefore elaboration is not expected

if the setting serves as sufficient explanation.

4.3.1.5. Purpose

The communicative purpose of UNGA addresses is multi-faceted. As illustrated in Texts

4.1 and 4.2, one of the primary goals of the address is to raise UN issues of particular national

interest and to convince other member states of their importance and the need to take action to

address them. To achieve this general goal, speakers explain why the issue is important and what

work is being done, particularly at the national or regional level. Speakers also praise work that

has already been carried out, they ask addressees to take action, and occasionally they denounce obstacles to tackling the issue effectively (e.g., lack of support from another member state). When countries have been accused of flouting international law or even convention, they are likely to defend themselves against this criticism during their address as well.

One other communicative purpose colors all the goals cited above: diplomacy. Because maintaining harmonious relations with other member states is critical for the very functioning of the UN, this communicative purpose shapes how speakers attempt to achieve their other goals of informing, requesting, praising, criticizing, and defending. One of the most obvious implications of diplomacy is the need to mitigate criticism except in the most extreme circumstances (as shown in Texts 4.1 and 4.2). Another effect is a focus on being informative. When an opinion is framed as an objective fact, it is less contentious. When a request is phrased as a logical necessity, it is less imperious.

For instance, in Text 4.1 from San Marino, the Agenda 2030 and its sustainable development goals are presented as a positive development, but H.E. Bodini expresses indirect concern that countries will not see the goals through to completion. He therefore underlines, again indirectly, the need for the UN to ensure national implementation (see Text 4.3).

Text 4.3. Excerpt from the San Marino Text

> The 15-years goals at the core of the Agenda are ambitious: they chart our efforts in pursuit of poverty eradication, people empowerment, protection of human dignity, shared prosperity, decent job for all, and protection of the planet. They will promote cooperation on sustainable development and guide public and private action. But it is important that each Country, after adopting its national Agenda, will take responsibility towards its own citizens and the international community for its implementation. San Marino commends that the GA and the ECOSOC will assess the implementation progress. Therefore, it is critical to produce a comprehensive and effective mechanism for review and accountability.

The speaker's argumentation includes no expression of personal stance. His concern about national implementation of the Agenda is apparent only from the fact that he mentions its importance (*it is important that*). His request for *comprehensive and effective* UN oversight of national implementation is prefaced by praising the decision to assess implementation at all. The speaker then uses another impersonal structure: *it is critical* in his call for not just assessment but holding countries accountable to their national progress on the Agenda.

It is precisely because diplomacy is of utmost concern and persuasion is so indirect that communicative purpose can be difficult to detect. Moreover, relatively little has been reported in the literature on international organization discourse as to the linguistic structures used to persuade. Communicative purpose is therefore one of the primary situational characteristics explored in the analyses in Chapters 4 and 5.

4.3.1.6. Topic

The final situational feature is topic. As shown in Texts 4.1 and 4.2, some topics are raised just in passing (e.g., climate change in Text 4.1) and others are discussed in greater depth (e.g., the South China Sea maritime dispute in Text 4.2). Some topics are vast (e.g., human rights) and others specific (e.g., the Iran nuclear deal). Though a number of the broader topics (peace and security) feature in UNGA addresses year after year, specifics change based on current events and concerns (violence in Iraq and Syria). Specific topics also tend to reflect the theme of the General Assembly for that year. I chose UNGA texts from one year (texts delivered over a span of only six consecutive days) in order to introduce as little variation as possible in terms of topic.

Texts 4.1 and 4.2 refer to the broad topics of peace and security, human rights, development, territorial integrity, poverty, democracy, climate change, economic crisis,

migration, terrorism, UN and Security Council reform, nuclear non-proliferation, gender

equality, and violence against women and children. Specific topics include Iraq, Syria, the

Agenda 2030 for Sustainable Development, the Iran nuclear deal, relations between the US and

Cuba, the South China Sea maritime dispute, and elections for UN offices. Given that the theme

of the 70th session of the General Assembly was "The UN at 70: The road ahead for peace,

security, and human rights," it is not surprising that Texts 4.1 and 4.2 mention the history and

future of the UN as well.

Other likely general topics can be gleaned from the UN Global Issues Overview (United

Nations, n.d.), including conflict resolution and peacekeeping, human rights, international

justice, economic and social progress, climate change, refugees, and AIDS. More specific topics

for the year can be found in the United Nations Year in Review for 2015 (UN Web TV, 2015).

Pivotal moments in the Review include the Agenda 2030 for Sustainable Development and its 17

Sustainable Development Goals (e.g., quality education, gender equality, clean water and

sanitation); the Paris Climate Change Conference and importance of environmental protection;

violence in Yemen, Syria, Libya, South Sudan, Palestine, Nigeria, and the Central African

Republic, with the ensuing refugee crises; terrorism, the Islamic State of Iraq and the Levant

(ISIL, also referred to as ISIS and Daesh), and the role young people can play in countering

violent extremism; Burundi and human rights violations; Ebola; the nuclear agreement with Iran;

elections and floods in Myanmar; the earthquake in Nepal; and the cyclone in Vanuatu.

4.3.1.7. Summary

This review of the communicative context of UNGA addresses reveals a number of

commonalities across texts. The addresses are all written by speechwriters, spoken by high-

ranking politicians and diplomats, and potentially read or heard by the general public. The texts

122

are delivered at the United Nations General Assembly General Debate over a six-day period, so neither setting nor time should influence language. The narrow time span also helps control, to some degree, the topic. Many countries are referring to the same recent events or upcoming conferences. Finally, the purpose of these addresses remains similar, at least on the surface: to convey a country's position on issues relevant to the international community and to persuade other member states accordingly. In contrast, some situational characteristics vary: the countries represented, the specific topic, and the specific communicative purposes. This variation in situational characteristics will be explored further in Chapter 5.

4.3.2. UNGA situational characteristics compared to four other registers

Section 4.3.2 contrasts the situational characteristics of the UNGA overall to other types of prepared speeches, as well as spontaneous speeches, official documents, and conversation to further elucidate the situational features distinguishing these texts from other registers. Comparing the UNGA to other registers helps reveal what characteristics make it distinctive. Four registers from Biber (1988) were selected for comparison purposes: prepared speeches (sermons, university lectures, court cases, political speeches, and popular lectures), spontaneous speeches (court cases, dinner speeches, radio essays, and speeches in the British House of Commons), official documents (government documents, foundation reports, industry reports, and college catalogues), and face-to-face conversation (interactions between intimates as well as strangers). See Section 3.3.2.2 for a more detailed description of the texts in each register.

Table 4.2 summarizes the similarities and differences in situational characteristics, showing that all four registers offer distinctive combinations of characteristics. By comparing the UNGA to all four registers in the linguistic analysis, it becomes possible to tease out which lexico-grammatical differences are due to which situational characteristics.

123

Table 4.2. Comparison of the UNGA to Planned Speeches, Spontaneous Speeches, Official Documents, and Conversation

Feature	Prepared speeches	Spontaneous Speeches	Official Documents	Conversation
IA. Participants: Addressor	Similar	Similar	Similar	Different
IB. Participants: Addressees	Similar	Similar	Similar	Different
IIA. Relations among participants: Interactiveness	Similar	Similar	Similar	Different
IIB. Relations: Social roles	Similar	Similar	Similar	Different
IIC. Relations: Personal	Similar	Similar	Similar	Different
IID. Relations: Shared knowledge	Similar	Similar	Similar	Similar
IIIA. Channel: Mode	Similar	Different	Different	Different
IIIB. Channel: Specific	Different	Different	Different	Different
IV. Production circumstances	Same	Similar	Similar	Different
VA. Setting: Time	Same	Same	Different	Same
VB. Setting: Place	Same	Same	Different	Same
VIA. Purpose: General	Similar	Similar	Similar	Different
VIB. Purpose: Specific	Different	Different	Different	Different
VIC. Purpose: Factuality	Same	Same	Similar	Different
VID. Purpose: Stance	Similar	Similar	Different	Similar
VIIA. Topic: General	Different	Different	Different	Different
VIIB. Topic: Specific	Different	Different	Different	Different
VIIC. Topic: Status	Similar	Similar	Similar	Similar

4.3.2.1. Participants and relationships between participants

The participants for prepared speeches typically include one addressor and one audience. With sermons, classroom lectures, court cases, and popular lectures, the speech writer is generally the same person as the speaker. Only political speeches are likely to be written and

delivered by two different people. For the most part, speakers such as university professors and popular lecturers are representing themselves, though arguably with a sermon the priest is representing the church and in a court case the lawyer is representing a client or the state. The addressee is generally the audience present during the delivery: the congregation, the students in a class, the audience in the courtroom or at a lecture. Only in a few rare cases is the speech recorded and broadcast or published in written format for the general public. As for the relationship between participants, prepared speeches are monologues, so there is generally no interaction. The relationship between the addressor and addressee is generally similar to an expert-novice relationship, where the speaker has knowledge or information to convey to the audience. In some instances, addressor and addressee know one another (e.g., sermons, classroom lectures), but in others they do not (e.g., court cases, popular lectures).

Some of the spontaneous speeches (court cases and speeches in the British House of Commons) resemble prepared speeches (court cases and political speeches, respectively) in terms of participants. For radio essays, the speaker is generally an expert, but in contrast to many of the other speech texts, the audience is not present; speakers do not know whether the audience is small or large, informed or uninformed on the topic. For dinner speeches, the addressor and addressee are more likely to know one another and to share the same status and background information. Participants and relationships between participants vary a great deal from one type of text to another within the register of spontaneous speeches.

For official documents, the addressor is one writer or group of writers but they represent another entity (e.g., the government) and not themselves personally. The addressee is the general public. There is no interaction or relationship between addressor and addressee and no assumed shared knowledge.

Finally, for face-to-face conversation, the addressor and addressee share some background information, interact with one another, and often know one another personally. In some cases, as with an exchange between customer and clerk, there might be a slight difference in status, but to a lesser degree than with an expert-novice relationship.

The participants and relationships between participants for these four registers offer an effectual comparison for UNGA addresses. Prepared and spontaneous speeches are similar in that the addressor is an expert, but an important difference is that the addressee for the UNGA is also an expert with the same background knowledge. Thus, the relationship between participants is not the same. Official documents are similar in that the addressor represents another entity (e.g., the government) but different in that addressor and addressee have no shared knowledge. Most official documents are written for the general public; UNGA addresses may be read by the public, but they are presumably written for an audience of high-ranking politicians and diplomats. Conversation offers the biggest contrast; though addressors and addressees are of the same status, their relationship is much more personal and interactive than with the UNGA.

4.3.2.2. Channel of communication and production circumstances

The channel of communication and production circumstances vary greatly across these four registers. Prepared speeches are written to be spoken. The degree to which the written speech is revised, edited, and read in the oral production may not be homogenous across all the texts, but the production circumstances make it possible to have dense phrasal structures like with texts that are written to be read. Because they are intended to be spoken, however, the texts may also have features that facilitate oral comprehension. With spontaneous speeches, the speaker is thinking and speaking at the same time, so the texts are less likely to have dense information-packing and are more likely to be marked with hesitations, repetitions, and self-

126

repairs. Spontaneous speeches may receive some thought before they are delivered and are likely to be on topics in which the speaker is well versed, so these discourse dysfluencies will probably be less pronounced than with the most unplanned spoken register: face-to-face conversation. In addition to hesitations, repetitions, and self-repairs due to lack to planning, conversation is likely to have interruptions, unfinished sentences, and non-rhetorical questions due to its interactive, dialogic nature. Official documents are written to be read. They are carefully revised and edited so writers are likely to pack information into nominal structures. Because they are not delivered orally, discourse structure is likely to be signaled using very different lexico-grammatical features. Given that addressor and addressee are not face-to-face during the production circumstances, second person pronouns and other references to the addressee are less likely.

One of the most compelling situational characteristics for UNGA addresses is production circumstances because relatively little is known about linguistic variation based on whether a speech is scripted. These four comparison registers serve as a crucial contrast for channel of communication and production circumstances. At one end of the spectrum, official documents are carefully polished written texts. At the other end, face-to-face conversation is unplanned and because it is dialogic, co-constructed; it is impossible for any one speaker to predict what will be said in a few turns. In between official documents and conversation lie prepared and spontaneous speeches. Prepared speeches most closely resemble UNGA addresses. Both are written to be spoken, though UNGA addresses are available in written format as well so the audience present during the oral delivery of the address does not necessarily have to listen to the speech; addressees can read it instead. Another moderate difference is that UNGA addresses are recorded and published, while many prepared speeches are not on permanent record. This difference is worth pointing out because knowing that a speech will be on public record for a long time could

very well affect both content and form. Overall, these four registers are very different from one another in terms of channel of communication and production circumstances, so they should elucidate the ways in which planning and scripting a text influence linguistic features.

4.3.2.3. Setting

The setting for prepared speeches, spontaneous speeches, and conversation is shared: the addressor and addressee are in the same place at the same time. Prepared speeches that are recorded or published may be heard or read at a later date as well, but they are delivered with the immediate audience in mind. Deictics are likely to reflect that fact, with possible reference to *here* and *now* in addition to references to the addressee using second person pronouns. In contrast, for official documents, setting is not shared and in most cases the texts are available to addressees for a relatively long period of time, sometimes years after they are produced.

With UNGA addresses, the speaker and the audience at the General Assembly also share time and place. As discussed in previous sections, the speechwriter writes the address before and the general public may read the text after delivery. However, this is unlikely to affect the linguistic features because the addresses are written with the General Assembly setting in mind, with references to *at this podium* and *today*. Thus, prepared speeches, spontaneous speeches, and conversation can be said to be the same in terms of shared setting (time and place), but official documents are different.

4.3.2.4. Purpose

The communicative purpose of the comparison registers varies a great deal. Within the register prepared speeches, sermons are likely to narrate and persuade, to edify and sometimes to entertain. Expression of personal stance is not frequent; instead, the speaker cites religious texts

128

as support for ideas and arguments. University lectures are primarily to inform and edify, synthesizing information from numerous sources. Expression of likelihood might nuance a proposition's degree of certainty. Court cases narrate, analyze, and most importantly persuade. Stance is likely to be expressed through impersonal structures such as adverbs or adjectives rather than clauses using the first person, suggesting a more factual and objective account of events. Political speeches inform and persuade, and though information is presented as factual, some personal opinion is also expressed as well. Popular lectures inform, entertain, sometimes persuade. In some lectures, speakers narrate their lives, revealing thoughts and feelings. Radio essays, in the register spontaneous speeches, are similar to popular lectures in terms of purpose. Spontaneous political speeches resemble prepared political speeches for purpose; spontaneous court cases are like prepared court cases. Other spontaneous speeches such as dinner speeches primarily narrate and entertain with a great deal of personal involvement and stance. In conversation, speakers narrate, explain, and express their ideas, opinions, and emotions. This is the most involved register with the most expression of stance. Official documents are primarily to inform. Most are expected to be factual with little overt stance.

For communicative purpose, again the four registers provide a good deal of variation. Of all the registers, prepared and spontaneous speeches, with the primary purposes of informing and persuading, are most like UNGA addresses, but there is a good deal of variation even within the two registers (e.g., dinner speeches compared to court cases), so general purpose cannot be said to be the same, only similar to the UNGA. In terms of degree of factuality, they are by and large the same: factual with some expression of opinion. The two speech registers are similar to the UNGA in terms of stance, but they are likely to have more stance reflecting personal attitudes; in the UNGA speakers might instead express the official position of the country they represent. For

129

conversation, stance is similar (again, stance is personal rather than reflecting the views of a country), but all general and specific purposes and factuality are quite different. Official documents are not unlike the UNGA for general purpose because they, too, intend to inform. They are also similar in terms of factuality but not entirely the same because little to no opinion is expected in official documents. Specific purpose and stance are quite different. Overall, purpose is not an easy situational characteristic to tease apart because there is so much variation even within each register.

4.3.2.5. Topic

The greatest variation within each of the four registers lies in topic. Within the register of prepared speeches, for example, topics span from crime to chemistry to public policy. Even within one text type such as academic lectures there is a great deal of variation based on the academic subject. The number of texts likely to address topics such as human rights and climate change for any of the registers is quite low, however. Even in political speeches and official government documents, there is little chance of the same topic emerging given that the texts from Biber (1988) date back to half a century before the UNGA texts. The only similarity, for all the registers is "status" of the topic, which is generally important for both addressor and addressee. However, the importance of a topic such as violence against children in UNGA addresses is not comparable to the importance of a topic such as weekend plans in face-to-face conversation. Overall, none of the comparison registers offers similar characteristics in terms of topic. In fact, it is unlikely that any register other than international organization discourse would address similar topics. This is not, however, particularly problematic for this study because topic, of all the situational characteristics is perhaps least likely to affect grammatical structures. The

influence of topic will be seen in the lexical analysis, but the lexical analysis uses the BNC as a reference corpus rather than the four general comparison corpora.

4.3.3. Summary of situational characteristics

Sections 4.3.1 and 4.3.2 have highlighted some of the situational characteristics that distinguish the UNGA from other registers. Participants, relationships between participants, channel, production circumstances, setting, purpose, and topic are all likely to affect the linguistic features of the texts. Given the situational characteristics of the UNGA, addresses are expected to be formal and informative, like official documents, but with some expression of opinion and veiled persuasion, similar to other speech registers. UNGA addresses, as with other prepared speeches, are a "hybrid" register: They are carefully planned and edited like official documents so speechwriters can pack in a great deal of information on topics such as peacekeeping and development. At the same time, the texts are delivered orally before a live audience like a spontaneous speech so they might include some involved features such as second person pronouns and rhetorical questions in acknowledgement of the addressees listening to the text. Further reference to these situational characteristics will be made in Sections 4.4 and 4.5 as part of the functional explanation of linguistic patterns in UNGA texts.

4.4. UNGA Lexical Features

To investigate lexical features, a keyword analysis was carried out, comparing the most frequent words in the UNGA corpus to the BNC using WordSmith Tools 7.0 (Scott, 2017). As discussed in Section 3.7.1, a keyword analysis determines which words are statistically more frequent in a given text or corpus when compared to another. Keywords reflect what is lexically distinct in the target corpus. The results of the keyword analysis comparing the UNGA to the

BNC as a reference corpus uncovered distinct topics, text structure, and purposes in UNGA

texts. The top 100 keywords with frequency and dispersion details are shown in Table 4.3.

Table 4.3. UNGA Keywords when Compared to the BNC

N	Key word	Freq.	%	Texts	RC. Freq.	RC. %	Keyness	*P*
1	our	1942	1.10	92	10042	0.10	5579.00	0.00
2	nations	830	0.47	91	359		5328.33	0.00
3	UN	683	0.39	83	483		4007.08	0.00
4	president	840	0.47	91	1393	0.01	3944.46	0.00
5	global	487	0.27	90	117		3383.08	0.00
6	development	785	0.44	88	1770	0.02	3313.25	0.00
7	peace	626	0.35	90	714		3279.79	0.00
8	of	7494	4.23	92	214341	2.08	3016.18	0.00
9	we	2713	1.53	92	44158	0.43	2929.54	0.00
10	security	659	0.37	86	1335	0.01	2891.66	0.00
11	international	667	0.38	88	1808	0.02	2618.92	0.00
12	united	737	0.42	91	2804	0.03	2486.61	0.00
13	sustainable	332	0.19	79	42		2446.79	0.00
14	climate	369	0.21	69	148		2395.93	0.00
15	human	464	0.26	85	841		2115.24	0.00
16	agenda	368	0.21	76	291		2107.23	0.00
17	rights	409	0.23	75	650		1945.81	0.00
18	states	423	0.24	83	954		1784.41	0.00
19	countries	413	0.23	83	887		1773.98	0.00
20	and	7179	4.05	92	245610	2.39	1705.58	0.00
21	challenges	256	0.14	80	92		1689.24	0.00
22	the	12612	7.11	92	505546	4.92	1587.89	0.00
23	world	675	0.38	92	4786	0.05	1585.00	0.00
24	assembly	247	0.14	84	221		1374.87	0.00
25	cooperation	161	0.09	70	45		1098.39	0.00
26	efforts	243	0.14	81	462		1089.64	0.00
27	conflict	202	0.11	65	249		1035.76	0.00
28	conflicts	149	0.08	62	51		989.88	0.00
29	Mr	777	0.44	90	11225	0.11	965.74	0.00
30	its	673	0.38	89	9385	0.09	870.71	0.00
31	humanitarian	131	0.07	52	47		864.33	0.00
32	implementation	138	0.08	59	72		857.96	0.00
33	charter	163	0.09	64	192		846.24	0.00
34	Syria	117	0.07	36	25		822.92	0.00
35	must	488	0.28	81	5491	0.05	787.71	0.00
36	general	331	0.19	86	2395	0.02	765.82	0.00
37	economic	246	0.14	76	1133	0.01	751.85	0.00

N	Key word	Freq.	%	Texts	RC. Freq.	RC. %	Keyness	P
38	poverty	141	0.08	67	171		726.20	0.00
39	change	332	0.19	78	2632	0.03	719.01	0.00
40	organization	137	0.08	48	163		709.30	0.00
41	terrorism	111	0.06	50	52		702.95	0.00
42	support	307	0.17	79	2456	0.02	660.34	0.00
43	humanity	101	0.06	46	46		642.60	0.00
44	commitment	156	0.09	66	388		633.63	0.00
45	excellencies	71	0.04	51	0		579.05	0.00
46	principles	110	0.06	55	140		558.95	0.00
47	peaceful	100	0.06	56	91		554.32	0.00
48	peoples	91	0.05	47	55		550.62	0.00
49	to	5928	3.34	92	250078	2.43	546.63	0.00
50	developing	132	0.07	50	315		544.83	0.00
51	gentlemen	95	0.05	58	79		537.75	0.00
52	peacekeeping	74	0.04	36	13		530.58	0.00
53	islands	113	0.06	17	194		524.27	0.00
54	agreement	148	0.08	63	515		520.62	0.00
55	governance	65	0.04	37	2		512.19	0.00
56	goals	169	0.10	65	798		509.49	0.00
57	community	251	0.14	74	2201	0.02	502.90	0.00
58	country	289	0.16	76	3000	0.03	502.25	0.00
59	prosperity	83	0.05	49	52		498.73	0.00
60	gender	77	0.04	36	34		492.37	0.00
61	crises	71	0.04	36	19		486.92	0.00
62	address	145	0.08	65	566		482.68	0.00
63	inclusive	72	0.04	41	24		480.05	0.00
64	extremism	62	0.03	35	4		475.60	0.00
65	continue	184	0.10	69	1131	0.01	474.48	0.00
66	migration	77	0.04	35	45		468.88	0.00
67	progress	129	0.07	68	452		452.31	0.00
68	session	132	0.07	59	486		452.10	0.00
69	this	1408	0.79	92	43650	0.42	445.98	0.00
70	ensure	137	0.08	64	605		428.16	0.00
71	pacific	77	0.04	19	71		425.48	0.00
72	universal	80	0.05	52	90		420.44	0.00
73	collective	83	0.05	46	112		414.55	0.00
74	resolution	84	0.05	42	125		407.71	0.00
75	state	197	0.11	73	1676	0.02	404.19	0.00
76	stability	83	0.05	44	123		403.36	0.00
77	committed	118	0.07	62	439		402.12	0.00
78	Israel	93	0.05	13	198		400.59	0.00
79	Ukraine	66	0.04	20	40		399.13	0.00
80	effective	112	0.06	57	385		396.21	0.00
81	weapons	95	0.05	39	222		395.26	0.00

N	Key word	Freq.	%	Texts	RC. Freq.	RC. %	Keyness	P
82	SDGS	48	0.03	18	0		391.46	0.00
83	SIDS	53	0.03	17	7		389.25	0.00
84	regional	143	0.08	60	810		388.12	0.00
85	Cuba	58	0.03	25	20		384.90	0.00
86	anniversary	109	0.06	54	393		377.06	0.00
87	excellency	47	0.03	24	1		373.62	0.00
88	respect	116	0.07	59	485		372.99	0.00
89	refugees	89	0.05	44	206		371.64	0.00
90	Iran	73	0.04	17	93		370.83	0.00
91	equality	76	0.04	46	112		369.97	0.00
92	Africa	111	0.06	45	439		367.09	0.00
93	crisis	117	0.07	51	513		367.02	0.00
94	republic	96	0.05	40	290		359.84	0.00
95	Paris	97	0.05	59	301		359.35	0.00
96	member	159	0.09	67	1189	0.01	359.25	0.00
97	war	206	0.12	71	2135	0.02	358.45	0.00
98	island	102	0.06	35	374		350.02	0.00
99	future	207	0.12	77	2253	0.02	344.30	0.00
100	political	174	0.10	74	1554	0.02	343.50	0.00

Note. Freq. = frequency. R.C. = reference corpus (here, the British National Corpus).

As discussed in Section 3.7.1, the top 100 keywords were classified semantically. Five categories for the keywords were selected based on a review of UN literature and on a qualitative examination of the keywords: (1) UN global issues or "topical keywords," (2) UN procedure and functioning or "UN general functioning keywords," (3) Terms of address, (4) Words related to solidarity and joint efforts or "cooperation keywords," and (5) Function keywords. Sections 4.4.1 through 4.4.5 present each of these categories in turn, with an operational definition of the category, list of keywords, interpretation, and textual examples.

4.4.1. Topical keywords

As discussed in Section 4.3.1.6, the UN tackles what the organization calls "Global Issues." These are general topics such as refugees and climate change (United Nations, n.d.). Because keywords reflect the "aboutness" of a text, I expected a large number of keywords to be

"topical," i.e., to be related to UN "Global Issues." I created a list for 18 of the 19 Global Issues (ageing, AIDS, atomic energy, big data for the SDGs, children, climate change, decolonization, democracy, food, health, human rights, international law and justice, oceans and the law of the sea, peace and security, population, refugees, water, and women). I removed the Global Issue Africa because it overlapped with UN members (see Section 4.4.2). I also added one category: development. Given that the Agenda 2030 for Sustainable Development and the Sustainable Development Goals (SDGs) had been adopted by the General Assembly just three days before the General Debate of 2015 began, development was a key topic for the UNGA that year.

I then classified keywords that were related to each one. After classifying, I merged two related categories (women and human rights) because *equality* seemed to be appropriate for either one. Only one other keyword fell under the category women (*gender*); an examination of concordance lines confirmed that *gender* also fell under the broader domain of human rights because references were made primarily to *gender equality* and violence against women. Of the top 100 keywords, 39 were "topical," related to UN global issues, as shown in Table 4.4.

Table 4.4. Topical Keywords Organized by Category

Peace and Security	Refugees	Human Rights	Development	Climate Change	Democracy
peace	refugees	human	economic	climate	governance
security	migration	rights	prosperity	change	respect
stability		equality	development	SIDS	
peaceful		gender	progress	island(s)	
peacekeeping		humanity	sustainable	pacific	
conflict(s)		humanitarian	goals	Paris (Conference)	
terrorism			SDGs		
crises/crisis			developing		
extremism			poverty		
weapons					
war					

All of the "topical" keywords were semantically transparent to me because I was very familiar with the texts. For instance, I recognized *SDGs* as *Sustainable Development Goals.* When in doubt, I confirmed associations by examining concordance lines. I knew that references to Paris related to the *Paris Conference* or the *Paris Agreement on Climate Change*, but I verified this manually. Similarly, I checked references to *island, islands*, and *Pacific*: All occurrences were in passages about the environment and climate change (see Text 4.4), not in passages on maritime law or marine biodiversity, which would have fallen under the category of oceans and the law of the sea.

Text 4.4. Excerpt from Samoa (keywords in italics)

> Against the backdrop of the existential threat *climate change* poses especially to atolls and low lying *islands*, *SIDS* had long been advocating for ambitious mitigation efforts by member states with the capacity to do so, and for a global goal of limiting the rise in average global temperature to well below 1.5 degrees Celsius to prevent some of our low-lying *islands* from being submerged by sea level rise. We have also consistently called for a Loss and Damage mechanism to be anchored in a new *Paris* Agreement to be treated separately from "adaptation" to accord to *climate change* the same deadly capacity.

It is not surprising that over 30% of the top 100 keywords are about these Global Issues. What is more surprising is that over a dozen Global Issue categories had no keywords (e.g., decolonization, food, health). Equally surprising is the general consistency in what is said about these topics. For instance, *conflict* and *conflicts* are often associated with *extremism, terrorism, migration,* and *refugees* and countries in the Middle East and Africa are frequently mentioned as conflict zones. (See Texts 4.5 through 4.8.)

Text 4.5. Excerpt from Denmark (keywords in italics)

> We need a strong and unified response to violent *extremism* and *terrorism*. Not just in Syria and Iraq, but also in parts of Africa, where violent *extremism* and armed *conflicts*

are growing. Denmark will do its part. We plan to increase our contribution to the UN operation in Mali - MINUSMA. A very visible consequence of all these *conflicts* is the massive refugee and *migration* flows.

Though the text from Denmark does not create a cause-effect relationship between

terrorism and *extremism* on the one hand and *conflicts* on the other, the two are juxtaposed,

implying an association, and a direct causal relationship is specified between *conflicts* and

migration. Syria, Iraq, and Africa are named as conflict zones.

Text 4.6. Excerpt from Australia (keywords in italics)

> Australia is committed to defeating *terrorism* in all its forms. We are responding to the challenges of violent *extremism* and foreign terrorist fighters, taking on those who would exploit our openness and modern communications to commit violence and promote terror. The humanitarian consequences of the *conflicts* in Syria and Iraq are devastating. I commend neighbouring countries - Jordan, Lebanon, Turkey - which continue to bear the brunt of the population of displaced persons. Australia has provided $230 million in humanitarian assistance since the start of the *conflict*. The massive flows of displaced people across the Middle East and into Europe have stretched response capacities to breaking point.

The text from Australia, like Text 4.5 from Denmark, does not explicitly link *terrorism*

and *conflict*, but the juxtaposition implies a relationship. Again, like the text from Denmark, the

association between *conflict* and migration (*the massive flows of displaced people*) is much more

direct, and Syria, Iraq, and the Middle East are mentioned as areas of conflict.

Text 4.7. Excerpt from Botswana (keywords in italics)

> Deeply troubling is the increasing involvement of non-state actors and radical extremists in most of these *conflict* situations. Activities by these entities introduce additional layers of difficulty to conflict management and resolution processes. *Terrorism*, including the emergence of abductions as one of its many forms, further complicates the complexities of responses to *conflict* in areas of unrest.

The word *involvement* in the text from Botswana much more explicitly links extremism and conflict, and *terrorism* is noted as an additional complication for resolving conflict.

Text 4.8. Excerpt from the Bahamas (keywords in italics)

> We are facing the most pressing refugee crisis since World War II. In recent months, we have seen hundreds of thousands of migrants and *refugees* from *conflict-* and war-torn areas throughout the Middle East embarking on perilous journeys seeking refuge in Western Europe.

The text from the Bahamas also suggests that *conflict* results in migration, again with specific mention to *refugees* fleeing conflict in the *Middle East.*

Another common association in the "topical" keywords is between peace and sustainable development. A number of countries quote or paraphrase the 2030 Agenda for Sustainable Development, which states "there can be no sustainable development without peace and no peace without sustainable development" (United Nations General Assembly, 2015, p. 2). Though "peace, justice, and strong institutions" is just one of the 17 Sustainable Development Goals, this connection between peace and security on the one hand and development on the other is made again and again: *It is an all-inclusive agenda which makes it abundantly clear that there can be "no sustainable development without peace and no peace without sustainable development"* [Seychelles], *And without peace and security we will not achieve sustainable development* [Iceland], *This new Agenda recognizes that sustainable development cannot be realized without peace and security; and peace and security will be at risk without sustainable development* [Kenya], *Peace, security and human rights must go hand in hand with sustainable development* [Bhutan], *There is no development without security* [Croatia].

The association between *conflict(s)* with *terrorism, extremism, refugees,* and *migration* and the association between *peace* and *security* with *sustainable development* are just two

examples illustrating commonalities between UNGA addresses. Overall, the "topical" keywords show the importance of topic in the lexical features of UNGA texts. Not only do speakers address similar general and specific topics, with a particular penchant for six Global Issues (peace and security, refugees and migration, human rights, development, climate change, and democracy), but they also tend to say similar things about these topics.

4.4.2. UN general functioning keywords

Another highly productive group of keywords relates to the general functioning of the UN, with 37 out of the top 100 keywords in this semantic category. The initial reading of UNGA addresses showed this category to be important, with a great deal of discussion on UN operations generally. Regardless of the "Global Issue" being addressed, speakers refer to the work of the UN: to its bodies (e.g., *General Assembly*), to its members (e.g., *countries, Iran*), and to its decision-making (e.g., setting an *agenda* and the need to *ensure* the *implementation* of an *agreement*). This category also refers to the scope of its work (e.g., *regional, global*) and to its continuity (e.g., *future*). Table 4.5 lists the keywords classified in this category.

Table 4.5. UN General Functioning Keywords

Bodies	Members	Decision-making	Scope	Continuity
nations	state(s)	agenda	regional	future
UN	country/-ies	implementation	global	anniversary
united	Iran	charter	international	continue
(general) assembly	Cuba	agreement	world	
general (assembly)	Africa	resolution	universal	
organization	Syria	address	effective	
session	republic	ensure	political	
	member	principles		
	peoples	challenges		
		goals		

139

"UN general functioning" keywords are used throughout the texts but primarily in sections in which the speaker reviews the institution's past achievements and shortcomings or outlines future work. As shown in Texts 4.1 and 4.2 (Section 4.2), after the introductory greetings to the Assembly and congratulations to the President, many texts situate the UN in the world of 2015, mentioning the organization's origins, history, scope, and future objectives. Texts 4.9 through 4.11 illustrate the use of some of the "UN general functioning" keywords in such passages, again revealing many similarities across texts.

Text 4.9. Excerpt from Romania on the United Nations (keywords in italics)

> The 70th *anniversary* of the *United Nations* is yet another opportunity for a lucid reflection on the lessons learned and on the *future* of our *organization*. What we have seen is a visionary project born of the ashes of a terrible tragedy with the lofty goal to save succeeding generations from the scourge of war. We have seen a working *agenda* that has continuously expanded into new areas of cooperation, within and beyond its main pillars: peace and security, development and human rights. We have witnessed the continuous efforts of the *organization* and of its *Member States* to adapt to new *challenges* and to major shifts in the paradigm of *international* relations. As a consequence, the *United Nations* has steadily assumed new mandates, thus reaching a level of complexity never seen before.

In Text 4.9 from Romania, the speaker situates the United Nations at its 70th anniversary, with reference to the inception of the UN and in general terms to how it has grown through its history (e.g., *the continuous efforts of the organization, has steadily assumed new mandates*).

Text 4.10. Excerpt from Ethiopia (keywords in italics)

> As we mark the 70th *Anniversary* of our *organization*, it is only fitting that we take stock of the achievements and *challenges* over the past seven decades and draw appropriate lessons from history as we move forward. That the world has made tremendous progress in a wide range of areas through the advancement of the noble objectives enshrined in the *UN Charter* is not in doubt.

Like the text from Romania, Text 4.10 from Ethiopia mentions the 70th anniversary and the need to draw on past *achievements and challenges* when considering future steps. Just as Romania refers to some of the overarching goals of the organization (*peace and security, development and human rights*), Ethiopia refers to the *noble objectives enshrined in the UN Charter.*

Text 4.11. Excerpt from Pakistan (keywords in italics)

> Seventy years ago, the *United Nations* was created from the ruins of the most devastating war the world has witnessed. Its purpose was to build *universal* peace and prosperity on the basis of equitable principles, cooperation and collective action. Despite the constraints of the Cold War, the *United Nations* served the *international* community as the beacon of hope; the repository of freedom; an advocate of the oppressed; a vehicle for development and progress. But, we - the *peoples* of the *United Nations* - have not succeeded in beating our arms into ploughshares or promoting universal prosperity and larger freedoms.

In Text 4.11, Pakistan, too, speaks of the 70-year history of the UN and the fundamental goals of the organization: *peace, prosperity, freedom, development and progress*. Though the final evaluation of the UN is less positive: *we… have not succeeded*, like Texts 4.9 and 4.10, Text 4.11 includes some favorable appraisal (*the United Nations served the international community as the beacon of hope*) and the overall tone and content are not unlike the passages in the other texts.

These excerpts with "UN general functioning" keywords illustrate how terms related to UN procedure and operations often emerge in discussions on a range of topics, both general and specific (e.g., *political* and *Cuba*, respectively). As with the "topical" keywords, the "UN general functioning" keywords show that not only do UNGA speakers tend to use the same terminology, they also, by and large, use that terminology in similar ways.

Another set of keywords are terms of address (e.g., *Mr. President*). Though type count is very small with only five keywords (see Table 4.6), token count makes this category non-negligible. Out of a total word count of 174,994, just over 1%, 1830 tokens, are terms of address. No other keywords are related to genre features, so these terms of address appears to be a distinct category.

Table 4.6. Terms of Address Keywords

president
Mr. (President)
excellencies
gentlemen
excellency

These terms of address are used primarily at the beginning of the address. Ninety-one speeches begin with a salutation: *Mr. President* (21 addresses), *Ladies and Gentlemen* (2 addresses), or a combination of *Mr. President*, *Ladies and Gentlemen*, and other titles such as *Excellencies*, *Distinguished delegates*, *Majesties* (68). Only one address, given by the representative of the European Union, does not commence with any salutation. In addition, *Mr. President* is sometimes repeated later in the address to transition to a new topic, emphasize an idea, or conclude. Text 4.12 illustrates its use in stressing a key point.

Text 4.12. Excerpt from Saint Lucia (keyword in italics)

> We are firmly of the view, that countries in our Hemisphere committed to this perspective now have an opportunity, on the basis of the initiative taken by Pope Francis and the President of the United States, to promote the path of peaceful resolution of disputes, in full awareness of the fact that national or regional conflict are inhibitors to persistent economic growth, and therefore social stability. *Mr. President*, the normalization of relations within this Hemisphere to which I have just adverted, is an

indication that old conflicts and disputes kept frigid by the era of the Cold War, are giving way to new avenues of cooperation and collaboration between countries.

Overall, these "terms of address keywords" relate to discourse structure and show the importance of production circumstances. The texts are delivered before an audience, and speakers address that audience personally. The identification of these terms of address as keywords is not particularly surprising given that addressing the audience and congratulating the president have been shown in the literature to be common in UNGA addresses (Donahue & Prosser, 1997, see Section 2.4.2).

4.4.4. Cooperation keywords

Nine keywords relate to cooperation and joint efforts (see Table 4.7).

Table 4.7. Cooperation Keywords

cooperation
efforts
support
commitment
agreement
community
inclusive
collective
committed

These "cooperation" keywords appear distinct from the other keywords. They do not relate to any particular Global Issue or to UN general functioning, and they do not stem from the discourse structure of a formal speech like terms of address. Instead, they are linked to communicative purpose (maintaining diplomacy and harmonious relations). Most of the words are transparent; only *efforts* required manual examination of concordance lines to ensure it was

semantically related. A few of the many examples illustrating its relationship to cooperation include: *We hope that our development partners will share our sense of priority and urgency in this vital area and support our efforts to ensure that it is relevant to achieve the 2030 Agenda* [Bhutan], *Rest assured, we shall continue with our efforts and contributions* [Croatia], and *As a regional player keen on promoting security in the area, Greece has been making steady efforts to improve cooperation with Turkey* [Greece].

With nine types representing 1539 tokens (less than 1% of the total word count) on the "cooperation keywords" list, this is clearly not the largest semantic domain of keywords. However, the importance of expressing agreement and support in the interest of maintaining diplomacy has been noted in the literature on international organization discourse (e.g., Valentine & Preston, 2002, see Section 2.4.1), which supports the inclusion of cooperation as its own keyword category.

Two texts exemplify the prominent use of solidarity-related lexicon: Text 4.13 from the Seychelles and Text 4.14 from Jordan.

Text 4.13. Excerpt from the Seychelles (keywords in italics)

> The commemoration of the 70th anniversary of the United Nations is a unique opportunity to transform the world by our own *efforts* and *commitments*, and to ensure that *collective cooperation* and partnership triumph over rivalry and mistrust. In this spirit we salute the rapprochement between the United States of America and Cuba, This is the unique opportunity to choose the future we want for ourselves, for our youth for our children. We must give them a better deal. They deserve a better deal. Let us truly become what we are destined to be: *committed*, determined, united and empowered.

Text 4.13 makes use of a number of keywords listed in Table 4.7 as well as other semantically related words (*partnership, rapprochement, united*). The speaker from the Seychelles begins with a more general statement on the importance of cooperation in the UN and

then uses the example of the normalization of relations between the US and Cuba as an example of commitment to unity. Both the text from the Seychelles and Text 4.14 from Jordan associate cooperation with a brighter future. The text from the Seychelles speaks of *a better deal* for *our children*, and the speaker from Jordan talks about a *safer, stronger world.*

Text 4.14. Excerpt from Jordan

> In all these areas, we must act, collectively, for the future of our world. And connectivity is key. The values we share - equality, compassion and hope - need to be connected to everything we do. And we must keep ourselves connected to each other, for the good of all. These bonds are the power and promise of our United Nations. Here, together, we can and must create the future our people need: a safer, stronger world of coexistence, inclusion, shared prosperity and peace.

Though none of the words in Text 4.14 are on the "cooperation keywords" list, words such as *collectively, connectivity, share, equality, connected, bonds, coexistence, inclusion,* and *shared* are all part of the semantic domain of cooperation. These words are indicative of the lexical diversity of this semantic category. Speakers talking about cooperation have a myriad of words to choose from. This lexical diversity makes it harder for words to appear as keywords, as discussed in Section 3.7.1. When many different word forms (*cooperate, cooperates, cooperated, cooperating, cooperation, cooperative, cooperatively*) and synonyms (*partnership, collaboration, teamwork, coordination, joint action, collective efforts*) are possible, each type will inevitably appear less frequently than if only one term expressed the same meaning. Thus, the prominence of words related to cooperation is suggested by the nine keywords on the list in Table 4.7, but may in actuality be even greater than these nine words would imply. The "cooperation" keywords category is not the most numerous in terms of either type or token, but it remains a viable semantic domain, supporting with quantitative data the qualitative findings of

previous research (e.g., Valentine & Preston, 2002) indicating the importance of agreement and solidarity.

4.4.5. Function keywords

The remaining nine keywords were function words and categorized based on word class (see Table 4.8).

Table 4.8. Function Words Categorized

1st person plural	Modal auxiliaries	Coordinator	Prepositions	Impersonal determiners
our	must	and	of	the
we			to	its
				this

Also known as grammatical words, function words indicate relationships between lexical or content words. Many keyword analyses exclude function words from the analysis because the focus is on topical content (see Baker, 2006). Register analyses also tend not to examine function keywords; instead, grammatical structures are tagged and either individual features or Multi-Dimensional (MD) scores are examined (e.g., Biber, 2006; Staples, 2015a). In this study, too, grammatical structures were tagged and MD scores examined (see Section 4.5), offering a broad and deep analysis of these and other function words. Because many function words are interpreted in Section 4.5, some of the function words from the keyword analysis will receive only cursory review in Section 4.4.5.3.

Given the objective of Section 4.4 and the keyword analysis, to look at lexical rather than grammatical patterns, only function words that contribute to the overall picture of lexical patterns will be examined in Section 4.4.5. These function words include *we* and *our* in Section 4.4.5.1

and *must* in Section 4.4.5.2. Section 4.4.5.3 is a brief discussion of the remaining function keywords and how they relate to grammatical patterns explored further in Section 4.5.

As function words, the lexical importance of *we, our,* and *must* is not as transparent as with content words. In order to clarify the relationship of these function words to content words and thus demonstrate their importance on the keyword list, I undertook an analysis of their collocates or the words in the immediate co-text, as described in Section 3.7.1. This type of analysis reveals lexical patterns, in keeping with the objective of the keyword analysis in Section 4.4.

4.4.5.1. *First person plural*

The only personal determiners or pronouns on the keyword list are in the first person plural: the determiner *our* and pronoun *we.* It is interesting to note that other personal pronouns such as *I* and *you* are not exceptionally frequent. The discussion in Section 4.4.5.1 focuses on the ways in which *our* and *we* are used in order to better understand the reason they occur with much higher frequency compared to other spoken and written texts. The discussion also explains why it is determiners and pronouns in the first person plural that are so frequent rather than other person forms such as the first person singular.

Collocational analysis was carried out for both *our* and *we* to elucidate any lexical patterns with these two function keywords. As discussed in Section 3.7.1, collocational analysis identifies the words appearing with significant frequency within a five-word span of the target word. The collocational analyses reveals an association between the first person plural and content words related to the semantic domain of cooperation and joint efforts (see Section 4.4.4 on classifying words in the semantic domain of cooperation). Table 4.9 lists the noun collocates

that appear in the top 100 collocates of *our* with a checkmark next to the collocate if it was

categorized as being semantically related to cooperation and joint efforts.

Table 4.9. Noun collocates with *our*

Noun collocate	Texts	Total	Joint efforts semantic domain
people	45	107	
development	49	101	
world	53	100	
efforts	39	76	√
future	39	65	
commitment	36	61	√
security	34	59	
nations	36	58	
peace	35	55	
support	36	53	√
agenda	33	53	√
country	23	45	
partners	24	43	√
president	26	41	
region	24	40	
work	30	40	√
planet	25	40	
organization	16	35	
countries	27	35	
challenges	23	32	
UN	17	30	
peoples	20	29	
children	17	28	
goals	19	26	√
economy	13	26	
part	18	26	√
states	19	25	
rights	17	24	
years	21	24	
community	19	23	√
action	16	23	√
time	19	23	
climate	19	23	
nation	12	22	

All of these collocates have a positive or neutral evaluation, and many reflect the belief that all nations and people are one (*our world, our children*). Of the 34 nouns in the top 100 collocates of *our,* 10 are related to cooperation and joint efforts. Texts 4.15 and 4.16 illustrate these collocations and their association to collaboration.

Text 4.15. Excerpt from Tonga (keywords in italics and collocates underlined)

> I congratulate you on your assumption of the presidency of the 70th Session of the General Assembly and assure you of *our* full <u>support</u>.

Text 4.16. Excerpt from Brunei Darussalam (keywords in italics and collocates underlined)

> This is indeed a historic occasion to celebrate *our* achievements, as well as reflect on matters that require *our* urgent <u>action</u>.

Collocates of the pronoun *we* reveal similar findings. Table 4.10 lists the verbs, excluding auxiliaries, that appear in the top 100 collocates of *we*. Many of the verbs collocating with *we* are positive (*hope*) and of the 20 verbs in the top 100 collocates of *we*, nine imply joint efforts or cooperation (*support*, *committed*), as illustrated in Texts 4.17 and 4.18.

Table 4.10. Verb collocates with *we*

Verb collocate	Texts	Total	Joint efforts semantic domain
support	47	77	√
welcome	39	63	√
continue	44	63	
see	34	59	
believe	33	52	
call	34	51	√
face	34	50	√
want	27	49	
look	34	47	
hope	38	46	
work	25	41	√
committed	27	38	√
remain	26	37	
ensure	28	36	√
make	22	32	
change	21	30	
adopted	24	30	√
made	24	28	
take	22	28	
address	21	27	√

Text 4.17. Excerpt from Slovenia (keywords in italics and collocates underlined)

> *We* <u>welcome</u> the adoption of the Addis Ababa Action Agenda and the new sustainable development agenda.

Text 4.17 illustrates the very common collocate *welcome,* indicating support for the two

UN agendas.

Text 4.18. Excerpt from the United Kingdom (keywords in italics and collocates underlined)

> And alongside these challenges, *we* <u>face</u> the longer term threats of climate change, global pandemics and anti-microbial resistance – threats that *we* must tackle effectively now, if *we* are not to put the lives of future generations at risk.

Though the excerpt from the UK shows the collocate *face* being used with negative evaluation, to discuss a problem rather than a positive development, the passage nonetheless shows a focus on joint efforts – the need to cooperate in order to tackle the challenge of climate change and disease together.

Some of the collocates required manual investigation in order to determine their meaning. For instance, out of the 51 occurrences of the collocation *we call*, it is used only three times to mean *name*. All other instances of the collocation are in the sense of *request* (e.g., *call for, call upon*), and are used to signal a need for cooperation as shown in Text 4.19.

Text 4.19. Excerpt from Namibia (keywords in italics and collocates underlined)

> In this context, *we* <u>call</u> upon the United Nations to assume its full responsibility by implementing all its resolutions and decisions on Palestine with no pre-conditions. *We* equally <u>call</u> for the urgent implementation of all Security Council and General Assembly resolutions, with the aim of holding a free and fair referendum in Western Sahara.

Overall Section 4.4.5.1 has shown that *our* and *we* are used not exclusively but nonetheless frequently in passages related to joint efforts. This further supports the findings in Section 4.4.4 on the "cooperation" keywords, showing an emphasis in UNGA addresses on the communicative purpose of maintaining diplomacy through expression of solidarity. It also supports the findings on the first person plural in previous studies of political speeches using very different methods. For example, the use of *we* for polarization and legitimation has been studied at length (see Section 2.3.3). Many studies have suggested that using the inclusive *we* is a means of implying solidarity, support, and cooperation. These studies (e.g., Mulderrig, 2012) manually categorized occurrences of *we* as inclusive or exclusive. The inclusive *we* includes both addressor and addressee (see Text 4.20); the exclusive *we* does not include the addressee (see Text 4.21).

151

Text 4.20. Excerpt from Finland (*we* in italics)

>   The international community must now show that *we* do care.

Text 4.21. Excerpt from Ghana (*we* in italics)

>   In Ghana, *we* have made tremendous progress in achieving the Millennium Development Goal target on universal basic education.

The problem with classifying inclusive and exclusive *we* is the frequency of ambiguous cases. In fact, studies such as Mulderrig (2012) suggest that ambiguity in first person plural referents is sometimes deliberate. Even when *we* is factually exclusive, the addressee might be led to interpret it as inclusive. In Text 4.22, the first *we* might initially be interpreted as exclusive because not all member states may wish to reform the Security Council. However, the second *we* is factually inclusive because the Human Rights Council was created through a UNGA resolution. The juxtaposition of inclusive and exclusive *we* blurs the distinction between the two.

Text 4.22. Excerpt from Myanmar (*we* in italics)

>   *We* believe that the UN Security Council also has to be reformed to make it more representative, efficient, transparent and to strengthen its capacity to better discharge its responsibilities to maintain international peace and security as entrusted by the Charter. As part of the UN reform, *we* have set up the Human Rights Council in 2006 to redress the shortcomings of its predecessor, the Human Rights Commission.

Because classifying pronouns as inclusive or exclusive is highly subjective, particularly given the potential for ambiguity, it is not a particularly reliable way of determining the extent to which the speaker is implying solidarity, support, and joint efforts. Categorizing collocates based on semantic meaning can also be somewhat subjective, but overall, collocational analysis offers a more systematic means of examining associations made with keywords. When *our* collocates

with significant frequency with words such as *efforts, agenda, work,* and *action,* a lexical pattern clearly emerges.

4.4.5.2.   Modal *must*

Must can express either logical necessity or personal obligation, so understanding the high frequency of *must* in the UNGA required further exploration into its use. Concordance lines with the modal *must* were examined in order to identify the most frequent subjects and whether *must* was used for logical necessity or personal obligation. The three most frequent subjects were *we, international community,* and *United Nations.* Of the 488 occurrences of *must*, 187 were with the subject *we*, 21 were with *UN, United Nations, United Nations system, world's nations, or association of nations,* and 16 were with the *world* or *international community.* All of these occurrences used the personal obligation sense of *must.*

When *must* is used for personal obligation, it has strong directive force, so it is noteworthy that the most frequent subjects with *must* include the addressor. This mitigates the directive force, again implying joint efforts and cooperation rather than imposition on the addressee alone. In fact, *you must* is never used in the UNGA as a directive targeted at the audience. *You must* occurs only twice in the entire corpus. Israel uses *you must* as a modal of logical necessity rather than personal obligation: *So here's a general rule that I've learned and you must have learned in your life time - When bad behavior is rewarded, it only gets worse*; Tonga uses *you must* in the inclusive sense of people in general: *It is this language of morality that speaks to the privileged with the imperative, "you must leave no one behind."*

Another way to mitigate the directive force of *must* is through the use of the passive voice, to displace agency. The use of *must* in the passive to lessen the force of imposition has been noted in the literature on political speeches (see Section 2.3.3) and on international

153

organization discourse (see Section 2.4). Of the 488 occurrences of *must*, 81 are in the passive. The passive is used most often with *must* when the subject is implied to be the international community: *Extremist ideologies must be opposed* [Pakistan], *These people must be held accountable for their crimes and brought to justice* [Malta], *Violence must never be allowed to prevail* [The Netherlands]. However, it is also used to avoid direct imposition and implied blame on one country or group of countries. Texts 4.23-4.25 illustrate this use of the passive with *must* to take focus off the agent being asked to fulfil an obligation.

Text 4.23. Excerpt from Zimbabwe (passive voice *must* in italics)

> We are deeply concerned by the continued denial of this basic right to the Saharawi people. We urge the United Nations to expeditiously finalise what *must be done* to conclude the decolonisation of the Western Sahara.

In Text 4.23, the speaker from Zimbabwe refers to the contested annexation of Western Sahara by Morocco. Rather than make any direct imposition on Morocco or even name Morocco, the speaker uses the passive to focus instead on the action necessary to ensure full decolonization.

Text 4.24. Excerpt from Croatia (passive voice *must* in italics)

> On the European continent, in the Ukraine, the implementation and the viability of the peace process have to be ensured. The ceasefire *must be respected* and the Minsk peace plan *adhered* to by all parties.

In Text 4.24, the speaker from Croatia speaks about the Russian occupation of Ukraine. Again, rather than blaming Russia for not respecting the ceasefire or asking Russia to adhere to the Minsk peace plan, Croatia uses the passive to avoid agency.

Text 4.25. Excerpt from Samoa (passive voice *must* in italics)

> The Paris Agreement requires a new brand of cooperation and broad outlook. The narrow pursuit of self-interests, the use of economic and political expediencies should be set aside and *must not be allowed* to de-rail the goal of concluding a universal, ambitious and legally binding agreement.

In Text 4.25, the speaker from Samoa expresses concern that some countries will not implement the Paris Agreement for economic and political reasons. Again, Samoa does not name any countries in particular or explicitly state that any countries would voluntarily *de-rail the goal.* Instead, *the narrow pursuit of self-interests* is put in subject position.

This analysis of *must* has shown that the modal is used primarily for personal obligation, but rarely with strong directive force. Rather, it is used with *we* and other subjects that include the speaker (e.g., *the international community*) to indicate joint efforts. The use of *must* in the passive voice similarly mitigates the directive force by placing the emphasis on the action rather than the agent. The passive is used primarily when the agent is implied to be the international community, again stressing the need for joint efforts, as well as when the speaker wishes to avoid naming one state or group of states. Overall, the use of *must* can be linked to the communicative purpose of maintaining harmonious relations with other UN member states through the expression of cooperation and joint efforts and through the avoidance of direct confrontation with other countries.

4.4.5.3. Coordinator, prepositions, and impersonal determiners

The coordinator *and*, prepositions *to* and *of*, and impersonal determiners *the*, *its*, and *this* are noteworthy keywords because these grammatical forms are associated with the degree to which a text is informational (with more phrasal structures) or, on the contrary, involved (more clausal). The degree to which UNGA addresses are informational or involved is a compelling

155

area of investigation because it gets at the heart of one of the questions about scripted texts raised in Sections 2.3, 2.4, and 4.3.1.3: Does this "hybrid" register, texts written to be spoken, most resemble spoken or written texts? More features signaling informational production would imply that the UNGA is similar in that respect to written texts; more features signaling involved production would suggest the UNGA more closely resembles spoken texts. Informational production is also linked to communicative purpose and is expected from a register with a focus on conveying information.

This is fundamentally a grammatical rather than a lexical, issue, however, and the question of involved versus informational production is considered in greater depth in the Multi-Dimensional analysis of Dimension 1 (see Section 4.5.1).

4.4.6. Summary of keyword analysis

The keyword analysis reveals some compelling findings on the lexical features of UNGA addresses. First, the addresses are, on average, quite similar. Of the top 100 keywords, 75% appear in at least half the texts (46 texts total). The keywords could be classified into just a few very large categories, indicating that the most frequent words were serving just a few broad functions. These functions included discussing UN Global Issues, referring to UN general functioning, adhering to the discourse structure of a UN address with terms of address, expressing support and encouraging joint efforts, and mitigating directive force to save face in the interest of diplomacy.

Not only do UNGA countries tend to talk about the same issues (e.g., peace and security, climate change, UN decision-making), they tend to say the same things about the issues (e.g., associating peace and security with development, talking about past achievements and current challenges for the UN at the beginning of the address). It is not particularly surprising to find that

the 70th anniversary of the UN is a common topic or that the collocations *our world* and *we support* are frequently used. What is surprising is that over half the countries in the UN do. These similarities across texts suggest relative homogeneity in UNGA content and even form (all but one address beginning with *Mr. President*).

These keywords also support with quantitative data previous qualitative findings on lexicon related to agreement and support, on the first person plural, as well as on the modal *must.* Nine of the top 100 keywords were semantically related to cooperation, showing that the communicative purpose of diplomacy and maintaining good relations with other countries is reflected in word choice. In addition, the high frequency of the first person plural collocating with words related to cooperation and joint efforts indicates that, like with other political speeches discussed in Section 2.3.3, UNGA speakers imply they share goals and values with the audience. Finally, the modal *must* is used most frequently with subjects that include the speaker (*we, international community, UN*) or in the passive voice. This mitigates the directive force of *must.* Once again, this indicates a concern for maintaining diplomacy and avoiding direct imposition. Overall, the keyword analysis has shown that the situational characteristics topic ("topical" and "UN general functioning" keywords), production circumstances ("terms of address" keywords), and communicative purpose ("cooperation" keywords, *our, we*, and *must*) influence the lexical features of UNGA addresses, with notable similarities across texts.

4.5. UNGA Grammatical Features: Multi-Dimensional Scores

The goal of Section 4.5 is to explore UNGA grammatical features and text type by analyzing MD scores. UNGA MD scores were calculated for five previously analyzed dimensions found in Biber (1988): Dimension 1: Involved versus Informational Production, Dimension 2: Narrative versus Non-Narrative Concerns, Dimension 3: Explicit versus Situation-

157

Dependent Reference, Dimension 4: Overt Expression of Persuasion, and Dimension 5: Abstract versus Non-Abstract Information. (See Section 3.7.2 for additional information on these dimensions and the linguistic features that define them.) These scores were then compared to four previously analyzed registers from Biber (official documents, prepared speeches, spontaneous speeches, and conversation) so that UNGA grammatical features and text type could be analyzed relative to other registers. Score differences were interpreted based on the distinctive situational characteristics of the UNGA relative to the four comparison registers.

Sections 4.5.1 through 4.5.5 describe and interpret the dimension scores for the UNGA and the four comparison registers; Section 4.5.6 summarizes the findings.

4.5.1. Dimension 1: Involved versus informational production

Dimension 1 represents a range from involved to informational texts. Involved texts scoring high on Dimension 1 are distinguished by their high level of interaction and personal affect, with particularly high frequency of features such as first and second person pronouns, private verbs (e.g., *assume, believe*), *that*-deletion, present tense, and contractions. In contrast, informational texts show little interpersonal involvement. The focus is instead on conveying information, with more nouns, prepositions, and attributive adjectives. Table 4.11 lists the linguistic features associated with involved production and with informational production. For factor scores of these features, indicating the degree of their importance for Dimension 1, see Table 3.11.

Table 4.11. Linguistic Features Associated with Involved versus Informational Production

Involved Features (high Dimension 1)	Informational Features (low Dimension 1)
private verbs	nouns
that deletion	word length
contractions	prepositions
present tense verbs	type-token ratio
2nd person pronouns	attributive adjectives
DO as pro-verb	
analytic negation	
demonstrative pronouns	
general emphatics	
1st person pronouns	
pronoun *it*	
BE as main verb	
causative subordination	
discourse particles	
indefinite pronouns	
general hedges	
amplifiers	
sentence relatives	
WH questions	
possibility modals	
non-phrasal coordination	
WH clauses	
sentence final prepositions	

Biber (1988) found that generally informational texts are written because with spoken real-time production circumstances, the dense, informative noun phrases are both difficult to produce for the speaker and difficult to process for the listener. Spoken texts tend to be more involved because the clausal structures are easier to produce and process in real time. Thus, in Biber (1988), face-to-face conversation scored very high on Dimension 1 and official documents very low. Spontaneous speeches and prepared speeches fell between the two, with prepared speeches almost equidistant between spontaneous speeches and written official documents.

UNGA addresses scored much lower than the prepared speeches in Biber and nearly as low as

official documents (see Table 4.12).

Table 4.12. Scores for Dimension 1: Involved versus Informational Production

Register	Mean	SD	Minimum	Maximum
UNGA	-15.3	5.9	-25.1	4.3
Official documents	-18.1	4.8	-26.3	-9.1
Prepared speeches	2.2	6.7	-7.3	14.8
Spontaneous speeches	18.2	12.3	-2.6	33.1
Conversation	35.3	9.1	17.7	54.1

Note. SD = standard deviation.

Given the situational characteristics of the UNGA, these texts would be expected to fall

at nearly the same level as prepared speeches. Both prepared speeches and the UNGA are written

to be spoken. Thus, the production circumstances allow speakers to carefully craft the written

text and potentially create dense nominal structures, which is characteristic of informational

production. At the same time, the texts are meant to be delivered before a live audience, which

presupposes some involved features as well. The fact that the UNGA Dimension 1 score is so

much closer to official documents than to prepared speeches indicates that speechwriters give

preference to packing information into the text over creating a text that is easy to listen to orally.

This information-packing results in a high density of nouns, prepositions, and attributive

adjectives, as shown in the most informational address (D1 = -25.1), Text 4.26 from Suriname.

Text 4.26. Informational text from Suriname (informational features underlined, involved
features italicized)

> These <u>past</u> seven <u>decades</u> *have certainly brought* many <u>achievements</u>. The <u>United
> Nations, in</u> collaboration <u>with</u> its <u>Member States, specialized organizations</u> and other
> <u>relevant stakeholders,</u> including <u>civil society,</u> promoted and ensured <u>access to education
> to boys</u> and <u>girls around</u> the <u>world;</u> brought <u>humanitarian relief to</u> those <u>in need;</u> assisted
> <u>societies with</u> the <u>return to democracy;</u> supported those defending <u>human rights;</u> gave a

voice to the marginalized groups, *especially* the elderly, women, youth, indigenous peoples and persons with disabilities, *and* played a crucial role in the fight against racism, racial discrimination, xenophobia and related intolerance.

Text 4.26 illustrates the dense noun phrases of informational texts, with frequent nouns (e.g., *achievements*, *collaboration*), attributive adjectives (e.g., *specialized*, *relevant*), and prepositions (e.g., *in*, *with*). In contrast, there are almost no involved features: only one present tense verb (*have brought*), two amplifiers (*certainly, especially*), and one instance of non-phrasal coordination (*and*). Indeed, Text 4.26 reads more like a government report than a speech. It has none of the involved features of the planned, non-scripted speeches discussed in Section 2.3.1: There are no first or second person pronouns, no private verbs expressing stance (*think, guess*). There is no signposting: no questions to alert the listener to an important idea, no guiding the audience through the discourse structure of the text (e.g., *Now I will turn to*). This makes it more difficult for the listener to process when delivered orally.

This does not mean, however, that involved features are completely absent in UNGA texts, only that they are relatively rare in relation to other spoken registers, even other prepared speeches. Text 4.27 is an excerpt from the most involved address in the UNGA, the Maldives (D1 = 4.4).

Text 4.27. Involved text from the Maldives (informational features underlined, involved features italicized)

> Last week *we* adopted a new Agenda for Sustainable Development. *It recognises* at its core that development must *be* holistic. That poverty *is* a Multi-Dimensional problem. That, *what* matters *is* the human being, whose rights must be protected, *and* promoted. Yet here in the United Nations, *we remain* trapped in silos: hiding away, behind the excuse of mandates. *Why is it* that the Security Council must only discuss guns and bombs? *Why can't* the Economic and Social Council discuss war and peace? *Why can't* development, *why can't* war, have a human rights dimension? *Why must* issues be confined to one specific body?

161

Involved features in Text 4.27 include first and second person pronouns such as *we*, the pronoun *it*, *BE* functioning as a main verb (*must be holistic*), the present tense (*is*), non-phrasal coordination (*and*), WH- clauses (*what matters is*), questions (*why is it*), and contractions (*can't*). Many of the involved features discussed in Section 2.3.1 on scripted political texts are in the text from the Maldives as well: first person pronouns, factive verbs (*recognises*), and rhetorical questions. There are informational features as well, including nouns (*development*), prepositions (*for*), and attributive adjectives (*sustainable*), but not nearly as many as in Text 4.26, the highly informational text from Suriname. Together, these involved features form a text that is much more oral than literate. The greater number of clauses and less dense noun phrases make the text easier to follow for the listener. The use of *we* and the rhetorical questions draw the audience into the address. By using these involved features, the speechwriter from the Maldives creates a text allowing for engagement with the audience.

Involved speeches like the text from the Maldives are relatively rare in the UNGA, however. Only two texts (the Maldives and the US) have positive Dimension 1 scores, indicating on average more involved than informational production; the other 90 addresses are more informational.

Though UNGA texts are, overall, more informational than involved, Text 4.26 from Suriname and Text 4.27 from the Maldives illustrate the range in informational versus involved production in the UNGA. This range is also reflected in the large standard deviation (SD) for Dimension 1 relative to the UNGA SDs for the four other Dimensions (see Tables 4.12, 4.14, 4.16, 4.18, and 4.20). In other words, of all the dimensions of textual variation, Dimension 1, reflecting linguistic differences based on involved versus informational production, is where UNGA texts vary the most. This wide range of Dimension 1 scores is not unusual, and in fact,

the UNGA SD is actually smaller than the SD for the other spoken registers (see Table 4.12). This shows that while there is variation in Dimension 1 scores (discussed further in Chapter 5), on average UNGA addresses are more uniform in their degree of informational production than the comparison registers.

The analysis of Dimension 1 scores has shown that some UNGA texts have involved features, but are, on average, extremely informational. Though it may initially seem surprising that UNGA texts are so much more informational than prepared speeches, nearly to the point of official written documents, three explanations related to production circumstances and communicative purpose may account for the low Dimension 1 scores. In terms of production circumstances, UNGA speakers have a suggested 15-minute time limit to deliver their address. Therefore, if a speechwriter hopes to include a lot of information, it must be packed as densely as possible. In addition, UNGA addresses are available to the Assembly in written format as well. Some speechwriters might assume that the address will be read rather than heard, even by the audience present. If a text is written to be read rather than spoken, the condensed noun phrases and information density are not nearly as problematic for the addressees. This begs the question: Are UNGA addresses truly written to be spoken or actually written to be read, delivered orally only for the sake of appearances? Finally, one of the primary communicative purposes of UNGA texts is to convey information. The analysis of Dimension 1 scores shows to what extent this purpose is important: The dense noun phrases of these information-packed texts reveals a preference for conveying information rather than presenting a text that is easy for the listener to process.

4.5.2. Dimension 2: Narrative versus non-narrative concerns

Dimension 2 represents variation between texts that have a narrative or non-narrative focus, with higher scores indicating narrative and lower scores non-narrative texts. Narrative discourse, as defined in Biber (1988, p. 109), is more active, event-oriented; non-narrative is more static, descriptive, or expository. Grammatical features inherent in narratives include past tense, perfect aspect, third person pronouns, and public verbs (e.g., *claim*). No linguistic features were included in the analysis of negative Dimension 2 scores, meaning that a "non-narrative text" is one with a lower frequency of narrative features rather than a high frequency of non-narrative features (see Table 4.13).

Table 4.13. Linguistic Features Associated with Narrative versus Non-Narrative Concerns

Narrative Features (high Dimension 2)
past tense verbs
3rd person pronouns
perfect aspect verbs
public verbs
synthetic negation
present participial clauses

Biber (1988) found spontaneous speeches to be highly narrative, and official documents to be extremely non-narrative. Prepared speeches were nearly as narrative as spontaneous speeches; conversation was essentially non-narrative, less narrative than prepared speeches but somewhat more narrative than official documents. As shown in Table 4.14, UNGA addresses were even less narrative than official documents.

Table 4.14. Scores for Dimension 2: Narrative versus Non-Narrative Concerns

Register	Mean	SD	Minimum	Maximum
UNGA	-3.1	0.6	-4.3	-1.6
Official documents	-2.9	1.2	-5.4	-1.5
Prepared speeches	0.7	3.3	-4.9	6.1
Spontaneous speeches	1.3	3.6	-3.8	9.4
Conversation	-0.6	2.0	-4.4	4.0

Note. SD = standard deviation

Less narrative than even official documents, UNGA addresses do not frequently tell stories: They rarely recount past events (past tense, present and past perfect verbs), speak of third parties (third person pronouns), or refer to what has been said (public verbs). In contrast, court cases (in both spontaneous and prepared speech registers) are likely to narrate events, such as the details of a crime. Dinner speeches and some radio essays (spontaneous speeches), particularly those intended to entertain, are also narrative. Popular lectures and sermons (prepared speeches) can be too, with sermons often relying heavily on narration to illustrate precepts. In contrast, non-narrative texts are expository, intended to educate, to inform and explain. This is illustrated in the text from Sweden, which had a low Dimension 2 score (-4.2), indicating a non-narrative focus. Rather than tell a story, the address from Sweden is built around arguing a case. A premise is made, examples are given, and a rationale is provided (see Text 4.28).

Text 4.28. Excerpt from Sweden

> Working to keep the peace also means following through on our mutual promises of disarmament:
> - to pursue legal, practical and technical solutions to fully rid our world of its remaining 16 000 nuclear weapons;
> - to finally destroy what remains of other inhumane arms, such as biological and chemical weapons;
> - to realise the Arms Trade Treaty; and,
> - to fight the spread of small arms and light weapons, including ammunition.
> If we sway from these goals, humanity will suffer for our faults.

In Text 4.28, Sweden puts forward the premise that UN member states must follow through with their commitment on disarmament. Four examples of how they can do that are enumerated (e.g., *to realise the Arms Trade Treaty*) and the rationale explains that disarmament is for the good of humanity. The text is expository rather than narrative and has no narrative (high Dimension 2) features.

This is typical of most UNGA texts. Only rarely do UNGA texts refer to past events. When they do, it is often to describe past achievements, either of the UN (see Text 4.29) or of the speaker's country (see Text 4.30). However, these narrative sections describing past events are relatively short, lying within a larger passage that is expository. The past achievements are generally used for argumentation purposes. They can be used, for instance, in a section of the text that compares and contrasts (Text 4.29) or that illustrates a point with examples (Text 4.30).

Text 4.29. Excerpt from Greece (narrative features in italics)

> Seventy years *have passed* since the UN *was founded*, after the most disastrous war in history. In the wake of this War the UN's noble declarations *raised* hopes for a common peaceful future, based on universal values, social justice and respect of human rights. Much *has been* since *achieved*. But, regrettably, the vision of the founders of the Organization is far from being fulfilled. We witness today a large humanitarian crisis, unprecedented movements of displaced people and an increase of poverty, violence, extremism and human rights abuses. Nevertheless, the big question for us is not to identify the challenges we face. The peoples of this world know very well what *they* face every day. The big question that this great organization was set up to answer, is how do we DEAL with these challenges on the basis of our shared values of the UN Charter.

In Text 4.29 from Greece, Prime Minister Tsipras begins his address by commenting on the history of the UN and the importance of its achievements, using the perfect aspect (e.g., *have passed*) and the past tense (e.g., *raised*). This brief narrative portion is quickly followed by discussion of present challenges for the UN and implications for future action. The goals and

achievements of the past are contrasted with the challenges of the present in order to raise the

question: What should the UN do now and tomorrow?

Text 4.30 discusses national rather than UN achievements, but again in the context of a

present and future commitment to UN objectives.

Text 4.30. Excerpt from Myanmar (narrative features in italics)

> During the tenure of the present government, Myanmar *has stepped* up to raise the profile
> of Myanmar in the field of disarmament. We *have started* with the signing of the LAEA
> Additional Protocol in September 2013. It *was followed* by Myanmar's ratification of the
> Biological Weapons Convention-BWC in December 2014 and the ratification of the
> Chemical Weapons Convention. These actions demonstrate Myanmar's dedicated
> commitments to the cause of disarmament. Myanmar tables every year a resolution on
> Nuclear Disarmament adopted by the support of a great majority of the member states.
> The resolution reflects our firm conviction that our world will be secured and safer
> without nuclear weapons.

In the text from Myanmar, a narrative portion describes action the country has taken as

support for the argument that Myanmar is dedicated to the cause of disarmament. Past

achievements in the perfect aspect (*has stepped*) and past tense (*was followed*) are used to situate

the country and its position on a UN issue in the present.

The standard deviation for Dimension 2 shows little variation, with both a minimum

value (Suriname at -4.59) and a maximum value (Fiji at -1.63) signaling texts that are

particularly non-narrative. UNGA addresses are not focused on telling a story but on making a

case. Any narrative elements in the expository texts of the UNGA tend to be used as support for

an argument and are therefore only short sections in longer, non-narrative passages. The

emphasis on argumentation exemplified in low Dimension 2 scores illustrates the importance of

persuading the audience through logical reasoning. (See Section 4.5.4 for further discussion on

persuasive techniques.)

4.5.3. Dimension 3: Explicit versus situation-dependent reference

Dimension 3 represents the explicitness of reference in a text. An explicit text is one with elaborated, endophoric reference (e.g., *the process of implementation, which includes…*). On the opposite side of the spectrum is a situation-dependent text, with exophoric reference (e.g., *here*). High scores, with frequent WH- clauses, phrasal coordination, and nominalization but infrequent time and place adverbials indicate very explicit language, while low scores reveal more situation-dependent language. Table 4.15 shows the linguistic features associated with explicit and situation-dependent reference.

Table 4.15. Linguistic Features Associated with Explicit and Situation-Dependent Reference

Explicit Reference (high Dimension 3)	Situation-Dependent (low Dimension 3)
WH relative clauses on object positions	time adverbs
pied piping constructions	place adverbs
WH relative clauses on subject positions	adverbs
phrasal coordination	
nominalizations	

Biber (1988) found that conversation had more situation-dependent reference and official documents more explicit reference. Speeches fell between the two, with only slightly more explicit than situation-dependent reference (see Table 4.16). This is attributed to production circumstances. In face-to-face conversation, participants share the same time and place and therefore reference can be made to the physical and temporal context. For official documents, not only are time and place not shared, they are often unknown and unimportant. The topic of official documents rarely relates to the communicative situation. Instead, these documents are informative, elaborating on nominal structures to make reference explicit. Speeches fall between conversation and official documents. Like conversation, speeches involve shared time and place

so there is the possibility of making situation-dependent reference. However, shared time and place is only one situational characteristic of speeches. Speeches also tend to refer to events and ideas unrelated to the immediate communicative context. For example, in a court case, lawyers might refer to the hearing itself, which is a shared time and place, but it is also likely that they will refer to a past crime using explicit reference to elaborate on key events and players. Biber notes that explicit reference generally marks informational texts (p. 148), so informative speeches are likely to have more explicit reference than narrative speeches. This can explain the wide range in scores for spontaneous speeches in particular, as shown in Table 4.16. Dimension 3 scores for UNGA addresses are similar to those for the speeches in Biber, indicating on average more explicit reference, but nonetheless some situation-dependent reference.

Table 4.16. Scores for Dimension 3: Explicit versus Situation-Dependent Reference

Register	Mean	SD	Minimum	Maximum
UNGA	0.8	1.8	-3.4	5.3
Official documents	7.3	3.6	2.1	13.4
Prepared speeches	0.3	3.6	-5.6	6.1
Spontaneous speeches	1.2	4.3	-5.4	9.7
Conversation	-3.9	2.1	-10.5	1.6

Note. SD = standard deviation.

Dimension 3 is the only dimension in which the UNGA resembles the speech registers more than the other comparison registers, most likely due to the similar production circumstances and, to a lesser extent, purpose: UNGA addresses are delivered before an audience, resulting in some situation-dependent reference. However, again like many of the speech sub-registers (e.g., popular lectures, radio essays), they are also informative, and speakers elaborate on nominal structures to explain their importance (see Text 4.31).

Text 4.31. Explicit text from Mauritius (explicit features in italics)

> More than ever, Mauritius needs the *support* of the international community to untangle itself from the middle-income web and reach the high income country status. Financial *institutions and development* partners need to go beyond our relatively high GDP, *which* obscures the real *cost that* we have to pay for our *development* and precludes us from accessing vital *development finance and support*. *Development* partners should not forget that the specific *and* unique *vulnerabilities* of SlDS have been reiterated, highlighted and acknowledged by world *leaders* in the SAMOA Pathway and recently in the Post 2015 Agenda.

The text from Mauritius has a high Dimension 3 score (D3 = 3.28). This excerpt from the text illustrates how relative clauses serve to elaborate, both on subject positions (*our relatively high GDP, which obscures…*) and on object positions (*the real cost that we have to pay*). Phrasal coordination (*financial institutions and development partners*) and nominalization (*support*) are also used extensively, revealing a heavy focus on nominal structures in a very informative passage. This text, out of context, could have easily been written to be read rather than spoken.

In contrast, the text from Israel (see Text 4.32) has much more situation-dependent reference (D3 = -3.2). This text is clearly written to be spoken, with references to time (*thirty-one years ago, now*) and place (*at this podium*).

Text 4.32. Situation-dependent text from Israel (situation-dependent features underlined)

> Thirty-one years ago, as Israel's Ambassador to the United Nations, I stood at this podium for the first time. I spoke that day against a resolution sponsored by Iran to expel Israel from the United Nations. <u>Then</u> as <u>now</u>, the UN was obsessively hostile towards Israel, the one true democracy in the Middle East. <u>Then</u> as <u>now</u>, some sought to deny the one and only Jewish state a place among the nations. I ended that first speech by saying: Gentlemen, check your fanaticism at the door. More than three decades <u>later</u>, as the Prime Minister of Israel, I am again privileged to speak from this podium.

On average, UNGA addresses are only slightly more explicit than situation-dependent. This can be explained by the communicative context. The purpose of the address is primarily to

inform, raising key UN issues and discussing national or international action related to those issues. Speakers elaborate on nominal structures in these discussions to provide additional information or to explain their importance. As addresses are written in advance, language can be chosen carefully, including all explicit references necessary. On the other hand, speakers are present at the same time and place and have a great deal of shared background knowledge. Some references are therefore made to this shared time and place. It is interesting to note that addresses with more explicit reference sound like they could have been written to be read, whereas addresses with more situation-dependent reference, on average, sound more like they are delivered before a live audience.

4.5.4. Dimension 4: Overt expression of persuasion

Dimension 4 represents overt expression of persuasion, defined in Biber (1988) as "the speaker's own assessment of likelihood or advisability" (p. 148). High scores indicate frequent use of prediction and necessity modals, conditional clauses, and suasive verbs (see Table 4.17). No features are associated with texts showing low overt expression of persuasion.

Table 4.17. Linguistic Features Associated with Overt Expression of Persuasion

Persuasive Features (high Dimension 4)
infinitives
prediction modals
suasive verbs
conditional subordination
necessity modals
split auxiliaries

All four registers from Biber (1988) have relatively little overt persuasion, with the speech registers scoring only slightly higher than conversation and official documents on Dimension 4 (see Table 4.18). Though average Dimension 4 scores are close to zero for all four

registers, indicating neither highly frequent nor highly infrequent overt persuasion, the SD shows

a wide range of scores for the two speech registers and for official documents, with minimum

scores indicating texts that are not at all persuasive and maximum scores showing texts that are

extremely persuasive. This wide range of Dimension 4 scores may be due to differences in

communicative purpose for the distinct text types within the same register. For example, within

the register of prepared speeches, entertaining or informative popular lectures are unlikely to use

a great deal of overt persuasion, whereas persuasive court cases and political speeches should

have more. UNGA texts scored on average lower on this dimension than prepared speeches,

spontaneous speeches, and even official documents, with the lowest score from Sierra Leone (-

3.6) and the highest score from Nauru (6.9).

Table 4.18. Scores for Dimension 4: Overt Expression of Persuasion

Register	Mean	SD	Minimum	Maximum
UNGA	-0.3	2.0	-3.6	6.9
Official documents	-0.2	4.1	-8.4	8.7
Prepared speeches	0.4	4.1	-4.4	11.2
Spontaneous speeches	0.3	4.4	-5.5	7.4
Conversation	-0.3	2.4	-5.2	6.5

Note. SD = standard deviation.

The generally low persuasion scores are surprising because one of the purposes of UNGA

addresses would presumably be persuasion (see Section 4.3.1.5), but there are, in fact, a few

explanations for the low UNGA Dimension 4 scores. First, some texts are more informative than

persuasive. Reporting facts was one of the primary purposes of UNGA addresses (see Section

4.3.1.5), and this is reflected in some very factual texts or passages in texts such as Text 4.33

from Sierra Leone.

Text 4.33. Excerpt from Sierra Leone

> Since 2012, Sierra Leone proactively tailored its development framework to the 2030 Agenda for Sustainable Development. Our vision for socio-economic development as contained in my government's Poverty Reduction Strategy Paper (PRSP) III, and articulated in the "Agenda for Prosperity" (A4P), was launched in July 2013 as Sierra Leone's roadmap to the Post-2015 Development Agenda. The Agenda for Prosperity aims to build a sustainable future for all Sierra Leoneans. It demonstrates our firm commitment towards putting Sierra Leone on the path to resilience and sustainability.

Arguably, Text 4.33 is indirectly persuasive in its informational style. Countries that receive Official Development Assistance (ODA) need to persuade donor countries that they are putting the money to good use and making progress on economic and social development. This passage from Sierra Leone informs the Assembly of national policies in order to convince addressees of the country's commitment to sustainable development. (According to the Organisation for Economic Co-operation and Development, 2018, Sierra Leone received over US $900 million in ODA in 2015.)

Another reason for the low Dimension 4 scores is that many texts with low overt persuasion express opinion rather than advisability or likelihood, an important distinction noted in Biber (1988, p. 151). Several excerpts from the South African text (Dimension 4 score of -2.9) demonstrate expression of opinion (see Text 4.34).

Text 4.34. Excerpts from South Africa (features expressing opinion are underlined)

> Another critical matter that <u>needs</u> attention is the selection of the UN Secretary General….
> We <u>welcome</u> the recently signed Peace Agreement between the parties in South Sudan….
> We <u>acknowledge</u> the contribution of His Holiness Pope Francis to the Cuban-United States normalisation process.
> We <u>reiterate our support</u> of the people of Western Sahara….
> We <u>reiterate our call</u> for the lifting of the economic and financial embargo to help the Cuban people to gain their economic freedom.
> A review of this process is <u>necessary</u> to enable a more meaningful participation of the UN General Assembly in the process.

It is therefore, <u>critical</u> that the discussions of violent extremism and terrorism in parts of Africa and the Middle East, look into the root causes of the problem and not just the symptoms.

It is <u>unacceptable</u> and <u>unjustifiable</u> that more than one billion people in the African continent are still excluded as permanent members of the key decision making structure of the United Nations, the UN Security Council.

The failure of the Non-Proliferation Treaty Review Conference to reach an agreement in the year that marks 70 years since the first atomic bombs were detonated in Japan, is a major <u>setback</u> in our commitment to rid the world of weapons of mass destruction, and in particular nuclear weapons.

The outcome document represents a <u>victory</u> for developing world….

In the excerpts from the South African text, opinion is expressed through mental verbs (*needs*), non-suasive communication verbs (*welcome, acknowledge*), lexical phrases with suasive meaning (*reiterate our support, reiterate our call*), stance adjectives (*necessary, critical, unacceptable, unjustifiable*), and stance nouns (*setback, victory*). These excerpts show how expressing opinion can nonetheless be used to persuasive effect. *A review of this process is necessary to enable a more meaningful participation of the UN General Assembly in the process* could be rephrased with a modal of necessity such as *A review of this process must be undertaken to enable a more meaningful participation of the UN General Assembly in the process.* An important difference between overt persuasion and expression of opinion is degree of imposition. A modal of necessity expresses more imposition than a stance adjective. The fact that stance adjectives are preferred over modals of necessity in some texts is not surprising given the desire to maintain diplomacy: Countries may shy away from explicit directives so as not to appear overbearing. The use of these expressions of opinion are not necessarily less powerful, however. The evaluative adjectives *unacceptable* and *unjustifiable* are much stronger than many markers of overt persuasion such as *shouldn't*.

On average, UNGA addresses have low Dimension 4 scores but some addresses do express a great deal of overt persuasion. Text 4.35 from Nauru illustrates a persuasive passage from a text with a high Dimension 4 score of 6.9.

Text 4.35. Excerpt from Nauru (overt persuasion features italicized)

> Therefore, in the implementation of the 2030 Development Agenda, we *must* go beyond capacity building. We *must* engage in the difficult task of institution building. International and regional agencies and development partners *should* back these efforts with real resources that include long-term in-country engagement when appropriate, so that they leave behind durable domestic institutions run by a skilled national workforce. Nauru is *asking* for a lot from its development partners. We *want to* be better partners ourselves. And at the end of the day, implementation of the 2030 Development Agenda *will* mostly happen at the national and local level.

Text 4.35 from Nauru makes use of several linguistic features typical of persuasive texts. Advisability is expressed with modals of necessity (*international and regional agencies and development partners should back these efforts*) and infinitives (*We want to be better partners*). Likelihood is expressed with prediction modals (*implementation of the 2030 Development Agenda will mostly happen at the national and local level*).

The low Dimension 4 scores of UNGA addresses on the whole do not mean that there is no persuasion in these texts, only that fewer texts express persuasion overtly. Other texts are less direct in their persuasion, either conveying information or expressing opinion to persuasive effect, like in Text 4.33 and 4.34, respectively. The low frequency of features marking overt persuasion can be attributed to communicative purpose: reporting information and above all maintaining diplomacy.

4.5.5. Dimension 5: Abstract versus non-abstract information

The last dimension underlying variation, Dimension 5, represents the degree of abstraction, with high scores indicating very abstract, technical, and formal texts. Grammatical features frequent in abstract texts include conjuncts, passives, and past participial clauses (see Table 4.19). No features are associated with low Dimension 5 scores.

Table 4.19. Linguistic Features Associated with Abstract Information

Abstract Features (high Dimension 5)
conjuncts
agentless passives
past participial clauses
by passives
past participial WHIZ deletions
other adverbial subordinators

Biber (1988) found official documents to be the most abstract register. Official documents tend to be informational, conceptual, and abstract; human agents are often irrelevant or can be inferred so passive structures are particularly frequent. The other comparison registers are primarily non-abstract, with conversation the least abstract due to its interactiveness and focus on human agents and concrete topics. UNGA texts are, compared to conversation and the two speech registers, very abstract, closer to official documents for abstraction (see Table 4.20).

Table 4.20. Scores for Dimension 5: Abstract versus Non-Abstract Information

Register	Mean	SD	Minimum	Maximum
UNGA	1.7	1.8	-1.4	6.5
Official documents	4.7	2.4	0.6	9.4
Prepared speeches	-1.9	1.4	-3.9	1.0
Spontaneous speeches	-2.6	1.7	-4.5	0.7
Conversation	-3.2	1.1	-4.5	0.1

Note. SD = standard deviation.

It is noteworthy that UNGA addresses are so much more abstract than other prepared speeches in terms of degree of abstraction. This is likely due to three characteristics of UNGA addresses: (1) the focus on abstract principles, policies, and events, (2) the implication of consensus and joint action among member states, and (3) the preference for indirectness and saving face. All three of these characteristics result in the frequent use of passive structures, as shown in the text excerpts below.

The address delivered by the Holy See (Vatican) was a very abstract text with a Dimension 5 score of 5.5, and an excerpt from this address (see Text 4.36) displays a number of features inherent in abstract texts.

Text 4.36. Excerpt from the Holy See (abstract features in italics)

> The principle of a "responsibility to protect" all peoples from massive atrocities, instances of genocide, war crimes, ethnic cleansing and crimes against humanity *is* today *recognized* and *accepted* by all. This "responsibility", as *has been noted*, first obliges national governments, and then the international community or regional groupings of states, but always in accordance with international law. *Yet* it is not always easy to carry out this duty in practice, not least because its observance often conflicts with a strict literal interpretation of the principle of non-intervention as *sanctioned* by Paragraph 7 of Article 2 of the Charter of the United Nations; there is also the suspicion, historically *founded*, that under the guise of humanitarian intervention, the principle of the sovereign equality of the members of the United Nations Organization, *established* in Paragraph 2 of the same Article of the Charter, *is overridden*.

Text 4.36 includes abstract features such as *by*-passives (*is today recognized and accepted by all*), agentless passives (*has been noted*), conjuncts (*yet*), and past participle WHIZ deletions (*as sanctioned*). The text illustrates how these features add to the level of abstraction. Passives and past participial clauses put the emphasis on the patient, a concept such as a *principle* or a *suspicion*, rather than the agent. With these constructions, the agent is either understood, unknown, or unimportant. This focus on abstract concepts is common in the United

177

Nations because it is an organization whose work is based on ideals and principles: Peace, security, human rights, development are fundamental abstractions that are the very foundation for UN work. This is also the case for official documents, another register with abstract texts. Government documents and foundation or industry reports communicate conceptual information, data and trends. In contrast, the less abstract speech registers address, on average, more concrete matters: for a court case who did what during the accident, for a dinner speech who did what when the bride first met the groom. This is all the more true of conversation, the least abstract comparison register, where numerous references are made not just to events and places but also people known to both addressor and addressee. In non-abstract conversation, the agent is an important part of the information conveyed so passive voice is much less frequent.

In addition, passive structures are extremely frequent in the UNGA corpus because they can imply consensus and joint action among member states. No agent is named, suggesting first that the action is more important than the agent undertaking the action but also, in numerous instances, that the agent is the entire international community. In *the principle is today recognized and accepted by all* (Text 4.36), unanimity (*by all*) is explicitly stated. The text from Tuvalu is also explicit in including *one and all* as the agent: *These goals must be advocated as widely as possible to all global citizens, through technological advances that we have for information and any communication, so that these goals can be understood, owned and adhered to by one and all*. Even more frequently, the passive voice removes agency completely, indirectly implying unity in belief and action. For example, *At the same time, a greater focus should be put on addressing the root causes of increased illegal migration, including by strengthening the capabilities of prevention of conflicts* [Latvia] implies the entire international community should put greater focus on addressing the root causes. *Furthermore, the UN Security*

Council must be reformed to reflect current geo-political realities, through the commencement of text-based negotiations on the basis of UNGA Decision 69/560 which was adopted by consensus [Bhutan] suggests the entire UN must reform the Security Council. *Global challenges must be dealt with multilaterally, or they will not be dealt with at all. There is a duty on all of us to make the words "international community" actually mean something real* [European Union] indicates that the international community must work together on global challenges. In these passages, it is not necessary to name the agent, to indicate who must reform the Security Council or deal with global challenges, because the agent includes everyone present: speaker and audience.

This use of the passive again is possible in official documents. A report to shareholders might imply joint action taken by all those involved in the company. A government document can suggest the shared goals of the entire country. However, using the passive because addressor and addressee share agency is less likely in official documents compared to the UNGA because of the relationship between participants. For the UNGA, the speaker and the audience at the General Assembly are all high level politicians and diplomats who are present to advance their work together at the UN. In official documents, often the addressor and addressee have different roles (e.g., university and student for college catalogues). Prepared speeches, spontaneous speeches, and conversation are also less likely to imply shared agency between addressor and addressee, with the exception, perhaps, of some political speeches.

Finally, in the UNGA passives are a way to save face and avoid pointing fingers at member states (Texts 4.37-4.39). With abstract, passive constructions, the agent is implicit or unnamed so the "blame" is not overt. In *the principle of the sovereign equality of the members of the United Nations Organization… is overridden* (Text 4.36), no countries are explicitly named for having overridden the principle of equality. The same is true in Texts 4.37-4.39.

Text 4.37. Excerpt from Eritrea (abstract features in italics)

> Six decades ago we *were denied* of our inalienable right to self-determination and independence; for three decades, we *were savagely bombed* from the air and the ground, with the aim of crushing our liberation struggle; and today, we *are subjected* to unfair and illegitimate sanctions and baseless accusations.

In Text 4.37, the speaker from Eritrea avoids naming the parties responsible for denying the country and its people of their *inalienable right to self-determination,* for bombing them, or for subjecting them to *unfair and illegitimate sanctions.* Likewise, in Text 4.38, the speaker does not specify who was pointing guns.

Text 4.38. Excerpt from Albania (abstract features in italics)

> For the first time last year, after hundred [sic] years of frozen or live conflicts and confrontation in the Balkans, guns *were no longer pointed* against any neighboring window.

Text 4.39. Excerpt from South Sudan (abstract features in italics)

> It is thus, mandatory upon the government of South Sudan to remind the international donors about the many pledges they have made even earlier but *were never delivered.*

In Text 4.39, it is particularly interesting to note the active construction for the more positive act of making pledges, but the passive construction for failing to respect those promises.

In the UN, diplomacy dictates the need to avoid threatening face. This is unlike many other registers. In court cases and many national political speeches, direct accusations are the norm. In dinner speeches and popular lectures, the topics addressed generally preclude any need to raise points of contention, so no accusations are made, either direct or indirect.

These excerpts from UNGA texts show that the remarkably high Dimension 5 scores, in comparison with other speech registers and conversation in particular, are due primarily to the

use of the passive voice. The passive is frequent because it allows speakers to focus on an abstract concept and avoid agency. The situational characteristics of the UNGA help to explain this. First, UNGA texts are about abstract topics such as peace and development (see Section 4.4.1). Second, one of the communicative goals of the UNGA is to discuss UN policies and joint action to address these abstract issues; discussion of what needs to be done can often forgo reference to the agent because the agent is implied to be all present at the General Assembly: addressor and addressee. Finally, another communicative goal is to maintain diplomacy. Speakers can make accusations very indirect by avoiding agency. These situational characteristics are largely absent in the three spoken comparison registers, which explains at least in part why UNGA texts are so much more abstract.

4.5.6. Summary of dimension scores

Comparing the UNGA corpus to four other registers (prepared speeches, spontaneous speeches, conversation, and official documents) has revealed several interesting findings. First and foremost is the fact that UNGA addresses are largely dissimilar to spontaneous and even prepared speeches in terms of dimension scores. Overall, the co-occurrence patterns of UNGA grammatical features are more similar to official documents, meaning that on average the linguistic feature frequencies for UNGA addresses resemble a highly informational, abstract written register more closely than other texts that have been written to be spoken.

Dimension 1 scores revealed that UNGA addresses are much less involved than the other spoken or written-to-be-spoken registers and are almost as informational as official documents. One of the purposes of the UNGA is to report facts and the production circumstances allow speechwriters to carefully craft dense, information-packed prose. In contrast, UNGA speechwriters appear less concerned about producing a text that will be easy for the listener to

process when delivered orally. This may be due to the fact that UNGA speakers are encouraged to respect a 15-minute time limit so the speech must be concise; it may also be because UNGA addresses are distributed in writing so the audience does not necessarily have to listen to the oral delivery.

Dimension 2 scores showed UNGA addresses to be the least narrative register, even less narrative than official documents. The primary goals of the UNGA, to convey information about current policies and argue the importance of particular issues or agendas, lead to more expository rather than narrative texts. This focus on argumentation also reflects the persuasive style of many UNGA texts, emphasizing logical reasoning over personal directive.

Dimension 3 was the only dimension where the UNGA had similar scores to other speeches, with both explicit reference and situation-dependent reference. This is undoubtedly due to the fact that speeches are written (or at least reflected upon) in advance, but are delivered orally. For all speeches, there is a shared time and place, either real if the speech is live, or implied, if the speech is recorded. This allows for some situation-dependent reference (*here*, *now*). However, because the production circumstances, particularly for the UNGA and other prepared speeches, grant time for planning, elaborated endophoric reference is also possible. Thus, prepared speeches can easily show features of one or the other, and on average converge close to a neutral 0 for Dimension 3.

Dimension 4 scores indicated low overt persuasion. This does not mean that UNGA texts are not persuasive. Rather, in addition to the somewhat limited use of overt persuasion, UNGA addresses convey information and express opinion for persuasive purposes. This is done through the use of other lexico-grammatical features such as mental verbs, lexical phrases with suasive

meaning, and stance adjectives. Persuasion is also conveyed through argumentation and logic, as discussed in Section 4.5.2 on Dimension 2.

Finally, Dimension 5 scores showed the UNGA to be more abstract than all the comparison registers except for official documents. This is due in large part to the extensive use of the passive to avoid or displace agency. Avoiding agency is desirable when the agent is unimportant; the focus of a proposition is on an abstract concept. It is also possible when the agent is understood; both addressor and addressee know they are the implied agents. Finally, avoiding agency is preferable when the speaker does not want to name the agent and attribute direct blame for reasons of diplomacy.

Though the four comparison registers show some similarities with the UNGA, situational characteristics differ sufficiently to produce markedly higher or lower frequencies of grammatical features. As a result, the overall text type also differs. This finding is most surprising for the comparison with other prepared speeches, which share the most situational characteristics and most notably have very similar production circumstances. Clearly, participants, setting, purpose, and topic are different enough to see substantially divergent text type even in this closely related register. Additionally, the fact that UNGA addresses are published and can therefore be read rather than heard by the audience may also account for differences in linguistic features when compared to other prepared speeches.

4.6. Synthesis

The keyword and MD analyses in Chapter 4 have found the UNGA to be a distinct register in terms of lexical and grammatical forms. For lexical features, the comparison with the BNC, representing a wide range of registers in written and spoken British English, shows that UNGA addresses have an unusually high number of words related to peace and security,

refugees, human rights, development, climate change, democracy, and the functioning of the UN. The discourse structure, with salutations to the president and members of the Assembly, and communicative purpose, with a focus on cooperation, also mark the lexicon of the UNGA. Function keywords reveal a preference for the first person plural (*we* and *our*) as well as for the modal *must*, particularly with subjects including the speaker (e.g., *we, the international community*) or in the passive voice.

For grammatical features, the comparison with prepared speeches, spontaneous speeches, official documents, and conversation demonstrates that UNGA addresses are surprisingly different from other speech registers and in many instances UNGA addresses more closely resemble dense, abstract official documents. Overall, UNGA texts are more informational than involved, have an expository rather than narrative style, make use of both explicit and situation-dependent reference, use less direct persuasive techniques, and are highly abstract. UNGA dimension scores can be explained by the situational characteristics of the register: Production circumstances with texts written in advance allow for information-packed texts (low Dimension 1). Communicative purpose, with a focus on reporting and arguing a persuasive case, explains the lack of narration (low Dimension 2). Production circumstances, with texts written in advance but delivered before an audience, account for the neutral Dimension 3 scores (some explicit but also some situation-dependent reference). Communicative purpose, to inform and to maintain diplomacy, is the reason for low overt persuasion (low Dimension 4). Production circumstances (texts written in advance), topic (abstract concepts such as democracy) and communicative purpose (to report and maintain diplomacy) prompt abstract texts (high Dimension 5) with extensive use of the passive voice.

184

In addition, Chapter 4 has shown relatively little linguistic variation within the UNGA. The keyword analysis has revealed numerous words and semantic domains used in a large percentage of the texts, often over 75%. Moreover, many of the keywords are used in the same way, with similar associations and collocational patterns. The SD of dimension scores also indicates only limited variation within the UNGA. For Dimension 1, the SD was lower for the UNGA than all registers except for official documents. For Dimensions 2, 3, and 4, SDs were lower for the UNGA than any of the comparison registers. Only Dimension 5 showed a higher SD for the UNGA than the three other spoken registers.

Overall, the UNGA represents a relatively uniform register. The variation that does occur within the UNGA will be explored in Chapter 5.

CHAPTER 5: COMPARISON OF COUNTRY GROUPS WITHIN THE CORPUS OF UNITED NATIONS GENERAL ASSEMBLY ADDRESSES

5.1. Introduction

The purpose of this chapter is to examine linguistic variation within the UNGA corpus. Countries were categorized according to geographic, social, political, and economic factors, as described in Section 3.6. The linguistic features of each country group were then compared to the features in the remaining corpus. Section 5.2 describes variation in situational characteristics within the UNGA. Section 5.3 presents the results of the lexical feature analysis, Section 5.4 grammatical features, and Section 5.5 Multi-Dimensional scores. Section 5.6 contextualizes these findings with in-depth analyses of five countries, illustrating how lexical and grammatical features combine to create an overall discourse style in individual texts.

5.2. Variation in UNGA Situational Characteristics

Situational characteristics for the UNGA are relatively uniform (see Section 4.3). However, some variation can be attributed to country group. For instance, with the situational characteristic topic, a Small Island Developing State (SIDS) is more likely to talk about tropical storms and rising sea levels because, on average, these environmental issues affect SIDS more than other country groups. Section 5.2 briefly describes some variation in situational characteristics (participants, relationships among participants, channel of communication, production circumstances, setting, communicative purpose, and topic) based on country group.

Participant type is the same across the UNGA. As discussed in Section 4.3.1, participants can be viewed as the speechwriter and general public for the written text, as the speaker and General Assembly audience for the spoken text, or, symbolically, as the countries represented in the General Assembly. Because speechwriters and speakers are not communicating on their own

behalf but on behalf of their countries, the analysis of situational characteristics is based on the premise that the primary participants are the countries, UN member states or states with observer status (see Sections 4.3.1 and 4.3.1.1). The type of participants (speechwriter, speaker, but most importantly country represented) are the same across UNGA texts. Relationships between participants can vary by country group given their position within the international community. For example, the least developed countries that receive Official Development Assistance (ODA) have a different relationship to the Assembly than highly developed countries that provide that assistance. Recipient countries often report to the Assembly on development progress they have made, thank their development partners, or request additional support (see Text 4.33). Expression of gratitude and requests for further assistance reflect the relationship between the addressor recipient country and the addressee donor countries.

Channel of communication and production circumstances are the same for all texts. However, it should be noted that just because all UN member states are invited to deliver an address at the General Assembly does not mean that the addresses of all UN member states are listened to. As mentioned in Section 4.5.1, the texts are also made available in written form, and it is possible that the majority of the Assembly listens to addresses delivered by certain key players but only reads the addresses from others. Some UNGA texts are so dense that they sound as if they were written to be read rather than spoken. This may be due to the fact that speechwriters presume that their speeches will be primarily read rather than heard.

Setting is the same for all texts, but specific communicative purpose and topic can vary depending on country group. For example, SIDSs are particularly vulnerable to the effects of climate change. As a result, they are more likely to talk about climate change and try to persuade the Assembly of its importance to the international community. Countries in the low Human

Development Index (HDI) group are the least socially and economically developed and are therefore more likely to request development assistance or to defend their continued reliance on development assistance. Big military spenders tend to be involved or to anticipate potential involvement in armed conflict. They are therefore more likely to defend their position in a conflict, to criticize a party with whom they are disputing, or to express concern over a conflict if they are being asked by an ally to provide military or logistical support.

This overview of situational characteristics shows that while some characteristics (participants and setting) are the same across texts, others (specific topic and purpose) are highly variable. Still other situational characteristics are thought to be static (channel of communication and production circumstances) but might diverge in actual fact. The extent to which this variation in situational characteristics affects linguistic features is explored in Sections 5.3-5.6.

5.3. Variation in UNGA Lexical Features

5.3.1. Introduction

In order to examine systematic patterns of lexical variation within the UNGA, keyword analysis was carried out on groups of countries. Countries were classified based on 10 variables: region, area, population, HDI, Gross Domestic Product (GDP), Least Developed Country (LDC), SIDS, Landlocked Developing Country (LLDC), military spending, and religious non-affiliation for a total of 31 different country groups (see Section 3.6.1). Each group was compared to the remaining texts in the corpus. For example, to investigate lexical patterns for the low HDI group, the texts in the low HDI group were the target corpus and the texts for the medium, high, and very high HDI groups combined were the reference corpus. The keyword analyses identified which words appeared with significantly high frequency in each country group. That is, each keyword analysis showed what was lexically unusual about each country group.

For the keywords of a country group to be analyzed, at least one keyword on the list had to occur in 50% or more of the texts. Moreover, that keyword had to be "meaningful." A keyword was deemed "not meaningful" if it was a place name from within the group. Thus *Europe, European,* and *Ireland* would not be considered "meaningful" as keywords for the Europe group, but would be considered "meaningful" if they were identified in the Asian group. The objective of setting this distribution criteria was to prevent inaccurate generalizations from being made for an entire group due to a few idiosyncratic texts within the group. Details on the rationale and methods of identifying keywords and setting distribution criteria can be found in Section 3.7.1.

The vast majority of the category results did not meet this distribution criteria. Of the 31 country groups, only ten had "meaningful" keywords. It is remarkable that 21 country groups did not use any words with unusual frequency when compared to the other groups. This dearth of lexical variation can be explained in two ways. First, UNGA texts do not vary greatly in terms of lexical features. This explanation is supported by the keyword analysis for the UNGA as a whole (Section 4.4), which showed that overall, UNGA texts use many of the same terms. Second, the lexical variation that does occur within the UNGA may not be captured by the country categories. As discussed in Section 3.6.1, country categories are experimental; they were selected based on a review of the literature but no previous studies had revealed systematic linguistic variation based on country group.

Though no keywords were identified for most of the country categories, the keywords for ten country groups revealed sometimes striking differences that can be explained based on situational characteristics and the geographic, social, political, or economic situation that makes each group distinct. The findings for these groups are reported in Section 5.3. They include

189

Africa (Section 5.3.2), Europe (Section 5.3.3), SIDSs (Section 5.3.4), LDCs (Section 5.3.5), low military spending (Section 5.3.6), all four levels of GDP (Section 5.3.7), and very high HDI (Section 5.3.8). In each section, keyword frequency and distribution information is presented; then keywords are classified semantically. The keywords are discussed and illustrated using textual examples.

5.3.2. Africa

Africa was one of just two regions with meaningful keywords. As shown in Table 5.1, out of a total of 23 keywords, only four met the distribution criteria, occurring in at least 50% of the texts: *Africa, African, Heads,* and *Excellency.* Two of these were regional references, however (*Africa, African*), leaving only two meaningful keywords.

Table 5.1. Keywords in African Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
Mauritius	52	0.13	2	0		157.19	139.41	0.00
Africa	*76*	*0.19*	*19*	*35*	*0.02*	*108.81*	*2.94*	*0.00*
Somalia	36	0.09	5	5		80.91	4.67	0.00
Namibia	26	0.06	1	0		78.59	138.41	0.00
Kenya	30	0.07	1	2		76.72	5.73	0.00
Sierra	24	0.06	4	0		72.55	138.29	0.00
Botswana	23	0.06	1	0		69.52	138.23	0.00
African	*52*	*0.13*	*14*	*29*	*0.02*	*65.97*	*2.66*	*0.00*
Leone	21	0.05	4	0		63.48	138.10	0.00
Zambia	19	0.05	1	0		57.43	137.95	0.00
Archipelago	18	0.04	1	1		47.07	5.99	0.00
Chagos	14	0.03	1	0		42.32	137.51	0.00
Ghana	13	0.03	1	0		39.30	137.41	0.00
continent	22	0.05	9	9		33.64	3.11	0.00
Ebola	22	0.05	5	10		31.74	2.96	0.00
Africa's	10	0.02	7	0		30.23	137.03	0.00
Malawi	10	0.02	1	0		30.23	137.03	0.00
Heads	*20*	*0.05*	*16*	*9*		*29.02*	*2.97*	*0.00*
Ethiopia	11	0.03	4	1		26.87	5.28	0.00
Nigeria	12	0.03	2	2		25.79	4.41	0.00
Excellency	*32*	*0.08*	*11*	*31*	*0.02*	*24.86*	*1.87*	*0.00*
Tanzania	8	0.02	2	0		24.18	136.71	0.00
Eritrea	10	0.02	1	1		24.02	5.14	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, all texts except Africa: Americas, Asia, Europe, Oceania).

The full set of keywords were grouped into three semantic categories: places in Africa, formal titles, and other Africa-related (see Table 5.2).

Table 5.2. Keyword Categories for African Texts

Places in Africa	Formal Titles	Other Africa-related
Mauritius	*Heads* (e.g., of State)	Ebola
Africa	*Excellency*	
Somalia		
Namibia		
Kenya		
Sierra Leone		
Botswana		
African		
Zambia		
Chagos Archipelago		
Ghana		
(African) continent		
Africa's		
Malawi		
Ethiopia		
Nigeria		
Tanzania		
Eritrea		

Note. Keywords in at least 50% of texts are italicized.

The vast majority of the keywords fell into the category "places in Africa." Many of these keywords are self-reference (occurring only in the text by that country). These place names are not interpreted because they are not particularly meaningful (see discussion in Section 3.7.1). Nearly all speakers in the UNGA referred to their country and other countries in their region multiple times. This finding is revealing in terms of overall patterns in the UNGA: It shows a focus across texts to speak about places of geographical proximity. However, it does not help distinguish unusual lexical patterns for any country group, the objective of Section 5.3.

The only two keywords outside of the "places in Africa" category to appear in at least 50% of the texts were *Heads* and *Excellency*, two formal titles. Heads appeared solely in the phrases *Heads of State*(*s*), *Heads of Government*, and *Heads of Delegations*. In 17 out of 20

occurrences, it was used in the vocative at the beginning of the address following *Mr. President*.

Likewise, *Excellency* was used as a formal term of address. Both keywords are shown in Text

5.1 from Namibia.

Text 5.1. Excerpt from Namibia (keywords in italics)

> Your *Excellency* Mr. Mogens Lykketoft, President of the 70th Session of the United
> Nations General Assembly; Your Majesties; Your Excellencies *Heads* of State and
> Government; Your *Excellency* Mr. Ban Ki Moon, Secretary General of the United
> Nations; Distinguished *Heads* of Delegations; Ladies and Gentlemen; I am honoured to
> be here today, to deliver my maiden address to the United Nations General Assembly, as
> the 3rd President of the Republic of *Namibia*.

The text from Namibia illustrates a particularly long introduction, distinguishing by title

the president of the General Assembly, royalty, presidents and prime ministers, the Secretary

General of the UN, and the heads of delegations. Though it is not unusual to greet the Assembly,

the speaker from Namibia appears particularly careful to differentiate attendees based on position

and rank.

The remaining occurrences of *Heads* were not vocatives, but nonetheless referred to

groups of people by their formal titles as in: *However, almost no progress has been achieved on*

the commitment made by Heads of State and Government in 2005, to the early reform of the UN

Security Council [South Africa].

The texts from African countries, as with the texts from many other developing countries,

have an overall tendency to maintain the formal conventions of UN addresses and to show great

respect for people in positions of authority. The use of formal titles is one example; other

examples will be discussed in Sections 5.3.5, 5.3.7.1, and 5.4.2.

Ebola was the only other keyword to be identified outside of place names and formal

titles. It did not meet the distribution criteria, but it is noteworthy because it demonstrates the

focus speakers place on topics that affect their countries directly. Ebola is not surprising as a topic of interest among African nations in 2015 because the disease spread with unprecedented speed through west Africa primarily in 2014, but through 2015 as well (see Text 5.2).

Text 5.2. Excerpt from Sierra Leone (keywords in italics)

> At a time when *Sierra Leone* was being commended for its remarkable progress in peace stability and steady economic growth we were hit by the unprecedented *Ebola* Virus outbreak. The *Ebola* Virus disease outbreak has taken a heavy toll on the entire socio-economic fabric of *Sierra Leone*. But with support from our international friends, we fought back. Today, we have almost defeated the evil virus - only one case of *Ebola* was recorded in the country for the whole of August.

In the text from Sierra Leone, the Ebola virus becomes an opportunity to apprise the Assembly of progress made in the country and to show appreciation to *international friends* for their *support*. Discussing national progress and praising the international community were two specific purposes identified across UNGA texts in Chapter 4. Text 5.2 illustrates how these purposes can be achieved, even with a very specific topic such as Ebola.

The keyword analysis for Africa indicates that, aside from names of countries within the region, few terms are particularly frequent across texts. Only official terms of address are consistently frequent in African texts, suggesting a certain formality and an importance placed on titles and positions of power. *Ebola*, a disease that struck the African continent unlike any other region, also emerges as a keyword, though not in the majority of texts. It is interesting to note that even in the UNGA, an international body said to tackle the most critical issues affecting all nations and regions, this regional crisis was discussed significantly more frequently by representatives from the region itself. While other countries may have been sympathetic to the plight of Ebola victims and their governments, the issue was not one they necessarily thought to raise before the UNGA.

194

Europe was the only other geographical region with keywords that met the distribution criteria (see Table 5.3). Of the 23 keywords, nine occur in at least 50% of the texts. Three of these are regional place names (*Ukraine, European, Europe*), leaving six meaningful keywords.

Table 5.3. Keywords in European Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
Ukraine	57	0.12	17	10		103.07	4.02	0.00
European	59	0.12	17	23	0.02	75.44	2.87	0.00
Ireland	30	0.06	1	1		72.57	6.42	0.00
Europe	68	0.14	22	39	0.03	66.27	2.31	0.00
Romania	19	0.04	1	0		51.17	137.72	0.00
Croatia	19	0.04	1	0		51.17	137.72	0.00
Iceland	17	0.04	1	0		45.79	137.56	0.00
Moldova	17	0.04	4	0		45.79	137.56	0.00
Netherlands	16	0.03	1	0		43.09	137.47	0.00
crimes	45	0.10	15	27	0.02	42.20	2.25	0.00
Syria	62	0.13	20	54	0.04	39.26	1.71	0.00
Crimea	17	0.04	7	1		38.66	5.60	0.00
Russian	20	0.04	4	5		31.86	3.61	0.00
Malta	14	0.03	1	1		30.96	5.32	0.00
Greece	14	0.03	2	1		30.96	5.32	0.00
responsibility	56	0.12	18	55	0.04	30.09	1.53	0.00
situation	36	0.08	18	27	0.02	27.18	1.92	0.00
(San) Marino	10	0.02	1	0		26.93	136.79	0.00
Latvia	10	0.02	1	0		26.93	136.79	0.00
refugees	46	0.10	17	43	0.03	26.52	1.61	0.00
Ukrainian	12	0.03	5	1		25.87	5.09	0.00
migration	41	0.09	18	36	0.03	25.69	1.70	0.00
Donbas	9	0.02	1	0		24.24	136.64	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, all texts except Europe: Africa, Americas, Asia, Oceania).

All keywords were grouped into four semantic categories: places in Europe, places outside of Europe, refugees, and abstract nouns (see Table 5.4).

Table 5.4. Keyword Categories for European Texts

Places in Europe	Places outside of Europe	Refugees	Abstract nouns
Ukraine	*Syria*	*refugees*	*crimes*
European		*migration*	*responsibility*
Europe			*situation*
Ireland			
Romania			
Greece			
Croatia			
Iceland			
Netherlands			
Romania			
Latvia			
Iceland			
Malta			
(San) Marino			
Ukrainian			
Crimea			
Russian			
Donbas			
Moldova			

Note. Keywords in at least 50% of texts are italicized.

As with the texts from Africa, the category with the most numerous types grouped places in the region being investigated. Many of these were self-reference, country names that appeared only in the text itself (e.g., *Romania* in the Romanian text). As discussed in Section 3.7.1 and 5.3.2, these regional place names are not particularly meaningful and are therefore not discussed further.

In contrast, place names outside of the region are much less frequent, making the keyword *Syria* more noteworthy. An examination of concordance lines revealed that discussion on *Syria* revolves around violence and refugees. Just a few excerpts showing the association between *Syria*, conflict, and migration include: *As long as there is conflict in Syria, the refugee*

crisis will not go away [Bulgaria]; *Today the circle of proponents of the idea that Bashar al-Assad should be part of Syria's transition is growing. Yet we cannot forget that millions of people have fled his horrific methods of trying to secure stability in Syria* [European Union]; *We have before us, unfortunately, the immense tragedy of the war in Syria, with its thousands of dead, its millions displaced persons, and its tremendous consequences for stability in the region* [Holy See]. The unusually high frequency of *Syria* in European texts is not particularly surprising given that the violence there has had dramatic consequences in Europe, with the flood of refugees raising serious economic, social, and political issues.

The importance of the migration issue is reflected not just in the keyword *Syria*, but in two other keywords as well: *refugees* and *migration.* Though *refugees* and *migration* are referred to in association with *Syria* in numerous texts, several texts mention other countries as well (see Text 5.3) or speak of refugees globally (see Text 5.4).

Text 5.3. Excerpt from Austria (keywords in italics)

> We are facing a record high of over 50 violent conflicts, reaching from *Syria* and Iraq to Yemen, Sudan and Somalia. We are facing a rise of radicalisation and extremism. And we are facing the highest number of *refugees* since World War Two. Nearly 60 Million people worldwide are on the move as *refugees* or internally displaced persons. Alone 4 Million of them are Syrians who have left their country fleeing from civil war and terror. Every day over 40.000 people flee from their homes in search of peace, security or a new life.

The text from Austria mentions not just conflict in Syria, but also in other countries of the Middle East and Africa. In this passage, the speaker makes extensive reference to the number of refugees (e.g., *60 million people, 4 million of them, 40,000 people*). Other texts also refer to the high number of refugees: *We have all seen the figures - only this year some 500,000 refugees and migrants have crossed the Mediterranean to Europe and millions have fled to neighbouring*

197

countries [Iceland]; *One of the consequences is the massive refugee crisis we are seeing - with*

60 million refugees and displaced people [Norway]; *The number of refugees and displaced*

persons is growing at alarming speed [Denmark].

Interestingly, Austria is not the only European country to compare present-day refugee

numbers to those resulting from World War II: *Largely due to wars and conflicts, we are facing*

a new era of migration. Around sixty million refugees, largest number since the Second World

War, are a clear proof of that [Finland]; *It is a bitter irony that the seventieth anniversary of the*

United Nations Organization is accompanied by an exodus of peoples which is the greatest seen

since those caused by the Second World War [Holy See]. The comparison to World War II lends

an additional gravity to the discussion of the refugee crisis.

Other countries use negatively connoted words to underline the seriousness of the

migration issue. Examples of negative evaluation include: *the suffering of the refugees* [Iceland];

the present refugee situation in the Middle East, in many parts of Africa, and in the

Mediterranean is unbearable for all [Finland]; and *many refugees have been in a desperate*

situation for too long [Netherlands]. In Text 5.4 from Slovakia, negative words or phrases

include *illegal, gross human rights violations,* and *tragedies.*

Text 5.4. Excerpt from Slovakia (keywords in italics)

> Both immediate as well as long-term measures are absolutely necessary to address these
> challenges, which include illegal *migration*, smuggling and trafficking of human beings
> and gross human rights violations... We must offer our assistance to those in need and
> prevent further tragedies involving *refugees*.

The text from Slovakia refers to another side of the migration issue: human trafficking.

Other texts mention trafficking as well: *In the framework of the UN, we have to build the*

necessary resettlement mechanism from countries neighbouring Syria, while also supporting

198

them directly in hosting refugees and dismantling trafficking networks [Greece]; *We need to create the possibility for refugees to apply for asylum already in their countries of origin or neighbouring countries. The United Nations could help in the establishment of such reception and information centres which would allow to introduce resettlement programs. In my opinion trafficking in human beings is one of the worst crimes* [Austria].

The European focus on refugees can be explained in part by the migration crisis and the record number of new asylum applications received in Europe, particularly from Syria (United Nations High Commissioner for Refugees, 2016). The sudden influx of refugees has had serious political, social, and economic ramifications in Europe. It is therefore not surprising that European states would refer most often to Syrian refugees, emphasize the *alarming* numbers, liken the crisis to the refugee situation during and after World War II, and use negative evaluation, speaking of the *tragedies* of migration or the *crimes* of human trafficking.

However, the association between *refugees* and the conflicts that lead to migration is also part of a larger concern among European countries about violence and the responsibility of UN member states to protect civilians in conflict zones. Three other abstract keywords also reflect these concerns: *crimes, responsibility,* and *situation.*

These keywords tend to occur in the same passages about the *responsibility* of the international community to protect civilians in *situations* where the most heinous *crimes* are being perpetrated. In these passages, *crimes* (45 occurrences) are nearly always specific references to the most serious types of crimes, under the jurisdiction of the International Criminal Court: *crimes against humanity* (19 occurrences), *war crimes* (10 occurrences), and *atrocity crimes* (9 occurrences). *Situation* is often used euphemistically to mean a conflict situation or violent situation: *We call on all Security Council members to, in the words of the*

Charter, "unite their strength" to halt the bloodshed in Syria. The situation is already having consequences far beyond the region [Iceland]; *We call upon the Security Council, and especially its permanent members, to overcome divisions and find the way to deal more effectively with this worsening situation, in line with its primary responsibility for the maintenance of international peace and security* [Slovenia]; *While the situation in the Middle East region gives little cause for optimism, there has recently been a triumph for diplomacy over conflict* [Iceland]; *It is no coincidence that the UN General Assembly debate this year focuses mainly on the situation in the Middle East, especially in Syria* [EU]; *We are witnessing a fragile peace and security situation in many parts of the world* [Croatia]. The connection between the words *responsibility, situation,* and *crimes* is implicit in many of these excerpts (from Iceland and Slovenia in particular). Sometimes the relationship is made explicit as well (see Text 5.5).

Text 5.5. Excerpt from Romania (keywords in italics)

> At this juncture, let me reiterate Romania's support for the initiative of France and Mexico to propose a collective and voluntary agreement among the permanent members of the Security Council regarding the non-use of veto when action is needed to prevent, or bring to an end, *situations* of mass atrocities and war *crimes* at a large scale. Ten years after the endorsement of the *responsibility* to protect, we should continue to support it by identifying best ways to understand, to implement and to operationalize it.

In Text 5.5, the speaker from Romania refers to the "responsibility to protect." The "responsibility to protect" is a commitment made by the international community in adopting the 2005 World Summit Outcome Document, which states that all countries have the "responsibility to protect" their own citizens from genocide, war crimes, ethnic cleansing and crimes against humanity and to encourage other countries to do so. The "responsibility to protect" is the grounds for the French-Mexican initiative calling on Security Council Permanent Members to refrain from exercising their right to veto for *situations of mass atrocities and war crimes*, also

200

mentioned in Text 5.5. The reasoning is that the international community has a responsibility to protect civilians, one way of protecting civilians is by maintaining peace and security through the Security Council, but the Security Council is often paralyzed because one or two Permanent Members veto a resolution. It follows that any resolution on *crimes against humanity, war crimes,* and *atrocity crimes* should not be vetoed because the costs for humanity are too high.

A number of texts refer to this initiative put forward by France and Mexico and the related Accountability, Coherence and Transparency (ACT) Code of Conduct, asking Security Council members not to vote against any resolution in cases of genocide, crimes against humanity, and war crimes. Again, these passages show the relationship between the words *crimes, responsibility,* and *situation: The permanent members of the Security Council have a particular responsibility. Norway urges all States to join the proposed code of conduct to enable the Security Council to act decisively against mass atrocities. We support the French initiative to suspend the use of veto in such situations* [Norway]; *Meanwhile, Iceland has stated its support for the initiative by France and Mexico on regulating the veto and for the Code of Conduct on Security Council action against genocide and other crimes against humanity, drafted by the ACT group of states* [Iceland]; *Therefore, we support the initiatives of Accountability, Coherence and Transparency Group and of France and Mexico on the non-use of veto in cases of genocide, crimes against humanity, war crimes or other atrocity crimes* [Estonia]; *We strongly endorse the Code of Conduct on Security Council action on mass atrocity crimes developed by Liechtenstein and the ACT group, and we support the Declaration by France and Mexico on regulating the use of the veto* [Ireland]; *Latvia supports the proposal to voluntarily restrain the use of the veto at the Security Council in situations involving mass atrocity crimes. It also supports a Code of*

Conduct for any member of the Council not to vote against any action designed to end and prevent mass atrocity crimes [Latvia].

Overall, the keywords for European texts reveal a focus on conflict zones (e.g., *Syria*) and the responsibility of the international community to protect civilians in or from conflict zones (e.g., *refugees*). The focus among European countries on refugees can be explained by the political, economic, and social consequences of the refugee crisis in Europe. However, the other keywords and most importantly the ways the keywords are used together in the same passages suggest that the refugee issue is part of a broader concern about conflict and protecting civilians. European countries may be particularly disturbed about war crimes because of their history and experience during World War II. Text 5.3 from Austria, as well as the texts from Finland and the Holy See, compare refugee numbers during the Second World War with those in 2015. Similarly, the failure of the international community in their "responsibility to protect" civilians before and during the Second World War could be compared to the failure to protect civilians in Syria and a number of other conflict zones in 2015. Though World War II was tragic for countries around the world, civilians in Europe suffered greatly because neighboring states failed to respond in a timely manner. European countries may be particularly cognizant of the "responsibility to protect" because that responsibility was not met during World War II, with disastrous consequences.

5.3.4. Small Island Developing States

The analysis of texts from the 26 Small Island Developing States (SIDSs, e.g., Saint Lucia, Papua New Guinea, the Seychelles, the Maldives) identified 49 keywords, nine of which occurred in at least 50% of the texts, as shown in Table 5.5.

Table 5.5. Keywords in Small Island Developing State Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	P
islands	*113*	*0.21*	*14*	*3*		*249.64*	*6.49*	*0.00*
small	137	0.25	25	49	0.04	153.72	2.74	0.00
climate	223	0.41	25	144	0.11	152.77	1.88	0.00
island	87	0.16	24	15	0.01	137.35	3.79	0.00
SIDS	53	0.10	17	0		129.16	139.01	0.00
Mauritius	52	0.10	2	0		126.73	138.98	0.00
Pacific	69	0.13	12	8		122.40	4.36	0.00
Solomon	47	0.09	3	1		105.52	6.81	0.00
Caribbean	46	0.09	11	2		96.88	5.78	0.00
change	*181*	*0.34*	*25*	*150*	*0.12*	*90.30*	*1.52*	*0.00*
Guyana	31	0.06	7	0		75.55	138.24	0.00
Tobago	29	0.05	3	0		70.67	138.14	0.00
Trinidad	29	0.05	3	0		70.67	138.14	0.00
Fiji	27	0.05	5	0		65.80	138.04	0.00
Venezuela	27	0.05	7	0		65.80	138.04	0.00
developing	*84*	*0.16*	*25*	*48*	*0.04*	*65.31*	*2.06*	*0.00*
Saint	26	0.05	3	0		63.36	137.98	0.00
Bahamas	25	0.05	3	0		60.93	137.93	0.00
Papua	21	0.04	2	0		51.18	137.68	0.00
Tuvalu	20	0.04	4	0		48.74	137.61	0.00
Archipelago	19	0.04	2	0		46.30	137.53	0.00
Marshall	19	0.04	5	0		46.30	137.53	0.00
seas	22	0.04	11	1		46.09	5.71	0.00
ocean	*33*	*0.06*	*17*	*8*		*45.56*	*3.30*	*0.00*
Jamaica	18	0.03	4	0		43.87	137.45	0.00
lying	18	0.03	10	0		43.87	137.45	0.00
Barbados	18	0.03	2	0		43.87	137.45	0.00
Nevis	17	0.03	1	0		41.43	137.37	0.00
Dominica	17	0.03	9	0		41.43	137.37	0.00
CARICOM	17	0.03	5	0		41.43	137.37	0.00
Grenada	17	0.03	1	0		41.43	137.37	0.00
oceans	29	0.05	12	7		40.11	3.30	0.00
Taiwan	19	0.04	8	1		39.06	5.50	0.00
Cuba	*40*	*0.07*	*16*	*18*	*0.01*	*38.25*	*2.40*	*0.00*
Belize	15	0.03	2	0		36.56	137.19	0.00
Samoa	15	0.03	8	0		36.56	137.19	0.00
Tonga	15	0.03	2	0		36.56	137.19	0.00
Lucia	15	0.03	1	0		36.56	137.19	0.00

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	P
Chagos	14	0.03	1	0		34.12	137.09	0.00
Kitts	14	0.03	1	0		34.12	137.09	0.00
Micronesia	14	0.03	3	0		34.12	137.09	0.00
impacts	27	0.05	12	9		31.62	2.84	0.00
Vanuatu	12	0.02	6	0		29.24	136.87	0.00
Maldives	12	0.02	3	0		29.24	136.87	0.00
Grenadines	11	0.02	1	0		26.81	136.74	0.00
coastal	13	0.02	9	1		25.18	4.95	0.00
pathway	13	0.02	9	1		25.18	4.95	0.00
Kiribati	10	0.02	3	0		24.37	136.61	0.00
Federated	10	0.02	3	0		24.37	136.61	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, all texts except SIDS).

The complete list of keywords were grouped into the categories SIDS places, places outside of SIDS, other SIDS-related, the environment, and "other" (see Table 5.6).

Table 5.6. Keyword Categories for Small Island Developing States

SIDS places	Places outside of SIDS	SIDS-related (other)	Environment	Other
island(s)	Taiwan	CARICOM	*climate*	impacts
Small (i.e., Small Island Developing States)			*change*	
SIDS			*ocean*(s)	
developing (i.e., Small Island Developing States)			seas	
Cuba			(low-)lying	
Trinidad (and) Tobago			coastal	
Fiji				
Papua (New Guinea)				
Jamaica				
Barbados				
Caribbean				
Belize				
Tonga				
Federated (States of) Micronesia				
Maldives				
Mauritius				
Pacific				
Samoa Pathway				
(Saint) Kitts (and) Nevis				
Grenadines				
Chagos Archipelago				
(Saint) Lucia				
Grenada				
Guyana				
Venezuela				
Vanuatu				
Tuvalu				
Kiribati				
Solomon Islands				
Bahamas				
Marshall Islands				
Dominica				

Note. Keywords in at least 50% of texts are italicized.

As with the keyword analyses for Africa and Europe, the majority of SIDS keywords were names of SIDS countries (e.g., *Cuba*), regions (e.g., *Caribbean*), or oceans (e.g., *Pacific*). One keyword was a place name outside of SIDS (*Taiwan*), but interestingly, it related to SIDS, as it was mentioned primarily in recognition of Taiwanese financial assistance to SIDS (see Text 5.6).

Text 5.6. Excerpt from Kiribati (keywords in italics)

> We welcome the inclusive approach taken in developing the Post 2015 Development Agenda. We also welcome the inclusion of *Taiwan* in international processes of the World Health Assembly, in the fight against Ebola, and we would like to see similar inclusive approach prevail in respect of other international institutions and UN processes, in implementing the SDGs, in the call for urgent *climate* action, where *Taiwan* and all who can participate and contribute meaningfully for the good of humanity must be brought in to do so (….) We call for new and accessible financial resources to assist the most vulnerable to adapt and build resilience to *climate change*. We welcome the continued assistance of our partners, including *Taiwan*, but much much more needs to be done.

Text 5.6 from Kiribati shows how developing states often give thanks to other countries for support. In return, countries receiving developmental assistance tend to give their political assistance. In this text, for example, the speaker from Kiribati *welcomes the inclusion of Taiwan* in the World Health Assembly. For many years, Taiwan has sought recognition as an independent member of organizations but has generally been opposed by China. The support of other countries sometimes helps create enough political pressure to give Taiwan increasing recognition and power within international organizations. While this passages appears at first to be related solely to Taiwan and its position in the World Health Assembly, it becomes apparent later in the text that Taiwan is mentioned because it has provided economic assistance to Kiribati and other SIDS.

A passage from Nauru also shows the connection between developmental aid from Taiwan and SIDS political support: *All partnerships - be they multilateral, bilateral, or public-private - must be based on mutual respect - a respect that recognizes genuine development partners such as the Republic of China (Taiwan), a democracy that can contribute meaningfully towards development and prosperity in the world. Needless to say, Taiwanese technology and know-how can make a lot of difference towards helping the needy, and as such, I call for the inclusion of Taiwan in the development agencies of the UN*. The speaker from Nauru clearly does not want to completely estrange China, however, and has referred to Taiwan as *the Republic of China (Taiwan).* This shows the careful wording required to support developmental partners but at the same time maintain positive relationships with other countries. Additional calls for support of Taiwan's increased presence in international organizations include: *As such, we must move beyond our inexplicable exclusion of Taiwan from the work of the specialized agencies of the United Nations. The perspective, experience and example of Taiwan as an active and responsible global citizen are self-evident arguments in support of their greater inclusion and participation* [Saint Vincent and the Grenadines] and *We continue to call for the United Nations to engage with Taiwan whose 23 million people stand ready to contribute, especially through UN Specialized Agencies in resolving the world's problems and who have every expectation that their interests should be heard and represented* [Belize].

One other keyword was related to SIDS as a geographical group: *CARICOM* or the Caribbean Community, an organization of 15 Caribbean countries. An analysis of concordance lines showed that *CARICOM* was typically used to identify other countries of the Caribbean. Examples include: *We urge continued support for the people of Dominica, our sister CARICOM State, who suffered enormous damage and loss of life from Tropical Storm Erika in August*

[Bahamas] and *It was with especial pleasure that we received the news earlier this year that, forty three years after Barbados and three other CARICOM states, Trinidad and Tobago, Jamaica, and Guyana, had ended their diplomatic isolation of Cuba, the United States of America has decided to do the same* [Barbados]. Occasionally, *CARICOM* was used to refer to the region: *Indeed, my own country is on track to achieve zero mother-to-child transmission by the end of 2015 and we are committed to the CARICOM region becoming the first in the world to END AIDS by 2030* [Saint Kitts and Nevis]. The interpretation of *CARICOM* as a keyword is thus similar to other SIDS places: For all country groups, keywords are frequently place names in the region or regions represented. The presence of place name keywords in any one group does not indicate a unique pattern particular to that country group.

The most interesting keyword category is the environment. These terms relate to climate change, the resulting precariousness of living on the *low-lying* islands of SIDS, and the need to protect the *oceans* and *seas.* Many of the passages on the environment speak of disasters such as hurricanes in order to demonstrate the need for climate change legislation or financial assistance to countries that have been struck by natural disasters (Text 5.7).

Text 5.7. Excerpt from Jamaica (keywords in italics)

> The challenge posed by *climate change* is an existential one for the peoples of *Small Island Developing States (SIDS)*, whose lives and livelihoods too often hang in the balance. We therefore look toward the continued support of the international community for *SIDS*. (….) *Jamaica* joins the appeal to the international community to rally in support of the people of *Dominica*, by assisting, in the recovery, reconstruction and humanitarian efforts that are required. It is an unfortunate fact that such events will certainly occur again. In fact, another member of our *Caribbean* family, The *Bahamas*, is now being pummeled by Hurricane Joaquin. In spite of this fact however, the international community can assure the impact of the next natural disaster on our *small island* and low *lying coastal* countries is lessened by helping us to improve our resilience in the face of these every increasing calamities.

In Text 5.7, the speaker informs the audience of disaster relief efforts in Dominica and of the havoc being wrecked on the Bahamas by Hurricane Joaquin to persuade listeners to continue to support legislation on climate change and to provide assistance for the devastation that occurs on SIDS due to tropical storms. The text does not narrate past events (see Section 4.5.2 on narration), but instead focuses on the present: the current reconstruction of Dominica rather than the past destruction of Tropical Storm Erika that required it. This informative passage is used to persuasive effect without any overt persuasion (*we therefore look toward the continued support*; *the international community can assure the impact… is lessened*). This argumentation style is discussed further in Section 5.5.5. Other examples of texts that refer to environmental disasters to bolster support for climate change policy or for financial assistance include Texts 5.8 and 5.9.

Text 5.8. Excerpt from Trinidad and Tobago (keywords in italics)

> We recently witnessed the destructive effects of tropical storm Erika on our fellow *CARICOM* Member State, the Commonwealth of *Dominica*, which has set back that country's development by many years. It is undeniable, therefore, that *Climate Change* and development are closely inter-linked. Recognizing the need for bold and responsible action on the part of all Governments, albeit on the basis of common but differentiated responsibility, *Trinidad* and *Tobago* has adopted a proactive posture on this issue and has submitted its Intended Nationally Determined Contributions to the United Nations Framework Convention on *Climate Change* (UNFCCC), ahead of the 2015 Paris *Climate* Conference.

Text 5.9. Excerpt from Antigua and Barbuda (keywords in italics)

> I express Antigua and Barbuda's deep concern at the devastation of our neighbouring *Caribbean* island, *Dominica*, by tropical storm Erika. I also want to make the point Mr President, that countries in the *Caribbean* did not await an international response, before we rushed to the aid of *Dominica* and its suffering people. Within the limited capacity of every member-state of the *Caribbean* Community, we made immediate financial and other tangible contributions to *Dominica*. Despite our own challenges, we stood-up for the humanitarian needs of our neighbor. But, even as we did so, we were aware, that wider international support is necessary.

The last keyword category ("other") was found to relate to the environment as well. Concordance lines showed that *impacts* was nearly always associated with climate change. Examples include: *The impacts of climate change are already being felt the world over and without a successful legally binding agreement in Paris, we will see climate change continue unabated* [Grenada]; *The sea level rise, the impacts of El Nino, and unfavourable weather patterns points to a bleak future for humanity* [Vanuatu]; *Leaders were deeply concerned about the serious effects and impacts of climate change, particularly on national and regional development and security* [Papua New Guinea].

Not only does nearly every SIDS text (25 out of 26) refer to climate change, several of them speak of little else (see Text 5.10).

Text 5.10. Excerpt from Samoa (keywords in italics)

> I am therefore acutely aware of the diversity and gravity of issues confronting our organization and the need sometimes to refer to them in one's statement. However, partly in recognition of the new agenda we have now agreed to implement, and as a gesture to the 70th anniversary of our organization and the imminent danger of *Climate Change* for *Small Island Developing* States (*SIDS*) my statement in essence is devoted to the Sustainable Development Goal Number 13, "**Take urgent action to combat *climate change* and its *impacts***" [emphasis in the original].

The focus on climate change and environmental disaster in SIDS texts can be explained by the geographical and developmental position of SIDS, a group of countries particularly vulnerable to the effects of climate change and with limited financial and human resources to undertake reconstruction after storms.

5.3.5. Least Developed Countries

The analysis of texts from the 18 Least Developed Countries (LDC, e.g., Afghanistan, Kiribati, Somalia, South Sudan) identified 16 keywords, two of which occurred in at least 50% of the texts (*Excellency* and *SDGs*), as shown in Table 5.7.

Table 5.7. Keywords in Least Developed Countries Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*P*
Solomon	46	0.15	2	2		145.04	6.77	0.00
Myanmar	28	0.09	1	0		97.94	138.86	0.00
islands	63	0.20	4	53	0.04	80.68	2.50	0.00
Zambia	19	0.06	1	0		66.46	138.30	0.00
Sierra	22	0.07	2	2		63.95	5.71	0.00
Leone	19	0.06	2	2		54.02	5.50	0.00
Nepal	16	0.05	1	1		48.74	6.25	0.00
Tuvalu	17	0.05	2	3		43.70	4.75	0.00
Cambodia	14	0.04	1	2		37.68	5.05	0.00
Excellency	32	0.10	12	31	0.02	36.46	2.29	0.00
Malawi	10	0.03	1	0		34.98	137.37	0.00
SDGs	25	0.08	9	23	0.02	29.78	2.37	0.00
Eritrea	10	0.03	1	1		28.66	5.57	0.00
Kiribati	9	0.03	2	1		25.36	5.42	0.00
Vanuatu	10	0.03	4	2		24.93	4.57	0.00
MSG	7	0.02	1	0		24.49	136.86	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, all texts except LDCs).

All the keywords were grouped into four categories: places in LDCs, other LDC reference, formal titles, and development (see Table 5.8).

Table 5.8. Keyword Categories for Least Developed Countries

Places in LDCs	Other LDC place reference	Formal Titles	Development
Solomon	MSG (Melanesian Spearhead Group)	*Excellency*	*SDGs*
Islands			
Sierra Leone			
Tuvalu			
Kiribati			
Vanuatu			
Myanmar			
Zambia			
Nepal			
Cambodia			
Malawi			
Eritrea			

Note. Keywords in at least 50% of texts are italicized.

As with the other country groups, the majority of the keywords were place names related to the country group itself (LDCs). There was another keyword related to LDC places as well, the *MSG* (Melanesian Spearhead Group), not a place in and of itself, but an organization in which the Solomon Islands (the sole text in the group to use the term) is a member.

The only meaningful keywords are *Excellency* and *SDGs*, also the only two keywords to occur in at least half the texts. The keyword *Excellency* was identified for the African texts in Section 5.3.2 as well. However, it is not only the African texts in the LDC group that use this term. It occurs in 12 texts, six of which are not African countries: Cambodia, Kiribati, Myanmar, Nepal, the Solomon Islands, and Vanuatu. Like with the African texts, *Excellency* is used as a term of address and is part of a highly formalized introduction, greeting the Assembly and thanking officials for their service, as shown in Text 5.11.

Text 5.11. Excerpt from Vanuatu (keywords in italics)

> Your *Excellency*, Mr. Morgens Lykketoft, President of the Seventieth General Assembly of the United Nations; Your *Excellency* Mro Ban Ki-moon, Secretary General of the United Nations; Your Excellencies Heads of States, Governments and Delegations; Ladies and Gentlemen, Mr. President, Vanuatu is honoured to join other esteemed speakers who have spoken before me, to congratulate you for assuming the Presidency of the 70th UN General Assembly. Be rest assured of my delegation's full support. With your leadership, we are confident that you will guide us successfully through this session. I also wish to extend my warm gratitude to the last President of the assembly, His *Excellency* Sam Kutesa. His exemplary leadership has resulted in many important achievements which will continue to shape and strengthen the UN and enable it to execute its mandate effectively.

The keyword *Excellency* in LDC texts supports the findings from Section 5.3.2 indicating that developing countries on average tend to use more formal lexical features in their addresses. More will be said on formality and level of development in Sections 5.3.7.1 and 5.4.2.

The only other "meaningful" keyword for the LDCs relates to development: *SDGs*. The 2015 UN Sustainable Development Goals (SDGs) promote economic, social, and environmental well-being around the world (e.g., "no poverty," "gender equality," and "life below water"). It is not surprising that development goals would be a common topic in addresses delivered by representatives from the least developed countries of the world (see Texts 5.12-5.14)

Text 5.12. Excerpt from the Solomon Islands (keywords in italics)

> Our appreciation goes to your predecessor His *Excellency* Sam Kutesa for his strong leadership on this historic development framework - "Transforming our World: 2030 Agenda for Sustainable Development" - which this Assembly adopted last week. For developing countries, it means translating the Sustainable Development Goals (*SDGs*) into action. This can only happen if the *SDGs* are matched with resources and the Addis Ababa Action Agenda, to deliver on the needed means of implementation, and to convert the goals into action. Humanity's survival depends on the world coming together on our shared 2030 Agenda. It offers us the last hope of getting it right. Incremental action is not an option. Transformative change only comes about if we think big, smart and innovative to achieve the 17 global *SDGs* and the 169 targets.

Text 5.12 from the Solomon Islands takes a factual, informative approach. The passage begins with thanks to the President of the General Assembly's 69th session for his work on development. Then it enumerates a series of opinions on the SDGs that are presented as fact: *it means, can only happen if, humanity's survival depends on, is not an option, only comes about if*). The passage is highly persuasive, even if not explicit in its attempt to persuade, because it presents opinion as truth. (See Sections 4.5.2, 4.5.4, and 5.5.5 for more on persuasion.)

Text 5.13 from Myanmar is somewhat more overt in its persuasion with a direct *call for* lifting economic sanctions in order to allow developing countries to implement the SDGs. It is not surprising that Myanmar uses the SDGs as an opportunity to request the lifting of sanctions because the country has had numerous sanctions imposed on it.

Text 5.13. Excerpt from Myanmar (keywords in italics)

> Political commitment, enhanced partnership, strengthening of capacity and provision of adequate means of implementation will be keys to success in realizing ambitious *SDGs*. Unilateral economic sanctions imposed on developing countries will have a negative impact on their efforts for the people. It will also hinder in realizing the *SDGs*. We, therefore, call for early lifting of such measures.

Text 5.14 from Uganda also raises the issue of SDGs, but not with the intention to persuade. Instead, it is an opportunity to implicitly criticize. The SDGs are presented in a positive light, referred to as *enlightened self-interest*, but then the speaker laments the fact that it took so long for development to come to the fore as a key issue.

Text 5.14. Excerpt from Uganda (keywords in italics)

> Above all, the *SDGs* proclaim in bold letters the concept of universal prosperity by all societies for the first time in human history. While it is amazing that this enlightened self-interest has taken so long to dawn on all of us, the old saying that better late than never appropriately comes to mind in this case.

The keyword analysis for LDCs has shown that, like geographical region, level of development can influence the lexical features of a text. LDCs speak with unusual frequency about the Sustainable Development Goals because development is a topic of vital interest to them.

5.3.6. Low military spending

No keywords met the 50% dispersion criterion for the 47 high military spending countries, but the analysis of texts from the countries with low military spending or 0-1.9% GDP (e.g., Denmark, Ghana, Indonesia, Tonga) identified three keywords, two of which occurred in at least 50% of the texts (*climate, small*), as shown in Table 5.9.

Table 5.9. Keywords in Low Military Spending Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*P*
climate	*138*	*0.16*	*35*	*22*	*0.05*	*27.62*	*1.54*	*0.00*
small	*74*	*0.08*	*27*	*6*	*0.01*	*27.57*	*2.52*	*0.00*
Kenya	32	0.04	3	0		24.39	137.56	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, high military spending texts).

The three keywords were grouped into the categories low military spending places, the environment, and "other" (see Table 5.10).

Table 5.10. Keyword Categories for Low Military Spending States

Low military places	Environment	Other
Kenya	*climate*	*small*

Note. Keywords in at least 50% of texts are italicized.

As with the other groups, one keyword category relates to places in the low military spending group and is not particularly meaningful.

215

More interesting is the fact that countries with low military spending speak more about climate change. A qualitative examination did not reveal any textual evidence for a direct relationship between the two; military matters (e.g., nuclear proliferation) were never mentioned in the same context as environmental protection. It is possible that countries spending less on defense are more concerned about the environment, but without any textual support, it is impossible to suggest a direct causal relationship.

An investigation of *small,* on the other hand, did reveal some patterns. For the low military spenders, of the 74 instances of *small*, 32 co-occur with *states,* 10 with *weapons*, 10 with *arms,* and 9 with *large.* In contrast, the 18 texts from the biggest military spenders use *small* only six times: four times with *states* or *country* and once in the phrase *big and small.* Though it might initially seem unexpected for countries spending the least on defense to be talking about weapons the most, an examination of texts reveals that *weapons* and *arms* are usually in sentences with negative evaluation and the need to halt the proliferation of weapons, as shown in Texts 5.15 and 5.16.

Text 5.15. Excerpt from the Philippines (keywords in italics; collocates underlined)

> On conventional weapons, the Philippines will help sustain the positive momentum created by the Arms Trade Treaty's (ATT) progress, and continue to engage in discussions on *small* <u>arms</u> and light <u>weapons</u> and improvised explosive devices, particularly as used by armed non-state actors.

Text 5.16. Excerpt from Nigeria (keywords in italics; collocates underlined)

> We are witnessing a dreadful increase in conflicts fuelled by availability of *small* <u>arms</u> and light <u>weapons</u>. I call upon all member countries to demonstrate the political will needed to uphold the UN charter. For a start, a robust implementation of the Arms Trade Treaty will guarantee that *small* <u>arms</u> and light <u>weapons</u> are only legally transferred. Arms traffickers and human traffickers are two evil species which the world community should eradicate.

216

Both the Philippines and Nigeria speak of the Arms Trade Treaty, which regulates international trade in conventional weapons. It is not surprising that countries spending less on defense would be more concerned with the trade and proliferation of weapons seeing as their governments have fewer military resources to combat either state or non-state actors who have access to weapons (e.g., Boko Haram in Nigeria).

5.3.7. Gross Domestic Product

All four levels of Gross Domestic Product had at least one meaningful keyword that met the distribution criteria. Thus, all four levels are interpreted in Section 5.3.7.

5.3.7.1. Low Gross Domestic Product

The analysis of texts from the 57 countries in the low GDP group (e.g., Azerbaijan, Georgia, Namibia, Samoa) identified six keywords, though only one occurred in at least half the texts: *session* (see Table 5.11).

Table 5.11. Keywords in Low Gross Domestic Product State Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*P*
islands	97	0.09	10	10	0.02	37.10	2.36	0.00
session	166	0.15	56	31	0.05	35.30	1.50	0.00
Solomon	46	0.04	2	1		31.47	4.60	0.00
Excellency	58	0.05	25	4		28.04	2.9	0.00
SDGS	45	0.04	16	2		25.89	3.57	0.00
MDGS	30	0.03	14	0		25.45	137.11	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, medium, high, and very high Gross Domestic Product texts).

The full list of keywords were grouped into the categories low GDP places, formal language, and development (see Table 5.12).

Table 5.12. Keyword Categories for Low Gross Domestic Product (GDP) States

Low GDP places	Formal Language	Development
islands	*session*	SDGs
Solomon	Excellency	MDGs

Note. Keywords in at least 50% of texts are italicized.

As with the other country categories, two keywords related to places with low GDP:

Solomon and *Islands*. More meaningful are the keywords illustrating the use of formal language.

The keyword *session* appears in 56 of the 57 low GDP texts. *Session* most frequently occurs in

the opening remarks of a text and is part of the highly formalized discourse including

expressions of gratitude and congratulations (see Text 5.17).

Text 5.17. Excerpt from Zambia (keyword in italics)

> Your *Excellency*, Mr. Ban Ki-moon, Secretary-General of the United Nations; your
> Excellencies, heads of states and governments; distinguished delegates. I wish to begin
> by thanking the Secretary-General for convening this important *session* marking the 70th
> anniversary of the United Nations. Let me take this opportunity to congratulate his
> *Excellency* Mr. Morgens Lykketoft, for assuming the presidency of the high level
> segment of the *session*. Coming from the great country of Denmark, I am confident that
> he will preside over the *session* with great success. I also wish to pay tribute to his
> *Excellency* Mr. Sam Kutesa of the Republic of Uganda for successfully presiding over the
> last General Assembly *session*.

Because *session* occurs in these highly formalized introductory passages, its use in texts

from the lowest GDP states is yet another example of the least developed countries maintaining

many formal rhetorical conventions of UN discourse (see also Sections 5.3.2, 5.3.5, and 5.4.2).

This association is further strengthened by the appearance of *Excellency* on the keyword list. The

word is used in opening greetings and thanks, just as it was for African countries (Section 5.3.2)

and for LDCs (Section 5.3.5).

The last two keywords for the low GDP group reveal a tendency to discuss development issues, with reference to *SDGs* (Sustainable Development Goals) and *MDGs* (Millennium Development Goals), the precursor to the SDGs (see Section 5.3.5 for more on the SDGs). The focus on development for the low GDP countries, as with the LDCs, is understandable given the critical need among low GDP countries for developmental policy and assistance. Texts 5.18 and 5.19 illustrate the use of these terms.

Text 5.18. Excerpt from Bhutan (keywords in italics)

> Many of the *SDGs* in the 2030 Agenda will not be attainable if we do not relentlessly invest in building a dynamic and relevant education system in our countries - a system that nurtures our children and youth with the right values, knowledge and skills that reflect our aspirations and goals. In this regard, we are happy to note that ensuring inclusive and equitable quality education for all is a stand-alone goal in the 2030 Agenda. In moving forward, SDG 4 must remain at the forefront of our planning and implementation of national strategies to achieve the *SDGs*.

Text 5.18 from Bhutan illustrates three common persuasive techniques among UNGA addresses. One technique is to suggest a cause-effect relationship based on objective logic rather than subjective opinion (*will not be attainable if*). Another technique is to show appreciation for progress already made (*we are happy to note*). Only the last sentence in this passage includes a strong directive (*SDG 4 must remain*). It is interesting to note that *must* is used with a non-human subject to avoid agency (as opposed to *we must place SDG 4 at the forefront*). Thus, there is some overt persuasion, but it is mitigated by (a) the use of other persuasive techniques such as giving thanks and (b) the use of a non-human subject with a modal of strong directive force. (See Text 5.12 for an additional example of thanking before requesting; See Sections 4.4.5.2 and 5.6.5.4 for additional examples of non-human subjects with *must.*) Text 5.19 from Malawi illustrates two of these persuasive techniques as well.

Text 5.19. Excerpt from Malawi (keywords in italics)

> Despite varied progress among Member States, and developing countries in particular, the Millennium Development Goals - *MDGs* have proved that unity of purpose is critical to transforming the world to be a better place for our children and future generations. My government is grateful to development partners who worked with us for Malawi to achieve four of the eight *MDGS* namely; reducing child mortality; combating HIV and AIDS, Malaria and other diseases; ensuring environmental sustainability; and developing a global partnership for development.

Text 5.19 on development and the Millennium Development Goals presents the argument for *unity of purpose* in pursuing development as indisputable fact rather than opinion (*have proved*). The text also thanks *development partners* for their work with Malawi.

Section 5.3.7.1 offers additional support for the finding that countries in low development groups tend to respect the formal conventions of UNGA addresses, with long introductions including congratulations and thanks. It also provides further textual examples showing that countries in low development groups focus on the subject of development and in particular the Sustainable Development Goals and the Millennium Development Goals, a topic of vital interest to these countries given their economic and social situations.

5.3.7.2. Medium Gross Domestic Product

The analysis of texts from the 12 countries in the medium GDP group (e.g., Barbados, Croatia, the Seychelles, Trinidad and Tobago) identified 13 keywords, but only one that appears in at least 50% of the texts: *crimes* (see Table 5.13).

Table 5.13. Keywords in Medium Gross Domestic Product State Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*P*
Tobago	28	0.12	2	1		104.03	7.50	0.00
Trinidad	28	0.12	2	1		104.03	7.50	0.00
Croatia	19	0.08	1	0		76.30	138.74	0.00
Barbados	18	0.08	2	0		72.28	138.66	0.00
Nevis	17	0.07	1	0		68.27	138.58	0.00
Kitts	14	0.06	1	0		56.22	138.30	0.00
crimes	*32*	*0.14*	*7*	*32*	*0.02*	*49.01*	*2.69*	*0.00*
Greece	13	0.06	1	2		41.00	5.39	0.00
Latvia	10	0.04	1	0		40.16	137.81	0.00
St	14	0.06	1	7		31.50	3.69	0.00
Slovenia	7	0.03	1	0		28.11	137.30	0.00
Caribbean	21	0.09	4	27	0.02	26.32	2.33	0.00
Estonia	6	0.03	1	0		24.09	137.07	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, low, high, and very high Gross Domestic Product texts).

All the keywords were grouped into the categories Medium GDP places and crimes (see Table 5.14).

Table 5.14. Keyword Categories for Medium Gross Domestic Product States

Medium GDP places	Crimes
Trinidad (and) Tobago	*crimes*
Croatia	
Barbados	
Caribbean	
St Kitts (and) Nevis	
Greece	
Latvia	
Slovenia	
Estonia	

Note. Keywords in at least 50% of texts are italicized.

As with previous country groups, places in the group are not particularly meaningful, leaving only one keyword: *crimes*. This keyword was also identified for the European texts, and an examination of the texts in which *crimes* appears for the medium GDP group shows that five out of seven are in Europe: Croatia, Estonia, Latvia, Lithuania, and Slovenia. As with other European texts (see Section 5.3.3), of the 25 occurrences of *crimes* for the five European medium GDP countries, 21 are in the phrases *crimes against humanity, war crimes*, *atrocity crimes,* and *Rome Statute crimes* (which encompasses the three aforementioned crimes). All five European texts refer to one of these types of international crimes at least once. Text 5.20 illustrates the use of these words and the similar patterns to those discussed in Section 5.3.3.

Text 5.20. Excerpt from Estonia (keywords in italics)

> Under the first pillar of the Responsibility to Protect, all member states of the United Nations commit to protecting their people from genocide, war *crimes*, ethnic cleansing and *crimes* against humanity. On the occasion of the l0th anniversary of the Responsibility to Protect, I reiterate our support for this principle and call for a renewed commitment by the international community to prevent genocide, war *crimes*, ethnic cleansing and *crimes* against humanity.

The text from Trinidad and Tobago also refers to *crimes against humanity* and *war crimes.* In contrast, in the text from Saint Kitts and Nevis, *crimes* is not related to these international crimes, as shown in Text 5.21.

Text 5.21. Excerpt from Saint Kitts and Nevis (keywords in italics)

> At home, we instituted a Six-Point Plan in our fight to reduce violent *crimes* in particular. We are investing in new equipment and training to support our law enforcement professionals who prevent, detect and solve crime. We are working hard to build a new professional culture among and between our law enforcement agencies. We are encouraged by new statistics pointing to a decline in major *crimes* in the Federation, but much more must be done.

Text 5.21 focuses on crimes committed in Saint Kitts and Nevis and national efforts being undertaken to combat violence. These types of *crimes* are not comparable to the large-scale *war crimes* so frequently discussed in European texts.

This analysis of *crimes* for medium GDP countries generally supports the findings of Section 5.3.3 by showing that, of all medium GDP countries, it is primarily European countries that speak of these international crimes and their relationship to the "Responsibility to Protect." However, it also shows that European countries are not unique in their concern for these crimes. It is not the fact that European countries raise the issue of *war crimes* and *crimes against humanity* that is noteworthy but rather the fact that they do so with such high frequency.

5.3.7.3. High Gross Domestic Product

The analysis of texts from the four countries in the high GDP category (Bahamas, Brunei Darussalam, Cyprus, Malta) identified five keywords, two of which occurred in at least 50% of the texts, as shown in Table 5.15.

Table 5.15. Keywords in High Gross Domestic Product State Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*P*
Bahamas	23	0.29	1	2		128.21	7.91	0.00
Malta	14	0.18	1	1		79.16	8.19	0.00
Cyprus	16	0.20	1	8		68.95	5.38	0.00
Commonwealth	*9*	*0.11*	*2*	*7*		*34.28*	*4.75*	*0.00*
organisation	*14*	*0.18*	*3*	*46*	*0.03*	*25.62*	*2.67*	*0.00*

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, low, medium, and very high Gross Domestic Product texts).

All the keywords were grouped into the categories high GDP places, other high GDP related, and "other" (see Table 5.16).

Table 5.16. Keyword Categories for High Gross Domestic Product (GDP) States

High GDP places	High GDP related (other)	Other
Bahamas	*Commonwealth*	*organisation*
Malta		
Cyprus		

Note. Keywords in at least 50% of texts are italicized.

The term *Commonwealth* is related to the country category because in the text from the Bahamas it is used in the phrase *the Commonwealth of the Bahamas* and in the text from Malta it refers to the British Commonwealth, of which Malta is part. Like with the other keywords specifically related to the countries in the category under investigation, these words are not particularly meaningful.

The remaining keyword is *organisation.* It refers once to the Maritime Organisation and the remaining 13 times to the United Nations (see Text 5.22).

Text 5.22. Excerpt from Brunei Darussalam

> We hope to see further improvements to make our *organisation* more effective in responding to the ever-increasing demands of its members. On this note, I wish to reiterate our commitment to continue working with others to ensure that the UN remains a strong force for peace, security and human rights. The success of our *organisation* lies in our hands. After all, our collective efforts and strength are the best attributes of this *organisation*.

In the text from Brunei Darussalam, the United Nations is referred to as *our organisation* or *this organisation.* This is common across UNGA texts and the keyword is likely to have been identified only because the high GDP category is small. The use of *organisation* is therefore not particularly meaningful.

5.3.7.4. Very high Gross Domestic Product

The analysis of texts from the 14 very high GDP countries (e.g., Austria, Israel, US, New Zealand) identified 23 keywords, nine of which occur in at least 50% of the texts, as shown in Table 5.17.

Table 5.17. Keywords in Very High Gross Domestic Product State Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*P*
Israel	78	0.27	3	15	0.01	203.63	4.71	0.00
Iran	58	0.20	6	15	0.01	139.76	4.28	0.00
Iran's	29	0.10	5	0		104.24	139.04	0.00
Ireland	30	0.10	1	1		99.36	7.24	0.00
Syria	*60*	*0.21*	*13*	*51*	*0.04*	*81.01*	*2.57*	*0.00*
Iceland	17	0.06	1	0		61.11	138.27	0.00
it's	26	0.09	5	8		59.26	4.03	0.00
Australia	20	0.07	2	2		59.21	5.65	0.00
Netherlands	16	0.06	1	0		57.51	138.18	0.00
that's	17	0.06	2	1		53.74	6.42	0.00
Zealand	15	0.05	1	2		42.33	5.24	0.00
terror	*19*	*0.07*	*9*	*7*		*40.54*	*3.77*	*0.00*
deal	*31*	*0.11*	*7*	*30*	*0.02*	*37.75*	*2.38*	*0.00*
Iraq	*23*	*0.08*	*10*	*15*	*0.01*	*37.13*	*2.95*	*0.00*
refugees	*38*	*0.13*	*12*	*49*	*0.03*	*35.14*	*1.96*	*0.00*
ISIL	*18*	*0.06*	*7*	*9*		*33.59*	*3.33*	*0.00*
nuclear	*40*	*0.14*	*8*	*59*	*0.04*	*31.59*	*1.77*	*0.00*
Jewish	10	0.03	1	1		29.61	5.65	0.00
Finland	10	0.03	1	1		29.61	5.65	0.00
here's	8	0.03	1	0		28.76	137.18	0.00
Assad	9	0.03	6	1		26.21	5.50	0.00
Europe	*39*	*0.14*	*13*	*65*	*0.04*	*26.14*	*1.59*	*0.00*
help	*27*	*0.09*	*11*	*36*	*0.02*	*24.05*	*1.92*	*0.00*

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, low, medium, and high Gross Domestic Product texts).

All the keywords were grouped into the categories very high GDP places, Iran nuclear deal, Syria, conflict and its consequences, contractions, and "other" (see Table 5.18).

Table 5.18. Keyword Categories for Very High Gross Domestic Product (GDP) States

Very high GDP places	Iran nuclear deal	Syria	Conflict and its consequences	Contractions	Other
Europe	Iran('s)	*Syria*	*terror*	it's	*help*
Ireland	*deal*	Assad	*refugees*	that's	Iraq
Iceland	*nuclear*		*ISIL*	here's	Jewish
Netherlands					
(New) Zealand					
Finland					
Israel					
Australia					

Note. Keywords in at least 50% of texts are italicized.

In addition to the usual keywords that were place names from the country category, the keyword *Jewish*, initially categorized as "other," was ultimately identified as related to one member of the very high GDP group: Israel. It appeared only in the Israeli text and is not interpreted here because it reflects the aboutness of the Israeli text rather than the very high GDP texts.

References to places outside of the very high GDP zone include *Iran* (and the *nuclear deal*), *Syria* (and President *Assad*), and *Iraq* (categorized as "other"). An examination of concordance lines revealed that these keywords are related to conflict, either armed (*Syria, Iraq*) or political (*Iran*). Most speakers cast the nuclear deal with Iran in a positive light as an example of a successful resolution of a political challenge (e.g., *the Iran deal was a notable bright spot this year*, New Zealand; *this year also marks a major success of international diplomacy: the historic agreement reached on the Iran nuclear programme*, Finland; *I would like to congratulate all the parties which achieved the breakthrough agreement on the Joint Comprehensive Plan of Action on Iran's nuclear program*, Iceland). Only the text from Israel was blatantly negative (e.g., *I have long said that the greatest danger facing our world is the*

coupling of militant Islam with nuclear weapons. And I'm gravely concerned that the nuclear

deal with Iran will prove to be the marriage certificate of that unholy union). This can be

explained by the tense relationship between Israel and Iran and the position of Israel in the UN

(see also Texts 4.32, 5.34, 5.50, 6.11, 6.22, 6.27, 6.29, 6.30, and 6.36).

The two other places, Syria and Iraq, are both associated with armed conflict. In the very

high GDP texts, *Syria* frequently collocates with *conflict(s)* (a total of 11 occurrences in six

texts), *crisis* (a total of four times in three texts), and other conflict zones: *Iraq* (a total of 14

times in nine texts), *Yemen* (a total of five times in five texts), and *Sudan* (a total of four times in

four texts). Likewise, of the 23 occurrences of *Iraq,* it collocates with *Syria* a total of 14 times in

nine texts and with *conflicts* a total of four times in three texts. Texts 5.23 and 5.24 illustrate the

association for both *Syria* and *Iraq* with conflict.

Text 5.23. Excerpt from Australia (keywords in italics)

> As a non-permanent member of the Security Council, we experienced first-hand the
> difficulties the Council faces in responding to the crises and *conflicts* in *Syria*, *Iraq*,
> Libya, Yemen, South Sudan, Central African Republic and Mali. Nowhere is the
> devastation greater than in *Iraq* and *Syria*.

Text 5.24. Excerpt from Finland (keywords in italics)

> We are now facing an extremely dangerous *crisis* in *Syria*, *Iraq* and parts of North-
> Africa. We are also witnessing an unprecedented humanitarian catastrophe unlike any
> other since the Second World War. *ISIL* and its horrendous terror is a direct by-product of
> the *conflicts* in *Syria* and *Iraq*. It threatens the peace and stability in Middle East, in
> Africa and even in *Europe*.

Both of these texts use negatively connoted words to describe the crises in Syria and Iraq:

devastation in Text 5.23 and *dangerous, catastrophe, horrendous, terror,* and *threatens* in Text

5.24. Also of interest in Text 5.24 is the comparison to the Second World War. As discussed in

Section 5.3.3, some European texts liken the crises in Syria and Iraq to World War II. This association may be used as a warning of what could result if the international community does not step up and assume its "responsibility to protect."

The keyword *Assad* is related to conflict in Syria. Several texts speak of the violence of the *Assad regime* (e.g., *In Syria, civilians suffer from horrific abuses. Committed by the terror organisation ISIL. And by the Assad regime*, Denmark; *The twin evils of Assad's murderous regime and the brutality of ISIL or Daesh, have inflicted suffering on the Syrian people on an almost unimaginable scale*, UK; *The Assad regime continues to flout international law while ISIS visits gratuitous cruelty and suffering on the people of Syria and Northern Iraq. Its casual barbarism is a threat to every value the UN holds dear and therefore the UN has a duty to act*, Ireland). Though these excerpts do not explicitly link the Assad regime to ISIL (also referred to as ISIS and Daesh), the juxtaposition of the two creates an association between the president of Syria and the terrorist group.

The keywords *ISIL* and *terror* are often in passages about Syria and Iraq, as shown in the excerpts above. Some of the texts create an explicit cause-effect relationship between conflict in Syria and Iraq with ISIL and terror (e.g., *ISIL and its horrendous terror is a direct by-product of the conflicts in Syria and Iraq*, Finland; *The conflict in Syria has also allowed extreme terror groups like ISIL to gain a foothold*, Norway). The texts also use powerful language to talk of the need to defeat *ISIL* and *terror* (e.g., *stamp out ISIL*, US; *combat the scourge of ISIL*, New Zealand; *crush ISIL and banish its twisted ideology from the face of the earth*, UK). This strong language is extremely rare in UNGA addresses and can be attributed to the fact that ISIL is outside of the UN family. UN countries must mitigate criticism of the international community

and other member states, but diplomacy does not dictate that care be taken when discussing an entity outside of the international community.

The final keyword related to conflict is *refugees.* The majority of countries in the very high GDP group are from Europe and the association between *refugees* and conflict in the European texts was discussed in Section 5.3.3. In addition, the European texts and the three very high GDP texts outside of Europe that speak about *refugees* (Australia, New Zealand, the US) frequently speak of their country's assistance with the refugee crisis (see Text 5.25).

Text 5.25. Excerpt from the US (keywords in italics)

> There are no easy answers to *Syria*. And there are no simple answers to the changes that are taking place in much of the Middle East and North Africa. But so many families need help right now; they don't have time. And that's why the United States is increasing the number of *refugees* who we welcome within our borders. That's why we will continue to be the largest donor of assistance to support those *refugees.*

Other examples of referring to national assistance to refugees include *Consistent with our record over decades of resettling permanently thousands of refugees, Australia was pleased to announce recently that we would resettle permanently an additional 12,000 UNHCR registered refugees from Syria and Iraq, particularly those who are unlikely to have a home to which they might return* [Australia]; *In response to the humanitarian crisis we are helping fund programmes in Jordan, Lebanon and Turkey to assist their efforts to support thousands of displaced refugees* [New Zealand]; *We have agreed to take in a high number of refugees from Syria's neighbouring countries for resettlement under UN quotas* [Norway]; *We have deployed search and rescue ships to the Mediterranean. We are providing for the resettlement of over 4,000 refugees* [Ireland].

This emphasis on the need to assist is also reflected in the keyword *help*. Speakers stress the importance of helping civilians: *Everyone can offer help to the refugees* [EU]; *We are doing what we can to help the victims of the Syrian conflict* [New Zealand]. They also stress the importance of helping countries: *But as we help the poorest countries overcome conflict, poverty and instability and to move up the development curve, it is in all our interests that that development is sustainable over the long term* [UK]; *We recognise that the UN needs more tools for peacebuilding, to help vulnerable states emerge from crisis and prevent them from falling back into violence and disorder* [Australia].

Why do highly developed states use *help* more frequently than those with lower GDPs? It is true that some developing nations request help from the UN (e.g., *The current situation in Dominica represents what islands face when dealing with Climate Change, and we call on the international community to do more to help Small Island Developing States adapt and adjust*, Saint Kitts and Nevis). However, many others reject the term. Either they pride themselves on their self-sufficiency (e.g., *We have never, however, appealed for external help in dealing with the security problems of Uganda. We always emphasize building our own capacity at the earliest opportunity*, Uganda) or they call for partnership rather than "handouts" (e.g., *Africa is not looking for handouts. Rather, it is looking for partners in massive infrastructural development, in creating and exploiting the value chains from its God-given natural resources, and in improving the quality of life of the continent's citizens*, Zimbabwe).

The final keywords are contractions, which are indicative of involved language. Because contractions are a grammatical rather than a lexical feature, they are discussed in Section 5.5.2 on Dimension 1 (involved versus informational production).

The keyword analysis for very high GDP countries has identified a number of interesting keywords that appear to be frequent in these texts for similar reasons. *Syria, Assad, Iraq, ISIL*, and *terror* all relate to violent conflict. Most texts speak of *refugees* within the context of conflict, primarily in the Middle East and North Africa. The keyword *help* is used in the context of assisting refugees as well as the context of helping developing countries. Even the keywords related to the *Iran nuclear deal* show a focus on conflict (this time political), though in this particular instance the subject is, for all but Israel, considered a positive resolution of a conflictual situation. The fact that very high GDP countries speak with unusual frequency on conflict-related topics suggests one of two things: (a) that their vested interest in maintaining peace is greater than the interest of countries in the lower GDP groups or (b) that countries in the lower GDP groups are equally concerned about conflict, but have a number of other pressing concerns such as development and climate change, and as a result speak more about these topics than conflict.

Interestingly, according to the United Nations High Commissioner for Refugees (2016), in 2015 the top 10 countries receiving refugees were all in the Middle East and Africa. Sweden and Malta were on the list of the top 10 countries for per capita refugee populations, but overall, the countries that make up the very high GDP group are not among those receiving the most refugees. Many very high GDP countries are providing assistance to refugees, as indicated in the texts from New Zealand and the US, but they are by far not the only countries affected by conflict and the humanitarian consequences of conflict. Though low HDI countries such as Liberia also welcomed Syrian asylum seekers in 2015, the refugee crisis was a less imminent threat than developmental issues such as health and disease (e.g., malaria, Ebola). While Dominica, too, may find the nuclear deal with Iran a promising step toward peace, it remains

more concerned with how to develop economically and socially when faced with the ravages of tropical storm Erika.

Keyword analysis shows what is unusually frequent in a text, but this does not imply that these words or concepts are wholly absent in texts from the other groups. *Syria* is mentioned 51 times in the texts from the lower GDP groups and *refugees* 49 times. Clearly, these are important words for the other countries, too. However, the unusual frequency of these words in the very high GDP group shows that on average, very high GDP countries talk more about conflict and the consequences of conflict than about other topics such as development and climate change. This reflects not only a concern about conflict but also less of a concern, relatively speaking, about the other issues.

5.3.8. Very high Human Development Index

The low, medium, and high HDI groups had no keyword distributed over at least half the texts outside of place names from the group itself. In contrast, for the 25 very high HDI countries (e.g., Australia, Liechtenstein, Norway, UK) there were 26 keywords, six of which occurred in at least 50% of the texts, as shown in Table 5.19.

Table 5.19. Keywords in Very High Human Development Index Texts

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
Israel	79	0.16	4	14	0.01	127.22	3.80	0.00
Syria	*83*	*0.17*	*21*	*31*	*0.03*	*94.11*	*2.72*	*0.00*
Iran	60	0.12	8	13	0.01	89.68	3.51	0.00
Iran's	29	0.06	5	0		72.13	138.27	0.00
Ireland	30	0.06	1	1		66.46	6.21	0.00
it's	26	0.05	5	2		51.62	5.00	0.00
Cyprus	23	0.05	3	1		49.57	5.83	0.00
Croatia	19	0.04	1	0		47.26	137.66	0.00
Europe	*64*	*0.13*	*23*	*42*	*0.03*	*45.41*	*1.91*	*0.00*
that's	18	0.04	3	0		44.77	137.58	0.00
Iceland	17	0.03	1	0		42.28	137.50	0.00
Netherlands	16	0.03	1	0		39.80	137.41	0.00
Australia	20	0.04	2	2		37.70	4.63	0.00
ISIL	23	0.05	10	4		37.28	3.83	0.00
Iraq	*29*	*0.06*	*15*	*9*		*36.65*	*2.99*	*0.00*
deal	39	0.08	11	20	0.02	35.05	2.27	0.00
atrocity	14	0.03	8	0		34.82	137.22	0.00
crimes	*42*	*0.09*	*14*	*28*	*0.02*	*29.29*	*1.89*	*0.00*
Malta	14	0.03	1	1		28.15	5.11	0.00
refugees	*48*	*0.10*	*17*	*38*	*0.03*	*27.18*	*1.64*	*0.00*
Zealand	15	0.03	1	2		26.35	4.21	0.00
terror	20	0.04	10	6		25.74	3.04	0.00
nuclear	49	0.10	12	41	0.03	25.71	1.56	0.00
Latvia	10	0.02	1	0		24.87	136.74	0.00
Council's	10	0.02	5	0		24.87	136.74	0.00
prevent	*30*	*0.06*	*15*	*17*	*0.01*	*24.67*	*2.12*	*0.00*

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, low, medium, and high Human Development Index texts).

All the keywords were grouped into six categories: very high HDI places, Iran nuclear deal, Middle East, conflict and its consequences, contractions, and "other" (see Table 5.20).

Table 5.20. Keyword Categories for Very High Human Development Index Texts

Very high HDI places	Iran nuclear deal	Middle East	Conflict and its consequences	Contractions	Other
Europe	Iran('s)	*Syria*	*terror*	it's	prevent
Ireland	*deal*	Iraq	*refugees*	that's	(Security) Council's
Iceland	*nuclear*		*ISIL*		
Netherlands			*crimes*		
(New) Zealand			atrocity		
Israel					
Australia					
Croatia					
Malta					
Latvia					
Cyprus					

Note. Keywords in at least 50% of texts are italicized.

Many of the keywords identified for the very high HDI group are the same as those identified for the very high GDP or the European group because there is a great deal of overlap between these groups. All 14 very high GDP countries are also among the 25 very high HDI countries; 19 European countries are among the very high HDI countries as well. The keywords *Iran, Iran's, deal, nuclear, Syria, Iraq, terror, refugees, ISIL, it's* and *that's* were all discussed in Section 5.3.7.4 on the very high GDP group and are used in the same way for the very high HDI group. The keywords *Syria, refugees,* and *crimes* were also discussed in Section 5.3.3 on Europe and again are used in the same way for the very high HDI group.

The discussion here instead focuses on the keywords that are unique to the very high HDI group: *atrocity, prevent,* and *(Security) Council's.* All three of these keywords, like the other keywords previously discussed, relate to violence. Of the 14 occurrences of *atrocity*, 10 were in the phrase *atrocity crimes.* The others (e.g., *atrocity situations*) also related to crimes. For example, *For that reason, New Zealand supports the two proposals being put forward this year -*

by the A.C.T. Group of countries and by France and Mexico-to limit the use of the veto in mass atrocity situations [New Zealand]; *We will not let a Russian veto impede the efforts of countries grieving the loss of their citizens and demanding justice for the perpetrators of this atrocity* [Australia].

Prevent collocates primarily with words related to violence: *conflict(s), tragedies, threats, genocide, mass atrocity crimes, mass killings, crimes,* and *war(s).* Several of these occurrences are in passages criticizing the UN failure to prevent violence: *However, far too often international order and principles have been violated and the UN has been unable to prevent conflicts or build peace or to stop atrocities* [Latvia]; *But we cannot be complacent: on the contrary, in this anniversary year, we must critically appraise our failures as the international community to prevent war, genocide, forced displacement and deprivation* [Ireland]; *On the occasion of the l0th anniversary of the Responsibility to Protect, I reiterate our support for this principle and call for a renewed commitment by the international community to prevent genocide, war crimes, ethnic cleansing and crimes against humanity. The Security Council bears primary responsibility in this regard. Unfortunately, in most serious situations, the Security Council has failed to act due to the use of the veto* [Estonia]. Criticism is unusual in UNGA addresses and is reserved for only the most pressing issues. The fact that several very high HDI countries criticize the inability of the UN to prevent conflict and crimes underscores the importance of peace and security for this group of countries.

Similarly, *Council's* is used solely in reference to the Security Council and many occurrences are in passages about the Council's failure to meet its mandate of ensuring peace and security. Though it is difficult to generalize with only 10 occurrences, six of which are in the same text (Ireland), the use of *Council's* supports the analysis of the other keywords for the very

high HDI group. The most explicit examples are: *Some of the Council's most damaging failures occurred in the face of atrocity crimes: Rwanda, Srebrenica, and now Syria sadly stand out in this respect* [Liechtenstein]; *At the same time we see the consequences of the Security Council's failure to act over the past four years* [Ireland]. Australia is a little more nuanced: *Australia's recent experience as a Security Council member confirmed that the Council's role is more essential than ever. The Security Council can only perform its role if it has the tools it needs. The role of peacekeeping is fundamental*. Only later does Australia explain that the Security Council cannot act if Permanent Members exercise their right to veto in situations of mass atrocities: *we welcome proposals to restrain use of the veto where mass atrocity crimes are being committed*. These passages on the Security Council illustrate both explicit and indirect criticism. It is noteworthy that criticism in these very high HDI texts is directed against the international community, and therefore includes the speaker. Overall, criticism of the UN, its failure to act or the imbalance of representation in bodies like the Security Council, is much more common than criticism of any UN member state. Because the speaker is part of the UN, the criticism reads more as a request for reform rather than wholesale condemnation. Even when criticism of the UN is overt, it is less incriminating than criticism of ISIL (see Section 5.3.7.4).

Overall, the keyword analysis for the very high HDI texts supports the analysis of the very high GDP texts, showing a concern for conflict and the consequences of conflict. Speakers express the need to prevent violence and help refugees, the victims of violence. They also criticize, sometimes explicitly, sometimes indirectly, the failure of the international community to fulfil its mission of ensuring peace and security. Their focus on conflict can be interpreted in much the same way as for the very high GDP group. It is unlikely that these countries are more concerned about violence than countries in the lower development groups as many of those

236

countries lie at the very heart of conflict zones and have lost much due to violence within their own borders (e.g., Afghanistan, Pakistan, South Sudan). Nearly every UNGA country speaks of peace and security, but the most developed countries tend to speak more about conflict and less about climate change and development when compared to the less developed states.

5.3.9. Summary

The results of the keyword analyses show that states generally speak on topics of national or at least regional concern rather than issues that do not affect them directly. SIDS texts address environmental disasters, while European texts refer to the refugee crisis wreaking economic and social havoc on many of their nations. In addition, some groups of countries avoid issues for which their governments are being criticized. For example, the largest military spenders refer to *small arms* much less frequently than the low military spenders, denouncing the trade of weapons. Other country groups shy away from terms which may have a negative evaluation for their group: The most developed countries talk of *helping* vulnerable communities more frequently than developing nations. Instead, developing nations speak of the Sustainable Development Goals and the partnerships that need to be put in place in order to achieve them. Finally, in terms of rhetoric and discourse style, developing nations tend to use more formalized conventions for UN addresses, addressing members of the audience by official titles (*Excellency*) and expressing appreciation to UN leadership for their work in past *sessions.*

Of interest in Section 5.3 is not only what was identified through keyword analysis, but also what was not identified. In all, 31 country categories were investigated, but only 10 groups had one or more meaningful keywords occurring in at least half the texts. When no meaningful keywords can be identified for a country category, it indicates that the group exhibits no patterns of systematic lexical variation distinguishing it from the rest of the UNGA. This can be attributed

to an overall homogeneity of the UNGA texts or to a lack of homogeneity within the country group (or both). Findings from Section 4.4 support the former: In general, many words are used across UNGA addresses and in very similar ways. However, findings in Section 5.3 show some lexical patterns for ten country groups, quite strong patterns for Europe and countries with very high development. This would indicate that some lexical variation can be identified within the UNGA, but that many country groups do not exhibit a great deal of lexical similarity distinguishing them from other UNGA groups. Because little lexical variation was detected by country group, a post-hoc analysis of five separate texts was undertaken in order to determine whether patterns in word choice could be identified at the level of individual countries. (See Section 5.6 for the results of the analyses for these individual texts.)

5.4. Variation in UNGA Grammatical Features: Key Feature Analysis

5.4.1. Introduction

Section 5.4 describes the results of the grammatical feature analysis by country category. The UNGA corpus was tagged for grammatical structures and the normed frequencies of these grammatical structures were calculated for each text (see Section 3.5). Then texts were grouped into country categories (see Section 3.6) and the mean and standard deviation of each feature for each group was calculated (see Section 3.7.3). Appendices B-K present mean and standard deviation values for all 126 features analyzed for all 31 country groups. These values can be compared to the overall UNGA values reported in Appendix A.

Cohen's d scores were also calculated in order to determine which grammatical structures appeared with notable frequency for each country category (see Section 3.7.3). Section 5.4 discusses only features with large Cohen's d scores ($d \geq 0.80$), that is, only features that are unusually frequent or infrequent in a country group. Out of a total of 31 country groups, only 12

238

had any features that were unusually frequent or infrequent. Out of a total of 126 features, only 32 were unusually frequent or infrequent for any country group. This is perhaps the most notable finding from the key feature analysis: For nearly 100 features, there was no remarkable difference in frequency for any country group. The groups were, overall, surprisingly similar in their use of grammatical features. This finding supports the analysis in Section 4.5 showing great similarities across UNGA texts in grammatical structures. It also supports the analysis in Section 5.3 that very little of the variation that does occur in the UNGA is detected in the country categories.

Nevertheless, at least one unusually frequent or infrequent grammatical feature was identified for 12 country groups, and these groups are discussed in Section 5.4: Africa (Section 5.4.2), Asia (Section 5.4.3), Landlocked Developing Countries (LLDCs, Section 5.4.4), small area (Section 5.4.5), high military spending (Section 5.4.6), low religious non-affiliation (Section 5.4.7), medium, high, and very high GDP (Section 5.4.8), and low, medium, and very high HDI (Section 5.4.9). The majority of these groups had only one or two key features, limiting any possible interpretation. Only development, from the LLDC, GDP, and HDI categories, produced notable findings for grammatical forms. For these country categories, the analyses show that combinations of grammatical features can create a particular discourse style such as dense, informative texts or overtly persuasive texts, and that some styles are especially favored by certain groups of countries.

5.4.2. Africa

For the 21 African countries represented in the UNGA corpus, only one key grammatical feature was identified as having a large effect size: second person pronouns (see Table 5.21).

Table 5.21. Key Grammatical Features for Africa

Feature	Africa		UNGA average
	Frequency	**Cohen's *d***	
2nd person pronoun	4.7	0.91	3.0

Note. Frequencies normed per 1,000 words.

Interpreting one grammatical feature in isolation is not always possible, but the use of second person pronouns in African texts supports findings from the lexical analysis (see Section 5.3.2): Nearly all second person pronouns were in the highly formalized introduction, in expressions of gratitude and support (see Texts 5.26 and 5.27), demonstrating a strong adherence to the rhetorical conventions of the UNGA address. Of the 81 occurrences of *you*, 24 collocate with *thank*, ten with *congratulat**, nine with *support*, and two with *grateful*.

Text 5.26. Excerpt from Kenya (key feature in italics)

> Let me congratulate *you*, Mr. President, on your election as President of this historic session. We are confident that your vast diplomatic experience will steer *you* and contribute to the achievement of the objectives *you* have set. Kenya fully supports your leadership, and will constructively engage with *you* towards realizing the objectives of this historic session.

Text 5.27. Excerpt from Botswana (key feature in italics)

> On behalf of His Excellency the President of the Republic of Botswana, Lt. General Seretse Khama Ian Khama, allow me at the very outset, to join other delegations in extending our heartiest congratulations to *you* on your election to the Presidency of the 70th Session of this august body of the United Nations. Your election bears testimony to the excellent credentials *you* bring to the Office of the President of the General Assembly, and which should greatly support the execution of your mandate during this historic year on the UN calendar. We wish to assure *you* of Botswana's unequivocal support throughout your term of office.

In Texts 5.26 and 5.27, the speakers congratulate the President of the Assembly and assure him of the support of their delegations in rather lengthy and effusive introductions. This

was common in the vast majority of African texts: Out of 21 African texts, 15 congratulated the President (71%).

The act of congratulating the President is not unique to African countries. Rather, it is frequent for developing countries generally. The percentage of texts congratulating the President was nearly identical for the low GDP countries: 41 out of 57 countries, or 72%. The contrast to very high GDP countries is striking, with only one out of 14 very high GDP texts congratulating the President (7%). The reason second person pronouns were identified as key for African texts but not low GDP countries is the length of the passages, resulting in a higher frequency of the second person pronoun. Text 5.28 from Mongolia illustrates the much briefer congratulations from an Asian low GDP country.

Text 5.28. Excerpt from Mongolia (*you* in italics)

> My heartfelt congratulations to *you*, Mr. President of the General Assembly. *You* can count on my delegation's full support during the jubilee session and the year ahead.

Congratulating the President of the UNGA is a convention that may be increasingly infrequent among developed countries, but developing countries on average appear to be upholding these formalities. The respect for rhetorical conventions, reflected in the high frequency of the second person pronoun for congratulating, thanking, and showing support, is in keeping with the findings in Sections 5.3.2, 5.3.5, and 5.3.7.1.

5.4.3. Asia

The 19 Asian countries made up the only other regional group with a key feature: verbs of likelihood (e.g., *seem, presume, think*, see Table 5.22).

Table 5.22. Key Grammatical Features for Asia

Feature	Asia		UNGA average
	Frequency	**Cohen's *d***	
Verb complement clause with likelihood verb	1.5	0.82	0.9

Note. Frequencies normed per 1,000 words.

Verbs of likelihood express epistemic stance that is specifically attributed to the speaker (Text 5.29) or to both speaker and addressees (Text 5.30) through the use of the first person.

Text 5.29. Excerpt from Cambodia (key feature in italics)

> But we *believe that* the key will lie on how to ensure successful achievement with the lessons learnt from MDG's. We sincerely *believe that* during the next 15 years, all stakeholders, developed and developing countries would join hands to follow-up this paramount "Development Agenda".

The two occurrences of *we believe* mitigate the force of the propositions in Text 5.29, but the second *we believe* also serves to indirectly advise, encouraging all countries to support the "Development Agenda" in the future.

Text 5.30. Excerpt from Cyprus (key feature in italics)

> While we *assume that* only Jordan, Turkey, Lebanon and subsequently Europe have been affected by the current refugee crisis, we fail to acknowledge that if this crisis is to persist other countries and Continents are to be affected as well.

In Text 5.30, *we assume* juxtaposes belief with fact; *we assume that* the (refugee) crisis affects a relatively small number of states, but this ignores the suggested reality that it could affect many. Similar to Text 5.29 from Cambodia, in Text 5.30, the use of the verb of likelihood indirectly advises by showing the fallacy of making such an assumption.

The key feature analysis shows that Asian countries might advise more indirectly, through the use of verbs of likelihood. However, it is impossible to generalize based on one key feature, with no keyword in Section 5.3 for additional support.

5.4.4. Landlocked developing countries

The landlocked developing countries (LLDCs, e.g., Afghanistan, Nepal, Rwanda, and South Sudan) are marked by their infrequent use of the pro-verb *DO*, of modals, and in particular of modals of necessity relative to the other UNGA texts. In contrast, *that* complement clauses controlled by a factive adjective (e.g., *obvious*) are relatively frequent (see Table 5.23).

Table 5.23. Key Grammatical Features for Landlocked Developing Countries (LLDCs)

Feature	LLDC		UNGA average
	Frequency	**Cohen's *d***	
Pro-verb *DO*	0.1	-0.89	0.4
All modals	11.5	-0.83	14.4
Necessity modals	3.3	-0.82	5.0
Factive adjective complement clause	0.7	0.94	0.6

Note. Frequencies normed per 1,000 words.

Modals generally and in particular necessity modals are frequently used for making direct requests or suggestions. Compared to other UNGA texts, LLDC texts disprefer overt persuasion through the use of modals. In contrast, these countries use relatively more factive adjective complement clauses to present opinion as objective truth. Texts 5.31 and 5.32 illustrate how the factive adjective followed by *that* can be used as an indirect suggestion.

Text 5.31. Excerpt from Swaziland (key feature in italics)

It is *apparent that* military intervention in resolving conflict, has short term effect that leaves the people it is meant to serve, worse off.

In the text from Swaziland, the adjective complement clause *apparent that…* is used to present an opinion (military intervention leaves people worse off) as fact. Using the factive adjective *apparent* rather than a modal such as *should* is a much more indirect way of persuading the audience that military intervention should not be used to resolve conflicts.

Text 5.32. Excerpt from Malawi (key feature in italics)

It is *evident that* as a country we are embarking on the implementation of the post-2015 Development agenda with unfinished business.

In the text from Malawi, the adjective complement clause *it is evident that* also precedes an opinion: we should finish the business we have left to do, but the use of *evident* makes the opinion sound like indisputable fact.

Overall, factive adjective complement clauses are frequent in LLDC texts relative to the other UNGA texts. However, in absolute terms, they are not particularly frequent. As shown in Table 5.23, there are only seven per 10,000 words in LLDC texts (as opposed to six for all UNGA texts). This is not a particularly remarkable difference when taken alone. However, when analyzed in conjunction with the low frequency of modals and modals of necessity (again, relative to other UNGA texts), a pattern emerges with a stronger tendency in LLDC texts to express opinion as fact when compared to other UNGA texts. These three grammatical features (modals, modals of necessity, and factive adjective complement clauses) underscore the importance of examining both Cohen's *d* values and normed frequencies. Cohen's *d* values show the preference or dispreference of a structure for one group of texts (e.g., the LLDC) compared to another group (e.g., non-LLDC). Normed frequencies reveal the preference or dispreference of a structure in absolute terms.

One last key feature from the LLDCs is the infrequent use of *DO* as a pro-verb. The low frequency of this structure would generally be associated with explicit language and the avoidance of reduced forms, but no other key features from the LLDC texts support this analysis. Moreover, even though the Cohen's *d* value shows a large effect size, the frequency of pro-verb *DO* is extremely low for all UNGA addresses, not just the LLDCs. On average there are four instances of the pro-verb *do* for every 10,000 words in the UNGA. For the LLDCs, that number is one in 10,000. The difference is not particularly meaningful.

The most interesting finding from the LLDC group remains the preference for presenting opinion as fact through the use of adjective complement clauses rather than expressing overt persuasion through the use of modals and in particular modals of necessity (discussed further in Section 5.5.5).

5.4.5. Small area

One key grammatical feature was identified for the 22 countries in the small geographical area (e.g., Antigua and Barbuda, Brunei Darussalam, Liechtenstein, Tuvalu): progressive verbs (see Table 5.24).

Table 5.24. Key Grammatical Features for Small Area

Feature	Small Area		UNGA average
	Frequency	**Cohen's *d***	
Progressive verbs	5.6	-0.89	7.3

Note. Frequencies normed per 1,000 words.

Though progressive verbs had a large Cohen's *d* value, an examination of normed frequencies shows that, on average, small area states used 5.6 progressive verbs per 1,000 words

while UNGA texts overall used 7.3 progressive verbs. The contrast is not striking. Given that no

other key feature was identified for this group, any interpretation would be conjectural.

5.4.6. High military spending

Two key grammatical features were identified for the 18 top military spenders with 2% or

more GDP on defense (e.g., Botswana, Greece, Jordan, Ukraine): *that* clauses controlled by

verbs of likelihood and adverbs of place (see Table 5.25).

Table 5.25. Key Grammatical Features for High Military Spending

Feature	High Military Spending		UNGA average
	Frequency	**Cohen's *d***	
Place adverbs	5.8	0.86	4.2
Verb complement clause with likelihood verb	1.5	0.81	0.9

Note. Frequencies normed per 1,000 words.

That clauses controlled by verbs of likelihood were also found to be frequent in Asian

texts, but an examination of frequencies for individual texts in the high military spending group

showed that it was both Asian countries and non-Asian countries that used the structure. As

discussed in Section 5.4.3, verbs of likelihood express epistemic stance, most frequently the

belief of the speaker (e.g., *We believe that the UN and its Global Counter Terrorism Forum also*

have an important role to play, Estonia).

However, in other instances verbs of likelihood serve additional functions. The text from

the United States makes particularly frequent use of verbs of likelihood ($d = 2.70$); *I believ* that*

appears 14 times out of 4,751 words. These repetitions of *I believe* are inspirational and

persuasive (see Text 5.33; see also the in-depth analysis of the US address in Section 5.6.2).

Text 5.33. Excerpt from the United States (key feature in italics)

> I say this, recognizing that diplomacy is hard; that the outcomes are sometimes unsatisfying; - that it's rarely politically popular. But I *believe that* leaders - of large nations, in particular, have an obligation to take these risks - precisely because we are strong enough to protect our interests if, and when, diplomacy fails. I also *believe that* to move forward in this new era, we have to be strong enough - to acknowledge when what you're doing is not working.

Another idiosyncratic use of verbs of likelihood is in the address from Israel, which is one of the rare few to use verbs of likelihood with the second person. This is done in a confrontational way that is, again, uncommon in these diplomatic speeches and may be due to the fact that Israel is expressing its concern over the nuclear deal with Iran, which it perceives as having dire consequences regionally and even internationally (see Text 5.34).

Text 5.34. Excerpt from Israel (key feature in italics)

> You *think* Iran is doing that to advance peace? You *think* hundreds of billions of dollars in sanctions relief and fat contracts will turn this rapacious tiger into a kitten?

Comparing the use of verbs of likelihood in Texts 5.33 and 5.34 reveals an interesting difference in expression of epistemic stance. The use of *I believe* in the text from the United States mitigates the force of a proposition (e.g., moving forward requires acknowledging when what you're doing is not working). The evaluation is positive, suggesting that speakers should agree. In contrast, the text from Israel, with the repetition of rhetorical, skeptical questions expresses doubt. The negative evaluation implies Israel disagrees with addressees; it is antagonistic in tone. These examples of verbs of likelihood illustrate how expression of stance can result in very different styles. In spite of the difference in tone, however, certain generalizations can be made about the verbs of likelihood in Texts 5.33 and 5.34, as well as 5.29 and 5.30 for Asian countries. Because they are used with first or second person pronouns, these

complement clauses are an involved feature, involving either the speaker (*I*) or the listener (*you*) or both (*we*) and overall they can be used to persuasive effect. However, in the US text they attempt to persuade the audience of the legitimacy of the speaker's beliefs while in the text from Israel they attempt to persuade the audience of the fallacy of the audience's beliefs.

The texts from the highest military spenders also make frequent use of adverbs of place and in particular *here*, which is often used to emphasize the unity and collaboration of the UN member states present during the General Debate. This is illustrated in Text 5.35, excerpts from Jordan, where *here* appeared six times in just 1,358 words.

Text 5.35. Two excerpts from Jordan (key feature in italics)

> By gathering *here* today, we acknowledge that the power of working together far exceeds any individual effort.
> *Here*, together, we can and must create the future our people need: a safer, stronger world of coexistence, inclusion, shared prosperity and peace.

These two key features, *that* clauses controlled by verbs of likelihood and adverbs of place, serve different functions. Coupled with the fact that no keywords emerged for this group of countries, it would be misleading to propose any firm interpretations about overall discourse patterns for this group.

5.4.7. Low religious non-affiliation

Only one key feature was identified for the 29 countries with low religious non-affiliation (e.g., Afghanistan, Fiji, Georgia, Uganda): infrequent occurrences of effort verbs followed by the infinitive (see Table 5.26).

Table 5.26. Key Grammatical Features for Low Religious Non-Affiliation

Feature	Low Religious Non-Affiliation		UNGA average
	Frequency	Cohen's *d*	
Effort verb *to*	1.0	-0.97	1.7

Note. Frequencies normed per 1,000 words.

In order to interpret the low frequency of effort verbs followed by the infinitive, it is important first to understand how and why they are used. When effort verbs followed by the infinitive are used, it is generally to make calls for or express support for joint endeavors to attain UN objectives, as shown in Texts 5.36 and 5.37. This structure therefore serves persuasive purposes.

Text 5.36. Excerpt from Slovakia (key feature in italics)

> We should collectively seek practical means that will *enable* the current system *to* live up to its original purposes...That will *enable* the United Nations *to* meet the weighty demands which are made upon it today, tomorrow and in many years to come...

The text from Slovakia tries to persuade all UN member states to work together to meet the challenges facing the organization.

Text 5.37. Excerpt from the Bahamas (key feature in italics)

> We welcome the launch on 30th September of the SIDS DOCK initiative, which was jointly developed by the Caribbean Community Climate Change Centre (5C's) and the Secretariat of the Pacific Regional Environment Programme (SPREP), in co-operation with the Alliance of Small Island States (AOSIS). The initiative is meant to *assist* the Small Island Developing States (SIDS) in the Pacific and Caribbean regions *to* address the impacts of climate change, provide them with a collective, institutional mechanism to assist them in transforming their national energy sectors into a catalyst for sustainable economic development, and *help* generate financial resources *to* address adaptation to climate change.

The text from the Bahamas expresses support for joint efforts in helping SIDS tackle the effects of climate change and pursue sustainable economic development.

Overall, however, effort verbs followed by the infinitive are relatively infrequent for the low religious non-affiliation category. The low frequency of this grammatical structure may indicate less overt persuasion for this group, but the connection between less overt persuasion and low religious non-affiliation is not readily apparent from the texts. Moreover, with no other key lexical or grammatical forms supporting a relationship between persuasion and religious non-affiliation, it is difficult to definitively make a connection between the two.

5.4.8. Gross Domestic Product

Large Cohen's *d* scores were calculated for grammatical features in three levels of GDP: medium, high, and very high. Each level will be examined in turn.

5.4.8.1. Medium GDP

Two related key features were identified for the 12 countries in the medium GDP category (e.g., Antigua and Barbuda, Croatia, Greece, and Trinidad and Tobago): infinitives and effort verbs followed by the infinitive (see Table 5.27).

Table 5.27. Key Grammatical Features for Medium Gross Domestic Product (GDP)

Feature	Medium GDP		UNGA average
	Frequency	**Cohen's *d***	
Effort verb *to*	2.3	0.80	1.7
All infinitives	20.0	0.86	17.6

Note. Frequencies normed per 1,000 words.

As noted in Section 5.4.7, effort verbs followed by the infinitive are used to persuade, often in relation to achieving joint goals. Other infinitives (*need to, would like to, an opportunity to)* are used for a similar function, as shown in Texts 5.38 and 5.39.

Text 5.38. Excerpt from the Seychelles (key feature in italics)

> We have therefore a duty and an obligation *to change* the world we live in. *To make* it a better place. For the present generation and for future generations.

In the text from the Seychelles, the infinitives *to change* and *to make* are used in a passage on the responsibility of the entire international community to make the world a better place. The nouns *duty* and *obligation* preceding these infinitives present opinion as a factual imperative.

Text 5.39. Excerpt from Slovenia (key feature in italics)

> We continue *to* strongly *support* the International Criminal Court that plays an important role in establishing the rule of law. Slovenia will continue *to advocate* for a strong and independent Court, as well as universality of the Rome Statute and the Kampala amendments, and we call upon States *to offer* full and prompt cooperation with the International Criminal Court.

The infinitives in the text from Slovenia are also used for persuasive ends. The passage begins by asserting Slovenia's support of the International Criminal Court (*we continue to support, continue to advocate*) and ends by asking other UN states to support the Court as well (*we call upon States to offer full and prompt cooperation*). This excerpt illustrates a common rhetorical device in UNGA addresses whereby speakers first show steps taken by their country before asking the rest of the community to do the same. Because similar patterns were found for the high GDP group, these findings are discussed further in Section 5.4.8.2.

5.4.8.2. High GDP

Countries in the high GDP category include the Bahamas, Brunei Darussalam, Cyprus, and Malta. A number of stance structures were frequent for this group: effort verbs followed by the infinitive, suasive verbs, all stance *that* complement clauses, *that* complement clauses controlled by an attitudinal noun, *that* complement clauses controlled by an attitudinal verb, and more generally *that* complement clauses controlled by a verb (see Table 5.28).

Table 5.28. Key Grammatical Features for High Gross Domestic Product (GDP)

Feature	High GDP		UNGA average
	Frequency	**Cohen's *d***	
Discourse particle	0.8	0.93	0.3
Suasive verbs	2.7	0.93	1.4
Effort verb *to*	2.7	0.82	1.7
VCC all	4.7	1.10	3.2
VCC attitude verb	2.7	0.98	1.7
Attitude noun complement clause	0.5	1.16	0.1
Stance complement clause	8.5	0.90	6.1

Note. Frequencies normed per 1,000 words.

Many of these features serve similar functions, expressing opinion and persuading, and some of them overlap, as illustrated in Texts 5.40 and 5.41.

Text 5.40. Excerpt from the Bahamas (key features in italics)

We *ask for* the support of Member-States. We *hope that* we can count on the support of all Member-States for our candidature, both to the Human Rights Council and to the Council of the International Maritime Organisation.

In Text 5.40 from the Bahamas, *ask for* is a suasive verb; *hope that* is an attitudinal verb controlling a *that* complement clause and therefore also an example of both *that* complement

clauses controlled by a verb and all stance *that* complement clauses. Both are used in requests: *ask for* being more overt in its persuasion and *hope that* more indirect.

Text 5.41. Excerpt from Brunei Darussalam (key features in italics)

> However, I *believe that* more can be done to enhance conflict prevention. On the way forward, I *agree with* our Secretary-General's *view that* we have to strengthen the partnerships between the UN and regional or sub-regional organisations.

In Text 5.41 from Brunei Darussalam, *believe that* is an example of *that* complement clauses controlled by a verb and all stance *that* complement clauses. *View that* is a *that* complement clause controlled by an attitudinal noun and therefore also included in all stance *that* complement clauses. Both are used to express opinion: *believe that* the opinion of the speaker and *view that* the opinion of the Secretary General. The text also features a suasive verb, *agree with*, again used to express personal stance.

Text 5.42. Excerpt from Cyprus (key features in italics)

> Following the non-renewal of actions which violated the exercise by the Republic of Cyprus of its sovereign rights within its exclusive economic zone, and the change in the leadership of the Turkish Cypriot community, a window of opportunity opened that revived our *hope that* the new round of negotiations which resumed this past May will lead to the final settlement of the Cyprus problem. A *hope that* is based in my *conviction that* both I and the Turkish Cypriot leader, Mr. Mustafa Akinci, share the same political courage and resolve to decisively move forward in order to materialize the joint vision of our people who desire the solution of the Cyprus problem through a viable, lasting and functional settlement.

The text from Cyprus has three occurrences of noun complement clauses, which are also stance complement clauses: two instances of *hope that* and one *conviction that*. These clauses, rather than clarify the speaker's opinion, make it more difficult to detect, particularly when presented in a series of multiple clauses (*a hope that is based in my conviction that…*). It is not

until the sixth clause of the final sentence in Text 5.42 that the President of the Republic of Cyprus arrives at the true object of his hope: *the solution of the Cyprus problem.* Several of the verbs in that last sentence (*is based in, share, resolve, move forward, materialize, desire*) suggest joint efforts and cooperation and ultimately, this passage is an extremely indirect request to the Turkish Cypriot leader to work with the Republic of Cyprus in finding a solution to the *Cyprus problem*, a reference to the Turkish occupation of the northern part of the island. While Texts 5.40 and 5.41 illustrate how stance clauses can clarify opinions and requests, Text 5.42 shows how the use of multiple stance clauses can cloud opinions and requests and are used for negative politeness, to avoid threatening face.

Finally, discourse particles were relatively frequent in this group. Text 5.43 illustrates the discourse particle *after all* in the text from Brunei Darussalam.

Text 5.43. Excerpt from Brunei Darussalam (key feature in italics)

> The success of our organisation lies in our hands. *After all*, our collective efforts and strength are the best attributes of this organisation.

The discourse particle *after all* in the text from Brunei Darussalam is used again to present opinion as self-evident fact and to call for cooperation and collaboration. However, discourse particles remain in absolute terms not particularly frequent for the high GDP group in spite of their relative frequency when compared to the rest of the UNGA: Sixty-two of the 92 countries in the UNGA corpus had no discourse particles whatsoever; in the high GDP group, the Bahamas averaged 1.7 discourse particles for every 1000 words, Brunei Darussalam 0.7, Cyprus 0, and Malta 0.8. Like with progressive verbs for the small area group (Section 5.4.5) and (infrequent) effort verbs for the low religious non-affiliation group (Section 5.4.7), the difference in the frequency of discourse particles for the high GDP group is not particularly meaningful.

For stance features, however, the combined results from the medium and high GDP groups suggest that these groups prefer expressing opinion as personal belief, citing the opinion of others, or presenting an opinion as an objective fact in making a case for an important issue rather than making an explicit request. Sometimes these expressions of stance mitigate the force of the proposition and other times (e.g., Text 5.42) they come close to concealing the proposition altogether. Three of these stance features (effort verbs followed by the infinitive, all infinitives, suasive verbs) are associated with overt persuasion (see Section 5.5.5 on Dimension 4). However, the stance features used most frequently by these two GDP groups do not have nearly the same directive force as the stance features in the very high GDP group (see Section 5.4.8.3, see also Section 5.5.5 for a final synthesis of stance and persuasion for all country groups).

5.4.8.3. Very high GDP

A number of grammatical features were identified for the 14 very high GDP countries (e.g., Australia, Israel, New Zealand, US), as shown in Table 5.29.

Table 5.29. Key Grammatical Features for Very High Gross Domestic Product (GDP)

Feature	Very High GDP		UNGA average
	Frequency	**Cohen's _d_**	
Demonstrative pronoun	5.0	0.85	3.5
All verbs	124.8	1.19	113.0
Mental verbs	16.8	1.10	12.7
Action verbs	19.4	0.95	16.1
Present tense	89.5	1.21	77.3
Passive postnominal modifier	0.9	-0.94	1.7
All modals	18.8	1.03	14.4
Possibility modals	5.6	0.88	3.6
Common nouns	237.3	-0.83	250.4
Prepositions	115.4	-0.92	124.9
All adjectives	81.2	-0.83	88.9
Epistemic predicative adjective	0.9	0.84	0.4
Attributive adjective	60.9	-0.86	68.5
Phrasal coordinating conjunction	2.9	-1.02	4.8
Adverbial conjunct	3.1	-0.87	4.6

Note. Frequencies normed per 1,000 words.

For texts from the very high GDP group, stance is expressed with greater frequency relative to other groups through epistemic predicative adjectives and modals: modals generally and particularly modals of possibility. Modals, even modals of possibility, can be used as directives with oftentimes greater underlying force than the expressions of stance identified for medium and high GDP groups (e.g., complement clauses). Texts 5.44 and 5.45 illustrate the directive force of modals).

Text 5.44. Excerpt from Iceland (key features in italics)

> And without peace and security we *will* not achieve sustainable development. At the same time peace and security *cannot* be used as a justification for violating human rights. The imposition of the death penalty *should* not be justified under the appearance of maintaining order and security.

The text from Iceland uses modals (*will not, cannot, should not*) to express strong opinion about the need for peace and security. Similarly, in Text 5.45 from Austria, the speaker discusses the importance of unity in the Security Council to fight against Daesh (ISIL) and to protect civilians, thereby ensuring peace and security.

Text 5.45. Excerpt from Austria (key features in italics)

> We *will* only be able to make a difference if the United States and Russia pull in one direction. The UN Security Council *has to* demonstrate unity. We need a mandate not only for action against Daesh but also for the protection of civilians. This *should* include the establishment of safe and buffer zones which *would* increase the security of the people and make access for humanitarian aid much easier. Action by the Security Council *would* offer many countries the possibility to join in and offer their support.

Texts 5.44 and 5.45 support the findings from Sections 5.3.7.4 and 5.3.8 regarding the emphasis on preventing conflict in texts from the highest development groups.

Epistemic predicative adjectives also express stance, presenting opinion as factual necessity: *It is crucial that we make our joint efforts on peace and security more efficient and effective* [Netherlands]; *International cooperation at all levels is essential* [Norway]; *It is necessary to get everybody on board to make the commitments a reality* [Finland]. The adjectives *crucial, essential,* and *necessary* express necessity. Like modals of necessity, these adjectives are forceful. Their directness serves as a stark contrast to the evasiveness of the five embedded clauses in Text 5.42 from Cyprus.

These three stance-related features (epistemic predicative adjectives, modals, and modals of possibility) are much more direct in their expression of opinion and persuasion. As such, they are easier for the listener to process. In contrast, the text from Cyprus is difficult to follow. Even in the text from Brunei Darussalam (…*I agree with our Secretary-General's view that we have to strengthen the partnerships*…), the opinion being expressed does not come until after two

expressions of stance. *Agree with* and *view that* make the directive *we have to strengthen* more indirect. The modals in the text from Iceland and Austria are not prefaced with other expressions of stance. They appear within the first few words of the sentence, making them clear and obvious to the listener.

The remaining features in the very high GDP group are related to one another; they all signal involved (interpersonal) language. Highly frequent verbs overall, action verbs, mental verbs, present tense verbs, and demonstrative pronouns are typical of involved language. Conversely, many features that are typical of dense, abstract, informational texts are infrequent for this country category: prepositions, coordinating conjunctions, adverbial conjuncts (e.g., *however*), postnominal passives, attributive adjectives, and common nouns. Though adjectives are infrequent in attributive position, adjectives generally are frequent and so are epistemic adjectives in predicative position. Again, this is indicative of more clausal rather than phrasal structures, which results in a less dense and more involved style.

Overall, these features come together to create texts which sound like they were meant to be spoken (see Texts 5.46 and 5.47). In Texts 5.46 and 5.47 informational features (prepositions, nouns, phrasal coordinating conjunctions, attributive adjectives) are underlined and involved features (first and second person pronouns, *it*, contractions, present tense verbs, clausal coordinating conjunctions, *that* deletion, emphatics, sentence final prepositions) are italicized.

Text 5.46. Excerpt from the European Union (involved features in italics; informational underlined)

> *I am* here today to reassure *you* that Europe *is* as committed to *its* values *and* objectives now, as *it has* ever *been*: Europe will stay the course, even though *it is* now *confronting* challenges unseen *and* unheard of since for decades. Wars *are raging* both to the South and to the East of our borders. European leaders *are tackling* the consequences of borders being changed on our continent by force, like in Ukraine, in violation of the Charter of the United Nations. *We are* also *dealing* with the refugee crisis, with terrorist attacks

inside Europe and with economic difficulties in some Member States. As difficult as the situation is, *I am* sure that *we* will cope with it, *and* most importantly, *we* will, at the same time, remain dedicated to helping make the whole world a better place.

The text from the European Union sounds like a spoken text. The numerous first person pronouns (singular and plural) involve the speaker and the audience. Structures are primarily clausal with clausal coordinating conjunctions (e.g., *we will cope with it, and most importantly, we will…*) and subordinators (e.g., *Europe will stay the course, even though it is now confronting*) that are easy for the listener to process in real time. Noun phrases are not particularly complex, with only a few instances of multiple prepositional phrases and attributive adjectives (e.g., *with the refugee crisis, with terrorist attacks inside Europe and with economic difficulties in some Member States*).

Text 5.47. Excerpt from New Zealand (involved features in italics; informational underlined)

Since *we have been* on the Council, *we have found it* as hard as many of *you* warned. The Council's agenda *is* lengthy and contentious. The dynamics within *it are* difficult. When Council members *work* together, as they did with the recent agreement over Iran's nuclear capability, they can still deliver on the Council's role as set out in the Charter.

The text from New Zealand is also primarily clausal, again with subordinators (e.g., *since*) and short sentences (*The dynamics within it are difficult*) making it easy to process.

Text 5.48 from Kenya, a low GDP country, serves as a contrast; this highly informational text is easier read than heard because the dense phrasal structures require greater processing time.

Text 5.48. Excerpt from Kenya (involved features in italics; informational underlined)

The adverse effects of globalization, the debilitating consequences of the global economic and financial crisis and the negative effects of climate change *have increased* the vulnerability of the global economy *and affected* United Nations member states to

259

varying degrees. Today, just as in the last 70 years, the challenges and aspirations faced by developing countries, including peace and security, development, protection of human rights, environmental conservation and global governance *remain* as real and relevant as ever, *and* they *are expected* to continue to be of primary concern for a number of United Nations Member States in the coming years.

The text from Kenya has very long noun phrases which are difficult for the listener to process. The subject of the first sentence includes three noun phrases connected with a phrasal coordinating conjunction. Each noun phrase has at least one attributive adjective and one prepositional phrase. As a result the main verb (*increased*) does not come until the 24th word in the sentence. By the time it arrives, the listener is likely to have forgotten the first noun phrase of the subject. Similarly, in the second sentence, the main verb (*remain*) is the 31st word. The verb is separated from the first noun of its subject (*challenges*) by three phrasal coordinating conjunctions, two prepositional phrases, and two non-finite clauses for a total of twenty words between the subject and the verb. These dense, information-packed structures are common in written texts (see Biber, 1988), but are extremely difficult for a listener to follow.

Overall, the very high GDP group prefers more direct language, with more involved features and explicit expression of opinion and persuasion. This is discussed further in Sections 5.5.2 and 5.5.5.

5.4.9. Human Development Index

Three levels of the Human Development Index (HDI) showed large Cohen's *d* values: low, medium, and very high. Each level will be examined in turn.

5.4.9.1. Low HDI

The 19 countries in the low HDI category (e.g., Pakistan, the Solomon Islands, Swaziland, Zimbabwe) had infrequent modals overall and in particular infrequent modals of necessity (see Table 5.30).

Table 5.30. Key Grammatical Features for Low Human Development Index (HDI)

Feature	Low HDI		UNGA average
	Frequency	**Cohen's *d***	
All modals	10.8	-1.06	14.4
Necessity modals	3.3	-0.85	5.0

Note. Frequencies normed per 1,000 words.

Modals are generally used as directives, as discussed in Section 5.4.8.3, but the low HDI group dispreferred such strong directive force. Similar findings were presented in Section 5.4.4 on Landlocked Developing Countries (LLDCs), which is not surprising given the overlap between LLDC and low HDI. Overall, the results from the development categories (Sections 5.4.4., 5.4.8, and 5.4.9) show countries using progressively more overt persuasion as development level increases. This is discussed further in Section 5.5.5.

5.4.9.2. Medium HDI

The 15 medium HDI countries (e.g., Bhutan, Guyana, Moldova, South Africa) made frequent use of prepositions (see Table 5.31).

Table 5.31. Key Grammatical Features for Medium Human Development Index (HDI)

Feature	Medium HDI		UNGA average
	Frequency	**Cohen's *d***	
Prepositions	132.8	0.80	124.9

Note. Frequencies normed per 1,000 words.

Prepositions are associated with dense phrasal structures and more informative texts. With only one key feature to emerge from the analysis, it is difficult to draw any firm conclusions. However, it appears to be part of a larger general pattern whereby less developed countries opt for more informational language and more developed countries for more involved language. This is discussed in further detail in Section 5.5.2.

5.4.9.3. Very high HDI

Several features were identified for the 25 very high HDI countries (e.g., Australia, Malta, Norway, UK): frequent verbs and present tense verbs; infrequent prepositions and postnominal passives (see Table 5.32).

Table 5.32. Key Grammatical Features for Very High Human Development Index (HDI)

Feature	Very High HDI		UNGA average
	Frequency	**Cohen's *d***	
All verbs	120.8	0.88	113.0
Present tense	85.8	1.00	77.3
Passive postnominal modifier	1.1	-0.95	1.7
Prepositions	117.9	-0.81	124.9

Note. Frequencies normed per 1,000 words.

The two frequent structures, verbs and present tense, are associated with involved language. The two infrequent structures, passive postnominal modifiers and prepositions, are associated with more informational language. Thus, the same conclusions can be drawn for very high GDP and very high HDI countries: Texts from the most developed countries tend to be more involved and interpersonal, less informational. This is discussed and illustrated in further detail in Section 5.5.2.

5.4.10. Summary

For the vast majority of the country categories, no key features were identified. In this respect, the results of the grammatical analysis are similar to the results of the lexical analysis in Section 5.3. Overall, little linguistic variation occurs in the UNGA, and the linguistic variation that does occur is detected primarily in the analysis of texts from individual countries rather than country categories. (See Section 5.6 for results of individual country analyses.)

However, the analysis of key grammatical features did reveal a few patterns. First, some texts are more informational, with dense noun phrases, and others are more involved, with higher frequencies of verbs and clauses. For instance, the most developed countries (very high HDI and GDP) tend to use more involved language. This is discussed in further detail in Section 5.5.2. Another textual difference is in expression of stance. Some countries express opinion and requests indirectly. These tend to be less developed states (Asia, landlocked developing countries, low HDI, and to a lesser extent medium and high GDP). Only very high HDI and very high GDP groups frequently used modals for directives with somewhat stronger directive force. This is discussed in further detail in Section 5.5.5.

5.5. Variation in UNGA Grammatical Features: Multi-Dimensional Scores

5.5.1. Introduction

The objective of Section 5.5 is to further explore patterns of variation in grammatical structures by examining Multi-Dimensional (MD) scores (see Section 3.7.2 for more on MD analysis). MD analysis investigates patterns of co-occurring grammatical features so this section builds on, adds to, and synthesizes many of the key feature findings from Section 5.4. In Sections 5.5.2 through 5.5.6, I present dimension scores by country group and discuss the most notable scores. On the whole and not surprisingly given the findings from Sections 5.3 and 5.4, few

263

striking differences are identified in dimension scores. However, some interesting patterns do emerge, revealing some distinctions in discourse style based on country category.

All dimension scores are plotted in Figures 5.1-5.40. Each figure shows scores for country groups in the same category (e.g., region), along with scores for the UNGA as a whole and for the other registers from Chapter 4 for comparison purposes. Though the objective of this section is not to contrast country groups with the comparison registers, the scores from the comparison registers illustrate the importance of the difference between country groups. For instance, in Figure 5.1, the difference between regions along Dimension 1 might at first seem considerable, but when compared to the other registers, the difference is more evidently minimal.

5.5.2. Dimension 1: Involved versus informational production

Dimension 1 represents the degree to which the language of a text is involved or, on the contrary, informational (see Section 4.5.1 for a description of the dimension and the features associated with the dimension). Figures 5.1-5.8 show Dimension 1 scores for the 31 country groups compared to the UNGA as a whole as well as the comparison registers from Chapter 4.
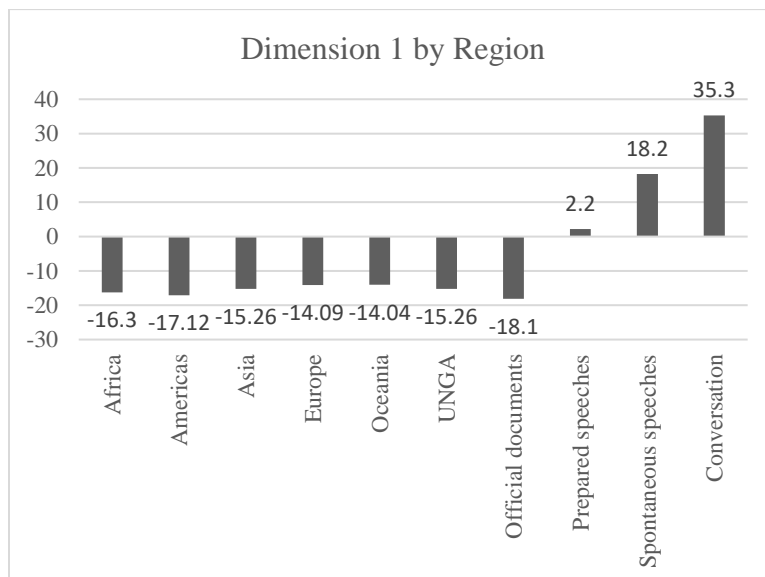
Figure 5.1. Dimension 1 scores by region.

Figure 5.2. Dimension 1 scores by area.

Figure 5.3. Dimension 1 scores by population.
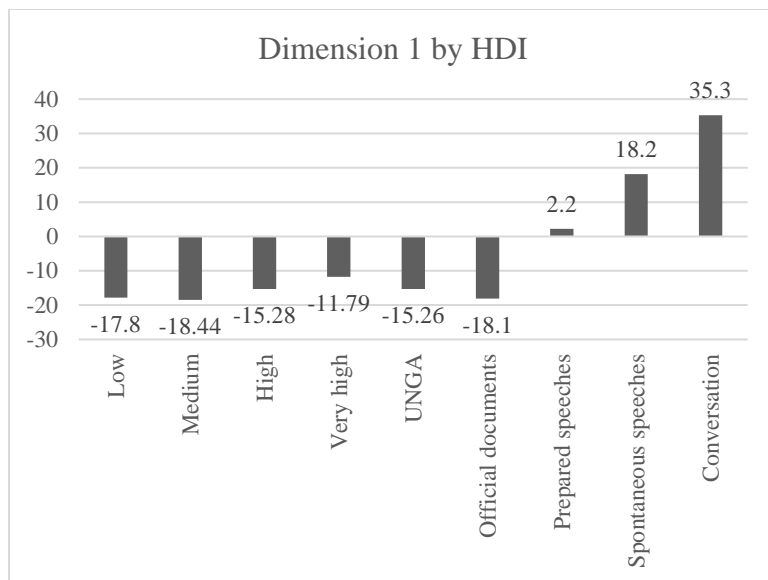
Figure 5.4. Dimension 1 scores by Human Development Index.
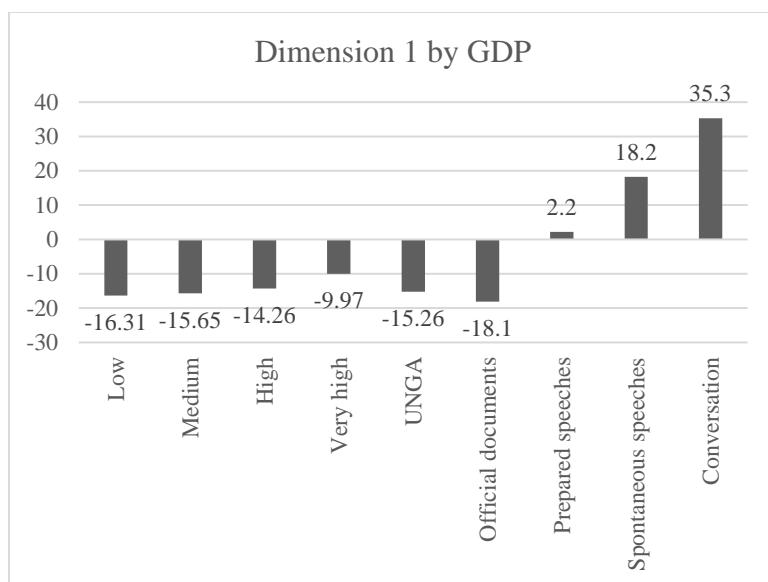
Figure 5.5. Dimension 1 scores by Gross Domestic Product.
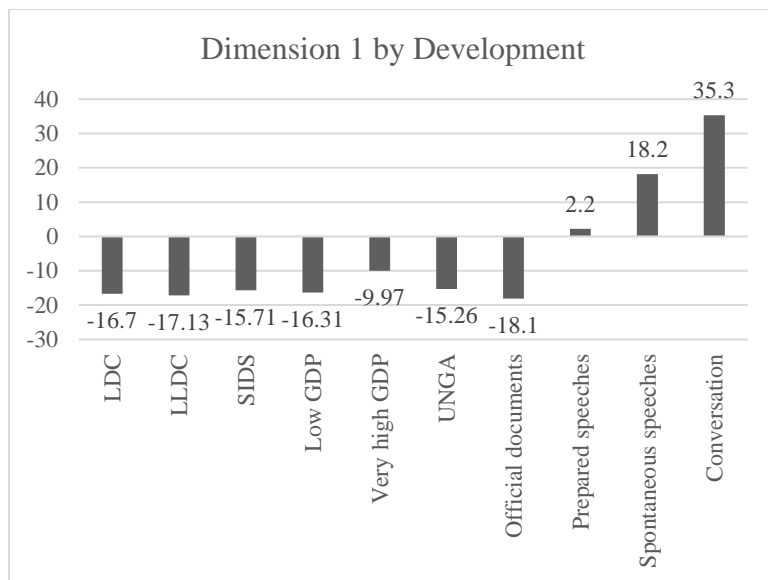
Figure 5.6. Dimension 1 scores by UN development group.
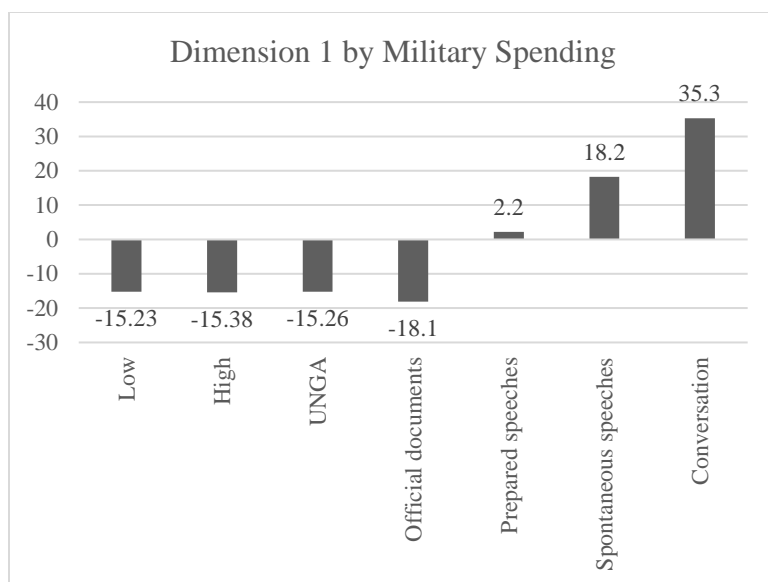
Figure 5.7. Dimension 1 scores by military spending.
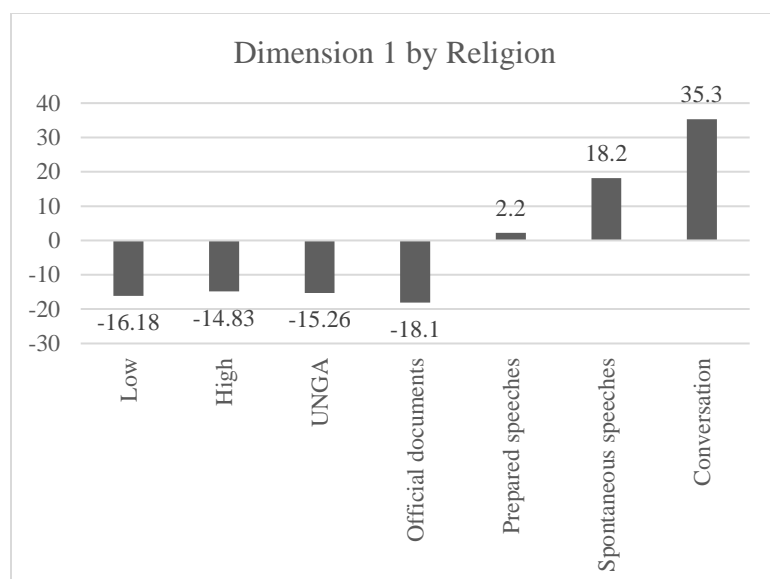
Figure 5.8. Dimension 1 scores by religious non-affiliation.

Few country groups had notably higher or lower Dimension 1 scores compared to the

UNGA average. The only pattern to emerge is for level of development. As shown in Figures

5.4-5.6, overall, texts from countries with lower development are more informational; texts from

countries with higher development are more involved. Thus, texts from low GDP countries more

closely resemble official documents; they sound like written texts. In contrast, texts from high

GDP countries begin to approach other prepared speeches in terms of involvement; they sound

more like spoken texts. Text 5.49 illustrates the highly informational style of a low GDP country,

Papua New Guinea; Text 5.50 illustrates the relatively involved style of a very high GDP

country, Israel. In these texts, involved features (first and second person pronouns, *it*,

contractions, present tense verbs, clausal coordinating conjunctions, *that* deletion, emphatics,

sentence final prepositions) are italicized and informational features (prepositions, nouns,

attributive adjectives) are underlined.

268

Text 5.49. Excerpt from Papua New Guinea (involved features italicized; informational underlined)

> The <u>Government of Papua New Guinea</u> *has taken* <u>enabling national policies, strategies</u> <u>and plans</u> to providing a <u>conducive environment</u> to achieving the <u>MDGs</u> and the <u>SDGs</u>. These <u>policies</u> <u>in</u> the <u>areas</u> <u>of</u> <u>population</u>, <u>water</u>, <u>sanitation</u> <u>and</u> <u>hygiene</u>, <u>free</u> <u>basic</u> <u>health</u> <u>services</u>, and <u>tuition</u> <u>fee</u> <u>free</u> <u>education</u>. <u>Free Education Policy</u> *has seen* an <u>additional</u> 2 million <u>children</u> enrolled <u>in</u> <u>schools</u>, <u>in</u> the last two <u>years</u>.

This informative passage is made up primarily of long phrasal structures with frequent nouns (*environment*), attributive adjectives (*conducive*), and prepositions (*in*). The text in no way involves either the speaker or the audience. It could very well have been published as a written government document intended for an unknown audience.

In contrast with this text from Papua New Guinea, the highly involved Text 5.50 from Israel sounds less formal, even conversational.

Text 5.50. Excerpt from Israel (involved features italicized; informational underlined)

> *I* ended that first speech by saying: Gentlemen, check *your* fanaticism at the door. More than three decades later, <u>as</u> the Prime <u>Minister</u> <u>of</u> Israel, *I am* again privileged to speak from this podium. *And* for *me*, that privilege *has* always *come* with a moral <u>responsibility</u> to speak the truth. *So* after three days <u>of</u> listening <u>to</u> world leaders *praise* the nuclear deal <u>with</u> Iran, *I begin* my speech today by saying: Ladies <u>and</u> Gentlemen, check your enthusiasm at the door. *You see*, this deal *doesn't make* peace more likely.

Text 5.50 includes first (*me*) and second (*you*) person pronouns, contractions (*doesn't*), present tense verbs (*begin*), emphatics (*you see*), and non-phrasal coordination (*and*). These features involve both the speaker (*I ended that first speech*) and the audience (*you see*). The overall result is a speech that is easy for the listener to process.

Dimension 1 scores thus support the keyword analysis from Section 5.3.7.4 identifying several contractions (an involved feature) for the very high GDP group. Dimension 1 scores also support the key feature analysis from Section 5.4.8.3 identifying numerous involved features for

269

the very high GDP group (e.g., frequent verbs and present tense; infrequent common nouns and prepositions), from Section 5.4.9.2 identifying frequent prepositions (an informational feature) for the medium HDI group, and from Section 5.4.9.3 identifying frequent verbs and present tense (two involved features) and infrequent prepositions (an informational feature) for the very high HDI group. Overall, countries in the lower development groups opt for more informational, less involved language. This makes their addresses sound more like written texts. The language is more formal but also harder for the listener to process. Speechwriters from countries with lower levels of development may assume their text is going to be read rather than heard (see Sections 4.3 and 5.2 for a discussion on production circumstances). Alternatively, they may give preference to text formality, making the address sound as though it were meant to be read, even at the expense of aural comprehensibility. This second interpretation is supported by the lexical analysis in Section 5.4, showing a preference for formal rhetorical conventions in texts from less developed countries.

Though this general trend can be seen in the average scores for these country groups, there is a large standard deviation for Dimension 1 (see Table 4.12 and Appendices B-K), making it difficult to assert a consistently clear pattern for all countries in these groups. Finally, it is important to bear in mind that while the more developed countries produce texts that are involved relative to the less developed countries, they are nonetheless less involved than the prepared and spontaneous speeches from Biber (1988) and much less involved than conversation.

5.5.3. Dimension 2: Narrative versus non-narrative concerns

Dimension 2 represents the degree of narration in a text (see Section 4.5.2 for a description of the dimension and the features associated with the dimension). Figures 5.9-5.16

show Dimension 2 scores for the 31 country groups compared to the UNGA as a whole as well as the comparison registers from Chapter 4. As shown in Figures 5.9-5.16, no patterns emerge by country group for narrative focus. All country groups converge around the UNGA average with very little difference in group scores.

Dimension 2 by Region

Category	Score
Africa	-3.05
Americas	-3.28
Asia	-2.95
Europe	-3.09
Oceania	-3.11
UNGA	-3.08
Official documents	-2.9
Prepared speeches	0.7
Spontaneous speeches	1.3
Conversation	-0.6
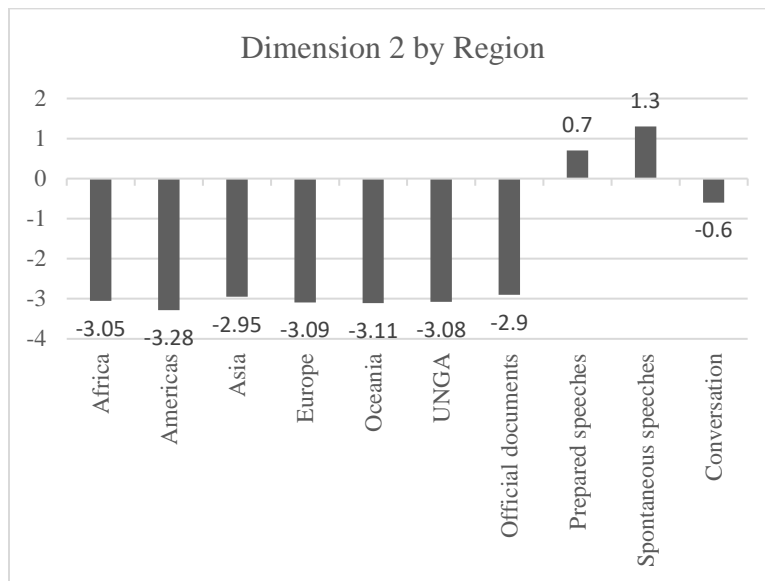
Figure 5.9. Dimension 2 scores by region.

Dimension 2 by Area

Category	Score
Small	-3.19
Medium	-3.04
Large	-3.08
Very large	-3
UNGA	-3.08
Official documents	-2.9
Prepared speeches	0.7
Spontaneous speeches	1.3
Conversation	-0.6
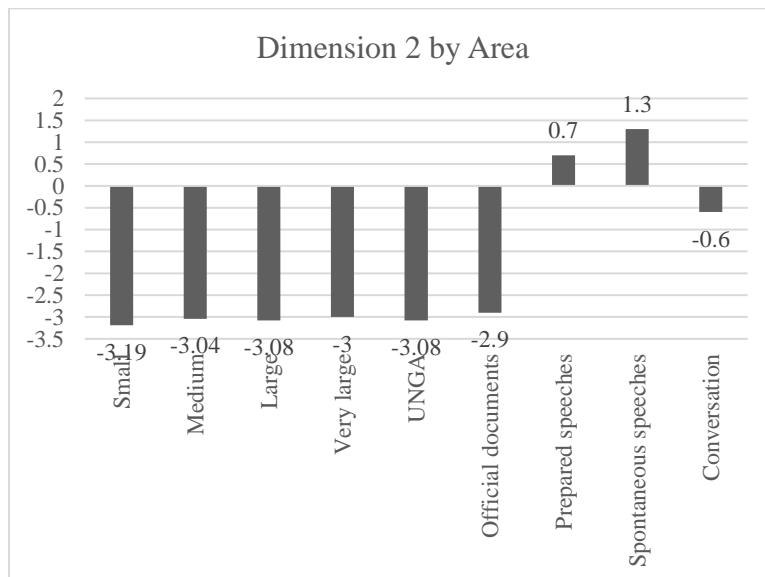
Figure 5.10. Dimension 2 scores by area.

271
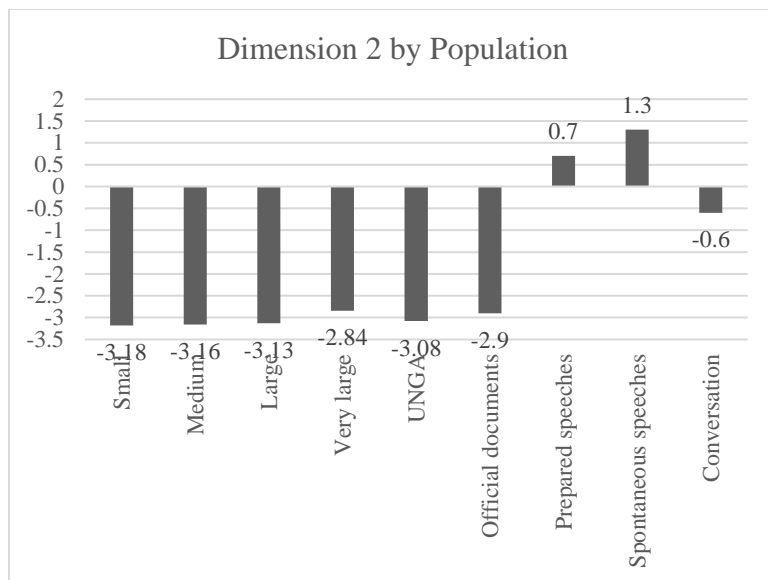
Figure 5.11. Dimension 2 scores by population.
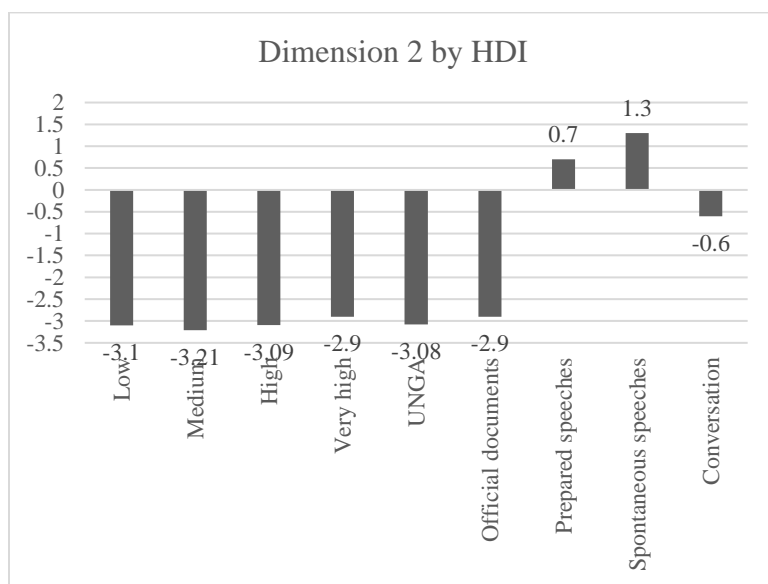
Figure 5.12. Dimension 2 scores by Human Development Index.
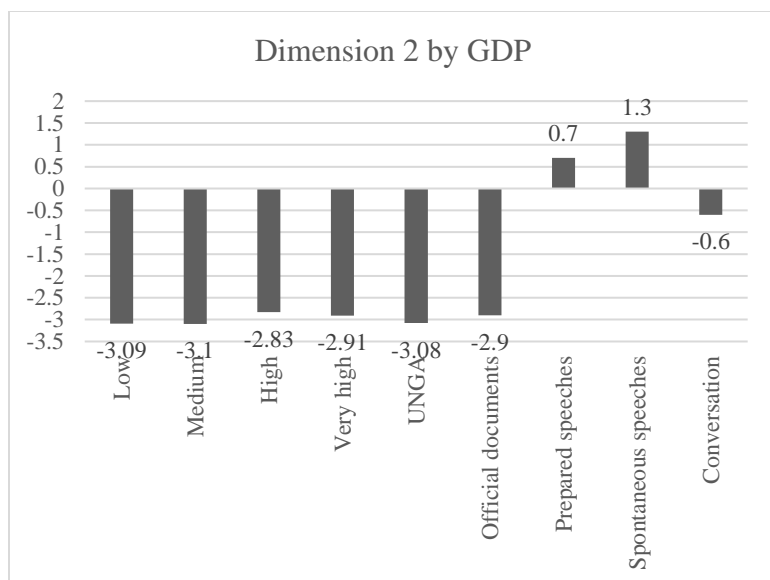
Figure 5.13. Dimension 2 scores by Gross Domestic Product.
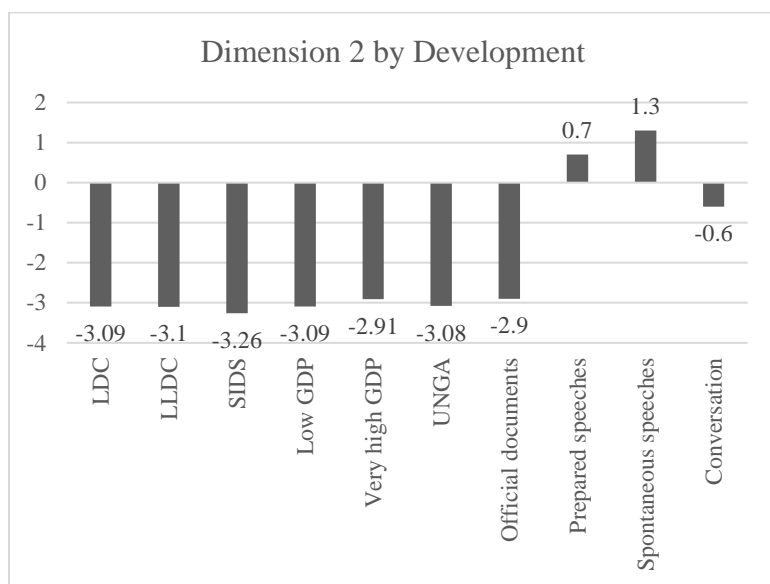
Figure 5.14. Dimension 2 scores by UN development group.
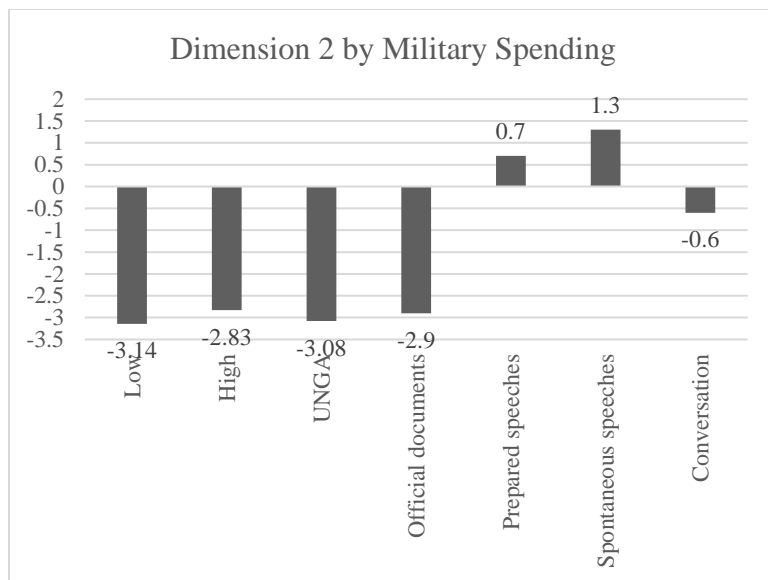
Figure 5.15. Dimension 2 scores by military spending.
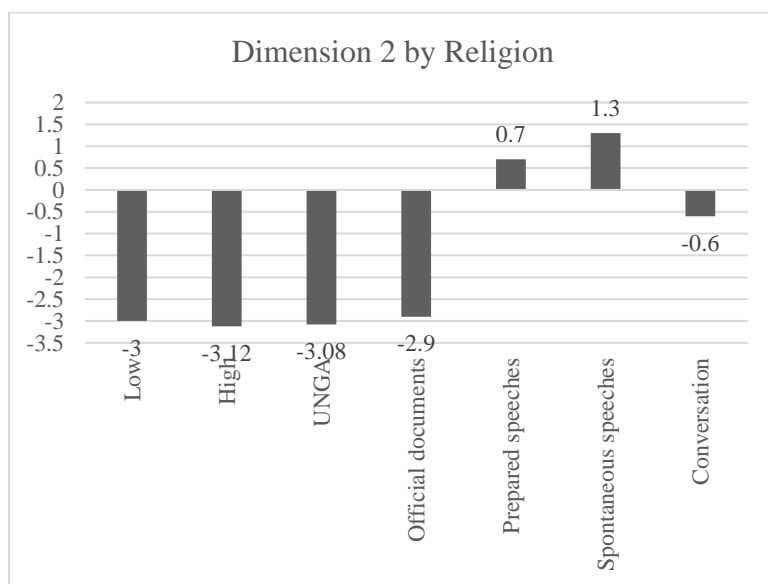
Figure 5.16. Dimension 2 scores by religious non-affiliation.

5.5.4. Dimension 3: Explicit versus situation-dependent reference

Dimension 3 represents the degree to which reference in a text is explicit or, on the contrary, situation-dependent (see Section 4.5.3 for a description of the dimension and the features associated with the dimension). Figures 5.17-5.24 show Dimension 3 scores for the 31

country groups compared to the UNGA as a whole as well as the comparison registers from

Chapter 4.
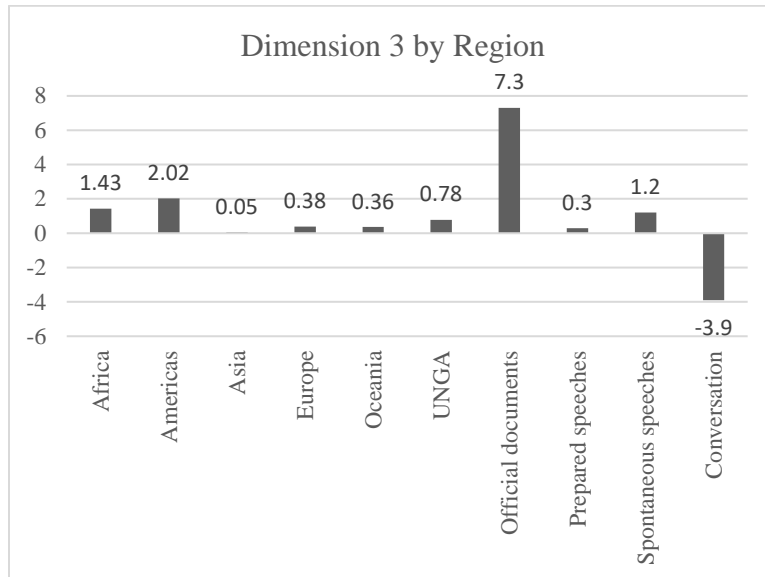
Figure 5.17. Dimension 3 scores by region.
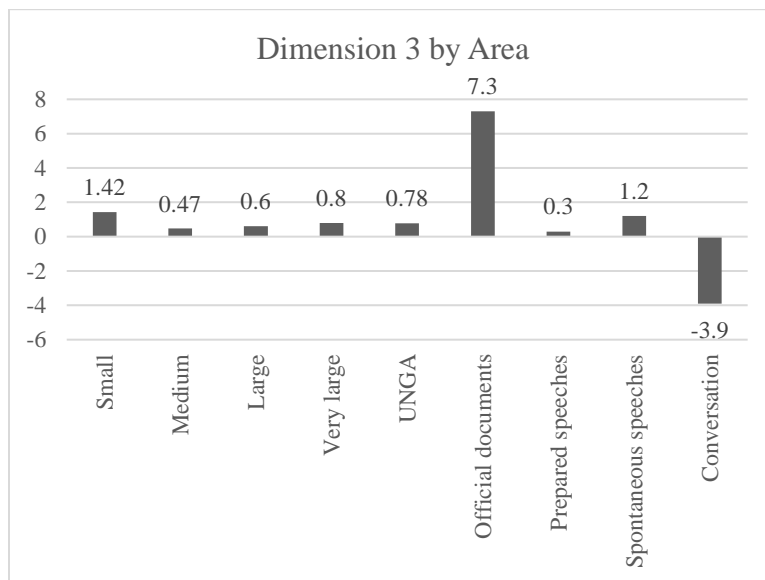
Figure 5.18. Dimension 3 scores by area.
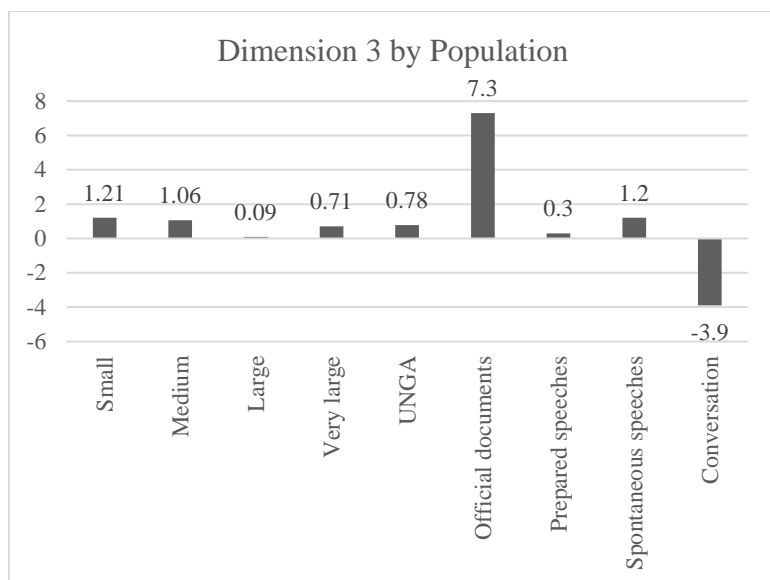
Figure 5.19. Dimension 3 scores by population.
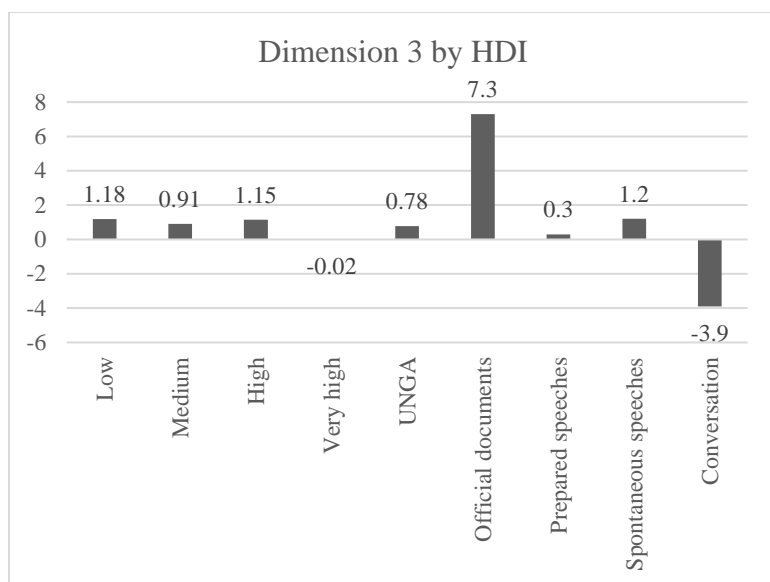
Figure 5.20. Dimension 3 scores by Human Development Index.
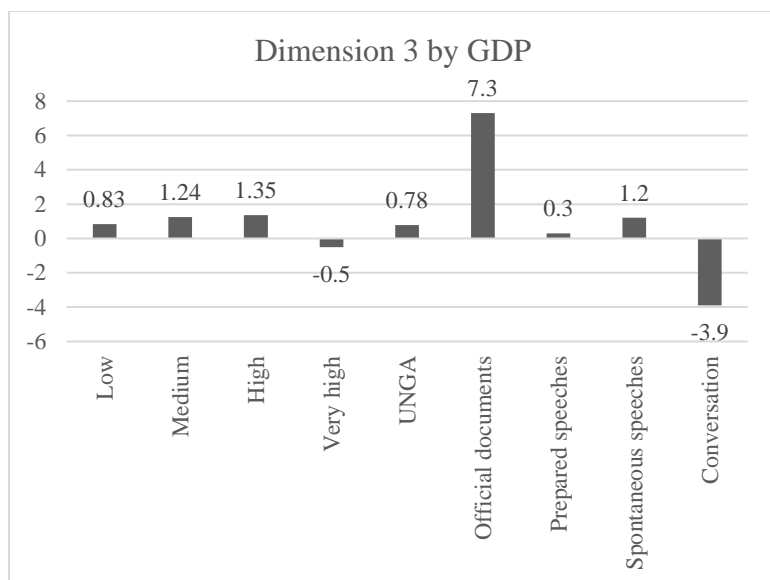
Figure 5.21. Dimension 3 scores by Gross Domestic Product.
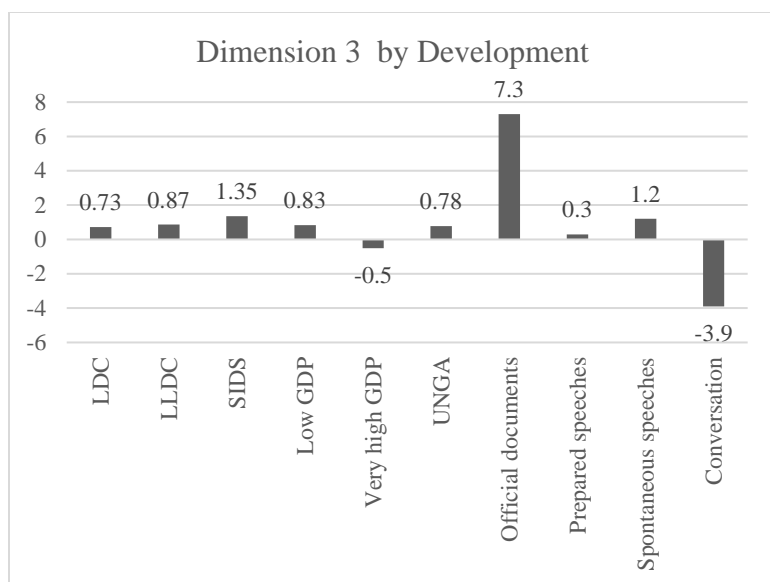
Figure 5.22. Dimension 3 scores by UN development group.
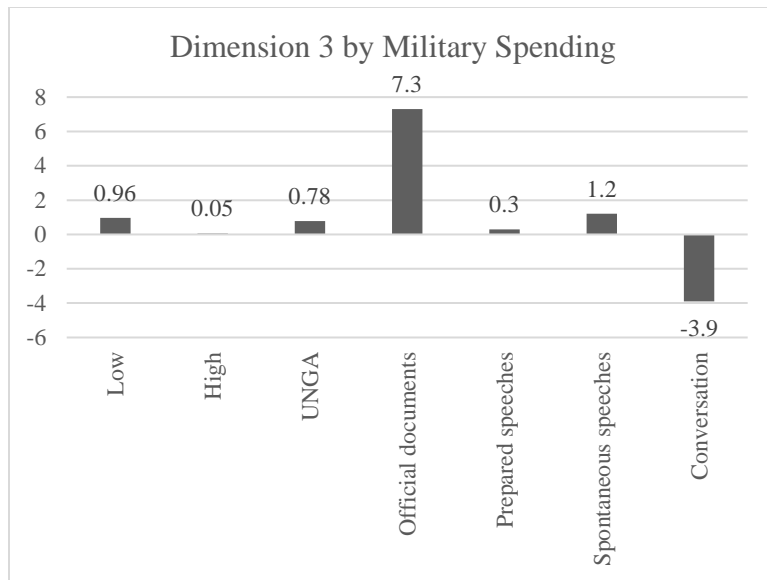
Figure 5.23. Dimension 3 scores by military spending.
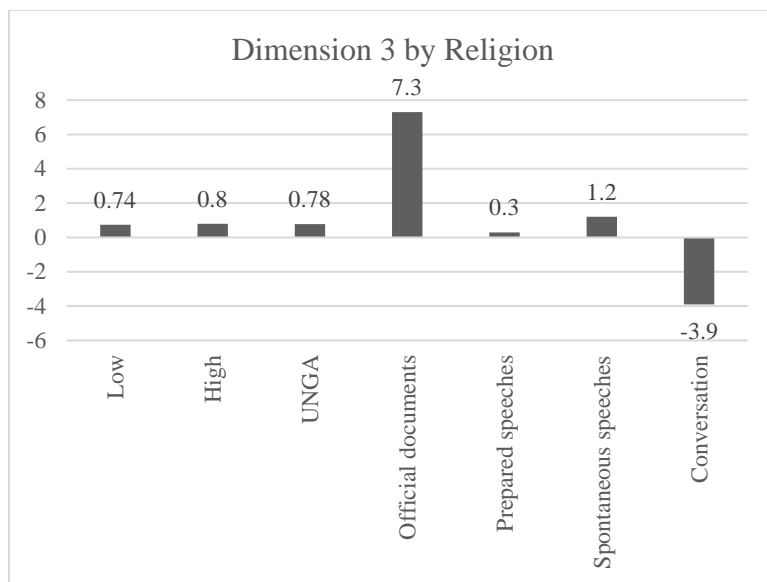
Figure 5.24. Dimension 3 scores by religious non-affiliation.

Dimension 3 scores show a few tendencies in terms of country groups. For the regional analysis, texts from Africa and the Americas have higher Dimension 3 scores, indicating more explicit reference. No one key feature associated with explicit reference is frequent enough for

either Africa or the Americas to emerge in the analysis in Section 5.4. However, when all

features associated with explicit language are examined as a whole, a pattern emerges. Overall,

there is more text-internal reference and information-packing in nominal structures, which is

typical of official documents and other formal writing, and less situation-dependent reference,

which is typical of conversation and other registers where participants share the same space and

time. The overall effect is a text that sounds more like a written document than a speech

delivered before a live audience (see Text 5.51). The result is similar to texts with informational

production, negative Dimension 1 scores. It is therefore not surprising that both Africa and Asia

have the lowest Dimension 1 scores (informational production, see Figure 5.1) and the highest

Dimension 3 scores (explicit reference, see Figure 5.17).

Text 5.51. Excerpt from Barbados (explicit features in italics)

> Seventy years on from the *creation* of the UN, it is appropriate that we reflect on the path *which* small states like mine have helped to prepare. It is a path in large part made possible by the principles this *organisation* has embraced from its *inception* and *which* are embodied in the Charter and in the UN's many treaties, *resolutions* and *plans* of *action*. A path from *exclusivity* to *universality*; a path *which* has seen a near *quadrupling* of this *organisation's membership* in 70 years; a path to an international legal *order which* seeks to balance the rights of states with their *responsibilities*, including *responsibility* to their citizens.

Explicit reference features in Text 5.51 include numerous relative clauses, both on object

positions (e.g., *which small states like mine have helped to prepare*) and subject positions (e.g.,

which are embodied in the Charter) as well as nominalizations (e.g., *creation*). There are no

situation-dependent features, no references to the setting shared by addressor and addressee, no

indication from this excerpt that the text is being delivered in real time before a live audience.

In contrast, the very high HDI and GDP groups have negative Dimension 3 scores,

indicating more situation-dependent reference, as illustrated in Text 5.52.

Text 5.52. Excerpt from Denmark (explicit features in italics; situation-dependent underlined)

<u>Now</u> we must act to show our grandchildren that we are able to deliver on our *promises*. <u>Today</u>. <u>Tomorrow</u>. And 70 years from <u>now</u>.

Text 5.52 sounds more like a speech intended for a live audience, not only with its adverbs of time (*now, today, tomorrow*), but in the short sentences with little nominal elaboration, making the passage easy for the listener to process.

Though no patterns emerge for all regions or development levels, the Dimension 3 scores for Africa, the Americas, and the very high GDP and HDI groups support the overall findings from Chapter 5 that low development countries, particularly in Africa and the Americas, produce texts that sound more literate than oral while high development countries produce texts that sound as though they were meant to be spoken not read.

5.5.5. Dimension 4: Overt expression of persuasion

Dimension 4 represents the degree to which a text is overtly persuasive (see Section 4.5.4 for a description of the dimension and the features associated with the dimension). Figures 5.25-5.32 show Dimension 4 scores for the 31 country groups compared to the UNGA as a whole as well as the comparison registers from Chapter 4.
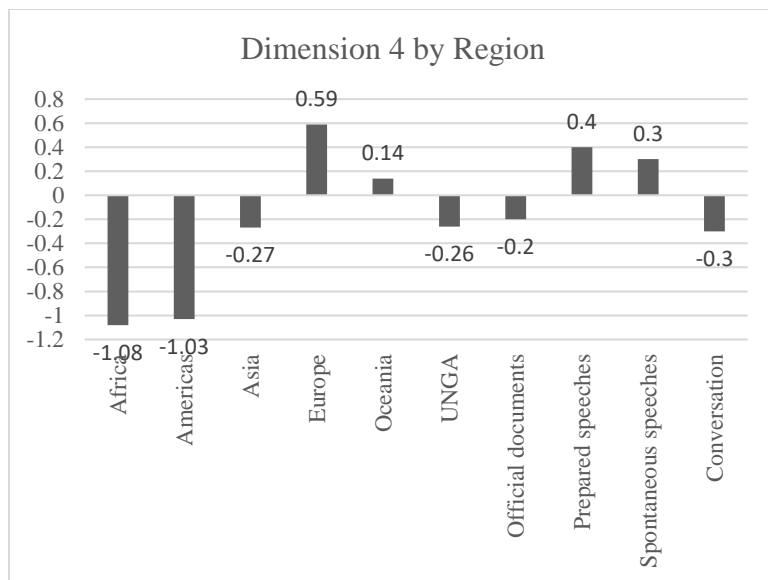
Figure 5.25. Dimension 4 scores by region.
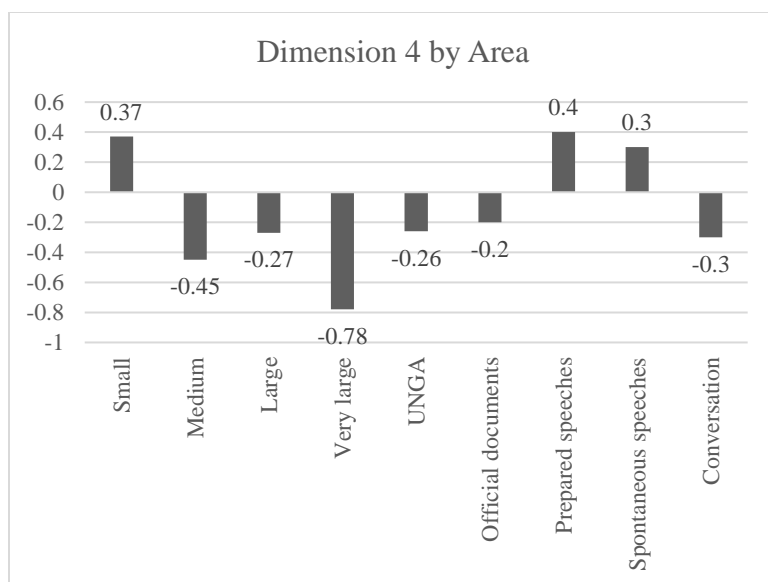
Figure 5.26. Dimension 4 scores by area.
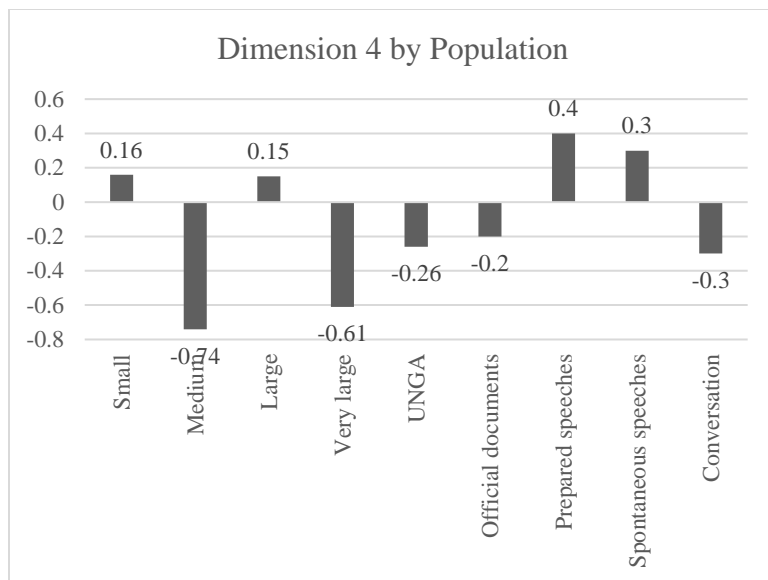
Figure 5.27. Dimension 4 scores by population.
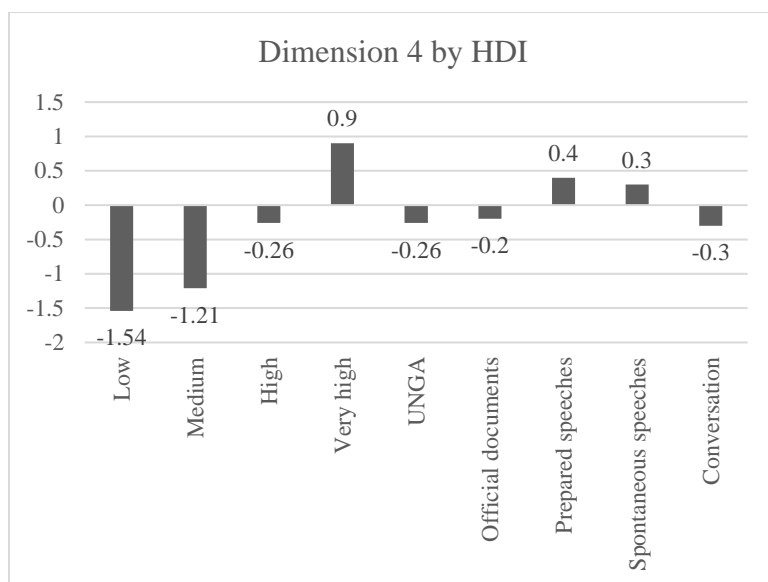
Figure 5.28. Dimension 4 scores by Human Development Index.
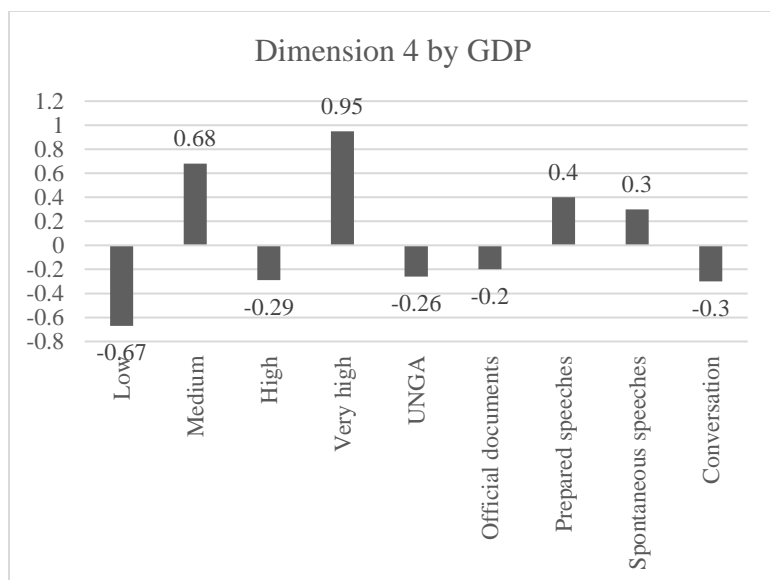
Figure 5.29. Dimension 4 scores by Gross Domestic Product.
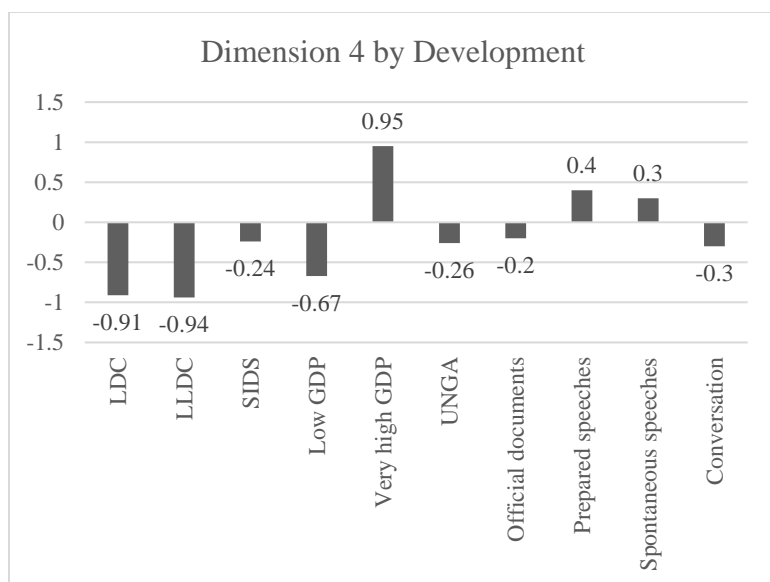
Figure 5.30. Dimension 4 scores by UN development group.
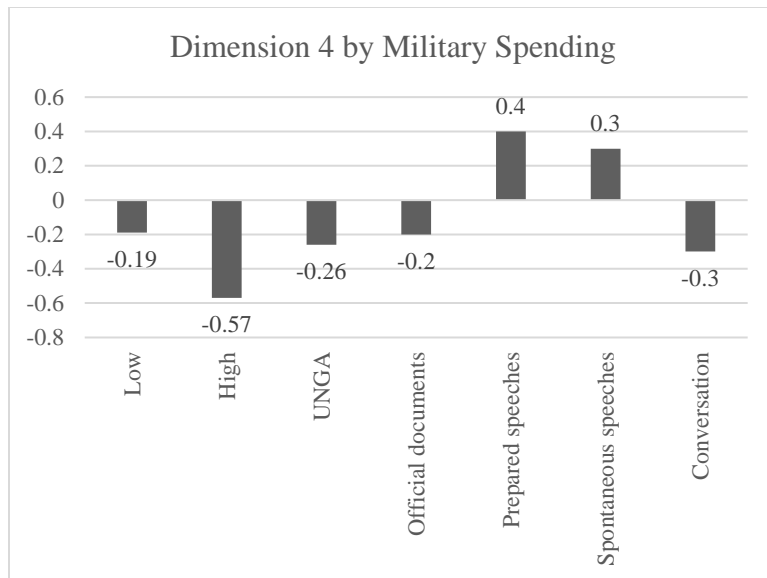
Figure 5.31. Dimension 4 scores by military spending.
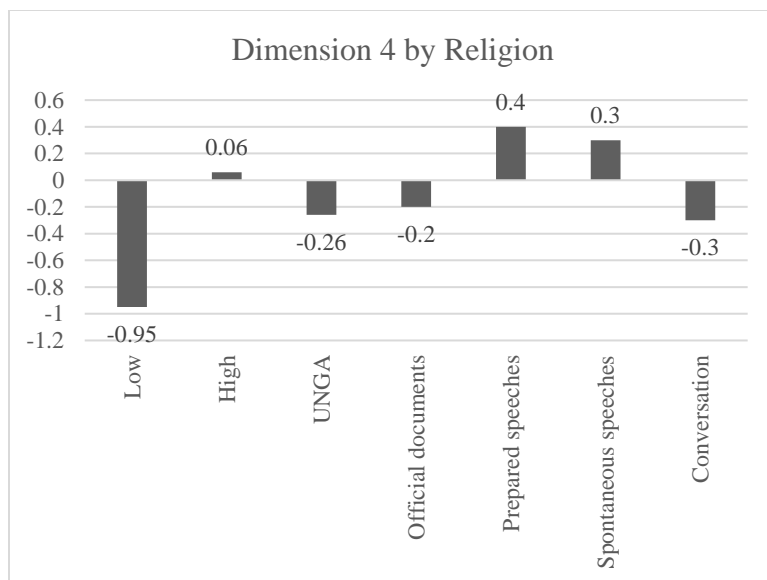
Figure 5.32. Dimension 4 scores by religious non-affiliation.

As shown in Figure 5.25, Africa and the Americas overall have much less overt persuasion and Europe much more. In addition, Figures 5.28-5.30 show a general pattern in which countries in low development groups (low HDI, low GDP, LDC, LLDC, SIDS) have

much less overt persuasion than countries in the highest development groups (very high HDI, very high GDP). This trend appears to hold for medium and high HDI as well. The pattern is not as consistent in the high GDP group, but this is likely due to individual differences skewing the data in this very small group of just four countries.

Interestingly, Dimension 4 scores also appear to vary based on military spending and religious affiliation, with lower overt persuasion for high military spending countries and low religious non-affiliation countries. This may, however, be due to the regional representation of countries in these groups. On average, there are considerably fewer countries from Europe and Oceania in the high military spending group (22% combined) and low religious affiliation group (24% combined) when compared to the UNGA overall (42% combined). Europe and Oceania are the regions with the most overt persuasion, so a lower representation from these regions in the high military spending and low religious non-affiliation groups explains the low overt persuasion scores for these two groups.

The high Dimension 4 score for the very high GDP group in particular is not surprising given the key features found in Section 5.4.8.3 (high frequency of modals for very high GDP). Texts 5.53 and 5.54 illustrate overt persuasion in two very high GDP texts.

Text 5.53. Excerpt from Sweden (overt persuasion features in italics)

> Climate change *can* work as a conflict multiplier, threatening peace as well as our chances of achieving all of our Sustainable Development Goals. We *need* a course of action to create a low-carbon and climate-resilient world economy. We *must* protect land and ocean ecosystems. It's not a choice, but a necessity for survival. The world *must* reach a fair, ambitious and legally binding agreement in Paris in December, which over time keeps the rise in global temperature as far as possible below two degrees Celsius.

In Text 5.53, the speaker from Sweden uses modals (*can*) and in particular modals of necessity (*must*) as well as an infinitive (*need to create*) to persuade addressees of the need to address climate change.

Text 5.54. Excerpt from Austria (overt persuasion features in italics)

> In this context we *have to* intensify our fight against Daesh and terrorist groups. They *have to* be neutralised and defeated! In the case of Syria it also means that we *have to* push harder for a closing of ranks in the region. We *will* only be able to achieve any meaningful progress *if* major regional players such as Saudi Arabia, Turkey, Egypt and Iran act in concert. And finally it means that we *have to* bring all relevant parties to the Syrian civil war on the table, as UN Special Envoy Staffan de Mistura *tries to do*. Let me be clear: that does not mean that the current political leadership of Syria *can* be part of a long term solution. It has brought endless harm to the people of Syria. But if *you want to make* peace, you don't only talk to your friends.

Text 5.54 from Austria illustrates the use of modals (*can*), modals of prediction (*will*), modals of necessity (*have to*), conditional subordination (*if*), and infinitives (*want to make*) to persuade the audience of the need to fight against terrorism and bring peace to Syria.

Such direct expression of stance is rare in the low development groups, but this does not necessarily imply that the lower development groups do not attempt to persuade listeners. Instead, they use less direct methods of persuasion. Some speakers opt for personal expression of stance, for instance through the use of verb complement clauses controlled by verbs of likelihood (e.g., *we believe that*, see Sections 5.4.3, 5.4.6, and 5.4.8.2); other speakers prefer more impersonal expressions of stance, such as adjective complement clauses controlled by a factive adjective (e.g., *it is evident that*, see Section 5.4.4) which imply a logical necessity rather than one imposed by the speaker. Text 5.55 illustrates more indirect persuasion.

Text 5.55. Excerpt from Zimbabwe (overt persuasion features in italics)

> Adaptation to change is the most crucial ingredient for the vibrancy and effectiveness of any organisation, including the United Nations. While the world has drastically changed since 1945, the United Nations, and indeed the global governance architecture, remains mired in a long bygone era. This archaic hierarchy among nations threatens *to erode* the confidence and support that the United Nations commands among the majority, but disadvantaged, of its membership. We are disappointed that we have lost the opportunity of this anniversary *to address* this burning issue of the reform of the United Nations Security Council in a manner that satisfies the just demands and expectations of the majority among us.

In Text 5.55 from Zimbabwe, the speaker argues for the need to reform the Security Council. Though the passage is intended to express disapproval and persuade, only one grammatical feature characteristic of overt persuasion is used, the infinitive (*to erode, to address*). Even that feature is not nearly as strong as the modals in the very high GDP Texts 5.53 and 5.54. Instead, President Mugabe sets up a logical argument. He begins with an assertion of fact: *Adaptation to change is the most crucial ingredient for the vibrancy and effectiveness of any organisation*, and proceeds by positing that the world has changed, but the UN has not. He then indicates a cause-effect relationship between lack of change and lack of support for the UN by the majority of the member states. Only at this point does President Mugabe express personal stance with an adjective complement clause controlled by an evaluative adjective (*we are disappointed that*). Overall, this passage illustrates a persuasive argument for Security Council reform but not once does the speaker use a directive. The argument is built primarily on an appeal to reason and logic.

As President Mugabe and numerous other speakers point out during the UNGA, the Security Council is a powerful body, but representation on the Security Council is not equally distributed across regions or levels of development. No country in a low development group has a Permanent Member seat on the Council, and of all the countries never elected to the Security

Council as a temporary member, more are in low development groups than high (see United

Nations Security Council, 2018). This leads to an imbalance in power in the UN. It is possible

that countries cognizant of their lack of power base their arguments on logical reasoning rather

than personal desire (*we hope*) or directives (*must*) because they believe only their logic, not their

wishes or demands will sway addressees.

5.5.6. Dimension 5: Abstract versus non-abstract information

Dimension 5 represents the degree to which the language of a text is abstract (see Section

4.5.5 for a description of the dimension and the features associated with the dimension). Figures

5.33-5.40 show Dimension 5 scores for the 31 country groups compared to the UNGA as a

whole as well as the comparison registers from Chapter 4.
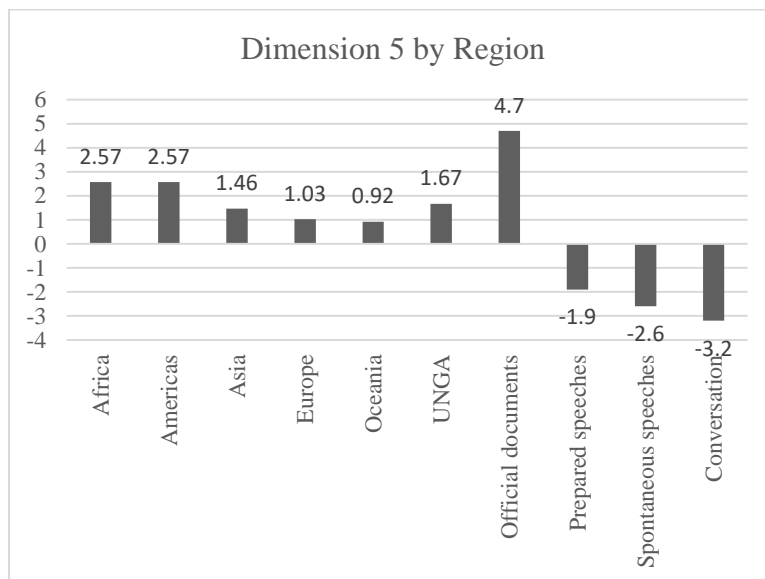
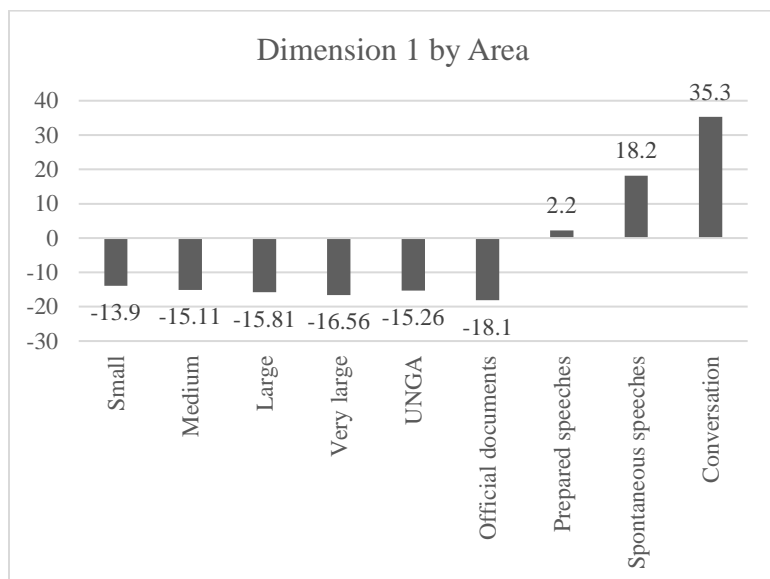

Figure 5.33. Dimension 5 scores by region.
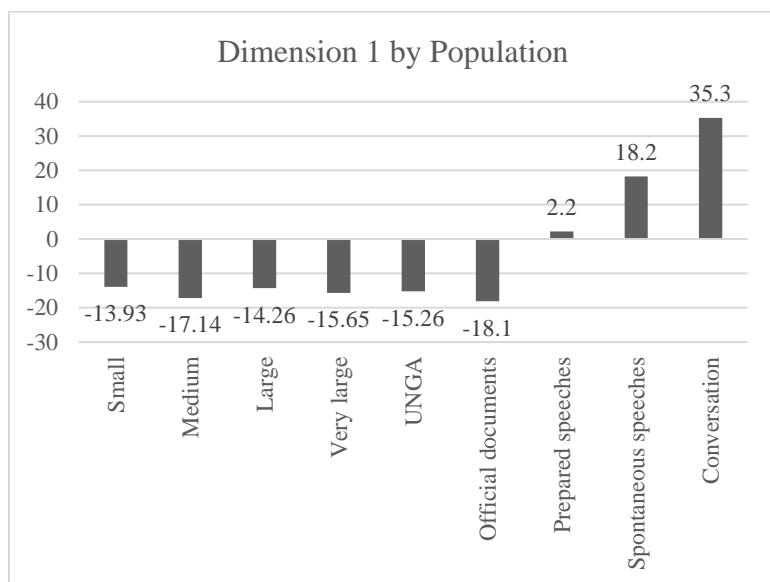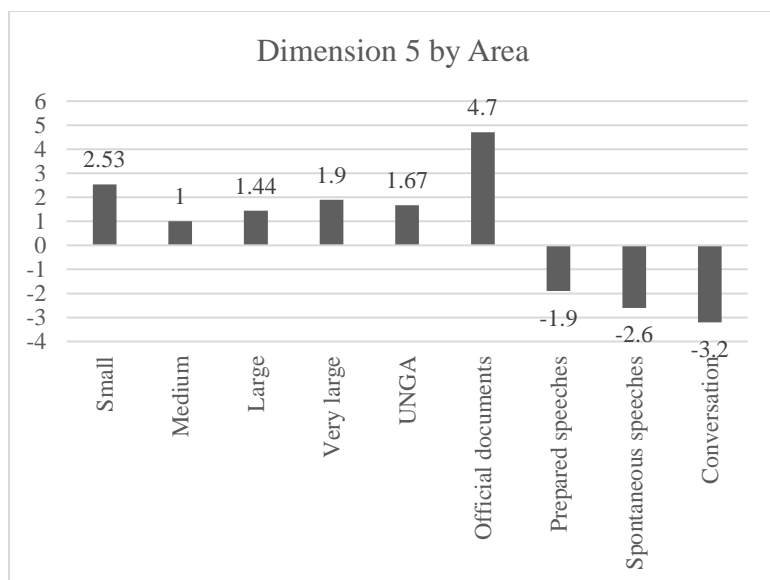
Figure 5.34. Dimension 5 scores by area.
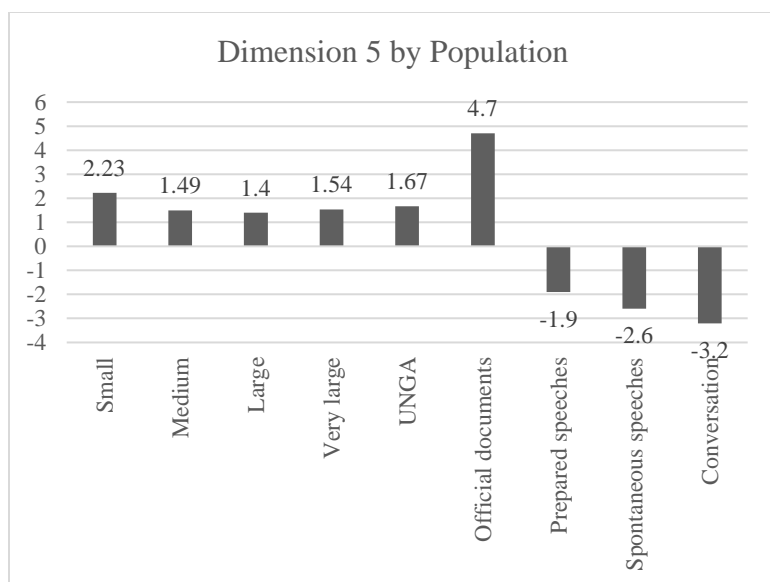
*Figure 5.35.* Dimension 5 scores by population.
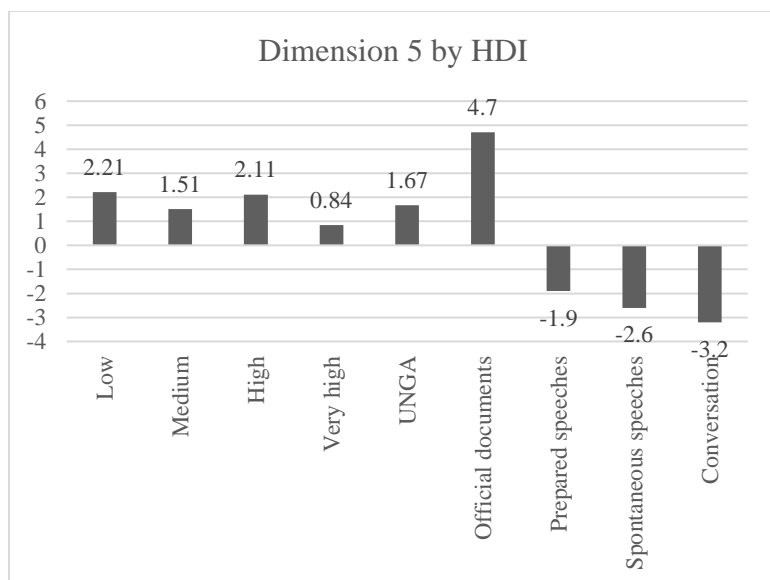
*Figure 5.36.* Dimension 5 scores by Human Development Index.
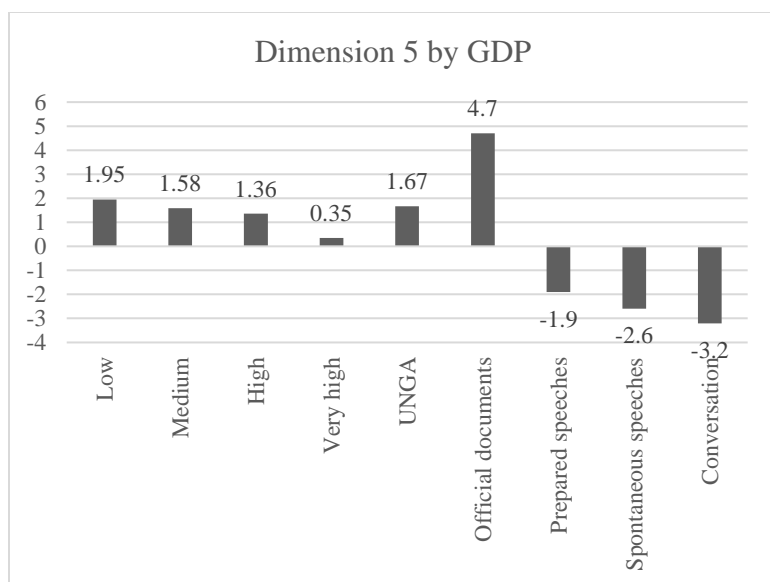
*Figure 5.37.* Dimension 5 scores by Gross Domestic Product.
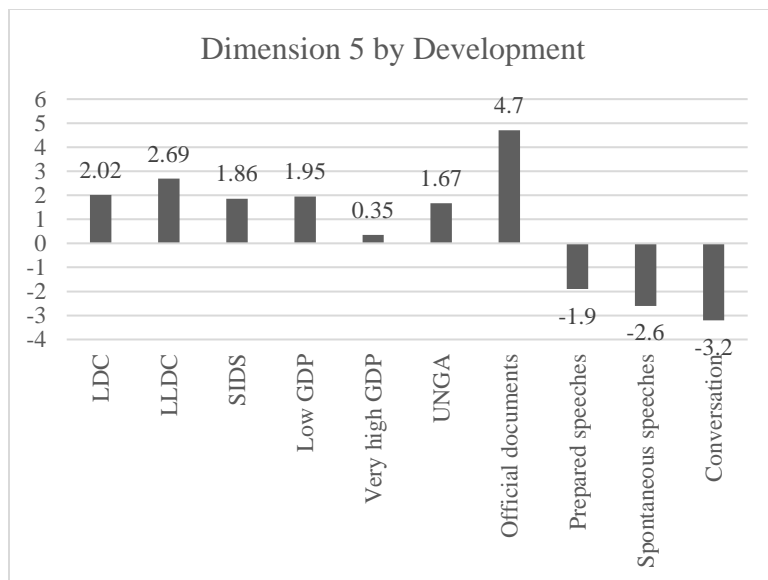
*Figure 5.38.* Dimension 5 scores by UN development group.
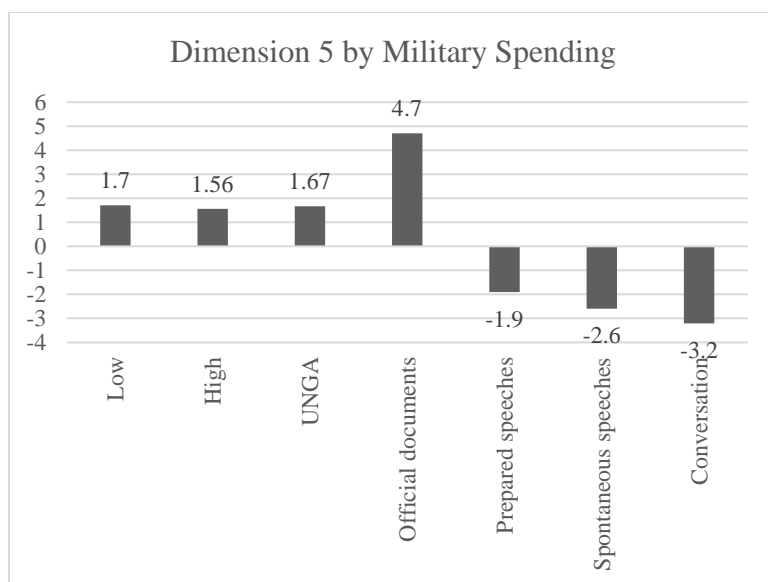

Figure 5.39. Dimension 5 scores by military spending.
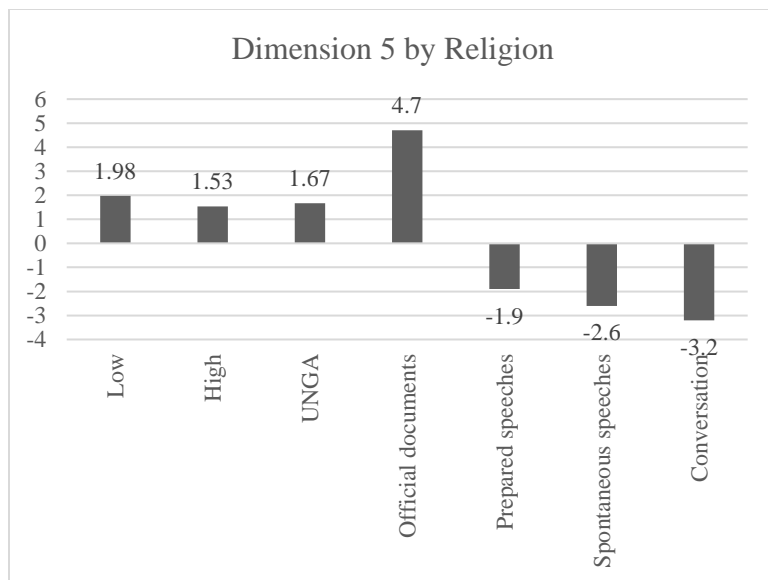
Figure 5.40. Dimension 5 scores by religious non-affiliation.

Again, patterns in Dimension 5 scores appear to be associated with region and development. Overall, the African and American groups are very abstract, Asia moderately abstract, and Europe and Oceania the least abstract. As for development, the least developed countries produce, on average, the most abstract texts and the most developed countries the least abstract texts. The primary reason for the high level of abstraction is the frequent use of passive structures. Passives allow the speaker to displace agency, to make requests and criticism much less direct. These texts are less confrontational but their level of abstraction also makes them more difficult for the listener to process. The texts sound as though they were meant to be read rather than spoken. In many ways, high Dimension 5 texts are similar in that respect to low Dimension 1 and high Dimension 3 texts. This is illustrated in abstract Texts 5.56, and 5.57. Text 5.58 serves as a contrast as a passage that is not abstract.

Text 5.56. Excerpt from Saint Vincent and the Grenadines (abstract features in italics)

> This year, possibly more than at any point in our modern history, our Assembly *is beset by* global threats and risks that force us to consider the ways in which our core principles

of sovereignty and non-interference can overcome today's challenges. Borderless menaces like terrorism, economic crises, contagious diseases and climate change heed neither geopolitical boundaries nor governmental jurisdiction. *Further*, the calamitous fallout of military adventurism, economic recklessness or environmental negligence *is* not *confined* to discrete national confines. *Rather*, the chickens often return to roost in far-flung, unexpected and often blameless locales. *As such*, more than ever, our international relations must *be defined by* cooperation, collaboration and decisive action.

The text from Saint Vincent and the Grenadines makes use of three conjuncts (*further, rather, as such*) as well as both short passives (*is not confined*) and *by*-passives (*is beset by, must be defined by*), which often put the focus on abstract topics rather than a human agent (*the calamitous fallout of military adventurism, economic recklessness or environmental negligence, our international relations*). In one instance, *our Assembly* is the subject of a passive sentence in order to respect the principle of end-weight: the agent is extremely long (*global threats and risks that force us to consider the ways in which our core principles of sovereignty and non-interference can overcome today's challenges*). This long agent does not necessarily have any features associated with high Dimension 5, but it does have many features associated with low Dimension 1 informational production (dense nominal structures with frequent nouns, prepositions, and attributive adjectives) as well as features associated with high Dimension 3 explicit language such as pied-piping: *the ways in which*, phrasal coordinating conjunctions (*and*), and nominalizations (*threats, risks, sovereignty, non-interference, challenge*). All of these features are more typical of written texts because they are more difficult to process in real time.

This text illustrates why there tends to be a relationship between Dimension 1, 3, and 5 scores. Overall, African and American texts are low on Dimension 1 and high on Dimension 3 and 5 because they are more literate. Texts from Europe and Oceania tend to be just the opposite, more oral with higher Dimension 1 and lower Dimension 3 and 5 scores. These differences are also (and primarily) reflected in development: Countries in the low development groups produce

293

texts with more written features (low Dimension 1, high Dimension 3 and 5); countries in the highest development groups produce texts with relatively more oral features (higher Dimension 1, lower Dimension 3 and 5). Because of the overlap between development and region, it is difficult to tease out which variable is most strongly associated with this pattern, but it would appear to be primarily development (e.g., the Seychelles, an African country with high HDI and medium GDP has more oral features than many other African countries with lower development; Romania, a European country with high HDI and low GDP, has more written features than many other European countries with higher development). Countries may opt for more literate features because of a preference for formality (also noted in the analysis of lexical features) or because of an assumption that the texts will be read rather than heard (as discussed in Section 5.2).

Text 5.57 from Rwanda illustrates another reason for the high Dimension 5 scores among low development groups: negative politeness.

Text 5.57. Excerpt from Rwanda (abstract features in italics)

> Cooperation is the only way forward. And *yet*, the new consensus on sustainable development is incomplete, because it lacks a shared definition of the political legitimacy, *required* to sustain this international order. The divergence of visions *is rooted* in history. When world powers created the United Nations seventy years ago, independence for the colonized peoples of Africa and Asia, was not on the agenda. We *were seen* as people who still needed to *be `looked after'*.

The text from Rwanda includes a conjunct (*yet*), a past participial WHIZ deletion (*required*), and three short passives (*is rooted, were seen, be looked after*). As with the text from Saint Vincent and the Grenadines, these passives displace the agent. However, in this case it is to avoid naming the responsible party, as the speaker from Rwanda is criticizing the disregard of the UN founders for *the colonized peoples of Africa and Asia*. The last two passives are particularly powerful because the colonized are put in object position: the people of Africa and

294

Asia were not active participants in the international community; they were *seen* and *looked after*.

African, American, and low development groups tend to be more indirect in their criticism and the passive allows them to accomplish this. The texts from these countries are less direct, less confrontational. This is also a function served by less overt persuasion, which explains why countries with high Dimension 5 scores tend on average to have low Dimension 4 scores as well. Overall, the high frequency of passive structures, associated with abstract language, and the low frequency of suasive verbs and modals, associated with overt persuasion, create texts with less personal expression of stance and more appeals to logical reasoning. Occasionally, when the subject matter warrants it, these texts are clear in their criticism (e.g., *for fifty years, our small country has been prevented from fully exploiting our rich natural resources. Venezuela has threatened and deterred investors and frustrated our economic development*, Guyana; *almost no progress has been achieved on the commitment made by Heads of State and Government in 2005, to the early reform of the UN Security Council*, South Africa; *Zambia is, however, concerned that the gears to advance the three pillars of the United Nations Charter are moving at a very slow pace, a pace with the potential of negating the gains we have achieved so far*, Zambia). However, this criticism is expressed as objective fact rather than personal opinion.

In contrast, Text 5.58 from the Bahamas, a high HDI and GDP country, illustrates a much less abstract passage.

Text 5.58. Excerpt from the Bahamas

The Bahamas is pleased at the rapprochement between the United States of American and Cuba. U.S. President Barak Obama has made the right, courageous decision. This enhances the peace in our neighbourhood. Our officials are working together with Cuba

on synergies on the development of our economies. This year, again, we will support an end to the economic embargo when the vote comes at the United Nations.

No abstract features occur in this excerpt from the Bahamas. All sentences are in the active. Sentences are relatively short; four of the five sentences have only one clause. These characteristics make the text much easier for the listener to follow.

The African and American groups and the lower development groups (low HDI, low GDP, LDC, LLDC, and SIDS) appear to be less concerned about whether their texts are easy for listeners to understand. A preference is given to subtlety, diplomacy, and formality but sometimes at the expense of aural comprehensibility.

5.5.7. Summary

The Multi-Dimensional scores have shown that generally texts from the most developed states and from Europe and Oceania in particular are more involved (higher Dimension 1), have more situation-dependent reference (lower Dimension 3), are more overtly persuasive (higher Dimension 4), and are less abstract (lower Dimension 5). Conversely, less developed states and especially countries in Africa and the Americas are more informational (lower Dimension 1), have more explicit reference (higher Dimension 3), have less overt persuasion (lower Dimension 4), and are more abstract (higher Dimension 5). Given the scores along Dimensions 1, 3, and 5, it is not surprising that texts from countries in the very high development groups sound more like spoken texts while texts from countries in the low development groups sound as though they were written to be read. Dimension 4 scores reflect the fact that overall, very high development groups use more personal stance and explicit directives while low development groups tend to opt for appeals to logical reasoning.

Though some patterns were detected in the MD scores by country category, as shown in Figures 5.1-5.40, there were few differences between country groups overall, and the differences that were identified were relatively minor compared to the much more pronounced differences between the UNGA overall and the comparison registers. MD scores, like the keyword analysis and key feature analysis, show overall little variation within the UNGA. Again, this may be due to two factors. First, there is not a great deal of variation within the UNGA, as shown in Chapter 4. Second, countries with similar geographic, social, political, and economic situations do not necessarily express themselves in a similar manner in their UNGA addresses. They may sometimes address the same topics, as discussed in Section 5.3, but not necessarily using the same grammatical structures and not necessarily for the same communicative purpose. Some patterns do emerge, but generally individual countries adopt their own discourse style.

The unique combination of lexical and grammatical forms that mark each text results in a particular discourse style and achieves a particular communicative purpose. Section 5.6 offers examples from five countries to show how this is accomplished.

5.6. Individual Country Analysis

5.6.1. Introduction

The lexical analyses in Section 5.3 have shown that some groups of countries speak on similar topics of interest given their geographic, social, political, or economic situation, but the majority of country groups do not share any distinctive lexical patterns. Moreover, as discussed in Sections 5.4 and 5.5, there are few instances where groups consistently use the same combinations of grammatical features in order to achieve a communicative function such as overtly persuade addressees. Rather, individual countries adopt certain sets of features in order to achieve very specific purposes based on their own unique situations. Section 5.6 illustrates how

297

they do this by presenting five studies of individual countries: the United States (Section 5.6.2),

Ukraine (Section 5.6.3), Eritrea (Section 5.6.4), Tuvalu (Section 5.6.5), and Bhutan (Section

5.6.6). The five countries were selected in order to present the starkest contrast in terms of key

country characteristics (e.g., geographical region, GDP) and of text type within the UNGA

corpus (e.g., highly involved language) based on the frequencies of grammatical features. Table

5.33 presents country characteristics for these five countries (see Sections 3.6.1 and 3.6.2 for

operationalization of categories).

Table 5.33. Key Country Characteristics

Characteristic	U.S.	Ukraine	Eritrea	Tuvalu	Bhutan
Region	Americas	Europe	Africa	Oceania	Asia
Area	Very large	Very large	Large	Small	Medium
Population	Very large	Very large	Large	Small	Medium
GDP	Very high	Low	N/A	Low	Low
HDI	Very high	High	Low	N/A	Medium
LDC	No	No	Yes	Yes	Yes
LLDC	No	No	No	No	Yes
SIDS	No	No	No	Yes	No

Note. GDP = Gross Domestic Product, HDI = Human Development Index, LDC = Least
Developed Country, LLDC = Landlocked Developing Country, SIDS = Small Island Developing
State, N/A = not available.

For each country, much like in Sections 5.3 through 5.5, keyword analysis shows lexical

features, key feature analysis grammatical features, and Multi-Dimensional analysis

combinations of grammatical features to achieve specific communicative purposes. All keywords

are presented in the lexical analysis. In contrast, not all 126 grammatical forms are discussed in

the key feature analysis, only those with large Cohen's *d* values, indicating a very strong

deviation from the UNGA average. However, normed frequencies and Cohen's *d* values for all

grammatical features can be found in Appendices L-P. Bar charts with Multi-Dimensional scores

in Section 5.6.2 allow scores for each country to be compared to the four other countries as well as to the UNGA overall and to the four comparison registers from Chapter 4. Comparing MD scores to these registers, as in Section 5.5, helps in determining the importance and meaningfulness of score differences between individual countries. The findings from the keyword, key feature, and MD analyses are then synthesized in order to present an overall picture of the discourse style for each text.

5.6.2. The United States

5.6.2.1. Introduction

The United States is a very large country in the Americas with a very large population. The country ranks very high for both GDP and HDI. As a Permanent Member of the Security Council, the US has veto power on Security Council resolutions, giving the country a great deal of political influence.

The text from the US touches on a number of different issues of international concern, from terrorism to the Sustainable Development Goals. However, most of the address is devoted to specific conflicts and argues that diplomacy solves international problems more effectively and responsibly than military might. The speaker, President Obama, asserts that while governments and political systems may not be perfect, ultimately human beings are good and deserve protection and support from the international community. This message is conveyed skillfully through a powerful discourse style. Most notably, the US address makes use of (1) expression of stance that appeals to logic, possibility, and likelihood over emotion, (2) highly involved language, and (3) rhetorical devices such as repetition and contrast. A number of lexical and grammatical features come together to achieve this discourse style. These features are examined in Sections 5.6.2.2 through 5.6.2.5: Keywords are analyzed in Section 5.6.2.2, key

299

grammatical features are discussed in Section 5.6.2.3, and Multi-Dimensional scores are interpreted in Section 5.6.2.4. Section 5.6.2.5 then synthesizes these results to present the overall discourse style of the text.

5.6.2.2. Lexical features

Keyword analysis was carried out to compare word frequencies in the US text to those in the remainder of the UNGA corpus. The keywords identified (see Table 5.34) reflect both the topics addressed as well as the discourse style of the US text.

Table 5.34. Keywords in the US Text

Keyword	Frequency	%	R.C. frequency	R.C. %	Log likelihood	Log Ratio	*P*
that's	13	0.27	6		70.83	6.30	0.00
can	41	0.87	324	0.19	58.15	2.21	0.00
Russia	13	0.27	25	0.01	46.73	4.24	0.00
but	36	0.76	390	0.23	35.21	1.75	0.00
people	39	0.82	467	0.27	33.06	1.61	0.00
who	25	0.53	218	0.13	31.92	2.06	0.00
strength	9	0.19	18	0.01	31.82	4.19	0.00
they	25	0.53	242	0.14	28.26	1.91	0.00
makes	7	0.15	10		28.23	4.67	0.00
impose	5	0.11	3		25.81	5.93	0.00

Note. R.C. = reference corpus (here, all texts except US).

The US address discusses several conflict situations at length and maintains that solutions to international problems should be sought through diplomacy rather than military action. The keywords *Russia, strength,* and *impose* stem from this topic and argument. One of the lengthiest discussions of conflict in the US text is on Russia (see Text 5.59).

300

Text 5.59. US excerpt illustrating *Russia* (keywords in italics)

> That same fidelity to international order guides our responses to other challenges around the world. Consider *Russia's* annexation of Crimea and further aggression in eastern Ukraine. America has few economic interests in Ukraine. We recognize the deep and complex history between *Russia* and Ukraine. *But* we cannot stand by when the sovereignty and territorial integrity of a nation is flagrantly violated. If that happens without consequence in Ukraine, it could happen to any nation gathered here today. *That's* the basis of the sanctions that the United States and our partners impose on *Russia*. It's not a desire to return to a Cold War.

In Text 5.59, President Obama presents the situation (*consider Russia's annexation of Crimea*). He explains the problem (*we cannot stand by when the sovereignty and territorial integrity of a nation is flagrantly violated*) and potential consequences of not finding a solution (*if that happens without consequence in Ukraine, it could happen to any nation gathered here today*). He also addresses US policy on the issue, asserting that US response to the conflict is not based on national interest or potential gain (*America has few economic interests in Ukraine*). He follows similar patterns in other portions of the text on conflicts with or in Syria, Iran, and China: presenting the situation (*nowhere is our commitment to international order more tested than in Syria*), explaining the problem and consequences (*when a dictator slaughters tens of thousands of his own people, that is not just a matter of one nation's internal affairs – it breeds human suffering on an order of magnitude that affects us all*), and addressing US policy (*but while military power is necessary, it is not sufficient to resolve the situation in Syria…. The United States is prepared to work with any nation, including Russia and Iran, to resolve the conflict*).

The keywords *impose* and *strength* are also used in passages emphasizing the importance of diplomacy over military might alone. Obama repudiates the claim that military action is sufficient to *impose* long-term stability or peace (Text 5.60). He further claims that the *strength*

of a nation is associated with international cooperation (Text 5.61) and the well-being of its

people (Text 5.62).

Text 5.60. US excerpt illustrating *impose* (keywords in italics)

> In Iraq, the United States learned the hard lesson that even hundreds of thousands of brave, effective troops, trillions of dollars from our Treasury, cannot by itself *impose* stability on a foreign land. Unless we work with other nations under the mantle of international norms and principles and law that offer legitimacy to our efforts, we will not succeed. And unless we work together to defeat the ideas that drive different communities in a country like Iraq into conflict, any order that our militaries can *impose* will be temporary. Just as force alone cannot *impose* order internationally, I believe in my core that repression cannot forge the social cohesion for nations to succeed.

Text 5.61. US excerpt illustrating *strength* and diplomacy (keywords in italics)

> The United States is not immune from this. Even as our economy is growing and our troops have largely returned from Iraq and Afghanistan, we see in our debates about America's role in the world a notion of *strength* that is defined by opposition to old enemies, perceived adversaries, a rising China, or a resurgent *Russia*; a revolutionary Iran, or an Islam that is incompatible with peace. We see an argument made that the only *strength* that matters for the United States is bellicose words and shows of military force; that cooperation and diplomacy will not work.

Text 5.62. US excerpt illustrating *strength* and the well-being of people (keywords in italics)

> Indeed, I believe that in today's world, the measure of *strength* is no longer defined by the control of territory. Lasting prosperity does not come solely from the ability to access and extract raw materials. The *strength* of nations depends on the success of their *people* -- their knowledge, their innovation, their imagination, their creativity, their drive, their opportunity -- and that, in turn, depends upon individual rights and good governance and personal security. Internal repression and foreign aggression are both symptoms of the failure to provide this foundation. A politics and solidarity that depend on demonizing others, that draws on religious sectarianism or narrow tribalism or jingoism may at times look like *strength* in the moment, *but* over time its weakness will be exposed. And history tells us that the dark forces unleashed by this type of politics surely *makes* all of us less secure.

Obama returns again and again to the theme of people and the need to ensure their well-being. He claims that while power corrupts, *ordinary people* are good (Text 5.63) and merit assistance, providing numerous examples such as Libya (Text 5.64) and Iran (Text 5.65).

Text 5.63. US excerpt illustrating *people* versus *power* (keywords in italics)

> Our systems are premised on the notion that absolute power will corrupt, *but* that *people* -- ordinary *people* -- are fundamentally good; that they value family and friendship, faith and the dignity of hard work; and that with appropriate checks and balances, governments can reflect this goodness.

Text 5.64. US excerpt illustrating the *people* in Libya (keywords in italics)

> Even as we helped the Libyan *people* bring an end to the reign of a tyrant, our coalition could have and should have done more to fill a vacuum left behind.

Text 5.65. US excerpt illustrating the *people* in Iran (keywords in italics)

> If Iran chose a different path, that would be good for the security of the region, good for the Iranian *people*, and good for the world.

The keyword *they* is often used in reference to people in powerful examples reflecting the fundamental goodness in human beings. Text 5.66 illustrates this use of *they* and again, *people.*

Text 5.66. US excerpt illustrating *they* and *people* (keywords in italics)

> Think of the Liberian doctor who went door-to-door to search for Ebola cases, and to tell families what to do if *they* show symptoms. Think of the Iranian shopkeeper who said, after the nuclear deal, "God willing, now we'll be able to offer many more goods at better prices." Think of the Americans who lowered the flag over our embassy in Havana in 1961 -- the year I was born -- and returned this summer to raise that flag back up. One of these men said of the Cuban *people*, "We could do things for them, and *they* could do things for us. We loved them."

The high frequency of the keyword *who* also reflects this focus on people, be they worthy of gratitude for what they have accomplished (Text 5.67) or worthy of assistance due to the

suffering they have endured (Texts 5.68 and 5.69). In addition to illustrating Obama's focus on the power of individual people to accomplish great deeds, Text 5.67 shows yet another excerpt where diplomacy is preferred over military action (*cooperation over conflict*).

Text 5.67. US excerpt illustrating *who* and the UN founders (keywords in italics)

> This institution was founded because men and women *who* came before us had the foresight to know that our nations are more secure when we uphold basic laws and basic norms, and pursue a path of cooperation over conflict.

Text 5.68. US excerpt illustrating *who* and the people of Syria (keywords in italics)

> And so Assad and his allies cannot simply pacify the broad majority of a population *who* have been brutalized by chemical weapons and indiscriminate bombing.

Text 5.69. US excerpt illustrating *who* and refugees (keywords in italics)

> And *that's* why the United States is increasing the number of refugees *who* we welcome within our borders.

Text 5.69 also shows the use of another keyword, *that's. That's* is used in expressions of stance, in this case to explain a reason (see also *that's the basis of the sanctions that the United States and our partners impose on Russia* in Text 5.59). *That's* and *that is* can be used to offer an opinion as well (Text 5.70) or to garner support (Text 5.71).

Text 5.70. US excerpt illustrating *that's* to express opinion (keywords in italics)

> I believe *that's* the future we must seek together. To believe in the dignity of every individual, to believe we can bridge our differences, and choose cooperation over conflict: that is not weakness, that is strength. Not everybody in America agrees with me. *That's* part of democracy.

Text 5.71. US excerpt illustrating *that's* to garner support (keywords in italics)

> *But* we also have to recognize that we must work more effectively in the future, as an international community, to build capacity for states that are in distress, before *they* collapse. And *that's* why we should celebrate the fact that later today the United States will join with more than 50 countries to enlist new capabilities -- infantry, intelligence, helicopters, hospitals, and tens of thousands of troops -- to strengthen United Nations peacekeeping.

Generally, stance is expressed through appeals to logic. In addition to *that's*, which presents an explanation or opinion as an objective definition (*that's part of democracy*), *makes* is sometimes used as a causal verb, suggesting a logic relationship, facts to support the speaker's assertions. In Text 5.72, Obama claims that democracy *makes us strong*; he attempts to persuade the audience to support democracy based on the logic that we will be stronger for it. Text 5.72 again focuses on the rights of individuals, with the keywords *people, who,* and *they* (*where people worship freely, everybody can participate no matter who they are, or what they look like, or who they love*). The strength brought by democracy comes through the freedoms granted to the people.

Text 5.72. US excerpt illustrating *makes* to appeal to logic (keywords in italics)

> Democracy -- inclusive democracy -- *makes* countries stronger…. I believe that the fact that you *can* walk the streets of this city right now and pass churches and synagogues and temples and mosques, where *people* worship freely; the fact that our nation of immigrants mirrors the diversity of the world -- you *can* find everybody from everywhere here in New York City -- the fact that, in this country, everybody *can* contribute, everybody *can* participate no matter *who they* are, or what *they* look like, or *who they* love -- *that's* what *makes* us strong.

Text 5.72 also illustrates several occurrences of *can*, the last keyword related to expression of stance. *Can,* a modal of possibility and ability, is used persuasively. Obama generally avoids strong directives like *should* or *must.* He does not command or dictate. He

encourages and motivates by showing what is possible, what *people* are capable of doing, what the UN is able to do, what *we can* do. Reminiscent of Obama's ubiquitous US presidential campaign slogan in 2008 (*Yes We Can*), the repetition of *can* in Texts 5.73 and 5.74 inspires possibility and promise rather than the obligation and accusation underlying a *should* or *must*.

Text 5.73. US excerpt illustrating *we can* to inspire possibility and promise (keywords in italics)

> We *can* be proud of our nations without defining ourselves in opposition to some other group. We *can* be patriotic without demonizing someone else. We *can* cherish our own identities -- our religion, our ethnicity, our traditions -- without putting others down.

Text 5.74. US excerpt illustrating *when… can* to inspire possibility and promise (keywords in italics)

> Democracy -- inclusive democracy -- *makes* countries stronger. When opposition parties *can* seek power peacefully through the ballot, a country draws upon new ideas. When a free media *can* inform the public, corruption and abuse are exposed and *can* be rooted out. When civil society thrives, communities *can* solve problems that governments cannot necessarily solve alone. When immigrants are welcomed, countries are more productive and more vibrant. When girls *can* go to school, and get a job, and pursue unlimited opportunity, *that's* when a country realizes its full potential.

Both Text 5.73 and Text 5.74 illustrate, in addition to the frequent use of *can*, the rhetorical style of parallel sentence structure and repetition. Text 5.74 in particular illustrates a powerful persuasive style with the repetition of *when* clauses in conjunction with *can.* These *when* clauses and the causal relationships they convey are again expressions of stance through logic. Obama claims: *When opposition parties can seek power peacefully through the ballot, a country draws upon new ideas*. There is no overt directive force as there would be with *opposition parties should seek power peacefully through the ballot.* Instead, the argument is founded on an appeal to reason. Persuasion is conveyed optimistically through the expression of possibility and promise.

306

Parallel structures such as these *when* clauses are extremely frequent in the US text. The keyword *they*, in addition to being used in passages related to the power and goodness of people as shown above, sometimes refers to abstract ideas. In such cases, *they* is used in short parallel sentences to make these abstract ideas easier for the listener to grasp (see Text 5.75).

Text 5.75. US excerpt illustrating *they* in short sentences (keywords in italics)

> The freedom to peacefully petition those in power without fear of arbitrary laws -- these are not ideas of one country or one culture. *They* are fundamental to human progress. *They* are a cornerstone of this institution.

Another stylistic device is juxtaposition of opposing ideas. *But* is used frequently in such instances of contrast. These phrases are also parallel in structure: *we know… but we also know* in Text 5.76, *not because we want… but because we want* in Text 5.77, and once again with the modal *you can… but you can't* in Text 5.78.

Text 5.76. US excerpt illustrating *but* for contrast (keywords in italics)

> We know that ISIL -- which emerged out of the chaos of Iraq and Syria - depends on perpetual war to survive. *But* we also know that they gain adherents because of a poisonous ideology.

Text 5.77. US excerpt illustrating *but* for contrast (keywords in italics)

> Not because we want to isolate Russia -- we don't -- *but* because we want a strong Russia *that's* invested in working with us to strengthen the international system as a whole.

Text 5.78. US excerpt illustrating *but* for contrast (keywords in italics)

> You *can* jail your opponents, *but* you can't imprison ideas. You *can* try to control access to information, *but* you cannot turn a lie into truth.

Overall, the keywords in the US text shed light on both topic and discourse style. The topical focus on conflicts and on the importance of diplomacy over military action gives rise to the keywords *Russia, strength,* and *impose.* The other topical focus on people over governments or political systems results in the keywords *people, who,* and *they.* In terms of discourse style, expression of stance that appeals to logic, possibility, and likelihood brings about *that's, makes,* and *can*. Rhetorical devices such as repetition and contrast lead to *but* and additional occurrences of *they.* As shown in Text 5.79, these lexical features often come together to express a positive message of solidarity and promise.

Text 5.79. US excerpt showing discourse style through lexical features (keywords in italics)

> The *people* of our United Nations are not as different as *they* are told. *They can* be made to fear; *they can* be taught to hate -- *but they can* also respond to hope.

One final note is in order before concluding this section on the lexical features of the US text. *The United States* or *US* is noticeably absent from the keyword list. It is the only text in Section 5.6 that does not have a "self-reference" term on its keyword list. It should be noted that this is not because Obama does not refer to *the United States*, but rather because the words *united* and *states* appear so frequently in other texts (e.g., *United Nations, Small Island Developing States*) that it is not key in the US text.

5.6.2.3. Grammatical features

For the US text, 65 grammatical features had large Cohen's *d* scores, revealing that the frequency of half of the 126 grammatical forms investigated deviated considerably from the norm. It would be prohibitive and not particularly productive to discuss all 65 features here, but frequencies and Cohen's *d* values are provided for the complete list of features in Appendix L.

Instead, Section 5.6.2.3 focuses on two patterns in types of grammatical forms that are especially

frequent or infrequent. The first pattern is for grammatical structures showing the US text has

more involved rather than informational language; the second pattern is for structures expressing

stance through logic, possibility, and likelihood rather than emotion or volition. These two

patterns of grammatical features result in a discussion of 39 features. Normed frequencies and

Cohen's *d* values for the 39 features discussed in Section 5.6.2.3 are shown in Table 5.35.

Table 5.35. Select Grammatical Features in the US Text

Feature	US frequency	UNGA Mean	Cohen's *d*
Verb (not including auxiliary)	143.5	113.00	2.30
Present tense verb	111.3	77.31	2.69
Pro-verb 'do'	1.9	0.43	2.71
Verb 'be'	3.7	2.10	1.06
Private verb	15.5	7.40	3.25
Mental verb	19.2	12.67	1.51
Causative verb	6.2	3.98	1.20
Modal of prediction	8.0	5.83	0.85
Modal of possibility	14.2	3.59	5.22
1st person pronoun	46.9	35.58	0.93
Demonstrative pronoun	9.5	3.49	2.92
All nouns	262.1	303.03	-1.65
All definite articles	45.4	64.76	-1.24
Preposition	107.2	124.88	-1.34
Attributive adjective	52.7	68.46	-1.33
Predicative epistemic adjective	1.1	0.45	1.15
All adverbs	38.7	30.22	1.23
Adverb – time	3.7	4.03	-0.16
Adverb – place	5.2	4.19	0.47
Emphatic	3.4	2.11	0.95
'That' deletion	3.0	0.93	2.57
Contraction	9.0	1.20	4.37
Subordinating conjunction – causative	1.7	0.48	1.60
Subordinating conjunction – conditional	2.4	1.17	1.07
Coordinating conjunction – clausal	24.7	7.80	3.37
Coordinating conjunction – phrasal	0.2	4.81	-2.09

Feature	US frequency	UNGA Mean	Cohen's d
Wh pronoun – object relative clause	1.7	0.84	1.09
Wh pronoun – subject relative clause	3.2	2.24	0.67
Wh pro. – rel. cls. with preposition fronting	0.2	0.56	-0.64
Wh CCC communicaton verb	0.2	0.05	0.97
'That' CCC verb	9.3	3.23	4.05
'That' CCC communication verb	2.6	1.29	1.37
'That' CCC factive verb	4.1	1.11	3.65
'That' CCC likelihood verb	3.2	0.91	2.70
All stance 'that' CCC verbs	11.8	4.97	3.55
All stance 'that' complement clauses	15.1	6.14	3.52
'To' CCC verb of desire	0.9	2.39	-1.07
'To' CCC verb of effort	3.0	1.66	1.24
'To' CCC probability verb	0.4	0.08	1.53

Note. Frequencies normed per 1,000 words. CCC = complement clause controlled by.

The first pattern in grammatical features shows that the language in the US text is more "involved," and less "informational" compared to other UNGA texts. Features that signal involved language are generally frequent in the US text; they include first person pronouns (e.g., *we, d* = 0.93) and emphatics (e.g., *of course, d* = 0.95): <u>*Of course,*</u> *there have been too many times when, collectively,* <u>*we*</u> *have fallen short of these ideals.* They also include present tense verbs (e.g., *believe, d* = 2.69), *that* deletion (e.g., *believe, d* = 2.57) and contractions (e.g., *that's, d* = 4.37): <u>*I believe that's*</u> *the future we must seek together.* Other involved features are demonstrative pronouns (e.g., *that, d* = 2.92), verbs (e.g., *has pursued, d* = 2.30) and *BE* as a main verb (e.g., *is, d* = 1.06): <u>*That*</u> <u>*is*</u> *the ideal that this body, at its best,* <u>*has pursued*</u>. Involved features are causal subordinate conjunctions (e.g., *because, d* = 1.60), private verbs (e.g., *want, d* = 3.25), and *DO* as a pro-verb (e.g., *don't, d* = 2.71): *Not* <u>*because*</u> *we want to isolate Russia -- we* <u>*don't*</u> *-- but* <u>*because*</u> *we want a strong Russia that's invested in working with us to strengthen the international system as a whole.* Finally, involved features comprise non-phrasal coordinating conjunctions (e.g., *and, d* = 3.37) and conditional subordinators (e.g., *if, d* = 1.07): <u>*And*</u> <u>*if*</u> *we*

cannot work together more effectively, we will all suffer the consequences. In contrast, features that mark informational language are generally infrequent: nouns (e.g., *system*, $d = - 1.65$), prepositions (e.g., *in, d = - 1.34*), definite articles (*the, d = - 1.24*), and attributive adjectives (e.g., *international*, $d = - 1.33$). The overall style that results from these frequent involved features and infrequent informational features is illustrated in the discussion of Dimension 1 scores in Section 5.6.2.4.

The second pattern of grammatical features shows that the US text has numerous examples of stance through logic, possibility, and likelihood rather than emotion or volition. Features expressing stance include modals of possibility ($d = 5.32$), modals of prediction ($d = 0.85$), causal verbs ($d = 1.20$), mental verbs ($d = 1.51$), verbs of probability and fact ($d = 1.53$), effort verbs followed by the infinitive ($d = 1.24$), WH- clauses controlled by communication verbs ($d = 0.97$), verb complement clauses ($d = 4.05$), and more specifically verb complement clauses controlled by factive verbs ($d = 3.65$), by verbs of likelihood ($d = 2.70$), and by communication verbs ($d = 1.37$). Predicative epistemic adjectives also express stance ($d = 1.15$), and generally there are high scores for overall stance features, all stance verb complement clauses ($d = 3.55$) and all stance complement clauses, be they controlled by verbs, nouns, or adjectives ($d = 3.52$). In contrast, expression of stance through emotion or volition is infrequent, with low frequencies of verbs of desire followed by the infinitive ($d = - 1.07$). Positive stance features are illustrated in the excerpts in Text 5.80.

Text 5.80. US excerpts illustrating stance (key features in italics)

> Change *won't* come overnight to Cuba, but I'm *confident* that openness, not coercion, *will* support the reforms and better the life the Cuban people deserve, just as I *believe that* Cuba *will* find its success if it pursues cooperation with other nations.
> Our objective was to *test whether* Iran could change course, accept constraints, and *allow* the world *to verify that* its nuclear program *will* be peaceful.

But we are called upon to offer a different type of leadership -- leadership strong enough to *recognize that* nations share common interests and people share a common humanity, and, yes, there are certain ideas and principles that are universal.

I *realize that* in many parts of the world there is a different view -- a *belief that* strong leadership must tolerate no dissent.

No matter how powerful our military, how strong our economy, we *understand* the United States *cannot* solve the world's problems alone.

We *can* roll back the pollution that we put in our skies, and *help* economies lift people out of poverty without condemning our children to the ravages of an ever-warming climate.

Let us carry forward that faith into the future -- for it is the only way we can *assure* that future *will* be brighter for my children, and for yours.

The excerpts in Text 5.80 exemplify modals of possibility (*can*), modals of prediction (*won't*), causal verbs (*allow*), mental verbs (*realize*), verbs of probability and fact (*understand*), effort verbs followed by the infinitive (*help*), *WH-* clauses controlled by communication verbs (*test whether*), verb complement clauses controlled by factive verbs (*recognize that*), verb complement clauses controlled by verbs of likelihood (*believe that*), verb complement clauses controlled by communication verbs (*verify that*), predicative epistemic adjectives (*I'm confident*), all stance verb complement clauses (*assure that*) and all stance complement clauses, be they controlled by verbs, nouns, or adjectives (*a belief that*).

These excerpts also demonstrate how Obama appeals to logic and likelihood rather than emotion and volition. He asserts beliefs, realizations, and recognitions rather than desires. Rather than frame his claim as a wish: *the United States doesn't want to solve the world's problems alone*, he frames it as a fact: *the United States cannot solve the world's problems alone*. Similarly, he appeals to possibility and promise rather than obligation. Instead of *we must roll back the pollution that we put in our skies*, he states *we can roll back the pollution that we put in our skies*. He urges action not because the international community is forced to do something but because it is able to do it.

Overall, this creates a text that is optimistic and encouraging, fostering cooperation and solidarity, rather than opinionated and demanding. This discourse style may, in fact, stem from a conscious choice to avoid sounding like an uncompromising superpower; the United States is often perceived as controlling the UN (Puchala, 2005). While other countries, particularly smaller, less powerful countries, may be able to use *we must* without appearing dogmatic, the US could come across as controlling if there is too much directive force in expressions of stance.

5.6.2.4 Multi-Dimensional scores

Multi-Dimensional scores show how these key features come together to create a particular text style. For each dimension, scores for the US text are compared to the four other individual countries investigated in Section 5.6 as well as to the UNGA overall and to the four comparison registers from Chapter 4. For dimension scores with particularly high or low values, text excerpts illustrate the style that results. For more information on dimensions, including the features associated with each dimension, see Sections 4.6 and 5.5.

The Dimension 1 score (involved versus informational production) revealed a sizeable difference between the US text and the other texts in the UNGA. Overall, the US text was much more involved (see Figure 5.41).
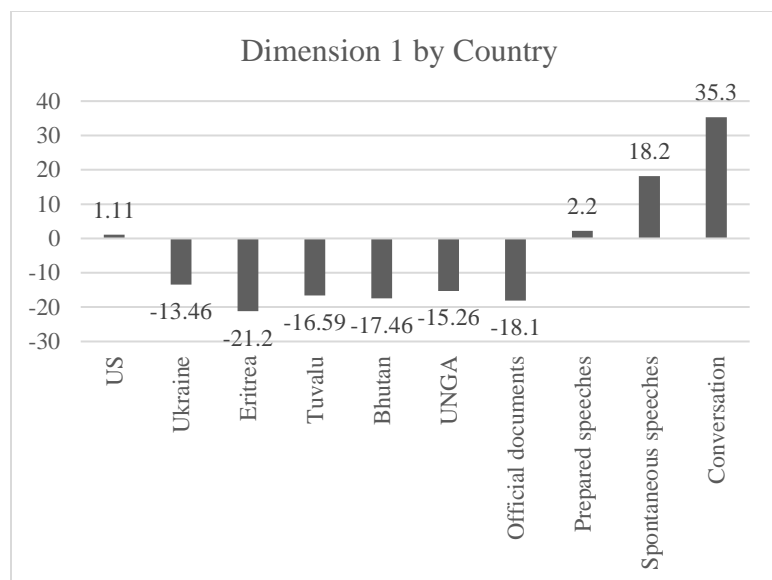
Figure 5.41. Dimension 1 by individual country.

The high Dimension 1 score for the US is not surprising given the results of the key feature analysis in Section 5.6.2.3, which showed that a great many features frequent in the US address are characteristic of involved rather than informational language. As discussed in Section 5.6.2.3, involved language favors interpersonal discourse and makes reference to the addressor and addressee; the dispreference for dense noun phrases makes the text easier for the listener to follow. Text 5.81 illustrates this involved rather than informational language, with involved features (first and second person pronouns, *it*, contractions, present tense verbs, clausal coordinating conjunctions, *that* deletion, emphatics, sentence final prepositions) italicized and informational features (prepositions, nouns, attributive adjectives) underlined.

Text 5.81. US excerpt illustrating involved language (involved features in italics; informational features underlined)

> *I say* this, recognizing that <u>diplomacy</u> *is* hard; that the <u>outcomes</u> *are* sometimes unsatisfying; that *it's* rarely politically popular. *But I believe* that <u>leaders</u> <u>of</u> <u>large</u> <u>nations</u>, in particular, *have* an <u>obligation</u> <u>to</u> take these <u>risks</u> -- precisely because *we are* strong enough to protect our <u>interests</u> if, and when, <u>diplomacy</u> *fails. I* also *believe* that to move

forward in this new era, *we have* to be strong enough to acknowledge when what *you're doing is* not *working*.

In contrast, the Dimension 2 (narrative versus non-narrative concerns) score showed very little difference between the UNGA average and the US text (see Figure 5.42). This is not particularly surprising given the overall lack of variation within the UNGA for Dimension 2 (see Section 5.5.3).
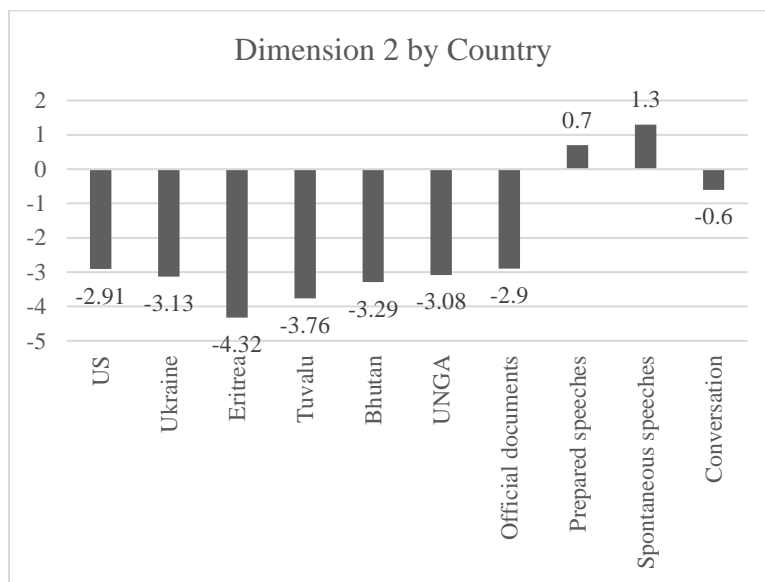
Figure 5.42. Dimension 2 by individual country.

The Dimension 3 (explicit versus situation-dependent reference) score showed much more situation-dependent language for the US text (see Figure 5.43), but the analysis of this dimension is far from straightforward because the US text does not have strong Cohen's *d* scores for most of the features related to Dimension 3.
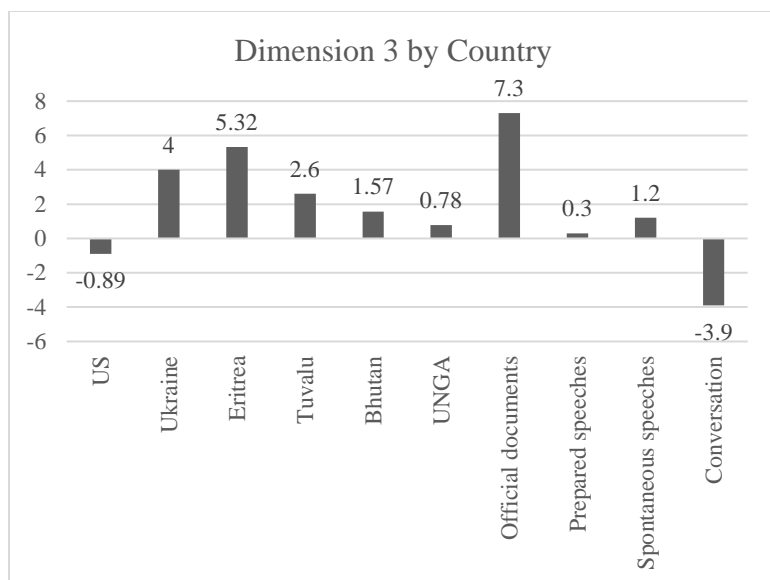
Figure 5.43. Dimension 3 by individual country.

Features that emerged in previous studies (e.g., Biber, 1988) as representative of Dimension 3 include, for explicit reference: *WH-* relative clauses on object positions, pied piping, *WH-* relative clauses on subject positions, and phrasal coordination; for situation-dependent reference: time adverbs, place adverbs, and all adverbs. However, only the high frequency of adverbs ($d = 1.23$) and the low frequency of phrasal coordinating conjunctions ($d = - 2.09$) explain the negative Dimension 3 scores (situation-dependent reference) in the US text. Pied piping is, indeed, relatively infrequent, but the Cohen's *d* value is only moderate ($d = - 0.64$). *WH-* relative clauses on object and subject positions are actually positive ($d = 1.09$ and $d = 0.67$, respectively), which is indicative of explicit reference, the contrary of what would be expected from a text with an overall negative Dimension 3 score. Likewise, adverbs of place are only moderately frequent ($d = 0.47$) and adverbs of time are infrequent, though Cohen's *d* is very low ($d = - 0.16$). Arguably, the negative Dimension 3 score is not due to particularly strong situation-dependent reference, but rather two separate phenomena: a low frequency of phrasal coordinating conjunctions and a high frequency of adverbs.

The low frequency of phrasal coordinating conjunctions can actually be explained by the overall low frequency of noun phrases and subsequent low frequencies of nouns, prepositions, definite articles, and attributive adjectives (see Section 5.6.2.3). This again reflects the involved rather than informational language used in the US address.

The high frequency of adverbs is related to stance, including adverbs of attitude (see Text 5.82) and adverbs of certainty (see Text 5.83).

Text 5.82. US excerpts illustrating adverbs of attitude (key features in italics)

> Most *ominously*, we see the fears of ordinary people being exploited through appeals to sectarianism, or tribalism, or racism, or anti-Semitism; appeals to a glorious past before the body politic was infected by those who look different, or worship God differently; a politics of us versus them.
> But from big cities to rural villages around the world, we also know that prosperity is still *cruelly* out of reach for too many.

Text 5.83. US excerpts illustrating adverbs of certainty (key features in italics)

> Democracy in the United States is *certainly* imperfect.
> As these contacts yield progress, I'm confident that our Congress will *inevitably* lift an embargo that should not be in place anymore.
> Now, within Russia, state-controlled media may describe these events as an example of a resurgent Russia -- a view shared, by the way, by a number of U.S. politicians and commentators who have always been deeply skeptical of Russia, and seem to be convinced a new Cold War is, *in fact*, upon us.
> *Indeed*, I believe that in today's world, the measure of strength is no longer defined by the control of territory.

Dimension 4 (overt persuasion) scores showed that the US text had somewhat more overt persuasion than the UNGA average (see Figure 5.44). The relatively high Dimension 4 score is likely due to the frequent use of modals of possibility and prediction in the US text (see Section 5.6.2.3). This results in an address full of optimism and promise, as shown in Texts 5.73, 5.74, and 5.79.
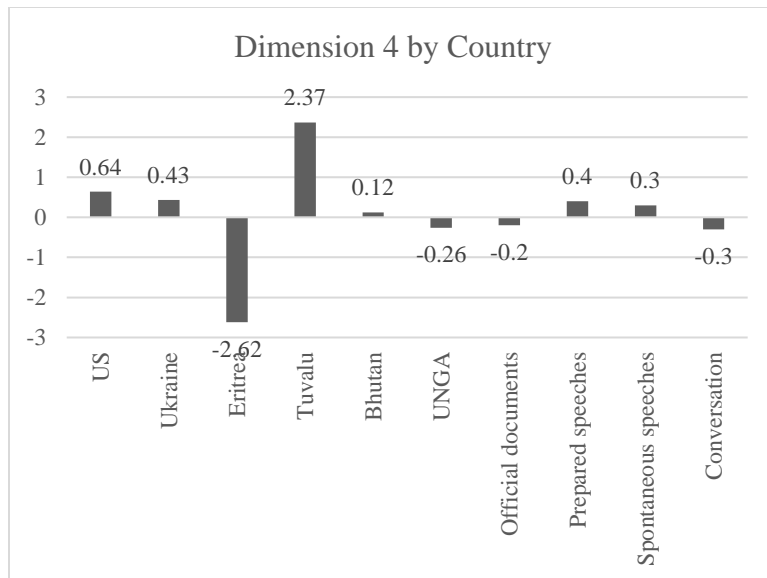
Figure 5.44. Dimension 4 by individual country.

Dimension 5 (abstract versus non-abstract information) scores showed that the US text was nearly the same as the UNGA average (see Figure 5.45). Thus, like with Dimension 2 scores, the overall style of the US address is not particularly marked by the degree of abstraction in the text.
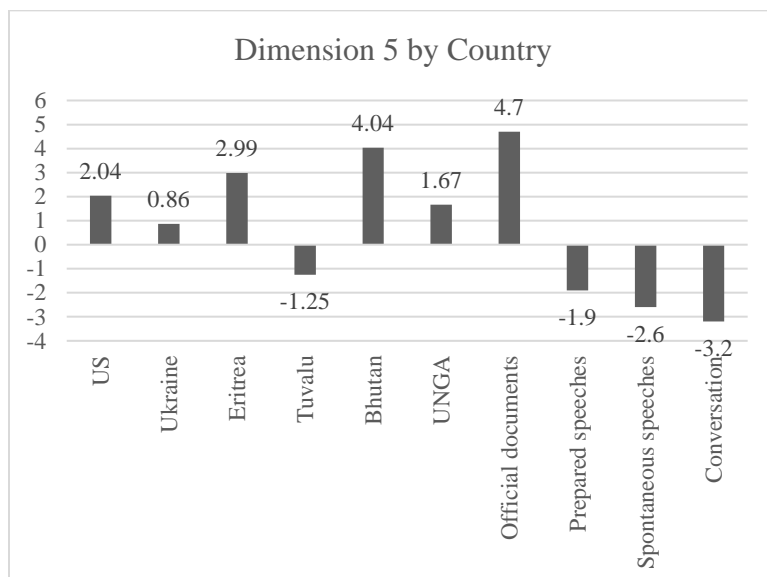
Figure 5.45. Dimension 5 by individual country.

Overall, Multi-Dimensional scores showed the US text to be unusually involved (Dimension 1), resulting in a text that engages the audience and is relatively easy to process. In addition, the analysis of Dimension 3 showed that while the US text did not consistently use all situation-dependent features frequently, the text does make use of an unusually high number of adverbs of attitude and certainty. These adverbs express stance with minimal directive force because they are impersonal. It is not the speaker imposing his view on the audience but rather an objective reality (e.g., *more ominously*). Finally, Dimension 4 scores reflected the frequent use of modals of prediction and especially modals of possibility, giving a sense of optimism and hope to the address.

5.6.2.5. Summary

The lexical, grammatical, and Multi-Dimensional analyses in Sections 5.6.2.2 through 5.6.2.4 have shown notable characteristics of the US address, in terms of both topic and discourse style. Again and again, the US text emphasizes diplomacy over military might and the goodness of the people over the imperfections of their governments. These arguments are presented in a discourse style that is in stark contrast with the styles of most other UNGA texts and that combines three major elements: (1) expression of stance that appeals to logic, possibility, and likelihood over emotion, (2) highly involved language relative to the other UNGA texts, and (3) repetition and contrast.

5.6.3. Ukraine

5.6.3.1. Introduction

Ukraine is a very large and very populous country in eastern Europe. Though its low GDP reveals a weak economy, high HDI scores indicate stronger social development (e.g.,

education levels). A former Soviet Republic, Ukraine has had a conflictual relationship with Russia since independence, and at the time the 2015 General Debate address was delivered, that conflictual relationship had turned particularly heated, with the Russian occupation of Ukrainian territory. The UN General Assembly had adopted Resolution 68/262 in March 2014 denouncing Russian annexation of Crimea, so Ukraine enjoyed the support of the majority of UN member states in its struggle to maintain, in the words of the UN Resolution, its "territorial integrity." The conflict between Russia and Ukraine shaped the Ukrainian address notably, in terms of both lexical and grammatical features. Keywords related specifically to the Russian occupation of Ukraine and grammatical structures signaled a text focusing on the description and overtly critical evaluation of past events and their effect on the people and places of Ukraine. Thus, the discourse style is strongly marked by the situation in Ukraine and by the condemnation of that situation by both the addressor, Ukraine, and addressees, the majority of UN member states.

5.6.3.2. Lexical features

Keyword analysis compared the Ukrainian wordlist with the wordlist for all other texts in the UNGA corpus. The keywords identified (see Table 5.36) reveal the topical focus in the Ukrainian address: the Russian occupation of Ukraine. Some of the keywords are more obviously related (*Russian*), while others require manual examination of the texts (*veto*).

Table 5.36. Keywords in the Ukrainian Text

Keyword	Frequency	%	R.C. frequency	R.C. %	Log likelihood	Log Ratio	P
Ukraine	27	0.86	39	0.02	129.71	5.26	0.00
Russian	15	0.48	10		87.61	6.37	0.00
Donbas	9	0.29	0		72.54	140.55	0.00
Russia	14	0.44	24	0.01	63.68	5.01	0.00
Ukrainians	6	0.19	1		42.65	8.37	0.00
Crimea	8	0.25	10		40.11	5.47	0.00
Ukrainian	7	0.22	6		38.69	6.01	0.00
aggression	10	0.32	28	0.02	37.80	4.30	0.00
veto	9	0.29	23	0.01	35.34	4.43	0.00
equipment	5	0.16	2		32.00	7.11	0.00
military	10	0.32	41	0.02	31.59	3.75	0.00
Ukraine's	5	0.16	4		28.08	6.11	0.00
hybrid	4	0.13	2		24.67	6.79	0.00

Note. R.C. = reference corpus (here, all texts except Ukraine).

Aggression and *military equipment*, for instance, are related to Russia: *For over 20 months, Russia's aggression against my country has been continuing through financing of terrorists and mercenaries, and supplies of arms and military equipment to the illegal armed groups in Donbas.* *Hybrid* appears in the phrase *hybrid war* and refers to the Russian use of the military and propaganda to create regional instability: *For over 24 years that have passed since the questionable procedure of transfer of the permanent Security Council membership of the former Soviet Union to the Russian Federation, it is not the only "hybrid" war that Russia has unleashed. In fact, in order to preserve its influence in neighboring countries, Russia for decades has deliberately created around itself the belt of instability.* President Poroshenko also criticizes Russia's abuse of its veto power on the Security Council, particularly on matters related to Ukraine: *Since the beginning of the aggression, Russia used its veto right twice, while the UN Security Council was considering questions related to Ukraine. At the outset, Russia blocked a draft resolution condemning "fake referendum" on Crimea's annexation in March 2014.*

The appearance of *Ukraine, Ukraine's, Ukrainian,* and *Ukrainians* on the keyword list is not surprising because it is not uncommon for speakers to refer to their country name when discussing national events and policies. What is more unusual is that all of the "self-reference" in the Ukrainian text inevitably leads back to the topic of the Russian occupation. This is true when discussing the complex history between Russian and Ukraine: *This country used to be a guarantor of Ukraine's security under the Budapest Memorandum, whereby security guarantees were provided to my country in exchange for a voluntary renunciation of the world's third nuclear arsenal.* It is likewise the case when discussing the current climate in Ukraine: *Millions of Ukrainians have found themselves under occupation.* Even in discussions of environmental protection, Ukrainian President Poroshenko finds a connection to Russian misdeeds (see Text 5.84).

Text 5.84. Ukrainian excerpt on environmental protection (keywords in italics)

> The path towards the achievement of the Sustainable Development Goals will not be successful without overcoming and preventing the consequences of environmental and technological disasters. As a result of the *Russian aggression*, *Ukraine* faces another challenge - the protection of the environment in *Donbas*. Irresponsible and criminal flooding of mines by terrorists led to the poisoning of drinking water, soil, flora and fauna of the region. The atmosphere is polluted due to explosions and shelling of sensitive industrial infrastructure. In fact, we can speak about the risk of environmental disaster.

The constant repetition of *Ukraine*, particularly when coupled with the repetition of *Russia(n)*, is a reminder of Ukraine's separate identity and sovereignty (see Text 5.85).

Text 5.85. Ukrainian excerpt illustrating repetition of *Ukraine* (keywords in italics)

> Full access of OSCE monitors to all occupied territories, withdrawal of the *Russian* military forces, military equipment as well as mercenaries from the territory of *Ukraine*, restoration of full control by *Ukraine* over the state border with *Russia* must be secured.

Freedom, peace, respect for the sovereignty and territorial integrity - *Ukraine* doesn't demand more. However, it will not settle for less.

Thus, the highly frequent and infrequent words in the Ukrainian text can be explained by the topical focus of the address: the Russian occupation of Ukraine.

5.6.3.3. Grammatical features

The analysis of grammatical features supports and further elucidates the findings of the lexical analysis in Section 5.6.3.2: The text from Ukraine is primarily an enumeration of objections to Russian behavior, first and foremost in relation to Russia's occupation of Ukrainian territory. Section 5.6.3.3 describes how grammatical structures accomplish this communicative purpose as well. The set of key features is initially surprising (see Table 5.37, see also Appendix M for Cohen's *d* scores of all grammatical features). The text is marked by frequent proper nouns and place nouns but infrequent adverbs of place; frequent use of *it* and *its*, but infrequent use of other third person pronouns; frequent complement clauses controlled by stance verbs including factive and communication verbs, as well as frequent past tense, but relatively infrequent public, causal, and existence verbs; frequent evaluative adjectives and prepositions and infrequent first person pronouns, markers of informational language, but also frequent second person pronouns, amplifiers, emphatics, conditional subordinate conjunctions, and *WH-*clauses, including *WH-* questions, which generally signal involved language. This section will explain these seemingly contradictory results and show how they all come together to achieve the primary purpose of this text: to describe and denounce Russian aggression against Ukraine.

Table 5.37. Select Grammatical Features in the Ukrainian Text

Feature	Ukraine frequency	UNGA Mean	Cohen's *d*
Verb (not including auxiliary)	101.8	113.00	-0.82
Past tense verb	14.5	10.71	0.96
Public verb	1.3	2.31	-0.80
Causative verb	1.6	3.98	-1.29
Existence verb	3.9	7.44	-1.22
All personal pronouns	30.8	45.39	-1.03
1st person pronoun	21.2	35.58	-1.18
2nd person pronoun	6.1	3.01	1.27
3rd person pronoun	3.5	6.80	-0.96
Pronoun 'it'	13.2	9.45	1.15
Proper noun	51.7	39.95	0.92
Place noun	10.6	6.46	1.40
Preposition	144.8	124.88	1.51
Attributive adjective – Evaluative	2.2	1.30	0.86
Adverb – place	1.6	4.19	-1.22
Amplifier	2.9	1.77	0.82
Emphatic	3.5	2.11	1.03
Subordinating conjunction – conditional	3.2	1.17	1.78
All wh- words	1.6	0.42	1.90
Wh question	1.6	0.28	2.36
All wh- relative clauses	7.4	3.65	1.83
Wh pronoun – subject relative clause	5.1	2.24	2.04
'That' CCC verb	4.2	3.23	0.59
'That' CCC communication verb	2.6	1.29	1.37
'That' CCC attitudinal verb	1.0	1.67	-0.65
'That' CCC factive verb	1.9	1.11	0.91
'That' CCC likelihood verb	1.3	0.91	0.45
All stance 'that' CCC verbs	6.7	4.97	0.85

Note. Frequencies normed per 1,000 words. CCC = complement clause controlled by.

The Ukrainian text is devoted in large part to describing the Russian occupation. As a result, the places and people affected by Russian aggression are highly frequent. *Ukraine, Russia, Donbas, Crimea, Russian,* and *Ukrainian(s)* all emerged as keywords (see Section 5.6.3.2). It is therefore not surprising that place nouns and proper nouns emerged as frequent

grammatical features ($d = 1.40$ and $d = 0.92$, respectively). In addition to place names in Russia

and Ukraine, some reference is made to other regions that are perceived by President Poroshenko

as victims of Russia: *In fact, in order to preserve its influence in neighboring countries, Russia*

for decades has deliberately created around itself the belt of instability. These are: Nagorny

Karabakh, Transnistria, Abkhazia, South Ossetia, Crimea and Donbas. Reference is also made

to specific people affected by the Russian occupation through the use of proper nouns: *I also feel*

obliged to mention the names of Nadiya Savchenko, Oleg Sentsov, Olexandr Kolchenko and

many other Ukrainians, political prisoners of the Kremlin, illegally detained and sentenced.

In contrast, adverbs of place are infrequent ($d = - 1.22$). Adverbs of place (e.g., *here,*

above, underground) are situation-dependent. That is, the physical surroundings described by the

addressor are well known by the addressee. The place names ubiquitous in the Ukrainian text are

not; they are specific, proper nouns, as shown above.

Given that most of the Ukrainian address centers around Russia and its occupation of

Ukraine, it is not surprising that the highly frequent pronouns *it* and *its* ($d = 1.15$) generally refer

to either Ukraine (see Text 5.86) or Russia (see Text 5.87).

Text 5.86. Ukrainian excerpt illustrating *its* reference to Ukraine (in italics)

> Unfortunately, not by *its* own free will today Ukraine is one of the areas of fight against
> terrorist threat.

Text 5.87. Ukrainian excerpt illustrating *its* reference to Russia (in italics)

> The second time Russia put *its* shameful veto on the draft resolution on establishment of
> the International Tribunal to investigate and bring to justice all responsible for Malaysian
> MH17 plane crush. By imposing *its* disgraceful veto on this draft resolution, Russia
> clearly demonstrated to the whole world *its* defiance in establishing the truth.

In contrast, other third person pronouns were infrequent ($d = -0.96$). Few sections of the text refer back to specific people (*he, she*) or discussed groups of people (*they*). Poroshenko does mention *Ukrainians* frequently, but tends to repeat the word or use near synonyms, as shown in Text 5.88.

Text 5.88. Ukrainian excerpt illustrating repetition of *Ukrainians* (in italics)

> Leading international human rights organizations are alerting about the radical deterioration of the human rights situation, which directly applies to *Ukrainians* and Crimean Tatars in the occupied Crimea. I am referring specifically to the practice, used by the occupation authorities of the Crimea: to enlist forcefully into Russian citizenship, as well as to systematic persecution, arrests, abductions and killings of pro-*Ukrainian residents* of the peninsula, and complete elimination of independent media. Ukraine reaffirms its commitment to the UN Declaration on the Rights of Indigenous Peoples. By all legal means, we will continue to defend the rights of the Crimean Tatars - *the indigenous people of Ukraine* – and the *Ukrainians*, who are suffering from the repressive policy of the occupation authorities in the Crimea.

In addition to illustrating the use of *it* and *its* in reference to Russia, Text 5.87 shows the powerful use of the evaluative adjectives *shameful* and *disgraceful.* Evaluative adjectives ($d = 0.86$) are often used to portray Russia as reprehensibly abusing its power. Other evaluative adjectives present Russia as a violent aggressor and Ukraine a tragic victim (*bloody invasion, brutal violation, illegal annexation, military reckless gamble, aggressive war, traitorous rhetoric*).

Again, because President Poroshenko is describing to a large extent what Russia *did* to Ukraine, the past tense is particularly frequent ($d = 0.96$). Text 5.89 illustrates the use of the past tense. It is also another example where the speaker begins a topic that does not initially appear to be related to Russia, but then ties it to Russian aggression.

Text 5.89. Ukrainian excerpt illustrating past tense (in italics)

> For the first time in 24 years of its independence, Ukraine *adopted* a National Human Rights Strategy. It *took* into account best international practices from the human rights perspective, including the EU Strategic Framework on Human Rights and Democracy. Russian aggression *exposed* the problem of securing the human rights in the Crimea and certain areas of Donetsk and Luhansk regions.

These sections, where President Poroshenko describes what Russia did to Ukraine, are highly informational. There are relatively few clauses compared to phrasal structures. As a result, overall in the text, there is infrequent use of first person pronouns ($d = -1.18$), verbs ($d = -0.82$) and in particular public verbs ($d = -0.80$), causal verbs ($d = -1.29$), and existential verbs ($d = -1.22$). In contrast, prepositions are frequent ($d = 1.51$). Text 5.90 illustrates the high density of phrasal structures and in particular prepositions.

Text 5.90. Ukrainian excerpt illustrating frequent prepositions (in italics)

> *For* over 20 months, Russia's aggression *against* my country has been continuing *through* financing *of* terrorists and mercenaries, and supplies *of* arms and military equipment *to* the illegal armed groups *in* Donbas. *Over* the last few days we have heard conciliatory statements *from* the Russian side *in* which, *in* particular, it called *for* the establishment *of* anti-terrorist coalition, or warned *of* fire danger to flirt *with* terrorists.

Though verbs overall are relatively infrequent, *that* verb complement clauses expressing stance are frequent ($d = 0.85$). That is, when verbs are used, they are used with greater relative frequency (compared to other UNGA texts) in complement clauses expressing stance. However, not all types of stance are equally frequent. *That* verb complement clauses controlled by a factive verb ($d = 0.91$) or controlled by a communication verb ($d = 1.37$) are common. Likelihood verbs are less frequent ($d = 0.45$) and attitudinal verbs actually infrequent ($d = -0.65$). Factive and communication verbs express more certainty than a likelihood verb (*believe*) and less emotion than an attitudinal verb (*condemn*). In the excerpt *However, we must recognize that in the 21st*

327

century our Organization lacks an effective instrument to bring the aggressor-country to justice,

which has stolen the territory of another sovereign state, the factive verb *recognize* implies a

truth that is simply being relayed to addressees. The same is true of the communication verb

state in: *Back to the situation in Donbas I have to state that here we are forced to fight proper,*

fully armed regular troops of the Armed Forces of the Russian Federation. Though such

complement clauses do not overtly convey opinion, they are frequently used to express

evaluation and specifically to criticize: *But, today, I have to recall that my country has become*

the object of external aggression.

The *WH-* relative clauses in the text often serve a similar function. Relative clauses are

either restrictive and identify the head noun (see Text 5.91) or are non-restrictive and provide

more information about it (see Text 5.92).

Text 5.91. Ukrainian excerpt illustrating restrictive relative clauses (in italics)

> The special peacekeeping mission in Donbas under the UN auspices could become a very
> useful instrument contributing to implementation of the Minsk Agreements. Ukraine is
> committed to follow the letter and the spirit of the Minsk deal. We demand the same
> approach from other signatories *that* have lately resorted to the language of blackmail.

Text 5.92. Ukrainian excerpt illustrating non-restrictive relative clauses

> By all legal means, we will continue to defend the rights of the Crimean Tatars - the
> indigenous people of Ukraine – and the Ukrainians, *who* are suffering from the repressive
> policy of the occupation authorities in the Crimea.

WH- relative clauses are frequent in the Ukrainian text ($d = 1.83$), particularly in subject

position ($d = 2.04$). As shown in Texts 5.91 and 5.92, they are used to add evaluative language.

In Text 5.91, Russia is indirectly accused of *blackmail* and in Text 5.92, Ukrainians are

described as *suffering from the repressive policy.* It is interesting to note than in both examples,

Russia is not named but sufficient information is given to identify the Russian government. *WH-*

words overall ($d = 1.90$) are used to this effect in the Ukrainian text. Text 5.93 shows how *WH-*

complement clauses can express stance, in this case by making the case for fighting the

information war.

Text 5.93. Ukrainian excerpt illustrating stance in a *WH-* complement clause (in italics)

> The ongoing hybrid war of Russia against Ukraine has demonstrated that the international
> community is facing another challenge, which requires consolidation of our efforts. The
> full-scale information war and propaganda campaign have become a particular
> destructive form of non-military aggression. Fake news, blatant lies spread to justify
> aggression, propaganda of intolerance and violence are phenomena of the same range,
> which undermine the principles of freedom of expression and poison human souls and
> minds. That is *why* the task of strengthening the role of information in the maintenance of
> peace and security is more important than ever.

WH- questions, too, frequently express stance in the Ukrainian address ($d = 2.86$) but of a

different kind. Five out of six of the *WH-* questions were in one passage of the text (see Text

5.94). These questions are used rhetorically, to express criticism and doubt, even irony.

**Text 5.94. Ukrainian excerpt illustrating impersonal *WH-* questions, conditional
subordinate conjunctions, and second person pronouns** (in italics)

> *How* can *you* urge an anti-terrorist coalition - *if you* inspire terrorism right in front of your
> door? *How* can *you* talk about peace and legitimacy - *if* your policy is war via puppet
> governments? *How* can *you* speak of freedom for nations - *if you* punish your neighbor
> for his choice? *How* can *you* demand respect for all - *if you* don't have respect for
> anyone? The Gospel of John teaches us: "In the beginning was the word". But *what kind*
> of a gospel do *you* bring to the world, *if* all your words are double-tongued like that?

This passage also illustrates the use of *if,* a conditional subordinate conjunction ($d = 1.78$)

and second person pronouns ($d = 1.27$). Some of the occurrences of the second person pronoun

refer to a member of the audience (e.g., *I wish you, Mr. President, every success in your activity*

in this crucial historic moment), but most were impersonal *you* such as those used in the

rhetorical questions of Text 5.94, which, given the context, appear to be addressed to Russia. The repetition of *WH-* questions is much more powerful than statements would be (*How can you urge an anti-terrorist coalition - if you inspire terrorism right in front of your door?* rather than *You cannot urge an anti-terrorist coalition if you inspire terrorism right in front of your door.*)

Likewise, amplifiers ($d = 0.82$) and emphatics ($d = 1.03$) are often used with rhetorical effect to express irony and criticize Russia, as shown in Texts 5.95 and 5.96.

Text 5.95. Ukrainian excerpt illustrating amplifiers (in italics)

Over the last few days we have heard conciliatory statements from the Russian side in which, in particular, it called for the establishment of anti-terrorist coalition, or warned of fire danger to flirt with terrorists. Cool story, but *really* hardly to believe!

Text 5.96. Ukrainian excerpt illustrating emphatics (in italics)

Heavy weaponry and military equipment are concentrated in the occupied territories in such quantities that armies of the majority of UN Member States can only dream about. These are, in particular, the state of art samples of military equipment of Russian production, which are unlikely, according to the well-known assumption of the Russian President might be purchased in an ordinary army store. Unless, *of course*, such a wholesale store, with free shipping, is from the Russian Federation.

While most of the text has an objective, informative tone, with place nouns, prepositions, and past tense used to detail events and present facts, these occasional phrases (*cool story, but really hardly to believe!*) and passages (Text 5.94) are highly involved and express skepticism and disdain. Such open criticism is very rare in UNGA texts. President Poroshenko may have felt at greater liberty to denounce the behavior of the Russian state given that Ukraine had clear support from the UN, with Resolution 68/262 condemning the Russian occupation of Crimea.

Overall, the key grammatical features in the Ukrainian text support some of the findings from the keyword analysis and elucidate other linguistic patterns as well. President Poroshenko

dedicates the bulk of his address to a description of Russian aggression toward Ukraine. Place

nouns and proper nouns are particularly frequent in order to describe where harm has been done.

In contrast, adverbs of place such as *here* and *above*, which imply a shared knowledge of place,

are not. *It* and *its* are used frequently in reference to Russia and Ukraine, but other third person

pronouns, which would relate to collective groups (*they*) or human beings (*she, he*) are not.

Evaluative adjectives, past tense, and prepositions are all used frequently in the description of the

Russian occupation of Ukraine. *WH-* relative clauses are also used to add details about places,

people, and events; these additional descriptions are phrased objectively but are in fact highly

evaluative. With the emphasis on recounting information and facts, verbs overall and first person

pronouns are infrequent. However, some clausal structures expressing stance were identified,

primarily showing certainty rather than likelihood. The only frequent features commonly

associated with involved language are second person pronouns, *WH-* questions, conditional

subordinate conjunctions, amplifiers, and emphatics, but these were concentrated primarily in

one passage of the text. Thus, the analysis of grammatical features reveals a combination of

structures that help detail the Russian occupation of Ukraine. Section 5.6.3.4 shows how many of

these features do not combine systematically in text varieties studied in the past (e.g., Biber,

1988), but rather are unique to the style of the Ukrainian president.

5.6.3.4. Multi-Dimensional scores

Section 5.6.3.3 presented the key grammatical features of the Ukrainian text. These

findings reveal a unique combination of structures. For instance, while most passages are

informational, a few are highly involved. As a result, Dimension 1 scores are moderate (D1 = -

13.46), only slightly more involved than the UNGA average (D1 = 15.26, see Figure 5.41).

Likewise, Dimensions 2 (D2 = -3.13), 4 (D4 = 0.43), and 5 (D5 = 0.86) are quite close to the

UNGA average (D2 = -3.08, D4 = -0.26, D5 = 1.67, see Figures 5.42, 5.44, 5.45).

The only dimension with a notable score is Dimension 3 (D3 = 4.00, compared to the

UNGA D3 = 0.78, see Figure 5.43). This positive value shows more explicit as opposed to

situation-dependent reference. With explicit reference, the addressor elaborates, identifying and

explaining rather than leaving inferences to be made by the addressee. Explicit reference makes

use of *WH-* clauses, phrasal coordination, and nominalizations. In contrast, situation-dependent

reference involves adverbs and in particular time and place adverbs. Frequent *WH-* clauses and

infrequent place adverbs are the most noteworthy features in the Ukrainian text and are discussed

in Section 5.6.3.3. Text 5.97 illustrates how these and the other features associated with

Dimension 3 are used for explicit reference.

Text 5.97. Ukrainian excerpt illustrating explicit reference (features associated with explicit
reference italicized; situation-dependent reference underlined)

> I call upon the UN *and* its Member States to launch a worldwide *campaign* to pressure
> Russian authorities to <u>immediately</u> release all Ukrainian citizens, *which* they hold
> hostage. We will be able to achieve our goal <u>only</u> if our *action* is global. Most of all,
> Ukraine needs solidarity *and assistance*, *which* is <u>really</u> a powerful instrument against
> *aggression and injustice*. Ukraine will win for sure because truth is on its side. But we
> will do it much <u>faster</u> if we feel *support and* solidarity of the whole international
> community.

Text 5.97 has a few occurrences of adverbs (*immediately, only, really, faster*), but no

adverbs of time or place. In contrast, there are several examples of *WH-* clauses (*which they hold*

hostage, which is really a powerful instrument), phrasal coordination (*and*), and nominalization

(*campaign, assistance*). Phrasal coordination and nominalization show how texts with explicit

reference often focus on noun phrases rather than clauses.

5.6.3.5. Summary

The analysis of the Ukrainian text reveals how lexical and grammatical forms combine to achieve two primary goals: to inform addressees of the consequences of the Russian occupation of Ukrainian territory and to persuade the UN to help Ukraine put an end to Russian aggression and to tackle the consequences. With UN support for Ukraine in a resolution condemning the Russian invasion in 2014, President Poroshenko was able to speak with more open criticism of the situation and occasionally used very direct structures such as the use of *you* and direct questions.

5.6.4. Eritrea

5.6.4.1. Introduction

Based on country classification outlined in Sections 3.6.1 and 3.6.2, Eritrea is a large country in the Horn of Africa. It has a large population and is one of the UN Least Developed Countries (LDCs) with a low HDI. After World War II, Eritrea passed from the hands of the Italians to the British. Then, in 1950, UN Resolution 390 declared Eritrea an autonomous entity under Ethiopian sovereignty. The Eritrean fight for independence from Ethiopia began less than a decade later and did not result in full independence until 1991. Independence did not, however, bring stability. From continued conflict with Ethiopia to UN sanctions, Eritrea's political, economic, and social situation has been tense. The UN has accused Eritrea of gross human rights violations and collaboration with terrorist organization Al-Shabab, which Eritrea claims is unfounded and unfair. The Eritrean address in 2015 was thus delivered in a context of poor development, poor prospects, and poor relationships with the international community.

The keyword analysis revealed only two keywords: *Eritrea* and *unfair* (see Table 5.38). Thus, the text from Eritrea tends to use the same words with a similar frequency compared to addresses overall in the UNGA. The country name *Eritrea* will not be examined in further detail because it is interspersed throughout the text and is considered self-reference (see discussion in Sections 3.7.1 and 5.3.1), but *unfair* merits further discussion.

Table 5.38. Keywords in the Eritrean Text

Keyword	Frequency	%	R.C. frequency	R.C. %	Log likelihood	Log Ratio	*P*
Eritrea	8	0.97	1		79.63	10.74	0.00
unfair	3	0.36	0		32.21	140.89	0.00

Note. R.C. = reference corpus (here, all texts except Eritrea).

The keyword *unfair* is used to describe inequality in the world and specifically in the United Nations (see Texts 5.98 and 5.99). It collocates twice with *unequal* in Text 5.98. Other related terms in Text 5.98 include *justice, undemocratic, marginalized,* and *bereft of real power and influence.* In Text 5.98, Foreign Minister Saleh laments the disparity between the rich and the poor, between possibility (treating curable diseases) and reality (death by those same curable diseases). He also denounces the discrepancy in decision-making within the UN, where a small minority of member states wield the greatest power.

Text 5.98. Eritrean excerpt illustrating *unfair* relating to equality (keyword in italics)

> Seventy years after the birth of the United Nations with its promise of peace, justice and development for all, of all nations acting in concert for the common good, it is undeniable that we continue to live in an *unfair* and unequal world, a world where conflicts and wars rage, extreme poverty persists in the midst of plenty, children die from easily preventable diseases and justice is routinely trampled. The United Nations Organisation itself is a reflection of this *unfair*, unequal and undemocratic global order. In the UN, the overwhelming majority of member states are marginalized, this assembly of nations

which should be the most powerful organ is bereft of real power and influence and power and decision making are dominated by a few among the few.

Text 5.99 narrows the lens from this discussion of overall inequality between the haves and the have-nots to a more specific discussion of what Foreign Minister Saleh considers to be the unfair treatment of Eritrea. He refers in this passage to the long fight for independence in Eritrea and to *baseless accusations* and *illegitimate sanctions* against Eritrea. This is most likely in reference to UN Security Council Resolution 1907, sanctions first imposed on the country in 2009 but which continued to be enforced to one degree or another through (and beyond) 2015. In addition to Security Council Resolution 1907, numerous UN documents denounce Eritrea over a range of issues, from accusations of Eritrean support of the terrorist organization Al-Shabab to gross human rights violations (United Nations Security Council Report, 2018).

Text 5.99. Eritrean excerpt illustrating *unfair* relating to sanctions against Eritrea (keyword in italics)

> Six decades ago we were denied of our inalienable right to self-determination and independence; for three decades, we were savagely bombed from the air and the ground, with the aim of crushing our liberation struggle; and today, we are subjected to *unfair* and illegitimate sanctions and baseless accusations.

The main focus of the Eritrean text is the unfair marginalization of less developed, less powerful nations in the world and specifically within the UN organization. This idea returns again and again, even in a passage of the text on environmental protection (see Text 5.100). The Eritrean emphasis on inequality is further shown through text excerpts in Section 5.6.4.3 on grammatical features.

Text 5.100. Eritrean excerpt on environmental protection and inequality

> We need to fight to avert an environmental catastrophe that awaits human kind and threatens human civilization. This will require more than human ingenuity and advances in science and technology. We must realize that there can be no technological fix of the environmental challenge. It will require radical change away from economic and social systems based on greed, maximization of profits for the few, massive inequality, unsustainable and unhealthy patterns of consumption and gross wastage.

5.6.4.3. Grammatical features

Unlike the lexical analysis, where only one meaningful keyword was identified, dozens of key features emerged in the analysis of grammatical features. Cohen's *d* scores for all grammatical features are reported in Appendix N. Section 5.6.4.3 analyzes the most compelling of these features (see Table 5.39), showing how the Eritrean address prefers dense noun phrases over verbs to create an information-packed text. The address also uses process nouns and passives to focus on an action and the patient of an action rather than the agent.

Table 5.39. Select Grammatical Features in the Eritrean Text

Feature	Eritrea frequency	UNGA Mean	Cohen's d
Verb (not including auxiliary)	91.4	113.00	-1.61
Present tense verb	64.6	77.31	-0.97
Past tense verb	6.1	10.71	-1.17
Perfect aspect verb	4.9	8.14	-1.31
Verb 'have'	0.0	1.43	-1.27
Private verb	4.9	7.40	-0.95
Public verb	0.0	2.31	-1.86
Suasive verb	2.4	1.45	0.94
Communication verb	4.9	8.06	-1.21
Causative verb	2.4	3.98	-0.85
Occurrence verb	1.2	2.67	-0.93
Existence verb	3.7	7.44	-1.29
All modals	7.3	14.39	-1.43
All passives	14.6	8.84	2.12
Passive verb + by	3.7	1.24	2.73
Passive postnominal modifier	3.7	1.70	1.93
1st person pronoun	21.9	35.58	-1.12
Common noun	285.0	250.37	1.86
Process noun	21.9	15.24	1.35
Determiner + stance noun	1.2	0.54	0.97
All adjectives	112.1	88.93	1.92
Attributive adjective	87.7	68.46	1.63
Coordinating conjunction – clausal	1.2	7.80	-1.25
Coordinating conjunction – phrasal	13.4	4.81	4.14
Wh pronoun – subject relative clause	3.7	2.24	1.02
'That' CCC verb	1.2	3.23	-1.26
All stance 'that' CCC adjectives	0.0	0.59	-0.87
'To' CCC verb of desire	6.1	2.39	2.76
'To' CCC evaluative adjective	1.2	0.11	4.39
All stance 'to' CCC verb	6.1	4.57	0.85
All stance 'to' CCC adjective	1.2	0.28	2.21

Note. Frequencies normed per 1,000 words. CCC = complement clause controlled by.

Overall, the Eritrean text uses more phrasal structures, fewer verbs. A dispreference was identified for verbs overall (d = - 1.61), and more specifically present tense verbs (d = - 0.97),

337

past tense verbs (d = - 1.17), perfect aspect (d = - 1.17), private verbs (d = - 0.95), public verbs (d = - 1.86), the verb *HAVE* (d = - 1.27), communication verbs (d = - 1.21), causal verbs (d = - 0.85), occurrence verbs (d = - 0.93), existential verbs (d = - 1.29), modals (d = - 1.43), verbs controlling *that* complement clauses (d = - 1.26), and effort verbs followed by the infinitive (d = - 1.56). Only three frequent verb features were identified: suasive verbs (e.g., *accord with*, d = 0.94), desire verbs followed by the infinitive (e.g., *wish to, d* = 2.76) and consequently stance verbs followed by the infinitive overall (e.g., *need to, d* = 0.85). These frequent structures result primarily from repetition of just a few phrases. *We need* to, both a stance verb and a desire verb followed by the infinitive, is repeated four times in the space of less than 100 words, all in reference to joint efforts within the international community (see Text 5.101).

Text 5.101. Eritrean excerpts illustrating *we need to* (in italics)

> *We need to* persist in our efforts to rebuild and revitalize the United Nations.
> *We need to* strengthen our solidarity (….)
> *We need to* strive and cooperate at the national, regional and global levels (….)
> *We need to* fight to avert an environmental catastrophe (….)

Generally, aside from these few exceptions, stance is expressed through adjectives rather than verbs. Adjectives overall are very frequent (d = 1.92) and in particular attributive adjectives (d = 1.63) as in *it will require <u>radical</u> change away from <u>economic</u> and <u>social</u> systems based on greed, maximization of profits for the few, <u>massive</u> inequality, <u>unsustainable</u> and <u>unhealthy</u> patterns of consumption and <u>gross</u> wastage*. Also frequent are stance adjectives followed by *to* (d = 2.21) such as *Eritrea welcomes the adoption of the Sustainable Development Goals of Agenda 2030 which accord with its development vision and priorities and is <u>determined to</u> achieve them ahead of the 2030 deadline*; evaluative adjectives followed by *to* are especially frequent (d = 4.39): *Eritrea is <u>committed to</u> striving for a fair, just, truly representative and effective United*

Nations. Attributive adjectives and adjectives followed by *to* complement clauses result in denser structures than finite clauses. In addition to the preference for finite clauses controlled by verbs, the Eritrean text shows a dispreference for the more elaborated *that* complement clauses controlled by stance adjectives ($d = - 0.87$).

Other features associated with dense noun phrases include common nouns ($d = 1.86$), determiner followed by stance nouns ($d = 0.97$), and phrasal coordinating conjunctions ($d = 4.14$) but not clausal coordinating conjunctions ($d = - 1.25$). *WH-* relative clauses also pack information into phrasal structures, and in subject position these, too, are frequent ($d = 1.02$). Finally, passive postnominal modifiers create dense phrases ($d = 1.93$). Text 5.102 illustrates these grammatical forms: common nouns (*unity*), determiner followed by stance noun (*the pitfalls*), phrasal coordinating conjunctions (*and*), *WH-* relative clauses in subject position (*that have infected the surrounding region and many other areas in the world*), and passive postnominal modifiers (*animated*).

Text 5.102. Eritrean excerpt illustrating dense phrasal structures (key features in italics)

> Achieving *unity* in *diversity*, with its *people animated* by a common *struggle* and committed to a common *future*, it has avoided the *pitfalls* of *sectarianism, radicalization and terrorism that* have infected the surrounding *region and* many other *areas* in the *world*.

These dense noun phrases create a more information-packed text, but noun phrases serve another purpose as well. Noun phrases can help emphasize the act itself or the patient of an action; they can even help avoid naming an agent altogether. In *the United Nations and its Security Council continue to countenance <u>the illegal occupation</u> of our sovereign territory in violation of international law and several Security Council resolutions*, the focus is on the

occupation itself rather than who is doing the occupying. *Occupation* is a "process noun" that

replaces a verb (*occupy*), another frequent feature in the Eritrean text ($d = 1.35$).

Passives, too, accomplish this function. In *justice is routinely trampled*, the critical

element is not the entity responsible for trampling, but rather the act of trampling and the patient,

justice. Text 5.103 shows a passage with particularly frequent use of passives ($d = 2.12$) to focus

on the patient, Eritrea (*we*).

Text 5.103. Eritrean excerpt illustrating the passive (in italics)

> Six decades ago we *were denied* of our inalienable right to self-determination and
> independence; for three decades, we *were* savagely *bombed* from the air and the ground,
> with the aim of crushing our liberation struggle; and today, *we are subjected* to unfair and
> illegitimate sanctions and baseless accusations.

Though short passives such as those in Text 5.103 leave the agent unnamed, even in *by-*

passives ($d = 2.73$) emphasis remains on the patient: *power and decision making* (see Texts

5.104).

Text 5.104. Eritrean excerpt illustrating the *by*-passive (in italics)

> In the UN, the overwhelming majority of member states are marginalized, this assembly
> of nations which should be the most powerful organ is bereft of real power and influence
> and power and decision making *are dominated by* a few among the few.

The use of noun phrases and passives in the Eritrean text serves as a compelling contrast

to the Ukrainian text. The representatives from both Eritrea and Ukraine express concern about

violated rights and violence against their countries. However, the Eritrean address avoids

focusing on the perpetrator of the aggression while the Ukrainian address repeatedly names the

culprit. This can be explained by the expected reactions of the addressees to these accusations.

Eritrea is accusing the UN of *unfair* treatment. The UN had repeatedly imposed sanctions on

Eritrea and was clearly not favorably disposed towards the country. In contrast, Ukraine enjoyed the support of the UN and in particular General Assembly Resolution 68/262 denouncing the Russian occupation (United Nations General Assembly, 2014). Political power and the status of the addressor in relation to addressees appear to influence choice of grammatical structures and discourse style.

The key feature analysis has shown that some grammatical forms have strongly influenced the overall discourse style of the Eritrean text. The dispreference for verbs in favor of noun clauses, with a particularly high frequency of common and process nouns, attributive adjectives, and WH- relative clauses creates an information-packed text. Complex noun phrases and passives also allow emphasis to be placed on the action or on the patient of an action rather than the agent. This serves a diplomatic purpose by helping the addressor focus on the action that is *unfair* rather than the agent who is being *unfair*.

5.6.4.4. Multi-Dimensional scores

This combination of key features results in strong scores along four dimensions: Dimension 1 (Eritrea = -21.2, compared to UNGA = -15.26, see Figure 5.41), Dimension 2 (Eritrea = -4.32, compared to UNGA = -3.08, see Figure 5.42), Dimension 3 (Eritrea = 5.32, compared to UNGA = 0.78, see Figure 5.43), and Dimension 4 (Eritrea = -2.62, compared to UNGA = -0.26, see Figure 5.44).

The low Dimension 1 score shows a more informational, rather than involved text. Unlike the text from the United States, the Eritrean text has fewer verb clauses, more noun phrases. Other features associated with informational texts include low frequencies of pronouns (d = - 0.88) and first person pronouns in particular (d = - 1.12). Text 5.105 illustrates this informational style, with involved features (first and second person pronouns, *it*, contractions,

present tense verbs, clausal coordinating conjunctions, *that* deletion, emphatics, sentence final prepositions) italicized and informational features (prepositions, nouns, attributive adjectives) underlined .

Text 5.105. Eritrean excerpt illustrating informational language (involved features in italics; informational features underlined)

<u>On</u> another <u>front</u>, <u>Eritrea</u> *is building* a <u>solid</u> <u>basis</u> <u>for</u> <u>sustainable</u> <u>development</u> <u>with</u> <u>social</u> <u>justice</u> by prioritizing <u>education</u>, <u>health</u>, <u>agriculture</u>, <u>industrialization</u>, <u>infrastructure</u>, <u>science</u> and <u>technology</u>, all <u>within</u> a <u>framework</u> <u>of</u> <u>regional</u> <u>cooperation</u> and <u>integration</u>, <u>in</u> the <u>Horn</u> <u>of</u> <u>Africa</u> and <u>the</u> <u>Gulf</u> <u>region</u> <u>across</u> the <u>Red</u> <u>Sea</u>.

The low Dimension 2 score indicates a less narrative text. This score is due to the infrequent use of past tense verbs, third person pronouns excluding *it*, perfect aspect, and public verbs as well as frequent present tense verbs, attributive adjectives and past-participial WHIZ deletions. Text 5.106 illustrates this non-narrative, more expository discourse style. No narrative features were identified in this excerpt. (See Texts 4.29 and 4.30 for contrasting examples of narrative texts.)

Text 5.106. Eritrean excerpt illustrating non-narrative discourse

We must realize that there can be no technological fix of the environmental challenge. It will require radical change away from economic and social systems based on greed, maximization of profits for the few, massive inequality, unsustainable and unhealthy patterns of consumption and gross wastage.

The very high Dimension 3 score signals more explicit reference rather than situation-dependent reference. Like the Ukrainian text, the high frequency of *WH-* relative clauses, phrasal coordination, and nominalizations and the low frequency of adverbs make reference in the Eritrean text more explicit. Nominalizations and *WH-* relative clauses are underlined in Text 5.107 to exemplify explicit reference. No features associated with situation-dependent reference

were identified in Text 5.107, but Texts 4.32 and 5.52 illustrate situation-dependent reference and can be referred to as a contrast to Text 5.107.

Text 5.107. Eritrean excerpt illustrating explicit reference (in italics)

> Two decades of *talk* to reform the *organization* has yielded zero *results*. The *resistance* of those *who* believe they benefit from the current *configuration* is so stiff and the ranks of those *who* seek *change* so divided by *approaches that* are driven by narrow national *interests*, the *enthusiasm* for *reform* is giving way to a disquieting *paralysis*.

Finally, the low Dimension 4 score signals less overt persuasion. Overt persuasion is carried out through infinitives, prediction modals, conditional subordination, and necessity modals. The only feature characteristic of overt persuasion in the Eritrean text is suasive verbs, as shown in Section 5.6.4.3. Little directive force is expressed in the Eritrean address. Rather, a description of events with evaluative adjectives conveys the addressor's opinion, defends a position, and suggests a course of action. Rather than demanding action with a modal of necessity (*we must fight human trafficking*) or showing the consequences of taking or not taking action with conditional subordination (*if we do not fight human trafficking…*), Foreign Minister Saleh informs addressees of the actions Eritrea is taking (*Eritrea is fighting human trafficking*). This not only demonstrates the goals deemed important by the Eritrean government but most importantly shows the contributions Eritrea is making toward achieving these goals, which helps to counter criticism against Eritrea. Text 5.108 illustrates this focus on Eritrean action rather than overt persuasion. No features characteristic of overt persuasion were identified in Text 5.108, but it can be contrasted with the highly overt persuasion in Text 5.115 from Tuvalu in Section 5.6.5.4.

Text 5.108. Eritrean excerpt illustrating focus on present action rather than overt persuasion (present tense in italics)

> Maintaining focus on development, Eritrea *is fighting* human trafficking, *stabilizing* illegal migration and *giving* youth and women adequate opportunities to pursue a high quality of life and build their nation. Eritrea *is* also *making* its contribution to regional peace, stability and security in the Red Sea and the Horn of Africa. Eritrea *sees* its own efforts for peace, justice, development, environmental protection as part of the global struggle for a better and more fair and equitable world; for a strong, effective and truly representative United Nations.

5.6.4.5. Summary

The analysis of the Eritrean text reveals a discourse style marked by highly informational, explicit language. Stance is rarely expressed directly and there is little overt persuasion. Instead, the address describes facts, past events, and present actions in order to defend the country's position. The lexical analysis did not reveal a great number of keywords, but the term that did emerge, *unfair*, articulates effectively the entire tone of the address. The message conveyed in the Eritrean text is that the decision-making process in the UN is not equitable for less powerful member states, including and particularly Eritrea. Foreign Minister Saleh perceives his country as being less influential and his discourse style reflects that. Rather than express strong directive force, he defends his country by informing his audience of how it has been misunderstood and mistreated.

5.6.5. Tuvalu

5.6.5.1. Introduction

Tuvalu is a Small Island Developing State in Oceania. With its small area and population and low GDP, it is classified as a least developed country by the UN. With its extremely low elevation, the highest point on the archipelago only 15 feet above sea level, it is particularly vulnerable to climate change and the resulting increasingly frequent cyclones. Each year,

cyclones destroy livelihoods and lives on the islands that make up Tuvalu. Over the longer term, rising ocean levels could very well submerge the archipelago, putting a definitive end to its existence. Foreign Minister Finikaso therefore calls climate change an "existential issue for Tuvalu" and focuses primarily on this issue in his address.

5.6.5.2. Lexical features

As with Eritrea, only one keyword apart from the country name itself was identified in the lexical analysis of the Tuvaluan address: *must* (see Table 5.40).

Table 5.40. Keywords in the Tuvaluan Text

Keyword	Frequency	%	R.C. frequency	R.C. %	Log likelihood	Log Ratio	*P*
Tuvalu	15	0.69	4		112.41	8.23	0.00
must	22	1.01	466	0.27	25.61	1.92	0.00

Note. R.C. = reference corpus (here, all texts except Tuvalu).

Must is a modal of necessity with the strongest directive power and initially may seem surprising from one of the smallest, least developed countries in the UN. Arguably, though, it is because Tuvalu is one of the smallest, least developed countries that Foreign Minister Finikaso perceives the need to use strong directive force when discussing climate change, which he calls an "existential issue for Tuvalu." Not only is climate change a critical topic for the country, it is also not very controversial. Therefore, the use of *must* is not particularly face-threatening, as shown in Text 5.109.

Text 5.109. Tuvaluan excerpt illustrating *must* (in italics)

The Paris Agreement therefore:
• *Must* reduce greenhouse gas emissions and keep the global average temperature rise to below 1.5 degrees Celsius;

• *must* recognize that climate change is a human right issue for Tuvaluans and many other millions; it is an urgent security and existential issue,
• *must* be committed to a low carbon future;
• *must* have a loss and damage architecture to recognize that preventing climate change is a national obligation for those who pollute and polluters
• *must* reduce emissions or pay to clean, mitigate, and aid those most vulnerable and without the means to adapt,
• *must* provide credible, timely public finance and clarity from developed countries to the pledges of $100 billion for climate change finances;
• *must* be reassuring for the private sector to transition and invest in clean energy and climate-resilient approaches.

Many of these directives do not actually bind addressees to anything tangible (*must be committed to a low carbon future*) or they restate previously agreed upon provisions of the Paris Agreement on Climate Change (*must reduce greenhouse gas emissions and keep the global average temperature rise to below 1.5 degrees Celsius*). *Must* will be discussed in further detail with other modals of necessity in Section 5.6.5.3 on grammatical features.

5.6.5.3. Grammatical features

The key feature analysis reveals clear expression of stance, particularly on environmental issues, with the use of modals of necessity, strong evaluative adjectives, suasive verbs, and some verb complement clauses (see Table 5.41, see also Appendix O for Cohen's *d* values of all grammatical features).

Table 5.41. Select Grammatical Features in the Tuvaluan Text

Feature	Tuvalu frequency	UNGA Mean	Cohen's *d*
Verb 'be'	4.6	2.10	1.67
Verb 'have'	3.3	1.43	1.68
Suasive verb	2.3	1.45	0.84
All modals	18.6	14.39	0.85
Modal of necessity	11.1	4.95	2.30
Agentless passive verb	7.9	5.91	0.93
Pronoun 'it'	6.5	9.45	-0.90
All nouns	320.5	303.03	0.70
Determiner + stance noun	0.0	0.54	-0.80
Predicative adjective	7.0	4.68	1.28
Predicative attitudinal adjective	2.8	1.33	1.40
Emphatic	3.3	2.11	0.88
'That' CCC attitudinal verb	2.8	1.67	1.10
'That' CCC factive verb	1.9	1.11	0.91
'That' CCC likelihood verb	0.0	0.91	-1.04
'To' CCC verb of desire	0.9	2.39	-1.07
'To' CCC verb of effort	0.0	1.66	-1.56
'To' CCC evaluative adjective	0.5	0.11	1.45
'To' CCC stance noun	0.9	3.28	-1.36

Note. Frequencies normed per 1,000 words. CCC = complement clause controlled by.

Foreign Minister Finikaso identifies essential UN goals for fostering environmental responsibility and overcoming obstacles to economic and social development and to the well-being of the people. This emphasis on an action (achieving goals) and on a patient (for development and the people) rather than on an agent (UN member states) results in structures such as the passive, existential *there*, and *HAVE* as a main verb. Section 5.6.5.3 shows how these structures create a discourse style that is explicit and direct but not face-threatening. The focus is on what needs to be done rather than who is responsible.

Expression of stance is strongest through the use of modals and in particular modals of necessity. As discussed in Section 5.6.5.2, *must* was extremely frequent in the text from Tuvalu,

and generally, the grammatical feature analysis identified many modals ($d = 0.85$) and modals of necessity ($d = 2.30$) in particular. Text 5.110 illustrates the use of modals.

Text 5.110. Tuvaluan excerpt illustrating modals (in italics)

> Every child, every woman, every citizen, every community, *should* know these Sustainable Development Goals and own the rights under these goals. These goals *must* be advocated as widely as possible to all global citizens, through technological advances that we have for information and any communication, so that these goals *can* be understood, owned and adhered to by one and all. These goals *must* bring accountability to our leaders, to our development partners, to the private sector, to religious bodies and to the youthful generation.

As with Text 5.109 in Section 5.6.5.2, Text 5.110 shows how these modals are used to suggest ideas that are not particularly contentious or binding. *These goals must be advocated as widely as possible to all global citizens* does not commit member states to anything concrete. Moreover, given that the Sustainable Development Goals had been adopted just prior to the delivery of this address, their benefit had already been widely recognized and was no longer under debate.

Text 5.110 also illustrates the use of short passives (*these goals must be advocated, these goals can be understood*). The short passive puts the focus on the patient rather than the agent in this and several other passages in the Tuvaluan text ($d = 0.93$). In Text 5.110, the goals are more important than the member states advocating the goals; it is assumed by the speaker that the agent, the member states, do not need to be identified. This is also true in *the new SDGs framework must ensure accountabilities are elaborated and enforced*. In other cases, the short passive is used when the agent is either unknown or less important than the patient. In *a significant number of "we the people", globally, are still homeless, jobless, many more are becoming displaced and stateless, many are deprived of opportunities to education and health*

348

services and many still are hopeless for a future, and as such extremism has become an alternative choice of existence for the disoriented if not, lost souls, the agent depriving people of opportunities is unknown or perhaps multiple: There may be several factors depriving people of opportunities.

Two other structures serve this same function in the text from Tuvalu: existential *there* and *we have*, with resulting high frequencies of the main verbs *BE* ($d = 1.67$) and *HAVE* ($d = 1.68$). Existential *there* puts the emphasis on the notional subject. In *Yet Mr. President, we must be honest with ourselves and note that there are persisting inequities, there are recurring insecurities and there are planetary distresses that require our urgent attention*, the focus is placed on *persisting inequities, recurring insecurities,* and *planetary distresses.* An alternative syntax: *We must urgently address persisting inequalities* or *It is urgent to address persisting inequalities* would not stress *persisting inequalities* to nearly the same degree. The phrase *we have* is used in a similar fashion in *we must have a good review mechanism to gauge whether cumulatively we are reducing emissions and not increasing them*; *there must be a good review mechanism* would not change markedly the meaning or focus in the sentence. Similarly, *we must have a United Nations answer, and our Agenda 2030 for Sustainable Development with its 17 goals and 169 targets, must deliver for the people* could be phrased *there must be a United Nations answer.* Neither of these structures results in the dense information-packing common to many UN addresses and other formal written texts. Instead, they are features more common in unplanned, oral language such as face-to-face conversation. Biber et al. (1999) do, however, note that existential *there* is used in academic prose for argumentation (pp. 953-954) and this would appear to be the case in the address from Tuvalu as well. Foreign Minister Finikaso is arguing for a *good review mechanism* and for a *United Nations answer*.

While the verb *HAVE* occurs as a main verb only in *we… have,* as shown above, the verb *BE* is also frequently used as a main verb with a predicative adjective ($d = 1.28$) and in particular with an attitudinal predicative adjective ($d = 1.40$) or an evaluative adjective followed by the infinitive ($d = 1.45$). These predicative adjectives express stance and in nearly all cases are used to argue the importance of climate change or developmental issues (see Text 5.111).

Text 5.111. Tuvaluan excerpt illustrating predicative adjectives to express stance (in italics)

> Mr. President, the 2030 Agenda will *be meaningless to* many of us low lying SIDS, if a credible Climate Change Paris Agreement in December 2015 *is* not *ambitious* and *action oriented*. Monitoring and reviewing our SDGs progress throughout the next fifteen years *is vital to* sustain and focus our transformative agenda, per country. Sea level rise continues to inundate many of our small island coastlines and inundate our food plantations. That *is* a security issue, an *urgent* one and an *inter-generational* one.

Other structures used to express stance include suasive verbs ($d = 0.84$) as well as *that* complement clauses controlled by attitudinal verbs ($d = 1.10$) and factive verbs ($d = 0.91$). Suasive verbs (e.g., *pledges*) imply plans for action or attainment, as in *Tuvalu pledges its full support for the new agenda, "Transforming Our World: The 2030 Agenda for Sustainable Development" and for the GA 70th Session "A New Commitment to Action."* Attitudinal verbs (e.g., *hope*) express viewpoints and feelings: *We sincerely hope that the embargo against Cuba will be lifted soon and for continued dialogue and smooth transition and resumption of economic partnerships in the near future*. Factive verbs (e.g., *remind*) express certainty: *Mr. President recent global social and economic tensions and market jitters remind us that we are of course not immune to a repeat of the global financial crisis; the cost and prolonged painful recovery even in the far remote, small economies like Tuvalu, are still vivid in our memories*. Overall, however, there is little expression of stance through lexical verbs, with negative Cohen's *d* values for stance verbs followed by the infinitive ($d = -2.07$), desire verbs followed by the

infinitive (d = - 1.07), effort verbs followed by the infinitive (d = - 1.56), and *that* complement

clauses controlled by a verb of likelihood (d = - 1.04). Similarly, stance is rarely expressed

through nouns, with highly infrequent occurrences of determiners followed by a stance noun (d =

- 0.80) and stance nouns followed by the infinitive (d = - 1.36).

The analysis of grammatical features has shown that stance is expressed with greater

frequency relative to other UNGA addresses through modals of necessity, predicative adjectives,

suasive verbs, and occasionally through *that* verb complement clauses. In addition, short

passives, existential *there*, and the main verb *have* are used to focus on the patient or the result of

an action rather than the agent. Textual examples showed how these structures emphasize UN

goals and actions and the people they are intended to serve rather than the member states

themselves.

5.6.5.4. Multi-Dimensional scores

Four dimensions emerged as strongly deviating from the UNGA average: Dimension 2 (-

3.76, compared to UNGA -3.08, see Figure 5.42), Dimension 3 (2.6, compared to UNGA 0.78,

see Figure 5.43), Dimension 4 (2.37, compared to UNGA -0.26, see Figure 5.44), and Dimension

5 (-1.25, compared to UNGA 1.67, see Figure 5.45). Though Dimension 1 did not emerge as

exceptionally low (-16.59, compared to UNGA -15.26, see Figure 5.41), it merits some

discussion as well.

Tuvalu did not have a strong overall score for Dimension 1, not because grammatical

features associated with this dimension were all similar to the UNGA average but because some

key grammatical features identified as particularly frequent or infrequent were involved while

others were informational. A positive Cohen's *d* value for emphatics (d = 0.88) and a negative

value for prepositions (d = - 1.37) signal involved language; a negative value for private verbs (d

= - 0.88) and for the pronoun *it* (*d* = - 0.90) indicate informational language. This mixture of involved and informational language is illustrated in Text 5.112, with involved features (first and second person pronouns, *it*, contractions, present tense verbs, clausal coordinating conjunctions, *that* deletion, emphatics, sentence final prepositions) italicized and informational features (prepositions, nouns, attributive adjectives) underlined.

Text 5.112. Excerpt illustrating both involved and informational language (involved features in italics; informational underlined)

> Mr. <u>President</u>, *as* <u>custodians</u> <u>of</u> the <u>Pacific</u> <u>Ocean</u>, <u>Tuvalu</u> fully *supports* <u>SDG</u> 14. *We are* all <u>oceanic</u> <u>States</u> *as* our <u>planet</u> and <u>mother</u> <u>earth</u> *is* 70 plus <u>percentage</u> Blue. The blue <u>oceans</u> *is* not a <u>sink</u> <u>for</u> radio-active <u>spillovers</u> <u>of</u> nuclear <u>wastes</u>, *it's* not a <u>dumb</u> <u>for</u> <u>industrial</u> and <u>general</u> <u>garbage</u>; *it's* not a <u>carpet</u> to *sweep* and *hide* our <u>dirt</u> <u>under</u>.

The negative Dimension 2 score reveals a less narrative text, with fewer instances of past tense verbs, third person pronouns excluding *it*, perfect aspect, and public verbs. Text 5.113 illustrates the more expository, non-narrative discourse. No narrative features were identified in this excerpt.

Text 5.113. Tuvaluan excerpt illustrating non-narrative discourse

> But as much as we need capacity enhancement, science, data and new technology, we must match this with Leadership integrity. The recent papal encyclical of His Holiness Pope Francis agrees that our leadership roles must be free of political wrangling, power squabbling, conflicts, trade imbalances, bad governance, race and gender differentiation, self-interests, greed and profit-only mentality and ignorance of nature's deterioration. We the people need SDGs championed by good accountable leaders.

The high Dimension 3 score indicates explicit reference rather than situation-dependent reference. Like the Ukrainian and Eritrean texts but contrary to the US text, the high frequency of WH relative clauses, phrasal coordination, and nominalizations and the low frequency of adverbs make reference in the Tuvaluan text more explicit.

Text 5.114. Tuvaluan excerpt illustrating explicit reference (explicit features in italics; situation-dependent underlined)

> Mr. President recent global social *and* economic *tensions and* market *jitters* remind us that we are of course not immune to a *repeat* of the global financial crisis; the *cost and* prolonged painful *recovery* even in the far remote, small economies like Tuvalu, are still vivid in our memories. The new SDGs framework must ensure *accountabilities* are elaborated *and* enforced, <u>especially</u> for the rich echelons, *who* are untouchable *and* ignorant of the <u>far</u> reaching repercussions of their *greed* games <u>especially</u> to the small genuine *investors*.

Finally, the high Dimension 4 scores indicate a great deal of overt persuasion. Overt persuasion is carried out through infinitives, prediction modals, conditional subordination, and necessity modals, which are italicized in Text 5.115.

Text 5.115. Tuvaluan excerpt illustrating overt persuasion (persuasive features in italics)

> Mr. President, Tuvalu *will* submit its INDC before Paris and we *must* all note that our obligation is not only *to submit* our INDCs but also *to ensure* that we achieve those set targets. There *must* be no backsliding. We *must* have a good review mechanism *to gauge* whether cumulatively we are reducing emissions and not increasing them.

The argumentation style is particularly interesting in Text 5.115. The speaker begins by promising national action before asking the international community to do the same. The subjects in this passage are *Tuvalu*, inclusive of the speaker but not the audience, then only later *we,* including both speaker and audience. Existential *there* is also used to avoid naming an agent. Though the three occurrences of *must* indicate strong directive force, this is mitigated by the choice of subjects.

Dimension scores reveal a text that combines both involved and informational language. Though brief references to past events such as Tropical Cyclone Pam are occasionally made, the Tuvaluan address is primarily non-narrative, with a focus on current action and goals. Reference

is explicit and persuasion is overt. Both are indicative of clear and direct descriptions and expression of stance.

5.6.5.5. Summary

Based on the analysis of lexical and grammatical features, the Tuvaluan text is an explicit discussion of two critical and not entirely unrelated problems: climate change and development. The address describes the rationale for supporting UN goals to tackle these issues, but focuses on the action to be taken rather than the agents who must carry out the action. This mitigates the strong directive force of forms such as *must.* Because directives are phrased to imply mutual concern and consensus, the text suggests that the international community is not being subjected to a litany of complaints and orders but rather that the other UN member states are willing participants in a plan to achieve a set of common goals.

5.6.6. Bhutan

5.6.6.1. Introduction

Bhutan is a medium sized, medium populated country in the eastern Himalayas of Asia. Classified as a Landlocked Developing Country and a Least Developed Country by the UN, it has a low GDP. Its slightly higher HDI, categorized as "medium" by the UNDP, is a result of its focus on social development. Bhutan prefers to measure national progress not through GDP but through its unique "GNH," or Gross National Happiness, with indicators such as life expectancy and level of education.

5.6.6.2. Lexical features

The lexical feature analysis identified three keywords in addition to *Bhutan*, the country name itself: *happiness, agenda,* and *whether* (see Table 5.42).

Table 5.42. Keywords in the Bhutanese Text

Keyword	Frequency	%	R.C. frequency	R.C. %	Log likelihood	Log Ratio	*P*
Bhutan	12	0.69	0		110.83	141.81	0.00
happiness	5	0.29	2		37.84	7.97	0.00
agenda	17	0.97	351	0.20	26.23	2.28	0.00
whether	6	0.34	25	0.01	25.45	4.59	0.00

Note. R.C. = reference corpus (here, all texts except Bhutan).

Rather than measure national success based on income with either GDP or Gross National Product, Bhutan has opted for an indicator known as Gross National Happiness (GNH), so it should come as no surprise that *happiness* is one of the keywords in the Bhutanese text. According to Foreign Minister Dorji, the GNH is an "all-inclusive approach to development that puts people and the environment at the center." Dorji praises the 2030 Agenda and its Sustainable Development Goals as in keeping with the values of the GNH. Texts 5.116 and 5.117 illustrate the focus on happiness and its relationship to development in the Bhutanese address.

Text 5.116. Bhutanese excerpt describing Gross National Happiness (keyword in italics)

Guided by our development philosophy of Gross National *Happiness* (GNH), which is essentially "development with values", we have consistently worked to ensure that the peace, security, *happiness* and the well-being of our people always remain at the center of development.

Text 5.117. Bhutanese excerpt relating Gross National Happiness to the 2030 Agenda (keywords in italics)

The pursuit of *happiness* is a fundamental human goal and embodies the spirit of the 2030 *Agenda*. We stand ready to share our experience with GNH and contribute to the growing discourse on holistic development paradigms and indicators.

The keyword *Agenda* appears solely in relation to the 2030 Agenda for Sustainable

Development, revealing the focus on developmental issues and the Agenda's Sustainable

Development Goals (SDGs) in particular (see Text 5.118).

Text 5.118. Bhutanese excerpt illustrating the use of *Agenda* (in italics)

> Many of the SDGs in the 2030 *Agenda* will not be attainable if we do not relentlessly invest in building a dynamic and relevant education system in our countries - a system that nurtures our children and youth with the right values, knowledge and skills that reflect our aspirations and goals.

The last keyword, *whether* is used in three instances as a post-predicate complement

clause (e.g., *in moving forward, the first litmus test for our commitment to action on the 2030*

Agenda will be whether or not we reach an ambitious and legally binding agreement at COP21).

The three other occurrences of *whether* are as a series of circumstance adverbials, as shown in

Text 5.119.

Text 5.119. Bhutanese excerpt illustrating the use of *whether* as a circumstance adverbial
(in italics)

> For every day we are confronted with compelling evidence that shows we are far from achieving our quest for peace, security, prosperity and human dignity for all.
> • *Whether* it is conflicts and acts of terrorism; or the unthinkable stockpiles of weapons of mass destruction;
> • *Whether* it is the inequalities that persist within and among nations or the rising expectations of billions living on the very margins of existence;
> • *Whether* it is climate change that threatens our very survival; Failure to address these interrelated global challenges is not an option.

In all six occurrences of *whether,* it is used ostensibly to present conditions, but also to

propose goals (*to reach an ambitious and legally binding agreement at COP21*) and key issues

(*conflicts and acts of terrorism, climate change*).

All three keywords do, in fact, highlight essential topics worthy of the attention of UN member states. *Happiness* is presented as a critical objective and a means of measuring the level of development in a country. The 2030 *Agenda* is discussed as a crucial plan of action for achieving sustainable development. Finally and more generally, *whether* is used as a structural device to introduce matters of greatest concern. Section 5.6.6.3 discusses in further detail grammatical features used to stress the importance of particular UN objectives and to show the steps being taken by Bhutan to help achieve those objectives.

5.6.6.3. Grammatical features

The analysis of key grammatical features reveals a unique discourse style in the Bhutanese text. As shown in Table 5.43, some of the most notable grammatical forms include first person pronouns, stance nouns, attitudinal adjectives and adverbs, progressive aspect, modals of prediction, and *by*-passives. (See Appendix P for Cohen's *d* values for all grammatical features.) Section 5.6.6.3 shows how a number of these structures are used to emphasize the importance of joint cooperation and solidarity in achieving "development with values."

Table 5.43. Select Grammatical Features in the Bhutanese Text

Feature	Bhutan frequency	UNGA Mean	Cohen's *d*
Progressive aspect verb	10.4	7.28	1.13
Verb 'be'	4.1	2.10	1.33
Suasive verb	2.9	1.45	1.44
Activity verb	20.3	16.13	0.89
Causative verb	8.1	3.98	2.27
Existence verb	9.9	7.44	0.84
Modal of prediction	10.4	5.83	1.81
1st person pronoun	51.6	35.58	1.32
3rd person pronoun	3.5	6.80	-0.96
Stance noun + preposition	3.5	1.84	1.15
Predicative attitudinal adjective	2.9	1.33	1.49
All adverbs	22.6	30.22	-1.10
All stance adverbs	3.5	1.81	1.09
Factive adverb	3.5	1.25	1.80
'That' CCC verb	4.6	3.23	0.84
'That' CCC attitudinal verb	2.9	1.67	1.20
'To' CCC stance noun	5.2	3.28	1.10

Note. Frequencies normed per 1,000 words. CCC = complement clause controlled by.

The style of the Bhutanese text is noteworthy in its emphasis placed on solidarity and cooperation. Interestingly, no keywords in Section 5.6.6.2 illustrate this focus on joint efforts, primarily because so many synonyms express this idea in the Bhutanese text (e.g., *support, collective will, commitment, together, shared goals and aspirations, collaboration, consensus, solidarity*) and thus no one word was key. However, many of these words are stance nouns (e.g., *agreement* in *this agreement must be firmly anchored in the principles of the UNFCCC, including the principle of common but differentiated responsibilities, d = 1.15*), and some were stance nouns followed by *to* complement clauses (e.g., *opportunity* in *opportunity to truly awaken our spirit of global solidarity*, d = 1.10), so these expressions of solidarity and cooperation emerge not in the lexical but in the grammatical analysis.

Another key grammatical feature signaling collaboration was first person pronouns ($d =$ 1.32), nearly exclusively in the form of the inclusive *we* and *us* (see Text 5.120). In contrast, third person pronouns excluding *it* were infrequent ($d = - 0.96$), indicating that the frequent use of *we* and *us* was not in an "*us* versus *them*" polarization but rather a consistent focus on the international community as an inclusive and unified body.

Text 5.120. Bhutanese excerpt illustrating first person plural pronouns (in italics)

> In the words of His Majesty the King of Bhutan: "Let *us* place the interest of humanity, not national populations and constituencies, above all else. Let *us* take political risks and strong decisions in addressing the needs of humanity. The answer to global problems will come closer at hand when *we* grasp that universal simplicity - that sense of a shared planet and a shared fate for those who walk on it. *We* need shared human endeavor not just negotiated change ".

A number of suasive verbs ($d = 1.44$), action verbs ($d = 0.89$), verbs of causation ($d = 2.27$), and existence verbs ($d = 0.84$) also show Bhutan's dedication to UN objectives in spite of the difficulty of doing so for a least developed country. In *a key challenge we* [Bhutan] *face is to build a strong and sustainable green economy that ensures gainful employment to our youth, inclusive growth and promotes self-reliance*, *build* is an action verb, *ensures* and *promotes* are verbs of causation. In Text 5.121, *pledged* is a suasive verb and *constitute* is an existence verb.

Text 5.121. Bhutanese excerpt illustrating suasive and existence verbs (in italics)

> In 2009, Bhutan *pledged* to remain carbon neutral for all times. We have kept that promise! In fact, our expectations are that by 2025, Bhutan through the sale of our clean hydro-power would have the capacity to offset approximately 35 million tonnes of carbon per annum in the region. This is no small feat considering Bhutan's emissions *constitute* under 7 percent of that figure.

Other grammatical features used to express support for joint efforts within the international community include *that* verb complement clauses ($d = 0.84$), in particular when

controlled by an attitudinal verb (e.g., *hope*, d = 1.20) such as *we hope that our development partners will share our sense of priority and urgency in this vital area and support our efforts to ensure that it is relevant to achieve the 2030 Agenda*. Attitudinal adjectives in predicative position (e.g., *pleased*, *d* = 1.49) also serve this function: *I am pleased to state that this week Bhutan ratified the Doha Amendment to the Kyoto Protocol and submitted its Intended Nationally Determined Contributions to the UN Framework Convention on Climate Change (UNFCCC)*.

Attitudinal adjectives in predicative position are partly responsible for the high frequency of *BE* as a main verb (*d* = 1.33). Additionally, *BE* is used in definitions or explanations: *The pursuit of happiness is a fundamental human goal and embodies the spirit of the 2030 Agenda*. *BE* is used twice as a main verb in the excerpt: *In this regard, we are happy to note that ensuring inclusive and equitable quality education for all is a stand-alone goal in the 2030 Agenda*. In the first clause, it is followed by an attitudinal adjective and in the final clause it serves to define. This passage illustrates how attitudinal adjectives are yet another structure used to show support of UN objectives.

Passives, too, are used to endorse development goals, most frequently with *by*-passives (*d* = 1.80). As with other UNGA texts (see in particular Sections 5.6.4.3 and 5.6.5.3), passives are used to focus on the patient or the result of the action (e.g., *our decisions in Paris must be driven by the latest science*). Thus, the subject of the sentence, the discourse focus, is on the importance of UN goals and the steps Bhutan is taking to attain them at a national level (see Text 5.122).

Text 5.122. Bhutanese excerpt illustrating *by*-passives (in italics)

> We remain a bastion of environmental conservation with 72 percent of land under forest cover. Our effort to safeguard the environment *is reinforced by* a constitutional mandate to maintain a minimum of 60 percent of our land under forest cover for all times.

The Bhutanese text describes current action and future plans to demonstrate the country's dedication to shared goals. This results in a high frequency of progressive verbs ($d = 1.13$) as shown in Text 5.123 and of prediction modals ($d = 1.81$), as shown in Text 5.124. Prediction modals are also used to encourage other countries to take action (see Text 5.124).

Text 5.123. Bhutanese excerpt illustrating progressive verbs (in italics)

> We remain a committed partner and are *working* to deepen and broaden our peacekeeping engagement with the United Nations. Today, it is a matter of great pride that Bhutanese peacekeepers are deployed and *serving* in 9 peacekeeping missions.

Text 5.124. Bhutanese excerpt illustrating prediction modals (in italics)

> Achieving the SDGs *will* require countries, both developed and developing alike, to take stock of their existing development strategies and embark on far reaching reforms to implement the Agenda. For a country like Bhutan, that is least developed and landlocked, addressing the challenges *will* require innovative and creative approaches to build on the progress we have achieved thus far.

Finally, some adverbs are used to express stance, either certainty or attitude, in Bhutan's commitment to shared objectives. Though adverbs overall were infrequent ($d = - 1.10$), the key feature analysis identified frequent stance adverbs (e.g., *necessarily*, $d = 1.09$) such as *the 2030 Agenda sets out a necessarily ambitious vision and SDGs*. Factive adverbs were particularly common (e.g., *in fact*, $d = 1.80$). Factive adverbs are used to express certainty, again in relation to the importance of UN objectives or in Bhutan's commitment to achieving them: *Despite our limited resources and competing demands of development, Bhutan remains committed to the conservation of our natural heritage and in fact has consistently done more than its fair share to contribute to global efforts to combat climate change*.

Overall, the key grammatical features reveal a number of structures used to emphasize the importance of joint cooperation and solidarity in achieving "development with values." The

inclusive *we,* stance nouns, attitudinal adjectives and adverbs, progressive aspect, modals of prediction, and *by*-passives are all used to stress the need to fulfil the 2030 Agenda and its objectives in support of human well-being and environmental protection as well as to highlight steps taken by Bhutan to reach those objectives.

5.6.6.4. Multi-Dimensional scores

The only strong Multi-Dimensional score to emerge was for Dimension 5 (4.04, compared to UNGA 1.67, see Figure 5.45). Texts with high positive Dimension 5 scores tend to be more abstract, technical, and formal. Features associated with positive Dimension 5 scores include conjuncts, passives, adverbial past participial clauses, adverbial subordinators, and attributive adjectives. These grammatical forms are italicized as illustration in Text 5.125.

Text 5.125. Bhutanese excerpt illustrating abstract language (in italics)

> A *careful* re-orientation of our economy is imperative *if* we are to ensure that our *future* development is inclusive and sustainable *as envisaged* in the 2030 Agenda. *In this regard*, the *continued* support and cooperation of our development partners will be crucial *as* we seek *innovative* approaches to balance our *environmental* aspirations and *developmental* needs. This assumes even *greater* importance *as* Bhutan progresses towards graduation from the LDC category, to ensure that *hard-earned developmental* gains *are sustained*.

Text 5.125 shows examples of conjuncts (*in this regard*), passives (*are sustained*), adverbial past participial clauses (*as envisaged in the 2030 Agenda*), adverbial subordinators (*as*), and attributive adjectives (*careful*). Once again, these grammatical forms are used to focus on the ways in which Bhutan is working toward UN objectives at the national level.

5.6.6.5. Summary

The text from Bhutan emphasizes the need for joint efforts to achieve UN goals, and more specifically developmental goals that promote the well-being and happiness of the people. This emphasis is achieved lexically and grammatically, from keywords such as the ubiquitous (2030) *Agenda* (for Sustainable Development) to the use of progressive aspect to describe steps Bhutan is carrying out in support of development goals. The discourse style is abstract and formal. Syntactically, focus is centered on an action or outcome by placing the objective rather than an agent in subject position (*our effort to safeguard the environment* in Text 5.122, *achieving the SDGs* in Text 5.124, *the continued support and cooperation of our development partners* in Text 5.125).

5.7. Synthesis

Chapter 5 began with a comparison of texts by country group. States had been categorized based on geographic, social, political, and economic factors that are commonly used in political science and in international organizations themselves. The objective was to determine whether and, if so, how groups of countries made distinctive use of lexical and grammatical forms. Section 5.3 showed that a few groups of countries such as Small Island Developing States typically address the same topics (e.g., climate change) and therefore certain patterns in lexical forms do emerge based on some country classifications. However, no meaningful keywords were identified for the majority of country groups. Even fewer patterns were identified in grammatical features and in Multi-Dimensional scores by country group. I therefore determined that geographic, social, political, and economic factors do not inherently lead to the preference or dispreference for a fixed set of lexical and grammatical forms.

Rather, individual speakers representing individual countries opt for unique combinations of words and grammatical structures. Thus, just because a country is in the Americas does not mean it will use possibility modals frequently. However, it is easy to explain the communicative purpose of these possibility modals if a country like the United States is trying to avoid strong directive force in expression of stance. The lexical and grammatical forms chosen serve communicative functions which can then be interpreted based on geographic, social, economic, and political factors, but membership in a category such as location in the Americas does not predict the use of a fixed set of forms.

From a methodological perspective, Chapter 5 illustrates the utility of both quantitative and qualitative analysis. The quantitative analysis shows some general patterns and tendencies, such as the more developed countries using more involved language and the less developed countries adhering to the traditional rhetorical conventions of General Assembly addresses. In contrast, the qualitative analysis reveals how each individual country achieves a particular discourse style based on its preference or dispreference for a unique set of lexical and grammatical forms. The structures adopted combine not only to accomplish a communicative purpose such as denouncing an invading "aggressor" but also to achieve an individual style.

Finally, it is important to consider the findings from Chapter 5 in light of the results of the analyses for the UNGA as a whole. Chapter 4 showed that overall, UNGA texts are relatively similar. The lexical analysis identified numerous keywords shared by all or virtually all UNGA countries; the standard deviations of grammatical feature frequencies and dimension scores are relatively small compared to previously analyzed registers. This indicates little variation within the UNGA. It is therefore not surprising that few consistent differences could be identified when comparing groups of countries within the UNGA; there was little variation *to be* identified. More

364

revelatory is the analysis of individual countries with more extreme frequency values. The final analysis in Section 5.6 shows that, while overall there is little variation within UNGA addresses, individual texts achieve striking discourse styles through their unique combinations of features.

Chapter 5 has examined UNGA texts in their entirety, investigating lexical and grammatical features across a full text. However, because countries do not address the same topics, it is not entirely clear whether any differences between groups and individual countries in Chapter 5 are related to generalizable patterns in the use of lexical and grammatical forms or whether the use of any given form can be attributed to the topic addressed. Chapter 6 therefore narrows the focus to explore three specific topics across UNGA texts.

CHAPTER 6: CASE STUDIES ON *TERRORISM, SECURITY COUNCIL,* AND *WOMEN*

6.1. Introduction

Corpus-Assisted Discourse Studies (CADS), as the name implies, uses corpus methods to facilitate discourse analysis, to explore how "language is used to (attempt to) influence the beliefs and behaviour of other people" (Partington et al., 2013, p. 5) and to "uncover ideologies and evidence for disadvantage" (Baker, p. 5). The primary focus of CADS is the word (or words) under investigation and its underlying concept. Of particular interest to the CADS researcher is how the use of the target word sheds light on political, social, or cultural identities, relationships, and ideologies. The objective of CADS is not to describe a register, but findings from CADS can add to what is known about a register, particularly with registers representing political, social, or cultural institutions. Chapter 6 presents three methodological case studies to provide additional depth to the register analysis of Chapters 4 and 5. The words for these three case studies were selected based on UNGA keywords (see Table 4.3): *terrorism, Security Council,* and *women.* As with Chapters 4 and 5, the investigation of how these words are conceptualized addresses both similarities across UNGA texts and variation within the register.

Generally, CADS compares two or more registers or distinct sub-registers, looking for both similarities and differences in how politically or socially loaded terms (e.g., *refugee, anti-Americanism*) are conceptualized in different types of texts. (See Section 2.2.2 for a more detailed description of CADS.) Politically or socially charged terms are chosen by the researcher and an investigation of collocates and/or keywords reveals how the terms are conceptualized and what this implies about the ideologies being promoted by the text authors. Texts are chosen with a view to finding interesting results. For example, in Baker and McEnery (2005), the use of the terms *refugee(s)* and *asylum seeker(s)* is examined in UK newspapers and United Nations High

Commissioner for Refugees (UNHCR) documents because these two institutions influence

discourses surrounding refugees in different ways: newspapers influence public opinion and the

UNHCR influences international policy.

The speech of politicians is of keen interest to CADS researchers because it contains a

wealth of politically and socially loaded terms (see Partington et al., 2013). Moreover,

politicians, it is thought, espouse ideologies and attempt to influence beliefs and behavior. The

UNGA corpus is no exception. As discussed in Chapter 5, countries raise issues of national and

regional interest to obtain international attention and support. Speakers may express stronger or

weaker directive force, but ultimately their objective is to solicit approval for a national agenda.

Though the term "ideology" itself is negatively evaluated and generally avoided in the UNGA,

there is little doubt that politicians speaking before the UNGA convey a message replete with the

ideas and ideals espoused by their governments.

After a careful examination of the texts, three terms were chosen for the CADS-inspired

analysis: *terrorism* (Section 6.2), *Security Council* (Section 6.3), and *women* (Section 6.4).

Terrorism and *women* are more typical of the types of words examined in CADS as they carry

strong political and social overtones. In contrast, *Security Council* may not initially seem like an

obvious choice, but a close reading of the texts revealed that *Security Council* is often linked to

political ideology: the need to expand membership, the misuse of the right to veto, and the most

important issues on which the Security Council should devote time and resources. The three

terms were expected to elicit somewhat different results due to the nature of associations

speakers might make with them. I anticipated that *terrorism* would be denounced by all but

perhaps defined differently, that *Security Council* would have both positively and negatively

evaluated associations depending on the relationship of a country to the Council, and that overall,

associations with *women* would be relatively homogenous, most likely calls for furthering women's rights, regardless of country group.

Each case study in this chapter begins with collocational analysis for the entire UNGA to explore how the terms are used in the corpus as a whole. Next, collocational analyses by country group explore differences within the corpus based on previously identified country categories (e.g., population size, Small Island Developing State, GDP, see Section 3.6 for more details). These collocational analyses examine co-occurrence patterns within a five-word span of the search term; they therefore reveal use at the micro level. Finally, keyword analysis shows lexical patterns at the macro level, for the entire text. The keyword analysis identifies unusually frequent words in texts that use the term under investigation (e.g., texts that use the term *terrorism*) when compared to the texts that do not. This shows other topics of interest to countries that speak about terrorism, which can then shed light on the reasons countries discuss terrorism.

Though collocational analysis and keyword analysis are common methods in CADS, there are two fundamental differences between the case studies in Chapter 6 and most CADS. First, the UNGA has relatively little variation in terms of situational characteristics (see Section 5.2) so the texts being compared are not necessarily expected to exhibit a great deal of lexical variation. Generally in CADS, the researcher chooses two or more corpora with distinctive situational characteristics, ranging from highly distinctive (White House press briefings compared to a judicial inquiry) to somewhat less pronounced (UK tabloids compared to UK broadsheets, see Partington et al., 2013).

In addition, the three case studies in Sections 6.2 through 6.4 differ from most CADS studies in that dispersion criteria, not just overall frequencies, have been set. CADS rarely includes dispersion criteria, and this greatly alters the conclusions that can be drawn from the

findings. (See discussion in Section 6.2.3.2.4). Typically, CADS identifies overall patterns that offer interesting political or social commentary. For instance, *racism* is associated with *police* in UK newspapers (Partington et al., 2013, p. 196). These findings are compelling because they offer insight into social attitudes. Information is provided on the number of times each collocate with *racism* occurs in each corpus (newspaper type), but not per text. *Racism* may collocate with *police* in dozens of UK newspapers or perhaps in just one newspaper or even in just one article. The focus of CADS is customarily to look at a register (UK newspapers) or sub-register (UK tabloids) and see what associations are made overall, not ensure that findings are systematic within a register or sub-register.

In this chapter, like in Chapter 5, both overall frequencies and dispersion across texts are important considerations in the analysis. If a keyword occurs dozens of times but only in one text, it is representative of only that one text rather than the group of texts under investigation and no generalizations can be made to the corpus or sub-corpus. In Chapter 6, as in Chapter 5, the objective of examining different groups within the UNGA corpus is to determine whether these groups are distinctive in systematic ways. If only five out of the 19 countries (about 25%) in the lowest HDI group make an association, is that truly indicative of the low HDI group as a whole? Can generalizations be made based on a pattern across just five texts? The dispersion criteria described in Chapter 3 and followed in Chapters 5 and 6 were set in order to find only systematic patterns across texts.

Eliminating collocates or keywords that do not occur in a sufficient number of texts would, logically, alter the results that can be reported, but this merits investigation and demonstration. Thus, to show the importance of dispersion criteria in the findings that emerge, in the first case study on *terrorism*, a collocational analysis for one country group is carried out

without the dispersion criteria and the results are contrasted with the results when including

dispersion criteria (see Section 6.2.3.2.3). A short discussion follows to argue the need for

dispersion criteria (Section 6.2.3.2.4) for the analysis of variation within the UNGA.

6.2. Case Study #1: Terrorism

6.2.1. Introduction

Terrorism is one of the keywords identified in the analysis of the UNGA corpus (see

Section 4.4). The concern UN member states express over terrorism in 2015 is not surprising

given the number of fatalities attributed to terror around the world: Attacks by ISIL in Iraq, the

Taliban in Afghanistan, Boko Haram in Nigeria, Al-Shabaab in Kenya, and Houthi extremists in

Yemen were some of the deadliest in the first half of 2015 (Institute for Economics and Peace,

2016).

These acts of extremist violence were quickly labeled "terrorism" by the international

community at large, but the definition, associations, and implications of the term are not

universally shared. European Council President Tusk used the term *state terrorism* to describe

violence in Syria. Would Syrian President Assad agree that his government was committing acts

of *terrorism*? Ukrainian President Poroshenko claimed that Russia was responsible for the

financing of terrorists in Ukraine. Would Russian President Putin also describe them as

terrorists? "The definition of 'terrorist' depends on who is throwing the bomb" (Baer, 2010, p.

ix). The definition of *terrorism* is a matter of perspective.

This first case study addresses the following research question: How do countries differ

in their conceptualization of terrorism in UN General Assembly addresses? In order to

distinguish differences in the concept of terrorism, it is important to first present an overall

picture of how terrorism is conceptualized. Groups of countries can then be compared to the

findings from the entire corpus. Thus, the first step was to analyze collocations of the lemma *terror** (*terrorism* and the related words *terror, terrorist,* and *terrorists*) in the UNGA corpus as a whole. Then collocational analyses were carried out for groups of countries. Collocational analysis, for the full corpus and for sections of the corpus, reveals how *terror** is used, its meanings, and the relationships between *terror** and other terms.

Also of interest are the discourses surrounding *terror*.* That is, what additional topics are raised by states that express a concern about *terror**? What is the relationship between *terror** and those topics? Keyword analysis is used to explore the "aboutness" of a text and was therefore carried out to compare all texts that frequently mention *terror** with those that do not. In contrast with collocational analysis, which investigates the associations made with *terror** at the micro level (within a five-word span), keyword analysis investigates discourses surrounding *terror** at the macro level. It does not identify direct relationships, but rather patterns in topics raised over the entire text. Identifying the unusually frequent or infrequent words in texts that discuss *terror** reveals topics of interest for countries concerned with terrorism in order to elucidate the reasons these countries raise the issue of terrorism.

Section 6.2.2 reports the methods and results of the collocational analysis for the entire UNGA corpus and Section 6.2.3 the collocational analysis by country group. Section 6.2.4 is a description of the methods and results of the keyword analysis. Some patterns were not detected in the quantitative analyses of Sections 6.2.2 through 6.2.4 but became apparent through a manual examination of individual texts; Section 6.2.5 presents these qualitative findings. Section 6.2.6 is a summary of this first case study.

6.2.2. UNGA collocational analysis for terror*

6.2.2.1. Methods: UNGA collocational analysis for terror*

The investigation began with an analysis of collocates for the lemma *terror** using WordSmith Tools 7.0 (Scott, 2017). The objective of this analysis was to explore meanings and associations of *terror** for the UNGA corpus as a whole. Collocates were operationalized as words occurring within five places to the left or right of *terror*.* Only content words were retained because the goal of the investigation was to examine topic-related differences. Collocates were grouped into semantic categories in order to find patterns in the types of associations UNGA countries make with *terror**. Most of the words were semantically transparent. For example, *war, conflicts,* and *destruction* were all clearly related to violence and could be grouped together. Other terms required a manual examination of the texts. *Comprehensive* is always used in the phrase *Comprehensive Convention on International Terrorism* and was therefore grouped with "UN counter-terrorism efforts."

6.2.2.2. Results: Terror* *collocates in the UNGA*

The objective of the UNGA collocational analysis was to look for general patterns in the associations UNGA speakers make with terrorism. The results of this collocational analysis reveal that *terror** is associated with acts of violence (e.g., *crime*), UN counter-terrorism efforts (e.g., *campaign*), religious extremism (e.g., *radicalization*), particular countries or regions (e.g., *Middle East*), terrorist networks (e.g., *Daesh*), as well as other UN issues (e.g., *climate change*). *Terror** also co-occurs with words meaning international (e.g., *global*), extensive (e.g., *mass*), and topical (e.g., *today*). Table 6.1 shows how the collocates of *terror** were grouped into these semantic domains with information on distribution ("texts") and overall frequency ("total occurrences").

Table 6.1. UNGA Collocation Results for *Terror** by Semantic Category

Collocate	Texts	Total occurrences
Acts of violence		**128**
violent	15	17
threat	10	10
violence	7	7
acts	7	7
crime	5	6
conflicts	5	5
threats	5	5
attacks	4	4
war	4	4
criminal	4	4
(at/into the) hands (of)	4	4
fighters	3	4
destruction	3	3
arms	3	3
committed	3	3
(violation of) human (rights)	3	3
fear	2	3
wars	2	2
criminality	2	2
attack	2	2
targeting	2	2
civil (war)	2	2
use	2	2
suffering	2	2
victim	2	2
death	2	2
ruthless	2	2
repression	2	2
soldiers	2	2
financing	2	2
activities	2	2
actions	2	2
organized	2	2
assault	1	2
proxies	1	2
UN counter-terrorism efforts		**56**
fight	7	9
counter (global counter-terrorism)	5	5

Collocate	Texts	Total occurrences
combating	4	4
condemns	4	4
UN (UN global counter-terrorism strategy)	4	4
anti (anti-terrorist coalition)	3	4
condemn	3	3
campaign	1	3
combat	2	2
strategy	2	2
efforts (counter terrorism efforts)	2	2
convention	2	2
comprehensive (Comprehensive Convention on International Terrorism)	2	2
tackling	2	2
lead	2	2
requires	2	2
coalition	1	2
defence	1	2
Extremism		**40**
extremism	16	19
religious (extremism)	5	5
extremists	3	3
radicalization	3	3
sectarianism	3	3
ethnic (intolerance)	3	3
extremist	2	2
radicalism	2	2
Countries and regions		**15**
Middle East	4	4
Syria	4	4
Iran's	1	3
Europe	2	2
Turkey	1	2
Terrorist networks		**36**
groups	8	8
organisations	6	6
Daesh	4	5
ISIL	3	3
networks	3	3
cells	1	3
elements	2	2
(terror/ist) outfits	2	2

Collocate	Texts	Total occurrences
Al-Shabaab	1	2
network	1	2
UN issues		**19**
climate change	5	5
diseases	3	3
challenges	3	3
problem	2	2
issue	2	2
phenomenon	2	2
values	2	2
International		**41**
global	9	14
international	12	12
state (inter-state; non-state actors, state terrorism)	5	5
foreign	4	5
transnational	3	3
anywhere	1	2
Extensive		**35**
forms (in all its forms)	10	10
(Mr.) President	5	5
mass	4	4
important	3	3
unspeakable	3	3
serious	2	2
strongly	2	2
multifaceted	2	2
unreservedly	2	2
largest	1	2
Topicality		**32**
rise	5	5
today	4	4
new	3	4
remain	3	3
proliferation	3	3
spread	3	3
emergence	2	2
growing	2	2
(recent) past	2	2
continuing	2	2
setting (up)	1	2

Note. Only content words were analyzed. Function words have been removed.

In many of the texts, *terror** (italicized in excerpts throughout this section) is associated with the semantic domain "acts of violence." Collocates (underlined in excerpts throughout this section) such as *wars, human rights (violations), conflicts,* and *crimes* are often presented in succession, as the speaker enumerates different forms of violence: *First we have to confront the root causes in the countries of origin: Civil wars, massive human rights violations and* terrorism [Austria]. *Failing states, frozen conflicts, wars and* terrorism *create insecurity around Europe* [Bulgaria].

Terrorism is repeatedly denounced and calls are made for the international community to take action. The semantic domain "UN counter-terrorism efforts" includes collocates such as *condemns, requires,* and *combating*: *Slovakia strongly* <u>condemns</u> *all* terrorist *acts as criminal and unjustifiable...No* terrorist *act should be left without adequate response of the international community, no perpetrators should be left unpunished, and no movement affiliated to any form of* terrorism *should be tolerated...* [Slovakia]. *International* terrorism *threatens us all, and* <u>requires</u> *concerted international cooperation* [Saint Vincent and the Grenadines]. *The current crisis, therefore, calls us to renewed efforts to apply the law in force and to develop new norms aimed also at* <u>combating</u> *the phenomenon of international* terrorism *in full respect for the law* [The Holy See].

Terrorism is also associated with religious extremism, particular countries and regions, and terrorist groups: *We have been witnessing ongoing turmoil,* <u>extremism</u>, <u>sectarianism</u>, *civil war and* terrorism *taking place in* <u>Middle East</u>, *North Africa and other regions of the World* [Cyprus]. *We have to intensify our fight against* <u>Daesh</u> *and* terrorist *groups* [Austria]. *While* <u>Al-Shabaab's</u> *capacity to launch* terrorist *acts inside Somalia has been diminished, Al-Shabaab's continued presence in Somalia has a negative impact on the security, stability and prosperity of*

Somalia [Kenya]. *The conflict in <u>Syria</u> has also allowed extreme* terror groups *like <u>ISIL</u> to gain a foothold* [Norway]. These brief excerpts illustrate the use of collocates in the semantic domain "extremism" (*extremism, sectarianism*), in "countries and regions" (*Syria, Middle East*), and in "terrorist networks" (*Daesh, groups, Al-Shabaab, ISIL).*

Finally, *terrorism* is a topical UN issue that is international and extensive in scope, with collocates such as *today* and *rise* for "topicality," *challenges* for "UN issue," *global, state,* and *international* for "international," and *serious* for "extensive." Terrorism *has become a <u>serious</u> <u>global</u> threat* and *we cannot talk about security <u>challenges</u> of <u>today</u> without mentioning* terrorism. [Tanzania]. *Violent extremism is on the <u>rise</u>, with* terrorist *groups demonstrating new levels of brutality and barbarity* [Lithuania]. *We can no longer confine peace and stability issues to differences between and within nations, as non-<u>state</u> actors, such as* terrorists, *have challenged the established <u>international</u> order, leading to religious fundamentalism, violent extremism, forcible displacement of people and forced migration* [Cyprus].

These collocates reveal certain patterns in associations made with *terror*.* However, it should be noted that several semantic categories are nearly always combined in one text, one paragraph, or even one sentence. The excerpt from Nepal (Text 6.1) illustrates this, with collocates from the semantic categories "acts of violence" (*threats, threat, violent*), "UN counter-terrorism efforts" (*condemn, comprehensive, convention*), "extremism" (*extremism*), "UN issues" (*problem*), "international" (*international*), "extensive" (*in all its forms, multifaceted*), and "topicality" (*emergence*).

Text 6.1. Excerpt from Nepal (*terror** in italics; collocates underlined)

> Mr. President, Nepal fully aligns itself with the spirit of the UN and the international community in dealing with all <u>threats</u> to peace and security. *Terrorism* is the biggest <u>threat</u> to peace, security, and development. It is shocking to see the <u>emergence</u> of several

terrorist outfits including <u>violent extremism</u> and religious fundamentalism in different parts of the globe. We unequivocally <u>condemn</u> *terrorism* <u>in all its forms</u> and manifestations. As *terrorism* is a <u>multifaceted problem</u>, its solution demands greater unity, solidarity, and continued and consistent collaborations among the nations to address its root causes. This is best done under the auspices of the United Nations, and we are in favour of an early conclusion of a <u>Comprehensive Convention</u> on <u>International</u> *terrorism*.

The results of this first analysis show some general patterns among all UN member states. That is, in the corpus as a whole, *terror** is associated with particular semantic domains (e.g., acts of violence, extremism), suggesting similarities in the ways in which UN member states conceptualize *terror**. Section 6.2.3 compares countries within the UNGA by category to determine whether associations made with *terror** vary systematically by country group.

6.2.3. Collocational analysis for terror* *by country group*

6.2.3.1. Methods: Collocational analysis for terror* *by country group*

Groups of countries were compared in order to explore any differences in how states conceptualize *terror**. For this more detailed analysis, only states referring with sufficient frequency to *terror** were included. Texts were divided into two categories: those in the "Terror+ group," which mention *terror** frequently (at least three times) and those in the "Terror— group," which mention *terror** fewer than three times. This threshold was based on a careful reading of the texts. Numerous countries mention *terror** once or twice in enumerations of problem areas to be addressed by the UN. As such, there is no true discussion of *terror** per se, but rather an acknowledgement that it, like global warming and poverty, is a concern for the international community, as illustrated in Text 6.2.

Text 6.2. Excerpt from Zimbabwe (*terror** in italics)

> The growing list of phenomena that neither respect nor know any borders, makes it imperative that we mobilise all mechanisms of cooperation to effectively overcome them. *Terrorism* and extremist violence, communicable diseases such as HIV and AIDS, Tuberculosis, Malaria, Bird Flu, Ebola, cannot be overcome single-handedly.

In this excerpt from Zimbabwe, *terrorism* is simply listed along with other alarming issues such as HIV and malaria without further elaboration. This is the only occurrence of *terror** in the address. The text was therefore put in the Terror— group with states that mention *terror** fewer than three times.

This classification of texts resulted in a Terror+ group made up of 28 countries using *terror** a total of 155 times and a Terror—group made up of 64 countries using the lemma 42 times. Table 6.2 lists countries in the Terror+ group and shows frequency information for each type by individual country.

Table 6.2. *Terror** **Frequencies in Terror+ group (Three or More Occurrences of** *Terror**)

Country	*Terror*	*Terrorism*	*Terrorist*	*Terrorists*	Total
Afghanistan	1	2	4	0	7
Australia	1	1	1	1	4
Austria	1	1	1	0	3
Brunei Darussalam	0	5	1	0	6
Bulgaria	0	1	1	1	3
Croatia	1	4	1	0	6
Cyprus	0	2	1	1	4
Estonia	0	2	1	0	3
EU	0	2	1	0	3
Holy See	0	3	1	0	4
Israel	10	1	0	0	11
Kenya	0	2	2	1	5
Lithuania	0	1	1	1	3
Mauritius	0	2	0	1	3
Moldova	0	1	2	0	3
Nepal	0	4	1	0	5
Nigeria	1	2	0	0	3
Norway	1	2	0	0	3
Pakistan	0	8	1	3	12
Romania	0	6	0	0	6
Saint Vincent and the Grenadines	0	2	2	0	4
Samoa	0	3	0	0	3
Slovakia	0	8	2	0	10
South Africa	0	3	0	0	3
Tanzania	0	4	0	0	4
Turkey	0	5	5	1	11
Ukraine	2	6	5	4	17
US	2	1	2	1	6
Total	**20**	**84**	**36**	**15**	**155**

Countries in the Terror+ group were then categorized based on geographic, social,

political, and economic features as described in Section 3.6 (e.g., population size, GDP) in order

to explore whether these country characteristics affected the ways states conceptualized *terror**.

A collocational analysis was carried out for each group (e.g., highest GDP, lowest GDP). As

with the analyses by country group in Chapter 5, only content words that occurred in at least

50% of texts were retained. As discussed in Section 5.3.1 and the related methods section

(Section 3.7.1), this minimum distribution was set in order to find systematic patterns of

variation by country group, research goal 2. If a collocate occurs in only one or two texts, it is

not representative of the group at large and therefore does not address the research goal. An

additional criterion was for the collocate to be "meaningful." A collocate was deemed not

"meaningful" if it was a place name from that group (e.g., *Ireland* or *European* in the Europe

group).

Results were interpreted through analysis of the geographic, social, political, and

economic situation of each country category in 2015. This sometimes required consultation of

reports published by international institutions and nongovernmental organizations such as the

Institute for Economics and Peace.

6.2.3.2. Results: Terror* *collocates by country group*

The investigation revealed few distinctions based on country group. Only content words

with a minimum dispersion of 50% of the texts in a group were retained for analysis, and most

country groups did not have any collocates that met this criteria. The only categories with

collocates shared in 50% of the texts were Asia and the biggest military spenders.

6.2.3.2.1. Terror+ Asian countries

Asian countries in the Terror+ group included Afghanistan, Brunei Darussalam, Cyprus,

Israel, Nepal, Pakistan, and Turkey. These countries represent different levels of economic and

social development (Afghanistan, Nepal, and Pakistan being lowest; Brunei Darussalam and

Israel being highest). They represent different religious majority groups (Cyprus predominantly

Greek Orthodox, Israel Jewish, Brunei Darussalam Muslim). Some have large defense budgets (e.g., Israel, Pakistan) and others small (e.g., Cyprus, Nepal). Some have large areas and populations (e.g., Afghanistan) and others small (e.g., Brunei Darussalam). The sole distinguishing feature of this group was region.

The collocational analysis for Terror+ Asian countries identified only one content word in at least 50% of the texts: *extremism.* This collocate was found in five texts (Afghanistan, Brunei Darussalam, Nepal, Pakistan, and Turkey) for a total of seven occurrences. Thus, these five countries conceptualize *terror** in part as it relates to extremism. Text excerpts from Nepal (Text 6.3), Turkey (Text 6.4), and Brunei Darussalam (Text 6.5) exemplify this relationship between terrorism and ideological radicalism. Though some texts associate fundamentalism with religion (see Text 6.3), speakers are generally careful not to link all followers of one religion with violent extremism (see Text 6.4). The one exception to this was the text from Israel, which refers to Iran as *a militant Islamic* terror *regime*.

Text 6.3. Excerpt from Nepal (*terror** in italics; collocates underlined)

> It is shocking to see the emergence of several *terrorist* outfits including violent extremism and religious fundamentalism in different parts of the globe.

Text 6.4. Excerpt from Turkey (*terror** in italics; collocates underlined)

> We must avoid alienation, exclusion, vilification of certain communities or religions if we want to bring down the walls that divide us. We must act together against all forms of racism and xenophobia, including Islamophobia, without exception. Only then can we collectively fight against extremism, radicalization, *terrorism* in an effective manner.

Text 6.5. Excerpt from Brunei Darussalam (*terror** in italics; collocates underlined)

> I believe that our organisation is well placed for the promotion of cooperation and partnership based on mutual respect, understanding and tolerance. These are necessary to prevent wars and conflicts; religious persecutions; ideological confrontations; and

terrorism. Since our last gathering, we have witnessed tragic loss of lives caused by inhumane *terrorist* acts in various parts of the world. These are intolerable and we join others in condemning *terrorism*, totally rejecting <u>extremism</u> and radicalism.

Conceptualizing *terror** as related to extremism is not surprising in a region such as Asia, which has known numerous *violent ideological confrontations* [Brunei Darussalam]. Though Asia is far from alone in that respect, these seven Terror+ countries have had to tackle particularly virulent and ongoing extremist violence domestically, from the Kurdistan Workers' Party (PKK) in Turkey to the Taliban in Afghanistan. Radical Hindu groups in Nepal and inter-ethnic conflict along the Nepal-India border are a growing cause for concern (United States Agency for International Development, 2013); Pakistan has been combatting armed religious groups for decades, with 14,953 deaths attributed to terrorism between 2000 and 2015 (Institute for Economics and Peace, 2016). In addition, these countries are in close proximity to other conflict zones with extremist violence (e.g., Syria). The association between *terror** and extremism is a reflection of the geographic, political, and social situations in these Asian countries.

However, it would be remiss not to point out that *extremism* is one of the most frequent collocates in the UNGA overall, with 19 occurrences in 16 texts. This overall high frequency of *extremism* may explain why it was also identified in the Asian group.

6.2.3.2.2. Terror+ biggest military spenders

The only other country category with collocates that met the distribution criteria was the biggest military spenders: Australia, Brunei Darussalam, Estonia, Israel, Pakistan, Turkey, Ukraine, and the United States. These countries are located on different continents; they differ in economic and social development, religious majority, area, and population size. The one variable

that unites them is that they spend 2% or more of GDP on defense. Collocates shared by at least half of the texts in this category include *threat*, *global*, and *forms*.

Threat co-occurs with *terror** five times and in five texts (one time each in the texts from Australia, Pakistan, Turkey, Ukraine, the United States). As shown in Texts 6.6, 6.7, and 6.8, *terrorist threat* or *threat of terrorism* was preferred over *terrorist act* or *attack.* An act or an attack is more tangible, while a threat is the fear of a future possibility. In Text 6.8 from the United States, *threat* occurs in a parallel sentence structure with *risk* and *danger.* This focus on potential risk rather than concrete event is not surprising in texts from the biggest military spenders because investing heavily in defense is an attempt to avert or better confront future danger. The word *threat* incorporates the violence of *attack* but also implies something to be feared.

Text 6.6. Excerpt from Turkey (*terror** in italics; collocates underlined)

> Any attempt to affiliate *terrorism* with any religion or ethnic group is patently mistaken and serves only to strengthen *terrorist* <u>threat</u>.

Text 6.7. Excerpt from Ukraine (*terror** in italics; collocates underlined)

> Unfortunately, not by its own free will today Ukraine is one of the areas of fight against *terrorist* <u>threat</u>.

Text 6.8. Excerpt from the United States (*terror** in italics; collocates underlined)

> No nation in this Assembly can insulate itself from the *threat* of <u>terrorism</u>, or the risk of financial contagion; the flow of migrants, or the danger of a warming planet.

Global appears in four texts for a total of eight occurrences. In each of these occurrences, *global* is used to underscore the urgency of terrorism (Texts 6.9, 6.10, 6.11) and legitimize the need for the United Nations to tackle the problem (Text 6.12).

Text 6.9. Excerpt from Pakistan (*terror** in italics; collocates underlined)

The <u>global</u> threat of *terrorism* cannot be defeated unless we address its underlying causes.

Text 6.10. Excerpt from Australia (*terror** in italics; collocates underlined)

Today, the world faces an unprecedented number of long-running and seemingly intractable conflicts, generating displacement on a massive scale and making humanitarian need more dire, than at any time since the Second World War. *Terrorism* today is a <u>global</u> threat.

Text 6.11. Excerpt from Israel (*terror** in italics; collocates underlined)

In 2013 president Rouhani began his so-called charm offensive here at the UN. Two years later, Iran is executing more political prisoners, escalating its regional aggression, and rapidly expanding its <u>global</u> *terror* network.

Text 6.12. Excerpt from Estonia (*terror** in italics; collocates underlined)

ISIL violates universal human values. No country is immune from the threat that it poses. Stopping it, and other *terrorist* organisations, requires a <u>global</u> effort. Estonia supports the international coalition against ISIL. We believe that the UN and its <u>Global</u> Counter *Terrorism* Forum also have an important role to play.

Finally, *forms* was used four times in four texts (i.e., one time per text), in expressions such as *terrorism in all its forms* to emphasize the extensiveness of the problem. Texts 6.13, 6.14, and 6.15 show three such phrases.

Text 6.13. Excerpt from Australia (*terror** in italics; collocates underlined)

Australia is committed to defeating *terrorism* in all its <u>forms</u>.

Text 6.14. Excerpt from Pakistan (*terror** in italics; collocates underlined)

We will fight *terrorism* in all its <u>forms</u> and manifestations, irrespective of who their sponsors are.

Text 6.15. Excerpt from Brunei Darussalam (*terror** in italics; collocates underlined)

> We reiterate our support towards all efforts by the international community to prevent and eliminate all <u>forms</u> of *terrorism*.

The biggest military spenders associate *terror** most frequently with *threat* (potential, feared violence), *global* (urgent and widespread), and *all * forms* (extensive). These countries present *terror** as an urgent, extensive problem that is to be feared, a conceptualization of *terror** that is coherent with their sizeable defense budgets.

These findings for the biggest military spenders are more robust than the findings for Asian countries because *threat, global*, and *forms* are less frequent collocates overall. Asian countries represent 25% of the Terror+ group and are responsible for 37% of the *extremism* occurrences. The number of Asian countries to use *extremism* is notable, but not overwhelming relative to their representation in the Terror+ group. In contrast, the biggest military spenders represent 29% of the Terror+ group and use 50% of the *threat* occurrences, 57% of the *global* occurrences, and 40% of the *forms* occurrences. The use of *threat* and *global* in particular are much greater than would be expected given their representation in the Terror+ group.

Though a few compelling collocates were found for Asian countries and the biggest military spenders, relatively little was found in the collocational analysis by country group. This was likely due to the restrictive nature of collocational analysis and the strict search criteria. Collocates are limited to a span of five words on either side of the search term, and a dispersion of 50% of texts is a conservative cut-off, though conservative by necessity to ensure that collocations are truly representative of the group and not a sub-category of the group.

A further difficulty in finding collocates is the use of synonyms (e.g., *groups, networks, cells, elements*), different grammatical forms (e.g., *condemn, condemns, condemning,*

condemned), or alternate phrasings (e.g., *fight terrorism, counter terrorism efforts, Global Counter-Terrorism Strategy, campaign against terror, anti-terrorist coalition*). Table 6.1 shows the lexical diversity of the semantic categories for *terror** collocates. This lexical diversity makes it more difficult to find the exact collocate within a five-word span, all in 50% of the texts.

6.2.3.2.3. High religious non-affiliation

The objective of Section 6.2.3.2.3 is to carry out a collocational analysis of one country group without any distribution criteria in order to compare the types of findings possible with and without a dispersion cut-off. As noted at the end of Section 6.2.3.2.2, collocational findings by country group were extremely limited: Only two out of the 31 groups had any collocations that met the distribution criteria. Results would surely be more interesting without the cut-off. Moreover, as discussed in Section 6.1, CADS generally does not report dispersion across texts and does not require a keyword or collocate to occur across a set number of texts. Section 6.2.3.2.3 therefore presents a more typical CADS analysis. The objective is to present findings with interesting political or social commentary.

For this collocational analysis, the same methods were used as for the analyses in Sections 6.2.3.2.1 and 6.2.3.2.2 except that no distribution criteria were imposed and texts from both the Terror+ and Terror— groups were used.

Results showed that countries with high religious non-affiliation tend to associate *terror** with religion and Islam in particular. Seven collocates were categorized as pertaining to religion: *religious, Middle East, Iran's, Syria, ISIL, Daesh,* and *Al-Shabaab.* The *Middle East*, *Iran's,* and *Syria* were categorized as related to religion because they are predominantly Muslim; *ISIL/Daesh*

and *Al-Shabaab* are terrorist organizations that associate themselves with Islam. Table 6.3 shows

20 expanded concordance lines with religious collocates underlined.

Table 6.3. Expanded Concordance Lines for *Terror** with Low Religious Affiliation Texts

1	religious extremism, violence and	**terrorism** are causing unspeakable suffering.
2	past decades, and still rage on today.	**Terrorism**, violent extremism, religious and
3	senseless violence and excesses of	**terrorists**, religious extremists, rogue soldiers,
4	violence, religious intolerance and	**terrorist** actions. There we are confronted with
5	concern about violent repression and	**terrorist** attacks on religious minorities. We have
6	sectarianism, civil war and	**terrorism** taking place in Middle East, North Africa
7	review progress made in countering	**terrorism** in the Middle East and North Africa
8	to you to tear down Iran's global	**terror** network. Ladies and Gentlemen, Israel is
9	keep arms out of the hands of Iran's	**terror** proxies. We agree on the need to stop Iran
10	response to violent extremism and	**terrorism**. Not just in Syria and Iraq, but also in
11	took hold in Syria spread death and	**terror** into the region and beyond. ISIL's warped
12	in Syria has also allowed extreme	**terror** groups like ISIL to gain a foothold. Now,
13	horrific abuses. Committed by the	**terror** organisation ISIL. And by the Assad regime.
14	Goals in 2015. With the rise of	**terrorists** like Daesh, the continuing depredations
15	committed by Daesh and other	**terrorist** groups and organizations. This must be
16	ground assault by Daesh and other	**terrorist** organizations. Ladies and gentlemen, You
17	regime and ground assault by the	**terrorist** organization Daesh has exceeded 4
18	Daesh and PKK. Our counter	**terrorism** efforts and our contribution to
19	Al-Shabaab's capacity to launch	**terrorist** acts inside Somalia has been diminished,
20	In the recent past, Al-Shabaab	**terrorists** have been targeting the civilian

Expanding these concordance lines further to reveal the larger co-text shows additional

associations with religion. Text 6.16 from Albania begins with reference to Pope Francis and *the*

joy he brought to the country as he spoke of the importance of coexisting peacefully with people

of other religions. This *religious fraternity* is then contrasted with *religious extremism, violence*

and terrorism. The passage ends by attributing *violence and extremism* with a forced migration

compared to the *Exodus in the Biblical times*.

Text 6.16. Excerpt from Albania (*terror** in italics; collocates underlined)

> One year ago Pope Francis started his visits in Europe from Albania. We saw last week
> here in America the hope and joy he could bring to the richest and most powerful country

in the world. Imagine the joy he brought to us, not least by the praise he bestowed upon Albania in saying we were proof "that the peaceful and fruitful coexistence between people and communities belonging to different religions is not only beneficial, but is concretely possible and practical." We are rightly proud of this heritage, of being, in Pope Francis's words, a religious fraternity, as a lived experience, at a time when <u>religious</u> extremism, violence and *terrorism* are causing unspeakable suffering. Violence and extremism that have put multitudes of people on the move, in scenes reminiscent of the Exodus in the Biblical times, of children, women and men arriving in their thousands on our doorsteps in the Balkans as they seek safety and protection in Europe.

The text from Albania associates *terrorism* with *religious* extremism, using *religious* to describe a type of violence. Similarly, in Text 6.17 from Cyprus, *religious fundamentalism* and *violent extremism* are the results of the work of *terrorists.*

Text 6.17. Excerpt from Cyprus (*terror** in italics; collocates underlined)

We can no longer confine peace and stability issues to differences between and within nations, as non-state actors, such as *terrorists*, have challenged the established international order, leading to religious fundamentalism, violent extremism, forcible displacement of people and forced migration.

In contrast, Ireland uses *religious* to characterize not a type of violence or extremism but rather the minority groups targeted by terrorists (see Text 6.18). The passage concludes with a reference to past *religious repression* in Ireland and the particular concern Ireland has for protecting religious rights.

Text 6.18. Excerpt from Ireland (*terror** in italics; collocates underlined)

In Ireland there is widespread concern about violent repression and *terrorist* attacks on <u>religious</u> minorities. We have seen an alarming rise in attacks on Christians, Baha'i, Jews and Muslims. Elsewhere people are oppressed and punished because of their gender, race or sexual orientation. In short, vulnerable minorities are under increasing attack, from repressive State regimes and from transnational *terrorist* networks including ISIS, al-Qaeda and Boko Haram. The United Nations, as champion of equality plurality and diversity, has an important role to play in defending and protecting vulnerable groups. We fully support the organisation in this essential task.

389

Interestingly, a number of states warn against associating *terror** with religion and in particular with Islam (see Texts 6.19-6.21). In spite of the stance taken against the association, the relationship between the two is recognized and speakers reinforce that relationship by providing yet another example of a text in which the discourse surrounding *terror** relates to religion.

Text 6.19. Excerpt from Turkey (*terror** in italics)

> Any attempt to affiliate *terrorism* with any religion or ethnic group is patently mistaken and serves only to strengthen *terrorist* threat.

Text 6.20. Excerpt from the US (*terror** in italics)

> We know that ISIL -- which emerged out of the chaos of Iraq and Syria - depends on perpetual war to survive. But we also know that they gain adherents because of a poisonous ideology. So part of our job, together, is to work to reject such extremism that infects too many of our young people. Part of that effort must be a continued rejection by Muslims of those who distort Islam to preach intolerance and promote violence, and it must also a rejection by non-Muslims of the ignorance that equates Islam with *terror*.

Text 6.21. Excerpt from Croatia (*terror** in italics)

> As to Syria, Iraq, Libya, Yemen and the other crisis areas of the Middle East and North Africa, we share a grave concern about the continuous trend of growing violence, religious intolerance and *terrorist* actions. There we are confronted with unspeakable barbarism and *terror* with the most appalling forms of murder such as beheadings, sexual violence against women and girls and other most heinous crimes against humanity, culture, religion, dignity and against dearest of all - human life. It is shocking and sickening. And it has nothing to do with Islam, a religion of peace and understanding.

Of particular interest in the text from Croatia is the reference to Syria, Iraq, Libya, Yemen, the Middle East, and North Africa as *crisis areas* in terms of *violence, religious intolerance and terrorist actions*. These countries and regions are predominantly Muslim. Syria, Iraq, and Yemen have, indeed, suffered serious terrorist attacks. However, according to the

Global Terrorism Index (Institute for Economics and Peace, 2016), no terrorist attack in North

Africa was deadlier than the attack on Cameroon in April 2015 or the attack on Ukraine in

August 2015. Similarly, the impact of terrorism was not as pronounced for Libya as it was for

Somalia or India in 2015. Yet central and eastern Africa, India, and Ukraine are rarely given as

examples of victims of terrorism outside of their own regions.

Again and again, associations are made between *terror** and predominantly Muslim

countries or regions. These associations are rarely as explicit as Text 6.22 from Israel, where Iran

is labeled *a militant Islamic terror regime,* but the repetition of references to religion in passages

on *terror** is noteworthy in the group with high religious non-affiliation.

Text 6.22. Excerpt from Israel (*terror** in italics)

> Under this deal, if Iran doesn't change its behavior, in fact, if it becomes even more
> dangerous in the years to come, the most important constraints will still be automatically
> lifted by year 10 and by year 15. That would place a militant Islamic *terror* regime weeks
> away from having the fissile material for an entire arsenal of nuclear bombs. That just
> doesn't make any sense. I've said that if Iran wants to be treated like a normal country, let
> it act like a normal country. But this deal, this deal will treat Iran like a normal country
> even if it remains a dark theocracy that conquers its neighbors, sponsors *terrorism*
> worldwide and chants "Death to Israel", "Death to America." Does anyone seriously
> believe that flooding a radical theocracy with weapons and cash will curb its appetite for
> aggression? Do any of you really believe that a theocratic Iran with sharper claws and
> sharper fangs will be more likely to change its stripes? So here's a general rule that I've
> learned and you must have learned in your life time - When bad behavior is rewarded, it
> only gets worse.

The text from Israel uses several terms related to religion (*Islamic, theocracy, theocratic*),

generally surrounded by words with negative evaluation (*dark, conquers, aggression, sharper*

claws and sharper fangs).

The findings from the high religious non-affiliation group are even more noteworthy

when contrasted with the low religious non-affiliation group, where none of these terms was

identified as a collocate. It may be that countries where most of the population is affiliated with a religion do not view or do not wish to view religion as associated with *terror**. As shown in Texts 6.19-6.21, even when countries warn again conflating terrorism with Islam or religion more generally, the association is nonetheless made, reinforcing the connection between the two.

6.2.3.2.4. Summary of collocational analyses by group and methodological implications

The collocational analysis of the high religious non-affiliation group without the dispersion cut-off in Section 6.2.3.2.3 yielded much more interesting findings than the two analyses with the distribution criteria in Sections 6.2.3.2.1 and 6.2.3.2.2. So why set a distribution cut-off? Why not carry out more analyses such as the one for the high religious non-affiliation group? Though Section 6.2.3.2.3 offers some interesting results that appear convincing, it does not effectively answer the research goal identified at the beginning of this study. That is, it does not distinguish the *systematic* patterns of variation within the UNGA.

Section 6.2.3.2.3 presented seven collocates, showing 20 concordance lines and six expanded passages. These examples were selected from 18 out of a total of 60 texts in the high religious non-affiliation group (both Terror+ and Terror— countries). That is, 42 (70%) of the countries with low religious affiliation did not associate *terror** with religion. In all, the analysis of the high religious non-affiliation group resulted in 132 occurrences of *terror** and 149 collocates. Can the findings truly be said to be representative of the entire group? It is unlikely, particularly given that three of the collocates in Section 6.2.3.2.3 have very limited distribution: *Iran's* appears only in the Israeli text, *Al-Shabaab* only in the Kenyan text, *ISIL* only from Northern European countries.

This is not to completely discount the analysis in Section 6.2.3.2.3. The findings are valuable in that they have elucidated a potential association that would require further

examination using additional texts. Though the analysis has not uncovered any systematic patterns for the high religious non-affiliation group, it has nonetheless shown a representation of *terror** in 18 texts. Whether there is a true relationship between high religious non-affiliation and that conceptualization of *terror** cannot be claimed, but that such a representation is made in individual texts can.

Analyzing all collocates without regard for distribution allows additional associations to be identified. In contrast, imposing a strict distribution cut-off ensures that the associations that are identified are, indeed, representative of the texts. The risk run when setting strict distribution criteria is not identifying important associations for some of the texts in a corpus or sub-corpus; the risk when not setting one is that some generalizations might be made erroneously. Crucially, interpretation of findings, claims, and implications must reflect the methods used. It would be misleading to present collocational analyses without the distribution criteria as evidence of systematic patterns of lexical variation by country group, the goal of this part of the analysis. As a result, all collocational analyses by country group in Sections 6.3 and 6.4 maintain the distribution cut-off of 50%.

6.2.4. Keyword analysis for the Terror+ group

The collocational analyses investigated associations with *terror** at the micro level, within a five-word span. In order to examine the discourses surrounding *terror** at the macro level, over the entire text, a keyword analysis was carried out comparing word frequencies for the Terror+ group with the Terror— group. The objective of the keyword analysis is to uncover unusually frequent topics in texts which discuss *terror**. Do these topics relate to *terror** and if so, how? Does a country's interest in these other topics help to elucidate the reason for its interest in *terror**?

6.2.4.1. Methods: Terror+ keyword analysis

Keywords were identified using WordSmith Tools 7.0 (Scott, 2017). As with the

collocational analysis, function words were removed from the list of keywords. As with Chapter

5, all keywords are reported and categorized semantically. Keywords occurring in at least 50% of

the texts are italicized and given particular emphasis because they are representative of the entire

Terror+ group. Keywords that do not occur across the majority of texts but that are in a

productive semantic category, with other keywords expressing the same idea, are also considered

relevant because they are one of several words sharing similar meaning (see Sections 3.7.1 on

semantic grouping).

6.2.4.2. Results: Terror+ keywords

The keyword analysis resulted in a list of 18 keywords (see Table 6.4). Only two

keywords occurred in at least half the texts, however, (*terrorism, terrorist*) and these two words

had been chosen to distinguish the Terror+ from the Terror— group; they are not meaningful and

not indicative of any new topic of interest to the Terror+ group.

Table 6.4. Keywords for Terror+ Group

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
Israel	81	0.13	4	12	0.01	110.06	3.67	0.00
Mauritius	51	0.08	1	1		98.85	6.58	0.00
terrorism	*85*	*0.13*	*28*	*25*	*0.02*	*83.22*	*2.68*	*0.00*
Ukraine	55	0.09	12	12	0.01	63.60	3.11	0.00
Iran	58	0.09	7	15	0.01	61.33	2.86	0.00
Kenya	30	0.05	1	2		50.20	4.82	0.00
archipelago	19	0.03	2	0		40.19	137.30	0.00
Croatia	19	0.03	1	0		40.19	137.30	0.00
Romania	19	0.03	1	0		40.19	137.30	0.00
Pakistan	22	0.03	2	1		39.16	5.37	0.00
Iran's	26	0.04	2	3		38.27	4.03	0.00
terrorist	*36*	*0.06*	*20*	*10*		*36.51*	*2.76*	*0.00*
Nepal	17	0.03	2	0		35.96	137.14	0.00
Moldova	17	0.03	4	0		35.96	137.14	0.00
Chagos	14	0.02	1	0		29.62	136.86	0.00
Russia	29	0.05	4	9		27.42	2.60	0.00
Russian	21	0.03	5	4		25.85	3.30	0.00
terrorists	15	0.02	10	1		25.10	4.82	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, Terror— group). All function words have been removed.

The full set of keywords were grouped into three semantic categories: places in the

Terror+ group, places outside the Terror+ group, and terrorism (see Table 6.5).

Table 6.5. Keyword Categories for Terror+ Group

Places in Terror+ Group	Places outside Terror+ Group	Terrorism
Israel	Iran	*terrorism*
Mauritius	Iran's	*terrorist*
Ukraine	Russia	terrorists
Kenya	Russian	
Croatia		
Romania		
Pakistan		
Nepal		
Moldova		
Chagos		
archipelago		

Note. Keywords in at least 50% of texts are italicized.

As with the keyword analyses in Chapter 5, the semantic category with places in the target group was the most productive. The majority of these places were named only by the country itself (e.g., *Chagos* by Mauritius) or by itself and just two or three other countries (e.g., *Israel*). One keyword in this category, *Ukraine*, was unlike the others in its high distribution, 12 out of the 28 texts (43%). Because of its high frequency overall (55 tokens) and particularly its wide distribution, it was deemed worthy of further investigation.

In the Terror+ texts, *Ukraine* is always used in reference to the Russian occupation of Ukraine. The use of *Ukraine* and *Ukrainian* in the Ukrainian text was discussed at length in Section 5.6.3.2. Examples from other texts of the association between *Ukraine* and the Russian occupation include: *The annexation of Crimea was a blatant violation of international law and its principles* [Bulgaria]; *We wish to see a political solution to the crisis in Ukraine based on Ukraine's territorial integrity and the principles of Minsk Agreements* [Turkey], *Consider Russia's annexation of Crimea and further aggression in eastern Ukraine* [US].

Likewise, *Russia* and *Russian*, in the category "places outside of the Terror+ group" are most frequently used in conjunction with the occupation of Ukraine (38 out of 48 occurrences), as shown in Text 6.23). In the remaining 10 occurrences, *Russia* and *Russian* relate to other examples of conflict and violence (see Texts 6.24 and 6.25).

Text 6.23. Excerpt from Moldova (keywords in italics)

> The annexation of the Crimean peninsula by the *Russian* Federation, in a flagrant violation of the international law, and the outbreak of the armed conflict in the eastern regions of the *Ukraine* have directly challenged the international peace and security.

Text 6.24. Excerpt from Ukraine (keywords in italics)

> The second time *Russia* put its shameful veto on the draft resolution on establishment of the International Tribunal to investigate and bring to justice all responsible for Malaysian MH17 plane crush. By imposing its disgraceful veto on this draft resolution, *Russia* clearly demonstrated to the whole world its defiance in establishing the truth. Not just the truth about perpetrators of this *terrorist* attack and arms, used to shot down that plane. What is most important is the truth about those, who organized this crime and from which country the mentioned arms had been transported.

Text 6.25. Excerpt from Australia (keywords in italics)

> Additionally, Australia remains determined to hold to account those responsible for the downing of Malaysia Airlines flight MH17 in July 2014. We will not let a *Russian* veto impede the efforts of countries grieving the loss of their citizens and demanding justice for the perpetrators of this atrocity. All states must uphold their responsibility to protect civilians from the most serious international crimes.

Iran and *Iran's* also co-occur with references to conflict (weapons) or violence. In seven different Terror+ texts, they relate to the nuclear agreement between Iran and the P-5 + 1 (five permanent members of the Security Council plus Germany). Six of the countries praise the deal (e.g., Text 6.26); only Israel expresses deep skepticism (Text 6.27). However, as shown in Text 6.26, even when a country applauds the nuclear agreement, calling it a *positive development* that

is hoped to *pave the way towards more stability in the broader Middle East*, *Iran* nonetheless

appears in a paragraph addressing terrorism and conflict.

Text 6.26. Excerpt from Estonia (keywords in italics)

> As we continue to tackle global challenges including *terrorism*, climate change, poverty and human rights violations, the number of conflicts and crises worldwide continues to grow. We also witnessed a positive development recently. Let us hope that the historic agreement on Iran nuclear programme will pave the way towards more stability in the broader Middle East. It is a region where conflicts in Syria and Libya have led to the radicalisation of an ever greater number of people and to the emergence of ISIL. It poses a serious threat to peace and security in Syria, Iraq and the broader Middle East. ISIL violates universal human values. No country is immune from the threat that it poses. Stopping it, and other *terrorist* organisations, requires a global effort. Estonia supports the international coalition against ISIL. We believe that the UN and its Global Counter *Terrorism* Forum also have an important role to play.

Text 6.27. Excerpt from Israel (keywords in italics)

> As the leader of a country defending itself every day against *Iran's* growing aggression, I wish I could take comfort in the claim that this deal blocks *Iran's* path to nuclear *weapons*. But I can't, because it doesn't.

Only two texts, from the United States and Israel, mention Iran in a context other than the

nuclear agreement, and there the surrounding context relates to tense relations and conflict (Text

6.28), even terrorism (Texts 6.29 and 6.30).

Text 6.28. Excerpt from the United States (keywords in italics)

> From Singapore to Colombia to Senegal, the facts shows that nations succeed when they pursue an inclusive peace and prosperity within their borders, and work cooperatively with countries beyond their borders. That path is now available to a nation like *Iran*, which, as of this moment, continues to deploy violent proxies to advance its interests. These efforts may appear to give *Iran* leverage in disputes with neighbors, but they fuel sectarian conflict that endangers the entire region, and isolates *Iran* from the promise of trade and commerce. The Iranian people have a proud history, and are filled with extraordinary potential. But chanting "Death to America" does not create jobs, or make *Iran* more secure. If *Iran* chose a different path, that would be good for the security of the region, good for the Iranian people, and good for the world.

Text 6.29. Excerpt from Israel (keywords in italics)

> Every few weeks, *Iran* and Hezbollah set up new terror cells in cities throughout the world. Three such cells were recently uncovered in Kuwait, Jordan and Cyprus. In May, security forces in Cyprus raided a Hezbollah agent's apartment in the city of Larnaca. There they found five tons of ammonium nitrate, that's roughly the same amount of ammonium nitrate that was used to blow up the federal building in Oklahoma City. And that's just in one apartment, in one city, in one country. But *Iran* is setting up dozens of terror cells like this around the world, ladies and gentlemen, they're setting up those terror cells in this hemisphere too.

Text 6.30. Excerpt from Israel (keywords in italics)

> President Obama and I agree on the need to keep arms out of the hands of *Iran's* terror proxies. We agree on the need to stop *Iran* from destabilizing countries throughout the Middle East. *Israel* deeply appreciates President Obama's willingness to bolster our security, help *Israel* maintain its qualitative military edge and help *Israel* confront the enormous challenges we face.

At first glance, the keyword list did not appear particularly noteworthy. *Terrorism, terrorist,* and *terrorists* were expected to emerge because they were the words distinguishing the two groups at the outset. All other keywords were place names. However, the analysis shows a focus on places in conflict and violence, with most frequent mention of the Russian occupation of Ukraine and the nuclear deal with Iran in the context of insecurity in the Middle East. This finding supports the UNGA collocational analysis showing an association at the micro level between *terror** and violence.

The collocational and keyword analyses have revealed some associations with *terror** across the UNGA as a whole (e.g., acts of violence, extremism) and for all countries in the UNGA that mention *terror** frequently (places in conflict). In contrast, as with the analyses by country group in Chapter 5, very few collocates were identified by country group. Once again, the lack of findings for groups of countries can be attributed to both similarities across texts and to individual differences. By and large, countries refer to events of keen national interest. Often

these topics are specific to only a few countries and therefore are not identified in the analysis by country category. The issues they raise take on a heightened sense of gravity when *terror** is also used in the passage. For instance, in Text 6.29, Israel accuses Iran of setting up *terror cells.* Though other countries may criticize Iran, they do not speak of the country setting up *terror cells*; only Israel, *a country defending itself every day against Iran's growing aggression*, (see Text 6.27) does so. Other countries also refer to *terror** in order to emphasize the magnitude of an issue, but because these issues are raised by one or just a few countries, they do not emerge in the quantitative findings and require manual examination of individual texts. Section 6.2.5 therefore presents brief analyses of twelve Terror+ text excerpts in order to investigate some of these individual associations with *terror*.* Though these texts show different ways in which *terror** is conceptualized, they also reveal an overarching pattern across texts in the Terror+ group: Countries speak about *terror** in relation to an issue of national concern in order to increase the magnitude of its importance before the international community.

6.2.5. Qualitative analysis of individual countries in the Terror+ group

The manual examination of Terror+ texts reveals that states make an association between *terror** and a topic of regional or national relevance that other countries might not consider as extreme as terror. The connection between terror and the chosen security or economic concern serves to underline the seriousness or urgency of the problem. *Terror** tends to be associated either with an entity (an individual, group, or country) or with a concept such as global warming.

Some countries refer to a group operating within their borders as a "terrorist" organization in order to emphasize the gravity and perilousness of the situation. Turkey, like many of the other states, focuses primarily on Daesh (also referred to as ISIS or ISIL) when discussing terrorist networks. However, the Turkish government also expresses concern about

the Kurdistan Workers' Party (PKK), a group that has been involved in armed conflict with the

Turkish state for decades. In order to lend credence to the perceived threat of the PKK, the

Turkish Prime Minister uses both the PKK and Daesh as co-occurring examples of terrorism (see

Text 6.31).

Text 6.31. Excerpt from Turkey

Today, as in the past, Turkey is combatting terrorism of all kinds, including Daesh and
PKK.

Similarly, the Nigerian speaker focuses on the violence perpetrated by Boko Haram, a

terrorist group that is implanted primarily in Nigeria (see Text 6.32).

Text 6.32. Excerpt from Nigeria

Peace, Mr. President, is close to the hearts of Nigerians, as we are in the front line in the
war on terror. Boko Haram's war against the people of Nigeria, Chad, Niger and
Cameroon may not attract as much worldwide attention as the wars in the Middle East
but the suffering is just as great and the human cost is equally high. This is a war about
values between progress and chaos; between democracy and the rule of law. Boko Haram
celebrates violence against the weak and the innocent and deplorably, they hide behind
their perverted interpretation of Islam. Boko Haram is as far away from Islam as anyone
can think of.

Kenya speaks of Al-Shabaab, a militant group based in East Africa (Text 6.33).

Text 6.33. Excerpt from Kenya

In the recent past, Al-Shabaab terrorists have been targeting the civilian population in
Kenya on the pretext that Kenya has deployed her troops in Somalia. The continued
threat by Al-Shabaab has had a great impact on Kenya's national security and economy at
large. Kenya therefore calls upon the international community to take strong, and
enforceable action in tackling Al-Shabaab and other terrorist groups including the
phenomenon of foreign fighters.

Afghanistan mentions the Haqqani network, a group that operates in Afghanistan (Text 6.34).

Text 6.34. Excerpt from Afghanistan

Allow me to dwell on the case of my own country as a prime example of a nation in transition, as well as of a country exposed to multiple risks and threats at the forefront of our combat against international terrorism and extremism. Afghanistan is suffering and its people demand solutions that are practical, verifiable and durable. The presence of terrorist sanctuaries and support networks in Pakistan continue to cause trouble inside Afghanistan. The Haqqani network has been identified as a main culprit and need to be dismantled as has been our demand in the past. Our demands are legitimate as our people continue to suffer at the hands of terrorist elements who cross into Afghanistan and indiscriminately victimize our citizens.

Other texts refer to countries rather than terrorist organizations. Ukraine explicitly accuses Russia of supporting terrorists (Text 6.35), and Israel condemns Iran for aiding Islamic Jihad and opening *terror fronts* (Text 6.36). No other states are so direct in their accusations against either country, but Ukraine and Israel have the most bellicose relationship with Russia and Iran, respectively. Their political situations are likely to have incited their strong language.

Text 6.35. Excerpt from Ukraine

During this period, more than 8.000 Ukrainians, of whom about 6.000 civilians, died at the hands of the Russian backed terrorists and occupiers in Donbas.

Text 6.36. Excerpt from Israel

Iran supplied Hezbollah with precision-guided surface-to-surface missiles and attack drones so it can accurately hit any target in Israel. Iran aided Hamas and Islamic Jihad in building armed drones in Gaza. Iran also made clear its plans to open two new terror fronts against Israel, promising to arm Palestinians in the West Bank and sending its Revolutionary Guard generals to the Golan Heights, from which its operatives recently fired rockets on northern Israel.

Similarly, Syria comes under sharp criticism by European Union (EU) member states. As discussed in Section 5.3.3, the EU is currently receiving a large number of Syrian refugees, which has had dire political and economic consequences in many EU countries. Text 6.37 from Norway accuses Syria of a *violent crackdown* resulting in the refugee crisis.

Text 6.37. Excerpt from Norway

> The crisis in Syria started with peaceful protests calling for freedom. These were met with a violent crackdown. Upholding human rights is one of the fundamental obligations of any government. The exodus from Syria today is a direct consequence of the violence unleashed by the government. The conflict in Syria has also allowed extreme terror groups like ISIL to gain a foothold. Now, both the government and non-state groups such as ISIL and the Nusra front are committing monstrous atrocities. The spread of these extreme groups needs to be stopped.

The text from the President of the European Council is even more direct in blaming Syria for *state terrorism* (Text 6.38).

Text 6.38. Excerpt from the European Union

> The fight against terrorism is no doubt important in this context; that is why we welcome the fact that this aim is gaining the support of new and quite unexpected allies. But we cannot overlook the fact that many refugees are fleeing state terrorism, of which Syria is a dramatic proof. Europe, which everyday witnesses the tragedies of millions of Syrian refugees, must be their advocate.

The texts from Norway and the European Union use harsh language in discussing the refugee crisis. In contrast, the Prime Minister from Samoa, also a Terror+ country, but one that has not received any Syrian refugees (United Nations High Commissioner for Refugees, n.d.) and that has few refugees more generally, is vague in his reference to *people fleeing from their countries*. Samoa is not directly concerned by the situation in Syria, and the language of the text reflects that fact (Text 6.39).

Text 6.39. Excerpt from Samoa

Over the last several weeks, we have witnessed the tragedy of people fleeing from their countries mired in the destruction from war and terrorism.

Finally, some countries raise topics of vital national interest very soon after using the term *terrorism* or *terrorists*. Mauritius adds *piracy* to the list of threats to world peace (see Text 6.40), an issue of particular concern to this island in the Indian Ocean.

Text 6.40. Excerpt from Mauritius

The spread of armaments, nuclear proliferation, terrorism and piracy remain important threats to world peace and require our constant vigilance.

The President of South Africa speaks at length about the need for greater representation of African countries in the United Nations. Terrorism in North Africa proves to be yet another opportunity to stress the importance of *heeding informed counsel from the African Union* (Text 6.41).

Text 6.41. Excerpt from South Africa

The 70th General Debate takes place in the context of growing international concern about the rise of violent extremism, terrorism and untold brutality which we strongly condemn. We welcome the meetings that will take place on the margins of the UNGA, to review progress made in countering terrorism in the Middle East and North Africa. We wish to emphasise that the UN Security Council must take into account the views of the African Continent and its sub-regional organisations when dealing with conflicts in Africa in future. The current situation in Libya and the Sahel region is a direct consequence of some members of the UN Security Council not heeding informed counsel from the African Union.

The text from Brunei Darussalam suggests that one way to tackle terrorism is to find a peaceful solution to the Palestinian situation (see Text 6.42). Brunei Darussalam is a staunch

supporter of a Palestinian State and does not recognize the state of Israel (Government

Publishing Office, 2008).

Text 6.42. Excerpt from Brunei Darussalam

> It is crucial for us to address the root causes of terrorism and extremism in a comprehensive manner. We need to ensure that irresponsible parties do not exploit others' despair and suffering to fuel their own extremist agendas. One way to approach this is to resolve the plight of our Palestinian brothers and sisters. It remains to be the core of the Middle East conflict. For over 60 years, they have struggled to achieve self-determination and I strongly believe that a lasting and peaceful solution is long overdue. Nevertheless, I am glad to see some positive developments where Palestine has been accepted as a member to various international bodies.

Thus, Terror+ countries tend not only to speak of *terror** frequently, they also use

terror* to emphasize issues of vital regional or national interest. Placing their concern side by

side with *terror** adds gravity and urgency. It focuses the attention on the speaker's country. The

Prime Minister from Pakistan perhaps states it best when he says: *Pakistan is the primary victim*

of terrorism.

6.2.6. Summary

The collocational, keyword, and qualitative analyses have shown that *terrorism* does

indeed mean different things to different people. Arguably, the term is used by design to mean

different things for different purposes.

Some notions are shared among the international community. The collocational analysis

of all 92 UNGA addresses shows associations between *terror** and concepts like acts of violence

and religious extremism. It is generally considered international in scope and of immediate

relevance to the world. Yet some countries (the Terror— group) shy away from the term or

simply name *terrorism* as one of many issues of concern to the UN. Others (the Terror+ group)

use it to create associations between extreme violence and personal enemies or to raise global sympathy for issues of national interest.

The most compelling finding is the fact that such few differences can be identified based on categories of states. Generalizations could be made only for Asia and the biggest military spenders; no other patterns emerged based on region, economic development, social development, percent religious affiliation, population, or area size. This suggests that categories of states do not inherently see all issues from the same perspective. Rather, certain categories of states see certain issues through a given lens because those groups have a vital interest or stake in those particular issues. Asia associates *terror** with *extremism* because one of the primary economic, social, security, and political concerns of 2015 was extremist violence throughout the region. While every region has had to face terrorism at one time, in 2015 four out of the five states most impacted by terrorism were in Asia and over half the deadliest terrorist attacks were in Asia (Institute for Economics and Peace, 2016). These acts of violence, when responsibility was claimed, were carried out in the name of extremist views, thus creating an inherent link between *terror** and *extremism.* Though Niger, Nigeria, Kenya, and Cameroon also suffered heavy losses due to terrorism, in Africa more fatalities were due to widespread poverty and diseases such as malaria and HIV/AIDS. Thus, *terror** is not necessarily one of the key issues defining the discourse community in Africa while it may be an integral part of the Asian discourse community in 2015.

This case study has shown that collocational and keyword analyses as well as the qualitative study of concordance lines and the broader co-text are useful tools in uncovering how addressors conceptualize topics of particular concern. The study has revealed patterns in associations with *terror** and the reasons for those patterns, but has also shown that not all

discourse communities are defined by *terror** and many individual players create unique associations with *terror** based on their vital interests.

Sections 6.3 and 6.4 offer two other case studies using the same methods and reveal similar findings.

6.3. Case Study #2: Security Council

6.3.1. Introduction

The Security Council is one of the six main organs of the United Nations. Responsible for maintaining international peace and security, the Council has the power to conduct an investigation into a conflict; mediate; dispatch a mission, military observers, or peacekeeping units; call for a ceasefire; impose economic sanctions and travel bans; suspend diplomatic relations; enforce a blockade; and organize collective military action (United Nations Security Council, n.d.).

The Security Council is made up of 15 members. The five permanent members are China, France, Great Britain, Russia, and the United States. The ten other members are elected by the General Assembly and serve two-year terms. The non-permanent members are selected on a regional basis to ensure that there are five from Africa and Asia, two from Latin America and the Caribbean, one from Eastern Europe, and two from Western European or other states. All permanent and non-permanent members have one vote and decisions require the assent of nine members. In addition, the five permanent members have the right to veto, meaning that their assent (or abstention) is required for any Council decision. All UN member states have the right to participate in discussions when the Council deems their interests to be at stake, but they do not have the right to vote if they are not members. Regardless of voting status, all UN member states are bound to the decisions made by the Security Council.

Since the Security Council first met in 1946, 66 UN member states have never served on the Security Council. As a result, General Assembly General Debate addresses include much criticism of Council representativeness and many calls for reform. Of interest is not only which countries raise the issue of Security Council representativeness, but more generally what associations they make with the Council. Therefore two additional collocational analyses were carried out, (1) for countries that had been members of the Council and (2) for countries that had never been members.

Section 6.3.2 is a description of the collocational analysis for the entire UNGA, Section 6.3.3 of the collocational analysis by Security Council membership, and Section 6.3.4 of the collocational analysis by country group. Section 6.3.5 reports the keyword analysis comparing all texts that mention *Security Council* to texts that do not, and Section 6.3.6 the keyword analyses by Security Council membership, comparing texts from countries that have been members to those that have not; Section 6.3.7 presents findings from individual texts. Section 6.3.8 summarizes the findings.

6.3.2. UNGA collocational analysis for Security Council

6.3.2.1. Methods: UNGA collocational analysis for Security Council

Collocates of *Security Council* were identified for the entire UNGA in order to investigate general patterns in the associations speakers make with the Security Council. The Security Council is sometimes referred to as the *Council,* and therefore the investigation began by sorting all occurrences of the word *Council.* All instances of *Security Council* were retained. Other *Council* references were examined manually; *Council* references that clearly meant *Security Council* were retained as well (e.g., *We welcome the unanimous support for the Security Council decision to terminate seven resolutions that inflicted sanctions on Iran upon receipt by*

the Council of a report from the International Atomic Energy Agency, Belize). No *Council* references were ambiguous: When the term was not directly preceded by *Security Council,* as shown above, *Council* was part of a phrase indicating a different council (e.g., *Human Rights Council*) and could be eliminated. Collocates of *(Security) Council* (hereafter *SC*) were identified using WordSmith Tools 7.0 (Scott, 2017). As with previous analyses (see Chapter 5 and Section 6.2), function words and "self-reference" terms (e.g., when the United Kingdom uses *UK*) were eliminated. Collocates were grouped into semantic categories in order to find patterns in the types of associations UNGA countries make with *SC*.

6.3.2.2. Results: Security Council *collocates in the UNGA*

The results of the collocational analysis for the full UNGA revealed that *SC* was associated primarily with reform, but also with terms related to the United Nations structure as a whole, with Security Council function and decision-making, as well as with the types of issues the body tackles. In addition, a number of terms were linked to extensiveness and time. Finally, some words had multiple meanings (e.g., *take*) and could therefore not be placed in one semantic category (see Table 6.6).

Table 6.6. UNGA Collocation Results for *Security Council* by Semantic Category

Collocate	Texts	Total occurrences
Representativeness and reform		**160**
reform	19	25
more	13	17
permanent	11	14
member	7	12
seat	8	10
reformed	10	10
members	7	10
representative	9	9
membership	6	9

Collocate	Texts	Total occurrences
better	6	8
democratic	4	4
reflect	3	3
realities	3	3
effectiveness	3	3
reforms	2	3
reforming	2	3
represented	3	3
Africa	2	2
expanded	2	2
enlargement	2	2
regional	2	2
inclusive	2	2
include	2	2
representation	1	2
United Nations structure		**88**
UN	28	47
United	11	13
Assembly	7	7
countries	4	4
body	3	3
states	2	2
general	2	2
ECOSOC	2	2
bodies	2	2
nations	2	2
Mr.	2	2
small (states)	2	2
Security Council function		**47**
resolutions	10	10
resolution	9	10
term	2	5
Code (of Conduct on Security Council action)	3	3
presidency	2	3
election	2	3
veto	3	3
sanctions	2	2
vote	2	2
open (debate)	2	2
entrusted	2	2
seek	2	2

Collocate	Texts	Total occurrences
Security Council decision-making		**72**
responsibility	8	8
action	4	6
negotiations	4	4
call (on)	4	4
opportunity	3	3
act (v.)	3	3
work	3	3
decisively	3	3
advancing	2	2
confirmed	2	2
accountable	2	2
challenges	2	2
failed	2	2
adapt	2	2
transparent	2	2
effectively	2	2
ending	2	2
enable	2	2
efforts	2	2
calls	2	2
implementation	2	2
implement	2	2
decision	2	2
situations	2	2
involving	2	2
initiative	2	2
question	2	2
Security Council topics		**24**
peace	5	5
women	3	4
mass (atrocity crimes)	3	3
rights	2	2
genocide	2	2
global	2	2
crimes	2	2
Iran	2	2
atrocity	2	2
Extensiveness		**36**
particular	7	7
only	5	6

Collocate	Texts	Total occurrences
relevant	3	3
also	3	3
all	3	3
international	3	3
including	3	3
intergovernmental	2	2
especially	2	2
genuinely	2	2
major	2	2
Time		**20**
time	2	4
recent	3	3
urgent	3	3
years	2	2
last	2	2
times	2	2
period	2	2
during	2	2
Other (multiple meanings)		**103**
is	17	25
be	14	16
make	4	5
able	4	5
need	5	5
take	5	5
do	2	3
continue	3	3
want	3	3
are	3	3
can	3	3
making	2	3
held	2	2
see	2	2
supports	2	2
clear	2	2
(first-)hand	2	2
because	1	2
mind	1	2
imperative	2	2
indeed	2	2
power	2	2

Collocate	Texts	Total occurrences
contribute	1	2
made	2	2

Note. Function words and self-reference terms have been removed.

The largest semantic domain of collocates for *(Security) Council* is reform and representativeness. Out of 550 total collocate tokens, 160 refer to reforming the system of Council membership. For many collocates (underlined in excerpts throughout Section 6.3), it was necessary to examine the co-text carefully in order to see the relationship between the collocate and the semantic category. For instance, *more* did not initially appear to be related to Council reform, but the text revealed a relationship: *I would like to add Afghanistan's voice in support of a reformed* Security Council *that is more inclusive, representative and transparent* [Afghanistan]. The same is true for the collocate *better*: *The* Security Council *in particular must better reflect the geopolitical realities of the world today, and therefore the list of permanent and non-permanent members should be expanded* [Nauru]. *Regional,* too, refers to the Security Council system of electing non-permanent members based on their regional affiliation: *For too long, reform efforts have fallen victim to the geopolitical ambitions of entrenched* Council *members and the regional rivalries of legitimate aspirants* [Saint Vincent and the Grenadines].

The second largest semantic domain is linked to the United Nations structure generally. The most frequent fixed phrase with *SC* is *UN Security Council* (44 occurrences). Other terms in this category include *General Assembly* and *countries. Small* was placed in this category because in both instances, it appears in the phrase *small states* in relation to their power within the UN. We *stood for the* Council *because we believe small states have a positive contribution to make and we want to provide a voice for those who go unheard too often* [New Zealand].

413

Other collocates relate to the functioning of the Security Council, such as the presidency of the Council and the right to veto. These terms frequently appear in adverbials to situate a topic: *During our first <u>term</u> on the* Security Council *in 1981-82, the Middle East was, as ever, prominent on the international agenda* [Ireland]. They are also used to specify details on Security Council decisions or actions: *It has been 15 years since the adoption of the landmark* Security Council *<u>Resolution</u> 1325 on Women, Peace and Security* [Slovakia].

Other semantic domains include Security Council "decision-making" and "topics." "Decision-making" is associated with broader notions such as the need to be *accountable* and *transparent* or to *work effectively* in *negotiations.* In Text 6.43, Iceland uses *responsibility* in reference to the Security Council making decisions and undertaking action for *maintenance of international peace and security. We call on* is used as a request for decision-making or action.

Text 6.43. Excerpt from Iceland (*Security Council* in italics; collocates underlined)

> The *Security Council* bears primary <u>responsibility</u> for the maintenance of international peace and security. We <u>call</u> on all *Security Council* members to, in the words of the Charter, "unite their strength" to halt the bloodshed in Syria.

The category "topics" includes references to the issues addressed by the Security Council: *Latvia supports the proposal to voluntarily restrain the use of the veto at the* Security Council *in situations involving <u>mass atrocity</u> crimes* [Latvia]. These "topics" are often part of longer passages stressing the importance of taking action on an area of concern (see Text 6.44).

Text 6.44. Excerpt from Norway (*Security Council* in italics; collocates underlined)

> The UN has established many global standards. Fifteen years ago, the UN *Security Council* adopted resolution 1325 on <u>women</u>, peace and security. It was a landmark resolution, but its implementation is taking too long. In several war-ravaged towns in Syria, groups of women are calling for a ceasefire and evacuation. They do this at great

personal risk. Their bravery should inspire us. We must intensify the implementation of the *Security Council* resolutions on <u>women</u>, peace and security.

As with the study of *terror** in Section 6.2, some collocates relate to the extensiveness of an issue, limiting, for instance, the reference or focus: *It is the case of the situation in Ukraine with respect to which the United Nations and the* Security Council*, in <u>particular</u>, did not take the action expected in terms of the UN Charter* [Romania] and *Being convinced that the security deterioration "in our region imposes an effective mechanism of small arms and light weapons control, we have also engaged in implementing the UN Programme of Action on Small Arms and Light Weapons, as well as, supporting the <u>relevant</u> UN* Security Council *resolutions 2117 and 2220* [Moldova].

Numerous references are made to time, often related to the urgency of a situation (see Text 6.45).

Text 6.45. Excerpt from New Zealand (*Security Council* in italics; collocates underlined)

> It is <u>time</u> for the *Council* to step up. It's <u>time</u> for it to stop talking about what's right and do what's right. It's <u>time</u> for the *Council* to do the job for which it was created. It's time for its members to set aside their vested interests and historical alliances in order to stop the violence and end the suffering. It's <u>time</u> for the *Council* to do its duty for those who have lost their lives and loved ones, and for the millions who have been displaced.

Other references simply situate an action in time: *Australia's <u>recent</u> experience as a* Security Council *member confirmed that the* Council's *role is more essential than ever* [Australia].

The last category, "other," includes terms that are ambiguous, have multiple meanings, or are used for different semantic domains. *Make*, for instance, is used in reference to topics addressed by the Security Council: *For this reason, Nauru calls on the* Security Council *to <u>make</u>*

climate change a permanent agenda item [Nauru] but also regarding Council reform: *The*

Philippines will support calls for continuing the conversations on the much-needed and long-

overdue reforms in the Security Council *to make the* Council *more democratic, more inclusive,*

more transparent and more accountable [Philippines].

      For the UNGA overall, the analysis reveals a relatively restricted set of semantic domains

associated with *SC*, with a primary focus on the functioning of the body and above all a reform

of the Security Council structure and membership. Sections 6.3.3.2 and 6.3.3.3 refine this

analysis, describing associations made by groups of countries within the UNGA.

6.3.3. Collocational analysis for Security Council *by Council membership*

6.3.3.1. Methods: Collocational analysis for Security Council *by Council membership*

      Given the large number of UN member states that have never held a seat on the Security

Council, the question as to whether current or past Security Council membership influences the

associations countries make with *Security Council* appeared compelling. I therefore carried out a

collocational analysis for the group of countries that had been SC members (the SC+ group) and

a collocational analysis for the group of countries that had never been SC members (the SC—

group). As shown in Table 6.7, of the 69 countries in the UNGA corpus that mention *Security*

Council (hereafter the SC group) 38 had at some point been members of the Security Council

and 31 had not (23 countries in the UNGA corpus did not mention *Security Council*). A

collocational analysis was carried out for both groups and collocations were compared.

Table 6.7. *Security Council* **Frequency by Country and Council Membership**

SC membership	Country	Total Occurrences
SC members (past or present)	**38**	**159**
	Australia	8
	Austria	5
	Azerbaijan	2
	Botswana	7
	Bulgaria	1
	Croatia	1
	Denmark	2
	Ethiopia	2
	Finland	1
	Guyana	1
	Ireland	9
	Jamaica	2
	Kenya	1
	Lithuania	5
	Malaysia	2
	Mauritius	4
	Namibia	3
	Netherlands	5
	New Zealand	22
	Nigeria	1
	Norway	5
	Pakistan	5
	Philippines	3
	Romania	4
	Sierra Leone	4
	Slovakia	1
	Slovenia	2
	South Africa	9
	Sweden	7
	Tanzania	4
	Trinidad and Tobago	2
	Turkey	2
	Ukraine	10
	United Kingdom	6
	United States	1
	Viet Nam	2
	Zambia	4
	Zimbabwe	4

SC membership	Country	Total Occurrences
Not SC members (past or present)	**31**	**94**
	Afghanistan	1
	Barbados	1
	Belize	2
	Bhutan	1
	Cambodia	1
	Cyprus	2
	Eritrea	2
	Estonia	8
	Georgia	2
	Grenada	6
	Holy See	1
	Iceland	6
	Israel	1
	Latvia	7
	Liechtenstein	9
	Malawi	3
	Maldives	1
	Marshall Islands	3
	Moldova	4
	Myanmar	1
	Nauru	3
	Papua New Guinea	3
	Saint Lucia	3
	Saint Vincent	4
	Samoa	4
	San Marino	2
	Seychelles	1
	Solomon Islands	7
	Swaziland	1
	Tonga	2
	Tuvalu	2
Total	**69**	**253**

6.3.3.2. Results: Security Council *collocates by Council membership*

The collocational analysis by Security Council membership did not reveal any

meaningful differences. A comparison of the top 10 content-word collocates for the SC+ and the

SC— groups identified four shared collocates (*UN, reform, members, United,* see Table 6.8).

Five other collocates were inflections (e.g., *resolution, resolutions*) or derivatives (*member, membership*) of those four shared collocates. Lemmas shared by the two groups are italicized in Table 6.8.

Table 6.8. Top 10 Collocates by Security Council Membership

Security Council + Group (Total texts: 38, Total occurrences: 159)	Security Council – Group (Total texts: 31, Total occurrences: 94)
UN (22-39)	*UN (8-11)*
reform (12-18)	*reform (8-8)*
permanent (10-10)	*members (6-6)*
members (7-12)	*reformed (5-5)*
member (6-11)	*resolutions (5-5)*
United (9-10)	*United (3-4)*
representative (7-7)	seat (3-3)
resolution (6-7)	responsibility (4-4)
Nations (8-9)	particular (3-3)
membership (4-7)	reflect (3-3)

Note. Function words have been removed. Italics indicates lemma shared by both groups. (Total number of texts, total number of occurrences).

The analysis by Security Council membership shows that, after the collocation *UN Security Council*, *reform* was the most frequent association with the Council for both groups. It is also the semantic domain with the most types (e.g., *reformed, membership, representative*). Though a few of the collocates did not appear in the top 10 for both groups, they were identified for the other group further down on the lists (e.g., *representative* content-word collocate 18 for the SC— group, *responsibility* collocate 21 for the SC+ group). Overall, no substantial differences in the collocates were detected between the two groups. Even frequencies, relative to the total number of texts for each group, were generally comparable.

It is surprising that countries never having served on the Security Council would not speak of reform more frequently than those who have. However, certain individual differences

appear in how speakers talk of reform and are discussed in Section 6.3.7 (see in particular Texts 6.57-6.60).

6.3.4. Collocational analysis for Security Council *by country group*

6.3.4.1. Methods: Collocational analysis for Security Council *by country group*

Collocational analysis was carried out for groups of countries to identify differences in how they conceptualized *Security Council*. Countries were categorized as described in Section 3.6 (e.g., area, HDI). As with the analysis on *terror** in Section 6.2, collocational analysis was carried out for each group (e.g., highest HDI, lowest HDI). Only content words were retained, and a minimum distribution of 50% of texts was set.

6.3.4.2. Results: Security Council *collocates by country group*

Very few country groups had meaningful collocates that met the 50% dispersion criteria. *UN Security Council* was a collocation for the groups Africa, Asia, Europe, Landlocked Developing Countries (LLDC), Least Developed Countries (LDC), medium HDI, very large area, and large population. However, the distinction between *Security Council* and *UN Security Council* is minute and not particularly compelling. This collocation emerged simply because *UN Security Council* is frequent enough in the corpus that it can meet the dispersion criteria and not for any functional reason related to country category.

In addition to *UN Security Council*, the collocate *permanent* was identified for the very high GDP group and the collocate *reform* for the medium population group. Again, these are highly frequent collocates in the corpus at large. In fact, *reform* is a collocate that generally appears in 25-40% of the texts for all groups. A qualitative analysis of concordance lines where the two collocations appear did not reveal any connection between the collocation and the

country category. Nor did Security Council membership status appear to have any effect on the collocate identified or how it is used.

The collocate *permanent* in the very high GDP group is used to identify whether a state is a permanent or non-permanent member of the Council: *We now seek your confidence to champion the perspective of small and medium-sized states, as a non-permanent member of the* Security Council [Sweden, SC+]. Twice the Secretary of State for Foreign and Commonwealth Affairs from the United Kingdom (SC+) uses the phrase *as a permanent member of the* Security Council. In addition, *permanent member of the Security Council* is used in reference to the use (and abuse) of the right to veto: *It is a matter of grave concern when a permanent member of the* Security Council *acts to undermine the territorial integrity of another* [Iceland, SC—].

The collocate *reform* in the medium population group is used by both Security Council members and non-members in the call for Council reform (see Texts 6.46 and 6.47). Ireland was a member of the Security Council three times between 1962 and 2002, while the Solomon Islands has never been a member of the Security Council.

Text 6.46. Excerpt from Ireland (*Security Council* in italics; collocates underlined)

> However, we need the Security Council to be more accountable, coherent and transparent. We support the Intergovernmental Negotiations on *Security Council* <u>reform</u> and want to see an expanded membership with wider geographical representation, especially for Africa.

Text 6.47. Excerpt from the Solomon Islands (*Security Council* in italics; collocates underlined)

> Solomon Islands supports a <u>reform</u> of the *Security Council* and calls for a SIDS dedicated seat in an expanded (Security) *Council*. This is to ensure SIDS security issues are meaningfully addressed in the *Council*. The *Council* will need to adapt to the ever changing realities we face in our world. We are pleased with the progress of negotiations on reforming the *Council* thus far.

Given the high frequency of these three collocates (*UN, permanent, reform*) for the UNGA as a whole and for most of the country categories, the results of the collocational analysis by country category did not appear to be identifying unique collocations for any category. Rather, the data suggest that geographic, social, political, and economic conditions do not greatly influence the associations made with *Security Council.*

6.3.5. Keyword analysis for the Security Council group

6.3.5.1. Methods: Keyword analysis for the Security Council group

The first keyword analysis compared the texts from all countries that use the term *Security Council* (the SC group) with the texts from all countries that do not use the term. The objective was to determine whether the "aboutness" of texts mentioning the Security Council is different from the "aboutness" of texts that do not. Keywords were identified using WordSmith Tools 7.0 (Scott, 2017). Again, function words were removed from the list of collocates.

6.3.5.2. Results: Keywords for the Security Council group

Only four keywords were identified for the *Security Council* group (see Table 6.9).

Table 6.9. Keywords for Full Security Council Group

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
Council	275	*0.20*	69	13	*0.03*	68.97	*2.56*	*0.00*
Israel	91	0.07	11	2		31.56	3.67	0.00
responsibility	107	*0.08*	42	4	*0.01*	30.40	*2.90*	*0.00*
Charter	151	*0.11*	55	12	*0.03*	25.15	*1.81*	*0.00*

Note. R.C. = reference corpus (here, all texts not mentioning *Security Council*). All function words have been removed.

422

The four keywords were categorized semantically, but they did not appear in any way related: *Council* is the word distinguishing the two groups, *Israel* is a place name represented by one of the texts in the target corpus, *responsibility* is an abstract noun shown previously to be associated with protecting citizens in conflict zones, and *Charter* a proper noun related to UN functioning. (See Table 6.10).

Table 6.10. Keyword Categories for Full Security Council Group

Security Council	Abstract nouns	UN functioning	Places in Group
Council	*responsibility*	*Charter*	Israel

Note. Keywords in at least 50% of texts are italicized.

The keyword *Council* appears at the top of the keyword list because it is the term distinguishing the target corpus (the SC group) from the reference corpus (all texts that do not mention *SC*). It is therefore not particularly meaningful.

The two other keywords that occur in at least half the texts do not necessarily appear in passages related to the Security Council, but they often do, and the connection between *responsibility* or *Charter* and the Security Council quickly becomes clear with the examination of the co-text in which they co-occur. Speakers tend to discuss the *responsibility* of the Security Council to ensure peace and protect civilians in conflict zones: *All states must uphold their responsibility to protect civilians from the most serious international crimes. Security* Council *members have a particular* responsibility *to do so* [Australia]. Several texts refer to abuse of the veto, either indirectly implying (see Text 6.48) or stating outright (see Text 6.49) that the Security Council has failed in its responsibility to protect because a permanent member state has vetoed a decision intended to stop mass atrocity crimes.

Text 6.48. Excerpt from Norway (keywords in italics)

> The permanent members of the Security *Council* have a particular *responsibility*. Norway urges all States to join the proposed code of conduct to enable the Security *Council* to act decisively against mass atrocities. We support the French initiative to suspend the use of veto in such situations.

Text 6.49. Excerpt from Estonia (keywords in italics)

> On the occasion of the l0th anniversary of the *Responsibility* to Protect, I reiterate our support for this principle and call for a renewed commitment by the international community to prevent genocide, war crimes, ethnic cleansing and crimes against humanity. The Security *Council* bears primary *responsibility* in this regard. Unfortunately, in most serious situations, the Security *Council* has failed to act due to the use of the veto. It is disappointing that in cases of Syria and Ukraine the *Council* debates have brought no result. *Council* members must not vote against actions aimed at preventing and stopping mass atrocity crimes.

Some texts addressing the inability of the Security Council to fulfill its responsibilities claim that this further supports calls for reform of the Council. Text 6.50 from Saint Vincent and the Grenadines suggests that while reform of the system of Security Council membership is critical, some countries oppose reform because the current system allows them to exercise greater power.

Text 6.50. Excerpt from Saint Vincent and the Grenadines (keywords in italics)

> The maintenance and restoration of international peace and security has been the *responsibility* of the UN Security *Council* for the last 70 years. The Security *Council*, more than any other body delineated in our *Charter*, is unmistakably in a period of doddering dotage, unable to act with the nimbleness or decisiveness necessary to meet modern challenges. The necessity of reform and rebirth, which is acute in any institution entering its eighth decade, is particularly pressing in the case of the Security *Council*. For too long, reform efforts have fallen victim to the geopolitical ambitions of entrenched *Council* members and the regional rivalries of legitimate aspirants. This must end.

Text 6.50 also illustrates the use of the keyword *Charter*, which is inherently linked to the idea of *responsibility*. In the SC texts, *Charter* refers to the United Nations founding Charter,

424

which outlines the principles, objectives, and responsibilities of the organization. References to *Charter* are often in passages expressing a concern that the United Nations is failing to uphold its responsibilities: *If we fail to deal with the enormous tasks facing us, there is a risk that we will undermine the key values and principles of the UN* Charter [Denmark].

Text 6.51 from Romania also criticizes the Security Council for not acting *responsibly* on the Russian occupation of Ukraine because it failed to *take the action expected in terms of the UN Charter.*

Text 6.51. Excerpt from Romania (keywords in italics)

> In addition, the non-action on such situations creates the wrong impression that unlawful territorial gains, to the detriment of independent and sovereign states, are possible and tolerated. It is the case of the situation in Ukraine with respect to which the United Nations and the Security *Council*, in particular, did not take the action expected in terms of the UN *Charter*. We call upon all members of the Security *Council* to act responsibly and decide in the interest of international peace and security and in accordance with the international law while addressing the situation in Ukraine.

Furthermore, a number of texts refer to UN principles delineated in the Charter to defend their actions or as evidence of the need for the UN to intervene in a conflict. The text from Azerbaijan, for instance, claims that the principles of sovereignty and territorial integrity ensure its right to self-defense against Armenia (see Text 6.52). Azerbaijan also requests the support of the international community based on these same UN principles.

Text 6.52. Excerpt from Azerbaijan (keywords in italics)

> However, if the negotiations fail to bring as an outcome the complete and unconditional withdrawal of the armed forces of Armenia from our occupied lands, Azerbaijan will be compelled to use its inherent right of self-defence guaranteed under Article 51 of the UN *Charter* to ensure the restoration of its sovereignty and territorial integrity within the internationally recognized borders. Azerbaijan highly appreciates the principled stance of the States Members of the United Nations that has been repeatedly expressed on issues

that are of utmost importance for Azerbaijan and pertaining to its sovereignty and territorial integrity. We count on the continued resolve of the international community in defending the purposes and principles of the *Charter* of the United Nations, as well as its strong solidarity with the just position of Azerbaijan.

Other mention of the UN Charter in calls for the peaceful resolution of conflict include the text from Guyana on its border dispute with Venezuela: *Guyana seeks a resolution of this controversy that is consistent with the* Charter *of the United Nations* [Guyana], the text from Lithuania on the Russian occupation of Ukraine: *In the middle of Europe the UN* Charter *has been breached with no consequences. Last year, Russia annexed Crimea* [Lithuania], and the text from Pakistan on its conflict with India: *We propose, that Pakistan and India reaffirm that they will not resort to the use or the threat of use of force under any circumstances. This is a central element of the UN* Charter [Pakistan].

Thus, there is an inherent link between the principles of the UN Charter and the Security Council's responsibility to maintain peace. Again and again, member states denounce the Council's failure to act and present the Council's shortcomings as evidence for the need to reform the Council, both in terms of the right to veto and of Council membership to more adequately represent the world's countries (see Text 6.53).

Text 6.53. Excerpt from Georgia (keywords in italics)

> The UN *Charter*, adopted 70 years ago, was designed to put an end to chaos and build a better world. Today, the world is still facing unprecedented challenges that require strong political resolve. Where necessary, we must reform international organizations to meet current demands. In this context, we commend the ongoing inter-state negotiations on Security *Council* reform, which aim at raising the effectiveness of the UN in crisis situations, as defined by the *Charter*. Increasing the role of small states in the Security *Council*, avoiding the misuse of the veto rights, as well as improving the working methods should be the cornerstone of the reform process. I express hope that the 70th anniversary of the UN will serve as a reminder to all of us to stand firmly for the principles enshrined in the UN *Charter* and the goal of peace and stability throughout the world.

426

The relationship of the keywords *responsibility* and *Charter* to *Security Council* was not apparent in the collocational analysis because in many instances the keywords did not appear within five words of the target term *Security Council.* This serves as yet another reminder of the need to undertake a qualitative analysis of the texts.

The final keyword, *Israel*, is a place name from the *Security Council* group. It does not occur in half the texts, but it is mentioned in eleven texts (as opposed to only two in the reference corpus) and its relationship to the *Security Council* group is not immediately transparent, making it worthy of further investigation.

Most of the text excerpts show that *Israel* is mentioned in calls for peace with Palestine. For instance, *Peace talks need to be resumed so that the State of* Israel *can live alongside a democratic, coherent and viable Palestinian State* [Sweden], *It is our wish that* Israel *and Palestine would resume the dialogue so that their long standing conflict can be resolved peacefully* [Tanzania], *Though the conflict between the states of* Israel *and Palestine are undeniably complex, they are not beyond the capacity of the parties and the international community to resolve* [Saint Vincent and the Grenadines].

In addition, several texts condemn Israeli policies. For example, *Currently* Israel *has forced its authority over Islam's Third Holiest Site - in defiance of the jurisdiction of King Abdullah of Jordan, the lawful Custodian* [Malaysia], *The tragedy of Palestine has intensified. The accepted avenue for peace between Palestine and* Israel *- a two state solution - appears further away today than ever before, due to the intransigent stance of the occupying power* [Pakistan], *We thought apartheid was dismantled; yet the world remains indifferent to the apartheid policies that* Israel *pursues in the occupied Palestine* [Maldives]. These texts speak of Israel in particularly harsh terms: *forced its authority, in defiance of, the intransigent stance of*

427

the occupying power, and *apartheid politics.* It is only in the most serious circumstances that UN countries speak of other member states with such overt criticism. The text from Israel also refers to the criticism it receives from the international community: *The UN should finally rid itself of the obsessive bashing of* Israel and *When will the UN finally check its anti-*Israel *fanaticism at the door?*

Israel does not co-occur with *Security Council*, so the relationship between the two is not explicit. However, the fact that *Israel* occurs in connection with peace shows the importance placed by countries that mention *Security Council* on one of the principal objectives of the Council: maintaining peace and security. This interpretation is supported by the interpretation of *responsibility* and *Charter*, which also showed the emphasis on peace and security.

One final point of interest in the keyword analysis is not what is on the keyword list but what is absent from it. Not one keyword relates to *reform*, in spite of the prominence of this semantic domain in the collocational analyses. In other words, the group that mentions *Security Council* does not talk about reform significantly more frequently than the group that does not. That is partly due to the number of references to UN *reform* that do not explicitly mention the Security Council, such as *We know the UN needs reform to make it more democratic and more representative* [Antigua and Barbuda], *We urgently need to realize the UN Reform. With its current 196 members, as compared to 60 in 1957, the UN system has to be even more inclusive and better reflect current global geopolitical conditions* [Indonesia]. Primarily, though, it is due to references to other types of reform. The text from Kiribati, for example, talks of national reform: *We have embarked on a major education reform programme with skills upgrading for our people in line with our program of migration with dignity* [Kiribati]. Other words in the semantic domain "reform" for the Security Council group are used with different meanings in

other texts (e.g., *permanent* in *Thus, we need tools and institutions to set out rules and tackle the phenomenon on a permanent basis*, Malta; *representative* in *Only a few women serve as special representative of the Secretary General and not a single woman has ever served as Secretary General*, Liberia).

The differences in the collocational and keyword findings for *Security Council* illustrate the utility of carrying out both types of analyses. The collocational analysis uncovered the association between *Security Council* and *reform* while the keyword analysis revealed the importance placed on peace and security for countries that discuss the Security Council.

6.3.6. Keyword analysis for the Security Council+ and Security Council – groups

6.3.6.1. Methods: Keyword analysis for the Security Council+ and Security Council – groups

The next keyword analysis examined texts within the subgroup of countries that mention *SC*, comparing word frequencies for past and present Security Council members (the SC+ group) to frequencies for non-members (the SC— group). The objective was to determine whether, of those countries that mention *Security Council*, there is a difference in the "aboutness" of texts based on whether a country has ever been a member of the Council. Keywords were identified using WordSmith Tools 7.0 (Scott, 2017) and function words were eliminated.

6.3.6.2. Results: Keywords for the Security Council+ and Security Council – groups

None of the keywords for either the SC+ group or the SC— group occur in at least half the texts and most importantly, none were meaningful. As shown in Table 6.11, all keywords in the SC+ group are names of countries in the group itself. They are mentioned either solely by the country in question or by, at most, two other countries. They are therefore not representative of the group at large.

Table 6.11. Keywords for Security Council+ Group

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
Mauritius	49	0.06	1	1		46.79	5.20	0.00
Azerbaijan	34	0.04	1	0		38.09	137.82	0.00
Ireland	31	0.04	2	0		34.73	137.68	0.00
Philippines	30	0.04	1	0		33.61	137.63	0.00
Kenya	29	0.04	3	0		32.49	137.59	0.00
Namibia	24	0.03	1	0		26.89	137.31	0.00

Note. R.C. = reference corpus (here, Security Council— group). All function words have been removed.

The keyword list for the SC— group was longer (see Table 6.12), but also not particularly meaningful.

Table 6.12. Keywords for Security Council— Group

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
islands	86	0.14	8	3		122.73	5.25	0.00
Israel	84	0.14	6	7		100.71	4.00	0.00
Solomon	47	0.08	4	0		79.58	138.70	0.00
Myanmar	27	0.05	1	0		45.71	137.90	0.00
impacts	30	0.05	14	1		43.08	5.32	0.00
Pacific	45	0.08	7	7		42.95	3.10	0.00
Saint	22	0.04	2	0		37.25	137.60	0.00
SIDS	39	0.07	10	7		34.64	2.89	0.00
Papua	20	0.03	2	0		33.86	137.46	0.00
Georgia	31	0.05	2	5		29.08	3.05	0.00
Iran's	26	0.04	2	3		28.09	3.53	0.00
Tuvalu	16	0.03	2	0		27.09	137.14	0.00
Barbados	16	0.03	2	0		27.09	137.14	0.00
developing	*67*	*0.11*	*23*	*30*	*0.04*	*27.06*	*1.57*	*0.00*
Grenada	15	0.03	1	0		25.40	137.05	0.00
island	53	0.09	14	21	0.03	24.98	1.75	0.00

Note. Keywords in italics denote distribution across at least 50% of texts. R.C. = reference corpus (here, Security Council+ group). All function words have been removed.

The keywords for the SC— group were categorized into four semantic domains: places in the Security Council— group, places outside the group, keywords related to Small Island Developing States (SIDS), and other (see Table 6.13).

Table 6.13. Keyword Categories for Security Council— Group

Places in Group	Places outside Group	SIDS-related	Other
Israel	Iran's	islands	impacts
Solomon		Pacific	
Myanmar		SIDS	
Saint (e.g., Vincent)		*developing*	
Papua (New Guinea)		island	
Georgia			
Tuvalu			
Barbados			
Grenada			

Note. Keywords in italics denote distribution across at least 50% of texts. SIDS = Small Island Developing States.

The keywords in the category "places in group" were, like with the SC+ group, primarily references in just that one country's text or in at most three other texts. *Israel* is the exception, but this keyword was discussed in Section 6.3.5.2. It was identified in the SC— group because of the frequency of *Israel* in the Israeli text, but it also occurs in seven of the SC+ group texts and is therefore not particularly representative of the SC— group but rather the full SC group (Section 6.3.5.2).

Likewise, *Iran's*, in the category "places outside group" was identified primarily because it occurs 25 times in the Israeli text (and only once in one other text, the text from Iceland). Again, it is not representative of the SC— group as a whole.

The SIDS related semantic category is also related to the SC— group members. The SC— group has a disproportionate number of SIDS compared to the SC+ group: Of the 31

countries in the SC—group, 14 (nearly half) are SIDS. In contrast, only four of the 38 countries

in the SC+ group (less than 11%) are SIDS. In order to verify the relationship between these

words and SC— group membership, the texts in which each "SIDS related" keyword occurs

were noted. This inspection revealed that *islands* and *SIDS* are used only by SIDS; *Pacific* is

used only once (out of 45 occurrences) by a non-SIDS; *island* is used 52 times by 13 SIDS and

only once by one non-SIDS. Of the 67 occurrences of *developing*, 38 (over half) are in the phrase

Small Island Developing State, all used by SIDS themselves. Of the remaining 29 instances of

developing, 15, over half, are in texts by SIDS. Thus, in the SC— group, *developing* is used

nearly 80% of the time by SIDS. An inspection of *impacts,* in the "other" category, also revealed

that it was primarily related to SIDS: *impacts* is used 25 times by 10 SIDS but only five times by

four non-SIDS. (See also Section 5.3.4.) Thus, all the keywords for the SC— group were related

to group membership and primarily to the fact that so many SIDS have never had a seat on the

Security Council.

Overall, UNGA countries that mention *Security Council* are relatively uniform in the

associations they make with the Council. Most surprisingly, few meaningful distinctions were

found between states that have served on the Security Council and those that have not (but see

discussion on differences in calls for reform in Section 6.3.7). The rare distinctions identified

between country categories also became doubtful upon further analysis: Collocates found in

some country categories are, in fact, highly frequent in nearly all country categories but were not

necessarily identified because they occur in just under the requisite 50% of texts.

6.3.7. Qualitative results of individual countries in the Security Council group

Previous analyses (see Sections 5.6 and 6.2.5) revealed interesting patterns when

individual texts were examined qualitatively after few differences had been identified based on

country categories. As a result, I decided to carry out a manual examination of individual texts for *Security Council* as well. As with the analysis of *terror**, the analysis of *(Security) Council* reveals that many countries use references to the Security Council and the principles it is responsible for safeguarding as opportunities to promote national or regional agendas. Certain patterns emerge in the types of passages *Security Council* occurs in; these patterns cannot be detected in the quantitative analysis by country category for three primary reasons. First, the number of countries within a group (e.g., Africa) that discuss the Security Council for the same reason is not always sufficient to meet the distribution criteria. Second, some words do appear to be associated with *Security Council,* but they do not occur within the requisite five-word span to be identified in the collocational analysis. Finally, some patterns emerge across rather than within groups. That is, the country categories do not capture the similarities that explain a country's focus on the Security Council.

Though overall most countries discuss the Security Council for reasons of national and regional interest, in two cases countries raise important issues for the international community at large or for another region: conflict resolution and reform. As discussed in Section 6.3.5.2 (see in particular Text 6.52 from Azerbaijan and the shorter excerpts from Guyana, Lithuania, and Pakistan), countries often cite the principles of the UN Charter to drum up support for conflict resolution. Most countries refer to conflicts they are directly affected by, either at the national (see Text 6.54) or regional level (see Text 6.55), but some countries do refer to conflicts outside of their region that have much less pronounced national impact (see Text 6.56).

Text 6.54. Excerpt from Cyprus

> And we also express our appreciation for a plethora of UN Resolutions and UN Security
> Council decisions which condemn the unacceptable status quo and the violent ongoing

division of Cyprus, calling for its re-unification and the withdrawal of occupation forces. Resolutions and decisions which unfortunately have yet to be implemented.

Text 6.55. Excerpt from Kenya

The maintenance of international peace and security continues to face increasing challenges due to the increased intensity and complexities in conflicts in the Middle East and Africa, especially in the Great Lakes region, Horn of Africa and in the Sahel. The continued grave security situation in Eastern Africa and further afield is of particular concern to Kenya. Kenya has continued to invest billions of shillings in the maintenance of regional peace and security, while many Kenyan lives have been lost in our quest for peace. Kenya therefore, implores the international community, in particular the United Nations Security Council to shoulder its responsibility in a manner that has a direct impact to the resolution of those conflicts.

Text 6.56. Excerpt from Slovenia

Over the past few years we have been witnessing war in Syria, raging conflicts in Iraq, Libya, Gaza, Yemen, South Sudan, Central African Republic and elsewhere, with their spill-over effects in their immediate neighbourhoods, and consequences felt around the world. The situation is further worsened by an escalation of violent extremism and crimes committed by Daesh and other terrorist groups and organizations. This must be stopped. We call upon the Security Council, and especially its permanent members, to overcome divisions and find the way to deal more effectively with this worsening situation, in line with its primary responsibility for the maintenance of international peace and security.

Regardless of Council membership, countries also frequently discuss the importance of

Security Council reform. Many calls for reform come from countries that have never served on

the Security Council (see Text 6.57).

Text 6.57. Excerpt from Cambodia

The UN Security Council should be more democratic and representative in both permanent and non-permanent membership to assure universal peace and security in the face of the dangerous evolving situation which prevails in the world today.

Some countries push for reform even though they have served, but this, too, is understandable because they are from an underrepresented region and are not likely to serve again soon (see Text 6.58).

Text 6.58. Excerpt from Botswana

> The UN Security Council is not an example of democracy where permanent membership is limited to five countries which further have a veto. There should be no permanent members and no veto at all.

Yet countries that have served relatively frequently and that are from a region with better representation (e.g., Western Europe) recognize the importance of reform, too (see Text 6.59).

Text 6.59. Excerpt from Austria

> We need a Security Council that is united and can act swiftly. We welcome the efforts to improve the work of the Security Council. We support the initiatives to suspend the use of the veto in case of mass atrocities. In order to be able to address global challenges we need a UN that is capable of taking action - and a Security Council that is more representative, accountable and transparent.

Even the UK, a permanent member, acknowledges the need to reform the Council (see Text 6.60).

Text 6.60. Excerpt from the UK

> The UN must strive to represent the new realities of our age, with a reformed Security Council. It must have the best possible leadership, with a transparent system for selecting the next Secretary General... ...and he (or dare I suggest, perhaps, she?) will have to head a more efficient organisation, ensuring that every cent it receives from its members states is used to maximum effect. As a founding member of the UN, and as a permanent member of the Security Council, the UK will champion that reform agenda.

Though Council membership does not affect the frequency of calls for Security Council reform, it does influence the language used in those calls for reform. Some are very vague. The

435

text from the UK, for instance, does not specifically mention expanding membership. In contrast, the text from Botswana is extremely direct and most radical, pushing not just for additional members, but for the full elimination of permanent membership. Also of interest is that the use of modals of necessity does not necessarily make the directive force and degree of imposition strong. Though the UK uses the strongest modal, *must*, in *the UN must strive to represent the new realities of our age, with a reformed Security Council*, there is little direct imposition on countries because nothing concrete is being asked of member states. Indeed, what <u>are</u> the "new realities of our age" and what kinds of reforms do they imply? Even the text from Cambodia is stronger: The use of *should* in *the UN Security Council should be more democratic and representative in both permanent and non-permanent membership,* is more specific in its recommendation. There is no question that Cambodia is asking for a reform of the system of membership. Any country could agree to the UK's call "to represent the new realities of our age" without committing themselves to anything, but expressing consent to Cambodia would mean supporting proposals to change the way Council members are selected. The use of the modal of necessity in the excerpt from Botswana (*There should be no permanent members and no veto at all*) is particularly interesting because Botswana is a Landlocked Developing Country (LLDC), a group which was shown to use modals and in particular modals of necessity very infrequently (see Section 5.4.4). The use of *should* in this excerpt underlines the importance of Security Council reform for Botswana.

One country group with frequent calls for reform using very explicit, strong language is Africa. Namibia criticizes the poor representation of Africa in the Security Council as part of *the long historical injustice endured by a Continent of over one (1) billion people*, and South Africa

condemns the absence of an African permanent member seat as reflective of colonial times (see Text 6.61).

Text 6.61. Excerpt from South Africa

> It is unacceptable and unjustifiable that more than one billion people in the African continent are still excluded as permanent members of the key decision making structure of the United Nations, the UN Security Council. A continent with a smaller population than Africa is represented by three countries on the UN Security Council as permanent members. The UN cannot pretend that the world has not changed since 1945. We are no longer colonies. We are free, independent sovereign states.

Numerous texts from Africa refer to the Ezulwini Consensus and Sirte Declaration. The Ezulwini Consensus is an African Union position paper calling for at least two permanent positions with right to veto and five non-permanent seats (African Union, 2005). African texts cite the Ezulwini Consensus and Sirte Declaration as support for their calls for reform: *Mauritius reaffirms its commitment to the African Common position enshrined in the Ezulwini Consensus and the Sirte Declaration* [Mauritius] and *In this context, it is only proper that we seize this occasion to, once again, echo Africa's call to be fully represented in all the decision-making organs of the 7 UN, particularly in the Security Council as encapsulated in the Ezilwini Consensus and the Sirte Declaration* [Ethiopia]. The use of the participles *enshrined* and *encapsulated* lends a certain authority and power to the African Union documents.

Many of these texts use strong language. In addition to the stance adjectives *unacceptable* and *unjustifiable* in Text 6.61 from South Africa, the text from Sierra Leone uses *overdue* and *imperative*: *The need to address the non-representation of Africa in the Permanent category and the under-representation in the Non-permanent category is long overdue and therefore now imperative*. Ethiopia, too, deems reform *imperative*, suggesting, in fact, that the UN as it currently stands is not entirely *legitimate*: *Comprehensive reform of the United Nations system,*

particularly that of the Security Council, is indeed imperative to reflect current geo-political realities and to make the UN more broadly representative, legitimate and effective. Zambia denounces *inequality* in the Council: *Goal 10, of the 2030 Agenda for Sustainable Development on reducing inequality among countries will not be achieved without eradicating the inequality among countries in the Security Council.* Zimbabwe talks of *abusing the authority* of the Council: *In the case of Libya, we are witnessing the results of abusing the authority of the United Nations Security Council and ignoring the opinion of regional organisations, in this instance, the African Union (AU), which are supposed to be the United Nations' partners in the maintenance of international peace and security.*

While no collocations were identified for the African group in Section 6.3.4.2, these excerpts from African countries show the clear association in many texts between the Security Council and reform, with references to colonialism and historical injustice and with the use of participles, stance adjectives and stance nouns to convey the importance of the issue and of the official African viewpoint on the issue. These excerpts from African countries illustrate the value of qualitatively studying texts. Some patterns may not be identified in a quantitative analysis because the difference lies in how (not simply whether) a text addresses an issue.

Another country category showing some patterns in associations with the Security Council is the SIDS group. Several countries *urge* the Security Council to discuss the *particular peace and security challenges facing small island developing states* [Barbados]. In the text from the Marshall Islands: *I urge the Council's establishment of a regular agenda item on security issues in small island developing states*, and in the text from Grenada: *In this regard, we strongly urge the Security Council to continue to give greater consideration to the special circumstances of SlDS in relation to both traditional and non-traditional security concerns.*

The text from Grenada continues by naming the environment as one of these "security concerns": *Climate Change is a critical issue not only for intellectual debate but, for its manifestation as a major multidimensional security threat, to Small Island developing states*. Grenada is just one of many SIDS texts to associate the Security Council with climate change. Others include Papua New Guinea: *We therefore, call on the UN Security Council to deal decisively with the implications of climate change and security issues*; Nauru: *In addition, the Security Council must continue to adapt and respond to the most pressing issues of today. For this reason, Nauru calls on the Security Council to make climate change a permanent agenda item*; Samoa: *Just a few months ago I was privileged to attend the special Security Council Open Debate on Security concerns of SIDS. It was an important opportunity to highlight in the Council the security threat of Climate Change for all our Small Island Developing States*; and Tonga: *we welcome the dialogues having taken place at the Security Council and a side event to be hosted by the Foreign Ministers of France and Germany tomorrow afternoon, touching upon the security implications of climate change, and support recognition of the links between our climate and international peace and* security.

The language in these passages contrasts sharply with that in many of the African texts. Adjectives such as *important* [Samoa], *critical* [Guyana], and even *pressing* [Nauru] are not nearly as strong as *unjustifiable* [South Africa] and *abusing* [Zimbabwe]. Suggestions are made with the softer *call on* [Papua New Guinea] and *welcome* [Tonga] rather than the more forceful *is imperative* [Ethiopia] and *will not be achieved without* [Zambia]. Even when SIDS call for a designated seat on the Council for their group (see Text 6.62), softer language is used: *concerned, uncertainties, could, supports.* Toward the end of Text 6.62, the Solomon Islands

439

becomes more direct: *to ensure, will need*, but this is mitigated by a positive evaluation of

progress already made (*we are pleased*) in the next sentence.

Text 6.62. Excerpt from the Solomon Islands

> We are concerned that the slow progress in climate negotiations causes uncertainties and possibilities for climate change induced situation that could trigger conflicts. This uncertainty prompts Solomon Islands to seek a seat in the Security Council for the period 2031-2032. This will be a year after we take stock of our 2030 Agenda. Solomon Islands supports a reform of the Security Council and calls for a SIDS dedicated seat in an expanded (Security) Council. This is to ensure SIDS security issues are meaningfully addressed in the Council. The Council will need to adapt to the ever changing realities we face in our world. We are pleased with the progress of negotiations on reforming the Council thus far.

Text 6.62 also illustrates one final association identified with *SC*: elections. Several

countries speak at length on the Security Council because they are seeking a non-permanent seat

on the Council. They refer to their dedication to peace and international security and ask for the

backing of other UN member states in their candidacy (see Texts 6.63, 6.64, and 6.65).

Text 6.63. Excerpt from Viet Nam

> At the international level, Viet Nam is prepared to further our active contributions to world peace and security and the well being of all. That is why we step up our participation in UN peace-keeping operations and has put forward candidatures for the Economic and Social Council for 2016-2018, UNESCO Executive Board for 2015-2019 and the Security Council for 2020-2021. We look forward to receiving your continued support.

Text 6.64. Excerpt from the Netherlands

> For the Kingdom of the Netherlands, the United Nations is, and will remain, the primary global organisation for peace, justice and development. That is why the Kingdom is a candidate for a non-permanent seat on the Security Council for the 2017 to 2018 term.

Text 6.65. Excerpt from Ireland

> In our wide-ranging work on conflict prevention, mediation, and state-building, we seek to share our experience. We hope to bring this to the table when we seek election to the Security Council again in 2020.

In Text 6.66, Sweden, too, highlights its contributions to international peace and security before announcing its candidacy to the Council. Like the other texts from countries seeking Council election, the language overall is very positive (*contribute(s), want, effective, transparent, fit-for-purpose, strengthen, cooperation, friends, stood by, confidence, champion, strive, representative, support*). Again, this is in stark contrast with the African texts on the whole. However, the situation of Sweden in the Council also stands in stark contrast to that of many African nations. Sweden served on the Council less than 20 years before delivering the address (and had served twice before that) while nine African countries have never served at all.

Text 6.66. Excerpt from Sweden

> Sweden contributes resources, but we also want to contribute reforms. We want to make the UN more effective, transparent and fit-for-purpose. We will work to strengthen cooperation between the UN and regional organisations, not least the African Union. Dear friends, In 2017 it will be twenty years since Sweden last took a seat on the Security Council. Over the years, we have stood by those fighting for independence and dignity, and against repression, colonialism, apartheid and inequality. We now seek your confidence to champion the perspective of small and medium-sized states, as a non-permanent member of the Security Council. We will strive tirelessly for a Council that can respond swiftly to the security challenges of our time; a Council that is more representative, transparent and effective. A reformed Council must reflect the realities of today, with adequate representation of Africa, Asia and Latin America, and we support efforts to limit the use of the veto.

The manual examination of passages in which *Security Council* appears shows that certain patterns do emerge in associations made with the Council. Two issues are raised by countries regardless of self-interest: Council action to resolve conflicts around the world and

441

Council reform. In addition, several African countries use very strong language to denounce the current system of Council membership and push for reform. One other pattern was identified based on country group, though like with the African countries, it was not strong enough to be detected in the quantitative analysis: Though SIDS generally make less reference to the Security Council than other country groups, when they do, they use it as an opportunity to raise an issue of immediate national interest: climate change. Finally one pattern surfaced across texts: When seeking a non-permanent member seat, countries mention the Security Council using relatively positive language, even when acknowledging the need for reform. This is true regardless of past membership on the Security Council or other country-related factors such as region.

6.3.8. Summary

Section 6.3 has demonstrated that, like with *terror**, clear associations are made with *Security Council* in the UNGA as a whole but that country groups do not systematically distinguish themselves in terms of their conceptualization of the Council. Nor does past or present membership on the Council greatly influence associations. In contrast, patterns do emerge in the qualitative analysis, showing that in addition to raising common concerns such as conflict resolution and Council reform, *SC* is sometimes mentioned to promote national agendas as with climate change among SIDS and Security Council elections among countries seeking a non-permanent member seat.

6.4. Case Study #3: Women

6.4.1. Introduction

Words denoting gender (e.g., *boys, girls*) have been the object of extensive CADS study, often with a view to examining representations of gender stereotypes (see Partington et al., 2013,

442

p. 310). In addition to revealing gender-related associations (e.g., *lovely girl, naughty boy*, Partington et al., pp. 315-316), the study of *women* in the UNGA could also uncover beliefs about *women* as an issue for UN action. In 2015, the UNGA was celebrating the 20th anniversary of the Beijing Declaration and Platform for Action, which established measures for advancing the rights of women nationally, regionally, and internationally. In addition, the 2030 Agenda for Sustainable Development, which had been adopted on 25 September, 2015, just three days before the start of the 2015 UNGA, makes specific mention of the importance of gender equality. Sustainable Development Goal 5 is to "achieve gender equality and empower all women and girls." Thus, women's rights and advancement were of topical interest during the UNGA in 2015.

The question remains as to what associations countries make with women's rights and *women* more generally and whether countries differ in their perceptions of gender equality and the empowerment of women and girls. Section 6.4 describes the investigation of *woman** (*woman*, *women*, *woman's*, and *women's*) in the UNGA, with collocational and keyword analyses elucidating the representation of *woman** and the discourses surrounding the term. Section 6.4.2 reports the collocational analysis for *woman** across all UNGA texts and Section 6.4.3 the collocational analysis by country group. Section 6.4.4 describes the keyword analysis for the texts that use *woman** compared to the texts that do not. Section 6.4.5 presents the qualitative analysis of individual texts, examining the larger co-text surrounding *woman**, and Section 6.4.6 summarizes the findings on *woman**.

6.4.2. UNGA Collocational analysis for woman*

6.4.2.1. Methods: UNGA Collocational analysis for woman*

        Collocates of *woman** were identified using WordSmith Tools 7.0 (Scott, 2017). Function words and self-reference terms (the speaker's own country) were eliminated. Collocates were grouped into semantic categories in order to find patterns in the types of associations UNGA countries make with *woman**.

6.4.2.2. Results: Woman* *collocates in the UNGA*

        The collocational analysis for the UNGA as a whole examined general patterns in the associations speakers make with *woman**. The results of this collocational analysis reveal several large semantic domains for *woman**. The largest domains in terms of tokens were "rights and equality" and "people." "Violence and its consequences" and "development" were much smaller domains. As with the other analyses in Chapter 6, the domains "UN action and decision-making," "extensiveness," and "time" were identified. Finally, some words had multiple meanings and were placed in an "other" category (see Table 6.14).

Table 6.14. Collocates by Semantic Category for *Woman in the UNGA**

Collocate	Texts	Total occurrences
Rights and equality		**145**
empowerment	15	22
rights	13	18
equality	13	16
gender	11	14
participation	6	7
equal	6	6
role	6	6
empowering	4	5
discrimination	4	4
able	3	3
public (e.g., service)	3	3

Collocate	Texts	Total occurrences
positions	3	3
political (participation)	2	3
advancement	3	3
efforts	2	3
advance	3	3
potential	2	2
place (of women)	2	2
served	1	2
representation	2	2
facilitate	1	2
barriers	2	2
active	1	2
better	2	2
human (rights)	2	2
opportunities	2	2
inequality	2	2
society	2	2
end (e.g., gender-based violence)	2	2
People		**110**
men	19	24
children	19	23
girls	16	22
youth	7	10
people	5	5
elderly	4	4
indigenous	3	3
boys	3	3
youths	2	2
mothers	2	2
civilian	2	2
civilians	2	2
peoples	2	2
disabled	2	2
innocent	2	2
individuals	2	2
Violence and its consequences		**38**
violence	6	8
vulnerable	6	7
conflict	3	3
trafficking	2	2
piracy	1	2

Collocate	Texts	Total occurrences
safety	2	2
victims	2	2
suffering	2	2
fleeing	2	2
migration	2	2
migrant	1	2
migrants	2	2
living (e.g., in vulnerable states)	2	2
Development		**21**
development	6	6
sustainable	3	4
poverty	2	3
services	2	2
employment	2	2
health	2	2
basic	2	2
UN action and decision-making		**108**
peace	11	20
security	9	16
UN	9	11
resolution	7	8
Nations	6	6
agenda	5	5
countries	4	4
conference	3	3
anniversary	3	3
processes	2	3
council	2	3
action	3	3
United	2	3
plan	2	2
proliferation	2	2
disarmament	2	2
goal	2	2
general (assembly)	2	2
Beijing (declaration)	2	2
implementation	2	2
meeting	2	2
deployed	2	2
peacekeeping	2	2
Extensiveness		**107**

Collocate	Texts	Total occurrences
all	8	12
many	8	9
including	7	7
world	6	6
most	5	5
more	4	5
every	3	6
particularly	3	5
still	3	4
millions	2	3
parts	3	3
large	3	3
high	3	3
especially	3	3
full	3	3
too	2	2
(not a) single	2	2
ever	2	2
extreme	2	2
global	2	2
national	2	2
municipal	1	2
number	2	2
particular	2	2
only	1	2
include	2	2
increasing	1	2
important	2	2
importance	2	2
key	2	2
Time		**19**
year	3	3
remains	3	3
today	3	3
time	2	2
already	2	2
continue	2	2
now	2	2
keep	2	2
Other (multiple meanings)		**37**
have	5	6

Collocate	Texts	Total occurrences
groups	4	4
has	4	4
well	3	3
under	2	2
post	2	2
so	1	2
see	2	2
realise	1	2
president	2	2
had	2	2
believe	2	2
other	2	2
impacts	2	2

Note. Function words and self-reference terms have been removed.

The largest semantic domain, women's rights and equality, includes both abstract concepts (e.g., *empowerment, discrimination, representation*) and concrete areas for women to be participating in (e.g., *public service, political participation*). Text 6.67 from Mongolia and Text 6.68 from Slovakia illustrate the use of many of these terms.

Text 6.67. Excerpt from Mongolia (*woman** italicized; collocates underlined)

> We believe in gender equality and *women's* empowerment. Mongolia is a strong supporter of the UN in this cause. For societies to advance, we need more *women* in public service at all levels - local and global. If *women* hold more positions of power, we will have less suffering and conflict, and more harmony and civic engagement.

Text 6.68 Excerpt from Slovakia (*woman** italicized; collocates underlined)

> The presence of *women* negotiators in high-profile United Nations peace and mediation processes, role of *women* deployed by the United Nations to major conflict zones and highstake inter-state negotiations, as well as in peacekeeping and peace-building in general, is irreplaceable and should continue to grow.

These passages suggesting the power of women contrast sharply with texts where women are associated with vulnerability. The term *women* was frequently identified in a list of

vulnerable populations, along with *children, the elderly,* and sometimes *the disabled*. Text 6.69 exemplifies such a passage.

Text 6.69. Excerpt from San Marino (*woman** italicized; collocates underlined)

> The Republic of San Marino has always paid special attention to the <u>most</u> <u>vulnerable</u> <u>groups</u>, such as *women*, <u>children</u>, the <u>elderly</u> and the disabled. Today, *women* are <u>still</u> the <u>victims</u> of discrimination and violence in many parts of the world, including in the <u>most</u> developed <u>countries</u>. <u>Trafficking</u> in *women* is far from being solved. *Women* <u>living</u> in <u>conflict</u> and <u>post</u>-conflict situations are often subject to sexual violence, torture and summary executions.

Text 6.69 also illustrates the association between *women* and violence (*trafficking, living in conflict*). *Women* are shown to be particularly vulnerable to violence and the consequences of conflict in Text 6.70 as well.

Text 6.70. Excerpt from Somalia (*woman** italicized; collocates underlined)

> Our *women* - <u>mothers</u>, daughters, sisters - also deserve limitless praise for enduring the worst of our brutal conflict. The survival instinct of the Somali people is <u>most</u> evident in *women*: raising families whilst their homes are being bombed, walking miles without water in search of safety and burying their children whilst continuing to work to feed the rest of their family. We know what war can do. And that is why I will ensure - as a father and a husband - that we will defeat terrorism. It is our duty, as a Government.

In Text 6.70, the Prime Minister of Somalia speaks *as a father and a husband* about *our women* who *deserve limitless praise for enduring the worst of our brutal conflict*. The government, represented by a father and husband, has a duty to defeat terrorism to alleviate the plight of women during conflict. In contrast with Text 6.67 from Mongolia and Text 6.68 from Slovakia, which speak of women in *positions of power* and as *negotiators in high-profile United Nations peace and mediation processes*, respectively, in Text 6.70 women are shown in a more passive role, *enduring… conflict* rather than working to end conflict as negotiators.

449

Many of the collocates in the "people" category occur in passages like Text 6.70, lamenting the vulnerability of civilians in conflict zones, but they can also occur in discussions on equality and rights. Of all the "people" collocates, *girls* is most frequently associated with vulnerability to violence: Women *and girls are sold on slave markets by the criminals of ISIS and abused in unspeakable ways* [Lithuania], *We are committed to combat terrorism and unreservedly condemn the atrocities perpetrated against* women *and girls in Nigeria by Boko Haram* [Belize]. Less frequently, *girls* collocates with *women* in a passage on equality and rights: *The declaration also contains a specific goal on achieving gender equality and the empowerment of* women *and girls* [South Africa] and *Social exclusions and inequalities still present a significant challenge in most countries, with* women *and girls; the youth bearing the heaviest weight of those circumstances* [Botswana].

Boys co-occurs only twice with *women*, notably in passages that also include the collocates *girls* and *men.* In one instance it suggests vulnerability to violence: *Today, we see similar pictures of girls and boys,* women *and men, seeking safety from war and certain death* [Maldives]. The other occurrence is in a passage on human rights: *The Agenda addresses the requirements for all humanity to be able to live decent lives free from poverty, hunger and inequality, with all men and* women*, girls and boys able to develop their full potential* [Kenya].

Women and *men* co-occur relatively rarely in passages on vulnerability to violence: *Violence and extremism that have put multitudes of people on the move, in scenes reminiscent of the Exodus in the Biblical times, of children,* women *and men arriving in their thousands on our doorsteps in the Balkans as they seek safety and protection in Europe* [Albania]. Only three out of 17 occurrences of the phrase *men and women* refer to vulnerability with the collocates *fleeing* or *suffer.* In contrast, in six occurrences of *men and women,* reference is made to equal rights

with one of the collocates *equal, equality,* or *inequality* also present: *Currently we create programs and policies to address this imbalance, yet regardless of how successful they may be, they are not permanent solutions. They do not solve the ultimate problem, which is the vast inequality between men and* women *that so many traditions have inculcated* [Ghana]. The remaining seven occurrences of *men and women* are in reference to heroes, peacekeepers (*men and women in uniform,* Sierra Leone), or simply good *men and women* [Jordan]. Because these and the other collocates in the "people" category represent different semantic domains in different texts, they were categorized in this broader category.

Though *boys* and *men* do appear in passages on vulnerability to violence, it should be noted that *men* never suggests vulnerability in the UNGA texts unless it co-occurs with *woman*.* A collocational analysis of the lemma *man** identified 49 occurrences of *man, man's, men,* and *men's* in the UNGA, 24 of which co-occurred with *woman*.* Of the remaining 25, nine occurrences are in phrases suggesting all of humanity (*the good of man*, Micronesia; *mortal men*, Saint Vincent and the Grenadines). Six are about specific men: *I remember the words of the young Muslim man Lassana Bathily who saved several innocent lives during the hostage crisis earlier this year in Paris* [Bulgaria]. Five are in the phrase *man-made.* The remaining five are in passages on equality: *And, after a succession of eight men in the position of Secretary General, it is high time for qualified female candidates to be seriously considered for this most important international position* [Iceland]. This analysis of *man** helps to clarify the collocate *men* in the analysis of *woman*.* It is not the addition of *men* that suggests vulnerability to *woman** but the contrary.

As with *terror** and *Security Council* in Sections 6.2 and 6.3, *woman** collocates frequently with terms relating to UN policy and decision-making. Some of these terms were

451

easily categorized (e.g., *resolution*), while others required examination of the co-text. Out of the 20 co-occurrences of *peace* and *woman**, 12 are in the phrase *Women, Peace and Security*, the title of a UN resolution. The other eight co-occurrences are related to the resolution, appearing in the co-text surrounding the phrase (see Text 6.71).

Text 6.71. Excerpt from the Philippines (*woman** italicized; collocates underlined)

> The Philippines shall continue on its path to successfully implement the <u>agenda</u> of *Women*, <u>Peace</u> and <u>Security</u> by highlighting at this session the important contributions of *women* in <u>peace</u> negotiations and <u>peace</u> building, and their roles in shaping the narratives of peace.

The semantic category "development" is related to UN decision-making because most of the references are about the 2030 Agenda and specifically about Sustainable Development Goal 5 on gender equality. However, it is a large enough category to merit separate analysis. In fact, the association between *woman** and development is greater than it first appeared because *development* and related terms were often not within the requisite five-word span. Only a manual examination of texts revealed the extent to which *development* and related terms occur in the larger co-text, with speakers explicitly linking women's rights to development. Text 6.72 illustrates *sustainable development* as a collocate in the requisite five-word span. In Text 6.73 *development* does not co-occur within the five-word span, but conceptually the text makes clear the association between *woman** and development.

Text 6.72. Excerpt from the Netherlands (*woman** italicized; collocates underlined)

> There is still not enough focus on the positive <u>role</u> that *women* can play. Even in the harshest and <u>most</u> hopeless circumstances, *women* often <u>have</u> the strength to go on. To find practical ways to improve the lives of their families and communities. Promoting <u>equal</u> <u>rights</u> and <u>opportunities</u> for *women* is actually <u>sustainable development</u> in action.

Text 6.73. Excerpt from Malawi (*woman** italicized; collocates underlined)

My government is committed to the "He-for-She Campaign" to address gender inequalities, and end gender-based violence, promote *women* political participation and facilitate the economic empowerment of *women*. It has therefore been singled out as one of the core priority areas of our revised Malawi Growth and Development Strategy, a blueprint of Malawi's development.

Other examples of *woman** associated with development, though without any

"development" collocates, are shown in Texts 6.74, 6.75, and 6.76.

Text 6.74. Excerpt from Turkey (*woman** italicized; collocates underlined)

We see a clear link between sustainable economic development and global stability. One way to ensure this, is through inclusive economic growth where no one, not least *women* or the vulnerable, in our societies is left behind.

Text 6.75. Excerpt from Denmark (*woman** italicized; collocates underlined)

Women are key drivers to ensure sustainable development and to end poverty.

Text 6.76. Excerpt from the Bahamas (*woman** italicized; collocates underlined)

There can be no economic and social development without the achievement of gender equality and the empowerment of *women* and girls.

The remaining semantic domains ("extensiveness," "time," "other") are similar to the

collocational analyses in Sections 6.2 and 6.3, identified most likely due to their high frequency

in the UNGA overall and not to particular associations made with *woman**.

6.4.3. Collocational analysis for woman* *by country group*

6.4.3.1. Methods: Collocational analysis for woman* *by country group*

The next collocational analyses examined groups of countries to identify differences in

how they conceptualized *woman**. Countries were categorized as described in Section 3.6 and

collocational analysis was carried out for each group (e.g., Africa, Americas). Only content words were retained, and a minimum distribution of 50% of texts was set.

6.4.3.2. Results: Woman* *collocates by country group*

Several categories had collocates that met the 50% dispersion criteria. *Children* co-occurs with *woman** in the Americas (five times in four out of six texts), Asia (six times in six out of eleven texts), SIDS (eight times in seven out of thirteen texts), the high HDI group (eight times in seven out of fourteen texts), the small population group (eight times in seven out of thirteen texts), and the small area group (nine times in seven out of eleven texts). Text 6.77, from the small population state of Liechtenstein, illustrates how women and children are perceived as being vulnerable.

Text 6.77. Excerpt from Liechtenstein (*woman** italicized; collocates underlined)

> Preventing mass atrocities is so important because their effects are irreversible. They have no remedy. How can one possibly compensate the slaughter of civilians, the mass rape of *women*, the brutalizing of <u>children</u>?

The results of these country groups should be interpreted cautiously, however. First, *children* may appear as a collocate in so many groups due to overlap in a few of the categories. For instance, SIDS by definition have a small area and population. Moreover, the high frequency of the collocate *children* in these groups is likely due to the frequency of *children* in the UNGA overall, particularly in lists of disadvantaged populations: *In many parts of the world, women, children, youths, persons with disability, the elderly, people living in conflicts and emergency situations remain marginalized and untouched by the progress of development* [Indonesia].

Like *children, girls* is a highly frequent collocate in the corpus overall so its co-occurrence with *woman** in the medium GDP group (six times in four out of seven texts) is not

454

particularly surprising. Text 6.78 from Croatia shows *women* and *girls* being presented as vulnerable and Text 6.79 from Estonia refers both to their vulnerability and to their rights.

Text 6.78. Excerpt from Croatia (*woman** italicized; collocates underlined)

> As to Syria, Iraq, Libya, Yemen and the other crisis areas of the Middle East and North Africa, we share a grave concern about the continuous trend of growing violence, religious intolerance and terrorist actions. There we are confronted with unspeakable barbarism and terror with the most appalling forms of murder such as beheadings, sexual <u>violence</u> against *women* and <u>girls</u> and other most heinous crimes against humanity, culture, religion, dignity and against dearest of all - human life. It is shocking and sickening. And it has nothing to do with Islam, a religion of peace and understanding.

Text 6.79. Excerpt from Estonia (*woman** italicized; collocates underlined)

> This year marks the 20th anniversary of the Beijing Declaration and the Platform of Action that have influenced how we address equality and *women's* rights today. We need to continuously stand for the <u>rights</u> of *women* and <u>girls</u> and to strive for the elimination of <u>gender</u> based <u>violence</u>. *Women's* <u>rights</u>, and the <u>empowerment</u> of *women* and <u>girls</u> should be promoted at all levels.

More compelling is the relatively high frequency of *gender* as a collocate with *woman** in LLDC texts (eight times in five out of nine texts). *Gender* co-occurs with *woman** only 14 times for the entire UNGA, so well over half of those co-occurrences are in LLDC texts, though LLDC countries represent only 16% of the countries using *woman**. Five of the eight co-occurrences are in the phrase *gender equality* and are in passages on development (see Text 6.80).

Text 6.80. Excerpt from Zimbabwe (*woman** italicized; collocates underlined)

> The promotion of gender equality is critical for the realisation of the Post-2015 Development Agenda. The African Union designated 2015 "The <u>Year</u> of the <u>Advancement</u> of *Women* <u>Empowerment</u> and <u>Development</u> towards <u>Agenda</u> 2063" in order to mobilise our individual and collective actions around this <u>key</u> issue. <u>Gender equality</u> and *women* <u>empowerment</u> are central to the achievement of human development, progress and the elimination of the scourge of poverty and deprivation. We must continue

to build upon the achievements that have been realised since the 1995 Fourth <u>United Nations</u> <u>Conference</u> on *Women* and the <u>Beijing</u> Declaration and its Platform for Action.

Only two texts use *gender* in reference to vulnerability (see Text 6.81).

Text 6.81. Excerpt from Zambia (*woman** italicized; collocates underlined)

Special attention should be paid to <u>gender</u> and *women* <u>empowerment</u>. The 2013 Global Review Report indicates that 35% of *women* had been <u>victims</u> of <u>gender</u> based violence.

The only group with more than one collocate to meet the 50% distribution criterion is the medium HDI group with *gender* and *girls* seven times in five out of nine texts and *equality* five times in five texts. Fifteen out of these 19 co-occurrences are in passages on development (see Text 6.82 and 6.83).

Text 6.82. Excerpt from the Philippines (*woman** italicized; collocates underlined)

The CEDAW, the Beijing Platform for Action and the Sustainable Development Goals on women, particularly Goal 5 on achieving <u>gender equality</u> and <u>empowering</u> all *women* and <u>girls</u>. Taken together, these instruments are powerful mechanisms for realizing national and international commitments to <u>advance</u> *women's* roles as enablers of <u>sustainable</u> development.

Text 6.83. Excerpt from the Vanuatu (*woman** italicized; collocates underlined)

I am delighted to see that the focus on gender remains an integral part of the 2030 Agenda. In my country, achieving gender equality continues to be a challenge given our tradition and cultural make up. However, the political will to further <u>advance</u> the interests of *women* remains steadfast. Notable progress has been made in terms of access to <u>basic</u> education and <u>health</u> <u>services</u>, *women* assuming leadership <u>positions</u> in Municipal Councils and that *women* today have better access to basic financial services than ever before. As a leader of my country, I want to see more tangible progress in *women* and <u>girls'</u> <u>advancement</u> and their active participation in national leadership. This progress can be expedited if more financial resources are channelled to assist my Government to <u>advance</u> <u>gender equality</u> and *women's* <u>empowerment</u>.

The findings of the collocational analyses by country group show overall a focus on violence against women and women in development for several groups.

6.4.4. Keyword analysis for the Women group

6.4.4.1. Methods: Keyword analysis for the Women group

The keyword analysis compared the texts from all countries that had used *woman** (the Women group) with the texts from all countries that had not used the terms. The purpose was to identify any differences in the "aboutness" of the 58 texts mentioning *woman**. (See Table 6.15 for a list of countries in the Women group with frequencies for each type in the *woman** lemma.) Keywords were identified using WordSmith Tools 7.0 (Scott, 2017). Function words were removed from the list of collocates.

Table 6.15. *Woman** Frequency by Country in Women group

Country	*Woman*	*Women*	*Woman's*	*Women's*	**Total**
Afghanistan	0	2	0	0	2
Albania	0	2	0	0	2
Australia	0	7	0	1	8
Azerbaijan	0	1	0	0	1
Bahamas	0	2	0	1	3
Belize	0	2	0	0	2
Botswana	0	2	0	0	2
Brunei Darussalam	0	1	0	0	1
Bulgaria	1	0	0	1	2
Croatia	0	2	0	0	2
Denmark	0	2	0	1	3
Eritrea	0	1	0	0	1
Estonia	0	2	0	2	4
Finland	0	1	0	0	1
Ghana	0	5	0	0	5
Holy See	0	1	0	0	1
Iceland	0	4	0	0	4
Indonesia	0	1	0	0	1

Country	*Woman*	*Women*	*Woman's*	*Women's*	Total
Ireland	0	8	0	1	9
Jamaica	0	4	0	2	6
Jordan	0	1	0	0	1
Kenya	0	1	0	0	1
Kiribati	0	3	0	0	3
Liberia	1	6	0	2	9
Liechtenstein	0	4	0	0	4
Lithuania	0	1	0	0	1
Malawi	0	6	0	1	7
Maldives	0	3	0	0	3
Malta	0	0	0	1	1
Mauritius	0	3	0	0	3
Mongolia	0	2	0	1	3
Namibia	0	5	0	0	5
Nepal	0	3	0	0	3
Netherlands	0	4	0	0	4
Nigeria	0	1	0	0	1
Norway	0	4	0	0	4
Pakistan	0	1	0	0	1
Papua New Guinea	0	1	0	0	1
Philippines	0	11	0	2	13
Rwanda	0	1	0	0	1
San Marino	0	5	0	0	5
Seychelles	0	1	0	0	1
Sierra Leone	0	6	0	1	7
Slovakia	0	6	0	0	6
Slovenia	0	3	0	0	3
Solomon Islands	0	2	0	0	2
Somalia	0	3	0	0	3
South Africa	0	4	0	0	4
Suriname	0	3	0	0	3
Sweden	0	3	0	0	3
Tanzania	0	1	0	0	1
Trinidad and Tobago	0	6	0	0	6
Turkey	0	1	0	0	1
Tuvalu	1	1	0	0	2
US	1	3	0	0	4
Vanuatu	0	5	0	1	6

Country	*Woman*	*Women*	*Woman's*	*Women's*	Total
Zambia	0	5	0	0	5
Zimbabwe	0	3	0	0	3
Total	**4**	**172**	**0**	**18**	**194**

6.4.4.2. Woman* *keywords*

Eight keywords were identified for the Women group, including three highly frequent

collocates (*gender, girls, equality,* see Table 6.16).

Table 6.16. Keywords for Women Group

Keyword	Frequency	%	Texts	R.C. frequency	R.C. %	Log Likelihood	Log Ratio	*p*
women	*173*	*0.15*	*57*	*0*		*156.14*	*139.65*	*0.00*
Mauritius	50	0.04	2	0		45.13	137.86	0.00
gender	*72*	*0.06*	*32*	*5*		*38.10*	*3.04*	*0.00*
girls	38	0.03	24	0		34.30	137.46	0.00
Solomon	46	0.04	3	1		33.86	4.71	0.00
Africa	*95*	*0.08*	*35*	*16*	*0.02*	*26.60*	*1.76*	*0.00*
equality	*68*	*0.06*	*38*	*8*	*0.01*	*26.43*	*2.28*	*0.00*
Kenya	29	0.03	3	0		26.17	137.07	0.00

Note. R.C. = reference corpus (here, texts that do not mention *woman**). All function words have been removed. Keywords in at least 50% of texts are italicized.

The keywords were categorized into three semantic groups: places in the women group,

gender, and equality, as shown in Table 6.17.

Table 6.17. Keyword Categories for Women Group

Places in Women Group	**Gender**	**Equality**
Mauritius	*women*	*equality*
Solomon	*gender*	
Africa	girls	
Kenya		

Note. Keywords in at least 50% of texts are italicized.

The keywords *Mauritius, Solomon,* and *Kenya* occur in only two or three texts and are therefore not representative of the Women group as a whole. *Women* was one of the terms distinguishing the target corpus from the reference corpus, and *gender, girls,* and *equality* were identified as highly frequent collocates, so their appearance on the keyword list is not surprising. However, it does merit pointing out that texts that do not mention *woman** are statistically less likely to talk about *gender, girls*, and *equality.* In other words, texts are generally not about *gender, girls,* and *equality* unless they are also about *woman*.* Though texts do mention equality in reference to the equality of nations or to other social groups, such as indigenous peoples, texts that are about equality are more likely to also be about *woman*.*

The remaining keyword, *Africa*, raises two issues. First, a disproportionate number of African states mention *woman*.* Seventeen of the 58 texts to use *woman** are from Africa (29%), whereas only four of the 34 texts that do not use *woman** are from Africa (12%). Overall, African states seem to be concerned about the issue of *woman*,* with 81% mentioning at least one of the types in the lemma. In addition, non-African states that mention *woman** are more likely to mention *Africa*. Of the 25 non-African countries to mention *Africa,* 18 also use *woman** (72%) while seven do not use *woman** (28%). *Africa* is not, however, a collocate of *woman*.* In fact, *Africa* does not even appear in the broader co-text surrounding *woman*.* African texts often speak about *woman** in an international context: *Most of the world's poorest people are women* [Ghana]. Likewise, non-African texts do not discuss the situation of women in Africa in particular. However, the keyword analysis shows that texts discussing *woman** also discuss *Africa,* indicating a shared discourse for *woman** and *Africa*. This is one of the most compelling findings in the analysis of *woman*.* Associations with *woman** include equality and development (see Section 6.4.2), both of which are also relevant for *Africa*. A concern for human

460

rights or for development may be the reason countries speak about *woman** and *Africa,* even though the two are not directly linked.

6.4.5. Qualitative results of individual countries in the Women group

Sections 6.2 and 6.3 provide numerous examples of states focusing on national and regional interests. Often, countries fail to mention topics that are not a direct concern to them or, if they do mention those topics, they relate them to a more pressing national issue. The same can be said of several texts when it comes to the use of *woman*.*

While in some groups, such as Africa, nearly all countries speak about *woman** (see Section 6.4.4.2), in other groups, such as small states and Small Island Developing States (SIDS) in particular, many do not mention the topic: Small states make up 32% of the reference corpus (11 out of 34 of the non-Women texts) but only 19% of the Women group (11 out of 58); there are similar numbers for SIDS, which make up 38% of the reference corpus (13 out of 34) but only 22% of the Women group (13 out of 58). This is yet another example of countries speaking on topics of gravest national and regional import. Small states, and in particular SIDS, focus on the particular challenges of small states. This is not to say that women's rights are not relevant to SIDS; rather, SIDS have more pressing issues such as climate change.

The SIDS that do mention *woman** sometimes manage to address the topic in passages about climate change. Some of these references to *woman** emphasize the vulnerability of women (and other disadvantaged populations) to the effects of climate change: *recognition of impacts of climate change on* <u>women</u> *and youth, the elderly, the disabled, and the indigenous people and other vulnerable and marginalized groups, and acknowledge their contributions to effective implementation of the Paris outcome* [Papua New Guinea] and *We must hear the cries of despair of islanders affected by the effects of climate change. We must feel the pain of* <u>women</u>

461

struggling to keep their children alive [Seychelles]. Others are calls to include women in the

process of addressing climate change (see Text 6.84).

Text 6.84. Excerpt from Kiribati

We must call with urgency on our development partners, on philanthropy on private businesses to assist those on the frontline of the climate calamity to deal with the impacts of climate change and sea level rise now being experienced in our countries, and in our efforts to build the resilience and preparedness our people for an uncertain future. It is high time we recognize that the new challenges require that we draw on all the resources available to the global community, and accept that sustainable development and global challenges such as climate change should not be confined to the sphere of Governments only. Let us call on those with the ability to assist and who have a contribution to make, to join in the global dialogue and more importantly, join urgent action to address this major challenge. Let us bring in our youth, let us bring in our women, civil society, the private sector, churches, universities, our traditional institutions, indigenous populations and everyone on board. Let us be inclusive.

Other small states use the term *woman** when comparing equality between men and

women to equality between large and small states (see Texts 6.85 and 6.86). They relate gender

equality to a national concern, the political equality of their small countries within the

international community.

Text 6.85. Excerpt from the Maldives

Our faith in fundamental human rights is reaffirmed in principle; yet, the equal rights of men and women, and of nations, large and small, are ignored.

Text 6.86. Excerpt from Liberia

It was this same vision, which, seventy years ago, motivated the founding of this global organization to promote peaceful coexistence, economic and social advancement of all peoples, respect for human dignity and equal rights of men and women, and nations large and small.

Several countries speak about progress their countries have made in terms of women's rights. Noting national progress on matters of interest to the UN is a common communicative purpose, particularly for developing countries and countries either receiving UN assistance or under scrutiny by the UN Human Rights Council. Texts 6.87-6.89 exemplify discussions of progress made on the issue of gender equality. All three of these countries received millions (in the case of Afghanistan billions) of US dollars in Official Development Assistance in 2015 (World Bank, 2018).

Text 6.87. Excerpt from Namibia

> To continue my intervention on poverty eradication, I would like to add that one of the most impactful interventions we can make in the war against poverty is through empowering women, who although representing half of the world's population, account for nearly 70% of the world's poor. In Namibia, gender equality is not mere lip service. We have benefitted from the SWAPO Party's internal reform of its party list system, to include, 50/50 representation of women.

Text 6.88. Excerpt from Mauritius

> Mauritius particularly welcomes Sustainable Development Goal 5 relating to women and girls empowerment. Mauritius has made steady progress on this agenda and continues to put in place appropriate strategies to further promote gender equality in all the spheres of development. I am proud to announce that for the first time in history, my country has three women in high positions, as President of the Republic, Vice President and Speaker of the National Assembly of Mauritius. Mr. President, There can be no democracy without human rights, and no human rights without development.

Text 6.89. Excerpt from Afghanistan

> Life expectancy has increased by an average of 20 years since 2001, with Afghans living well beyond the mere 40 years of age that had once been their norm. Improvements in the health of women and children is particularly notable, as illustrated by the 54 percent increase in the number of infants delivered, and the decrease of infant mortality by 62 percent. Through the focus on gender equality, equity, and equal opportunity adapted by the Government, female political participation has also been strengthened. During the 2014 election, voters comprised of 35 percent females. Women now claim 11 percent of judgeships, with an additional 20 percent in training.

Finally, some countries use the issue of *woman** to raise other topics that are highly specific in terms of national interest. Trinidad and Tobago, for instance, speaks at length about arms control, *given that Trinidad and Tobago is located in a region heavily impacted by the trafficking of small arms and light weapons, as well as its attendant ills* [Trinidad and Tobago]. Trinidad and Tobago is the sole text to refer to the Resolution on Women, Disarmament, Non-proliferation and Arms Control, tying *woman** to this topic of key national interest (see Text 6.90).

Text 6.90. Excerpt from Trinidad and Tobago

> Additionally, recognizing the progressive contribution of women to peace and development, Trinidad and Tobago has introduced and has been the main sponsor of the United Nations General Assembly Resolution on 'Women, Disarmament, Non-proliferation and Arms Control' since 2010.

A number of countries involved in conflict do not refer to *woman** in their texts, but when they do, some speak of *woman** in relation to the conflict. Like with SIDS focusing on climate change, when a country dedicates most of its address to a conflict (e.g., Ukraine, Guyana), little time is left for other issues, but linking *woman** to the conflict offers them another opportunity to return to the primary focus of their address: the conflict. Text 6.91 from Pakistan integrates *woman** and in particular the vulnerability of women and children into a discussion of ceasefire violations by India.

Text 6.91. Excerpt from Pakistan

> When I assumed office of the Prime Minister of Pakistan in June 2013, for the third time, one of my first priorities was to normalize relations with India. I reached out to the Indian leadership to emphasize that our common enemy was poverty and underdevelopment. Cooperation, not confrontation, should define our relationship. Yet today ceasefire violations along the Line of Control and the Working Boundary are intensifying, causing civilian deaths including women and children. Wisdom dictates that our immediate

464

neighbor refrains from fomenting instability in Pakistan. The two countries should address and resolve the causes of tension and take all possible measures to avert further escalation.

Finally, Bulgaria uses the issue of women's rights to promote the election of a Secretary General from Eastern Europe. UN member states vie for representation in the highest echelons of the organization, from seats in the Security Council to representatives in the UN Secretariat, with the Secretary General being the most prestigious. Text 6.92 shows how Bulgaria uses the need for gender equality in the UN Secretary General position as grounds for a representative from Eastern Europe in particular.

Text 6.92. Excerpt from Bulgaria

Among the eight UN Secretary-Generals, all of them men, there has never been a national from the Group of Eastern European States. The recently adopted resolution 69/321 by the General Assembly stresses the need to ensure equal and fair distribution based on gender and geographical balance while meeting the highest possible requirements, on the appointment of the Secretary-General. Time has come that a woman from Eastern Europe be entrusted with the highest position in the UN Secretariat.

This qualitative analysis of individual texts has shown that, like with *terror** and *Security Council,* countries use the topic of *woman** to raise issues of vital national importance. Some of these concerns are reflected in the country categories to which they belong (e.g., *climate change* for SIDS), but others span across country groups (e.g., which countries are seeking elected office or seats within the UN).

6.4.6. Summary

The analysis of *woman** has shown that overall UNGA texts associate women (rather than men) with gender equality and rights, but also with vulnerability. *Man* and *men* are associated with equality almost exclusively when co-occurring with *woman**; *gender equality* is

for *men and women* or *women and girls*, but never for *men and boys*. Moreover, *man* and *men* are never associated with vulnerability unless they co-occur with *woman**. Victims of conflict may be *women and men* and they may be *women and children,* but they are never *men and children.*

Some country groups also have strong associations with *woman**. Most compelling of these findings is the relationship between *woman** and *gender equality* for LLDCs and the medium HDI countries because *gender equality* is placed against the backdrop of development. The LLDC and medium HDI groups are, by definition, less developed and it is not surprising that they would be most concerned about issues of development. Also noteworthy is the fact that SIDS discuss *woman** less frequently, but when they do, they tend to bring up climate change in the same co-text. When small states mention *woman*,* they often associate the fight for gender equality with the fight for equality among states, large and small. Again and again in Chapters 5 and 6, analyses have shown that countries focus on key areas of national interest or relate more common UN issues such as terrorism or women's rights to those national concerns.

6.5. Synthesis

Chapter 6 has examined three politically and socially loaded terms: *terror*, Security Council*, and *woman**. The analyses in Sections 6.2 through 6.4 have shown that UNGA countries tend to represent these ideas in similar ways, with several lexically diverse semantic domains emerging for each term in the examination of all UNGA texts together. In addition, the qualitative analyses have shown that countries often refer to issues of vital national interest in conjunction with these terms, lending a sense of urgency and gravity to their individual concerns. However, few systematic patterns emerged for groups of countries in these sections.

Methodologically, Chapter 6 has also shown the importance of establishing search criteria based on the research question posed. Setting strict distribution criteria for the analyses by country category limited the findings that could be reported. However, it would have been misleading to present findings from a small minority of countries as representative of the entire group. Not only would this have suggested associations between topics raised in UNGA addresses and country categories where none exist, but it would have also obscured one of the most important overall findings from this study, a finding that resurfaces throughout Chapters 4, 5, and 6: UNGA addresses are, on average very similar lexically and grammatically. The variation that does occur is only rarely a recognizable pattern for some country groups. Instead, variation is primarily identified at the level of individual texts. In a broad sense, these individual texts show another pattern for the UNGA: Countries have an agenda with certain key issues of national interest that they raise again and again. One of the ways they highlight these key national issues is to associate them with other topics of common concern (e.g., terrorism, women's rights), even though they may not necessarily seem related in the eyes of other member states or of the outside observer (e.g., environmental protection and the Russian occupation of Ukraine, the Security Council and climate change).

Chapter 6 offers an additional methodological contribution by showing the utility of undertaking CADS case studies to complement findings from a register analysis. The primary goal of the register analysis in Chapters 4 and 5 was to describe the UNGA as a register, examining the lexical and grammatical features that characterize the UNGA. The goal of the CADS studies in Chapter 6 was to explore three concepts. The objective was not to describe the UNGA as a register and yet these case studies added to the register analysis by showing that even politically and socially charged words are used in remarkably similar ways across UNGA

texts. The three words explored in these three case studies did not reveal any striking differences in the expression of political ideologies, identities, and relationships across UNGA addresses. Speakers often raise subjects such as *terrorism* to highlight other issues of national importance, but overall similarities in associations made with these words reveal common political and social ideologies (e.g., *gender equality* is an important goal) for most if not all UNGA texts.

The results of Chapters 4, 5, and 6 have important implications for the understanding of political discourse but also for corpus methodology. Chapter 7 discusses these implications and concludes with suggestions for applying these findings.

CHAPTER 7: SYNTHESIS AND CONCLUSION

7.1. Introduction

The fundamental goal of this dissertation has been to describe United Nations General Assembly addresses as a register, both commonalities as well as variation across texts. To accomplish this, I carried out lexical and grammatical analyses in search of linguistic patterns. For similarities across UNGA texts, I contrasted the UNGA to the British National Corpus, a 100-million-word collection of spoken and written English, as well as four comparison registers: prepared speeches, spontaneous speeches, conversation, and official documents. For variation within the register, I contrasted groups of texts from the UNGA that I had categorized based on geographic, economic, social, and political factors such as population size and Human Development Index. I interpreted linguistic patterns based on the situational characteristics of the UNGA and of the countries represented by the texts. I also carried out three CADS case studies to investigate any systematic patterns of variation in how three politically and socially charged words are conceptualized in UNGA texts.

The results of the keyword, collocational, key feature, and Multi-Dimensional analyses have revealed a register with relatively little lexical and grammatical variation. Overall, UNGA addresses use similar words and grammatical structures in similar ways. The variation that does occur can occasionally be attributed to groups of countries and their geographic, economic, social, and political situations. For the most part, though, few patterns of linguistic variation were identified based on group membership. Rather, when individual countries deviate from rhetorical conventions and adopt a unique discourse style, it is to highlight certain topics or achieve specific objectives of national concern.

The objective of Chapter 7 is to summarize and further discuss the findings presented in Chapters 4 through 6 (Section 7.2), to suggest implications (Section 7.3), and to consider future directions for further research (Section 7.4).

7.2. Synthesis

Section 7.2 synthesizes the findings from Chapters 4 through 6. Section 7.2.1 describes the UNGA and commonalities across the register and Section 7.2.2 variation within the UNGA. Section 7.2.3 reviews findings from the case studies and outlines the ways in which these findings add to our understanding of the UNGA and variation within the register. Finally, Section 7.2.4 discusses results from a methodological perspective, showing how methodological choices made by the researcher can affect the nature of the findings, with implications for studies in political discourse analysis and corpus linguistics.

7.2.1. UNGA addresses

UNGA addresses share numerous lexical and grammatical features and overall represent a register with relatively little linguistic variation. The lexical analysis categorized keywords into several distinct semantic domains. Some are topical, related to issues that fall within the mandate of the UN (e.g., peace, human rights, development, the environment). Other keywords stem from setting: words related to the functioning of the UN (e.g., *agenda, countries*) and discourse markers typical of formal UN addresses (*Mr. President*). The communicative purpose fostering solidarity and diplomacy explains the high frequency of some content and function words related to cooperation (*committed, our*). It also accounts for how some keywords are used. For instance, *must* is used nearly exclusively with a subject that includes the speaker (*we, the international community*) or with a non-human subject, often in the passive voice (*the ceasefire must be*

respected). These subjects make *must* sound like a shared objective rather than a forceful directive. Finally, production circumstances, the fact that addresses are carefully planned and edited, result in dense, information-packed nominal structures, leading to high frequencies of some function words (*the, of*).

Multi-dimensional scores, too, can be explained by these same situational characteristics. Production circumstances and communicative purpose result in low Dimension 1 scores, indicating highly informational language. The dense noun phrases with frequent use of prepositions and attributive adjectives are difficult to produce without careful planning and editing. In addition, one communicative purpose of UNGA addresses is to inform, particularly about national policy as it relates to UN topics (peace, development, the environment). This focus on current events and future goals leads to heavy nominal structures explaining relevance and relationships (e.g., *instability in our region, a political solution that will lay the ground for…,* Greece).

Communicative purpose gives rise to low Dimension 2 scores (non-narrative concerns): Texts are primarily expository. In the rare instances when a speaker narrates, it is framed within the context of the present situation and future goals and is a component of the speaker's argumentation strategy.

Setting and production circumstances account for Dimension 3 scores, which veer only slightly toward more explicit rather than situation-dependent reference. Given that these addresses are written to be spoken, they can include both explicit and situation-dependent reference. Explicit reference is more typical of formal written speech, where grammatical structures such as relative clauses elaborate, providing additional details. Some situation-

dependent reference (*here, now*) is also possible, however, because addresses are delivered in a shared time and place.

Low Dimension 4 scores indicate relatively little overt persuasion. This initially seems surprising given the communicative purpose of persuading, but the goal of fostering diplomacy prompts speakers to search for less direct forms of persuasion. Though necessity, possibility, and prediction modals are used for overt persuasion at times, less direct expression of stance is preferred in many texts. Stance nouns (*our commitment to*), verbs (*we call upon*), adjectives (*it is critical*), and adverbs (*unfortunately*) allow for both communicative functions, persuading and maintaining diplomacy.

High Dimension 5 scores reveal very abstract language, with passives, conjuncts, and past participle WHIZ deletions. Again, highly abstract language is difficult to produce in real time and is more common in formal texts that are planned and edited. The use of passives in particular stems from the communicative purpose of fostering harmonious relations. Speakers tend to shy away from criticizing countries directly by name in all but the most extreme cases. The passive allows speakers to avoid pointing fingers. The passive also emphasizes a goal (*sustainable development*) rather than who is accomplishing the goal, which is often unnecessary to specify (*the international community*).

Most of these Dimension scores had relatively low standard deviations, nearly always lower than the comparison corpora, indicating a relatively homogenous register with linguistic features shared across texts more uniformly than in the other corpora. The UNGA corpus was compiled with a view to limiting variation due to situational characteristics to the extent possible. Thus, it is not surprising that texts show many similar linguistic features. The variation that does occur is discussed in Section 7.2.2.

One of the most noteworthy findings in the analysis of MD scores was the overall dissimilarity between the UNGA and prepared speeches from Biber (1988). On average, UNGA texts resemble official documents more than prepared speeches along most dimensions (Dimensions 1, 2, 4, and 5). This suggests that the UNGA is a truly unique register in which the oral delivery is seemingly less important than for other scripted texts. Most UNGA speechwriters include few lexico-grammatical forms that would assist listeners with real-time processing. Instead, they pack as much information as possible into dense, abstract nominal structures. This begs the question: Are UNGA addresses truly "written to be spoken?" When presidents and prime ministers deliver these addresses, do they expect the audience to be listening to and comprehending all the densely packed information or is the oral delivery only for the sake of appearances? Do speechwriters and speakers actually expect the audience to read the text instead? This remains an empirical question worthy of future investigation (see Section 7.4).

7.2.2. Variation within the UNGA

The secondary goal, after determining what features are prevalent in the UNGA as a whole, was to detect systematic patterns of variation within the UNGA. In order to identify variation, texts had to be categorized into smaller groups for comparison purposes. This classification process was carried out based on geographic, economic, social, and political divisions commonly used by the United Nations (e.g., Small Island Developing States, SIDS) and by political scientists more generally (e.g., Gross Domestic Product).

The number of country categories examined was extensive (31 country groups in 10 categories), and yet few systematic patterns were detected. In terms of lexicon, keywords generally showed groups of countries speaking on topics of direct interest to them. For instance, Europe, facing huge numbers of Syrian refugees, uses *refugees* and *migration* with significantly

high frequency; the Least Developed Countries (LDCs) talk about the Sustainable Development Goals. Somewhat less obviously, the very high GDP and HDI groups speak with greater frequency on subjects related to conflict (*terror, Syria, refugees, nuclear*). Arguably, more developed countries focus on short-term but very visible problems while the less developed countries see poverty and disease as more insidious because they are long-term issues that generally receive less attention from the international community.

The most striking pattern in terms of grammatical structures is in level of development, with less developed countries tending to prefer a more formal, written style (more elaborated noun phrases, fewer contractions) and indirect expression of stance (fewer modals); in contrast, more developed countries opt for a less formal, more involved style and more overt persuasion. Lexically, a few keywords support these findings on formality. Groups of developing countries use rhetorical conventions typical of formal UN addresses much more frequently (e.g., using *Excellency* as a term of address).

Though some generalizations could be made on the linguistic features preferred or dispreferred by groups of countries, overall, few patterns were detected. To a large degree, this can be explained by the relative similarity in UNGA texts discussed in Section 7.2.1. However, some remarkable differences were noted in individual texts during the manual reading of addresses. To demonstrate some of those differences, five studies of individual countries were carried out.

The five individual texts revealed notable differences in terms of topic and discourse style. The results of the keyword analyses were not particularly surprising: Countries focus on topics of vital national interest. Ukraine talks at length about the Russian occupation, while Bhutan speaks about development and happiness. The grammatical analysis, though, revealed

some unexpected distinctions in terms of discourse style. One of the most interesting contrasts is the involved language of the United States with the informational style of Eritrea. The result is an enthusiastic message of solidarity and promise from the United States and a more matter-of-fact account of events, claims, and objectives from Eritrea. Another compelling finding is that countries use very different types of persuasion. The text from the United States is persuasive in its use of *can* to express possibility and promise. Ukraine relies primarily on evaluative adjectives, rhetorical questions, emphatics, and amplifiers. Tuvalu uses the most overt form of persuasion with the modal *must,* but most frequently in passives or with nominalizations as the subject (e.g., *the Paris Agreement*) to focus on the objective rather than the agent of the action. Finally, the use of passive voice in the texts from Eritrea, Tuvalu, and Bhutan shows how these constructions, with the agent either eliminated completely or given less emphasis in the middle or at the end of the sentence, allow a speaker to focus on an objective or event rather than who is expected to meet the objective or who is responsible for the event. This is particularly important to avoid a strong degree of imposition when using a modal of necessity (Tuvalu) and when placing blame on another member state or the international community (Eritrea). In terms of voice and blame, the contrast between the Eritrean text and the Ukrainian text is notable. Eritrea blames the international community for unfair sanctions but often uses the passive voice to mitigate the accusation. Ukraine, on the other hand, denounces Russia and its *illegal occupation* of Ukrainian territory. Because Ukraine shares the support of the majority of member states, it uses the active voice, placing direct blame on Russia. Thus, the analysis of these five individual texts illustrates in greater detail how linguistic forms accomplish different communicative purposes.

This analysis of individual texts adds greatly to the investigation of variation within the UNGA. The analysis of country categories showed relatively few distinctions by country group, and while that can be explained partly by similarities across texts in the UNGA, the individual analysis shows that there are indeed differences within the UNGA. However, country categories only sometimes capture the reasons for variation within the UNGA. Topic and communicative purpose can be highly individualized. Tuvalu's emphasis on development and climate change is, indeed, shared by most SIDS and these keywords were identified in the analysis by country group as well. In contrast, Eritrea's focus on UN sanctions against the country and the US focus on the preeminence of people over political systems are highly specific. Overall, the analysis in Chapter 5 shows relatively few systematic patterns of variation. Rather, individual countries tend to have specific national agendas that they promote in their addresses.

7.2.3. Discourse analysis case studies

The final chapter to examine linguistic patterns across and within the UNGA was Chapter 6, with discourse analysis case studies of three pre-selected terms. The objective was to determine whether a different type of analysis might reveal additional information about the UNGA. The register analyses in Chapters 4 and 5 were bottom-up: Lexical and grammatical structures were investigated based on frequency. This meant that some structures I thought might be particularly informative were not analyzed due to low frequency and other structures that I had not anticipated to be potentially revelatory were discussed in detail. The objective of these chapters was to describe the UNGA as a whole. The case studies in Chapter 6, on the other hand, were designed with a view to exploring topics I expected to be noteworthy. The objective of these case studies was to determine whether UNGA countries express ideological differences through their use of these three words, thus adding an additional dimension to my study of the

476

UNGA register. All three search terms were selected based on the initial keyword analysis in Chapter 4, where *terrorism, security*, and *gender* appeared. After some manual examination of passages in which these words occurred, the words were modified to the search terms *terror*, (Security) Council,* and *woman*.* My expectation was that these three terms would produce different findings: that *terror** would be defined differently but denounced by all countries, that comments on the *(Security) Council* would diverge in large part based on past and present membership on the Council, and that the conceptualization of *woman** would show the most similarities, with all countries clamoring for gender equality.

Instead, I found that all three terms by and large supported the findings in Chapters 4 and 5: UNGA addresses produced relatively uniform results; the variation that does occur is primarily at the level of the individual text rather than based on country category. Collocational analyses for the UNGA as a whole showed very strong semantic categories; some semantic domains comprise over 30 collocate types, and a great many types occur twenty or more times in over a dozen texts. In contrast, very few patterns emerged by country category. Instead, a manual examination of individual texts revealed that countries raise these issues in order to highlight matters of particular importance to them. Some of the associations made were not surprising, such as when a country that has never served on the Security Council calls for reform of Council membership. Others were initially unexpected but could be explained by a country's situational characteristics, as when a SIDS emphasizes the need for the Security Council to recognize climate change as a security issue. While many associations are shared by several countries within a group, for most, the numbers were not large enough to confirm any pattern distinguishing one group of countries from another. These discourse analysis case studies showed that countries express little variation in ideology through their use of these three words.

Even politically and socially loaded terms are used relatively consistently across UNGA addresses.

7.2.4. Methodology

Chapters 4-6 have raised important methodological considerations. In this section, I discuss the limitations of keyword and collocational analyses as well as the importance of corpus size. I then point out the ways in which register analysis and CADS are complementary and the benefits of undertaking both types of studies for the same collection of texts.

Keyword analysis identifies the "aboutness" of a text. For instance, keywords tell the researcher that one text or group of texts makes significantly more references to the *Security Council* than another. However, it does not indicate what the texts say about the *Security Council*, and what a text says about the *Security Council* is often more important than the simple fact that the term is mentioned. When, for instance, Slovakia names the Security Council only in reference to *Security Council Resolution 1325 on Women, Peace and Security* in a discussion on gender equality, the discourses surrounding the *Security Council* cannot be said to be similar to those in the South African text, which states that it is *unjustifiable* that no African country has a permanent member seat on the Council.

Thus, keyword analysis alone gives very limited information. Collocational analysis, on the other hand, should reveal what countries say on a topic because it identifies the words surrounding a search term. Yet collocational analysis, too, has limitations. Generally, collocations are defined as being within five words of the search term, but software settings can be changed to include fewer or more words which will, in turn, modify results. Baker (2006) warns against the temptation to experiment with different spans in search of the most interesting findings (p. 103), so most researchers abide by the five-word span that is the default setting for

software such as WordSmith (Scott, 2017). The five-word span sometimes misses important associations, but increasing to even just a ten-word span can often identify words that are already on a different topic. The researcher is faced with the dilemma of potentially either omitting associations or mistakenly identifying incorrect associations.

Another problem with both keyword and collocational analyses is the use of different grammatical forms, synonyms, and alternate phrasings. A search term can be lemmatized (e.g., *contribut**) to include all forms: *contribute, contributes, contribution, contributions, contributive.* However, keyword and collocational analyses do not automatically detect such related words. Similarly, when synonyms or near synonyms abound (e.g., *violence, aggression, conflict, brutality, occupation, breaching the peace*), it can be difficult to identify the same word in multiple texts. Identifying semantic categories is a way to get around the problem of synonyms and alternate phrasings. The researcher classifies words with similar meanings as belonging to the same semantic category. Unfortunately, two problems remain. First, each of the different types must occur with sufficient frequency to be detected by the software. If each type is used only once or twice, it might not be identified by the software program being used. A type can only make it to the list of semantic categories when it is used sufficiently frequently to be identified as a keyword or collocate in the first place. Second, when distribution criteria are set, requiring that a keyword or collocate be used in a percentage of the texts, there may not be a sufficient number of occurrences to meet the cut-off.

This leads to another methodological consideration: setting a cut-off. This factor was shown to be decisive in the first case study (Sections 6.2.3.2.3 and 6.2.3.2.4). When distribution criteria are set, some patterns will not emerge because they do not occur in a sufficient number of

texts. In contrast, when distribution criteria are not set, patterns may be identified and attributed to a group of texts when in fact they arise only in a sub-group or simply by chance.

Another methodological consideration is in determining the requisite size of the corpus. The number of words needed depends to a great extent on the type of language and the research questions being investigated (see Baker, 2006, pp. 27-31). When a corpus is highly specialized, it need not be as large. Though it is true that a smaller corpus may not reveal words and grammatical structures that might be present in a larger body of texts, if a corpus is large enough to be representative of a register, the size should not prevent the researcher from identifying truly systematic patterns in prevalent linguistic features. Another advantage of a smaller corpus is that the manageable size allows the researcher to examine each text in detail, revealing patterns that do not emerge quantitatively.

For the present study specifically, the size of the UNGA corpus does not appear to be limiting results. The corpus includes all texts delivered at the UNGA in English in 2015, so it is representative of a very specialized register. Nor does the size of the sub-corpora examined (even the smallest group: high GDP with only four countries) appear to restrict findings: Chapters 5 and 6 identified a number of lexical and grammatical features for smaller groups representing fewer words but not necessarily the larger groups with more words.

However, the fact that the UNGA corpus does not represent all addresses delivered at the UNGA but only those delivered in English is an important consideration in the interpretation of findings. As discussed in Section 3.3.2.1, just over half the UNGA addresses were not delivered in English and were therefore not analyzed in this study. Some regions are particularly under-represented (e.g., Northern Africa, South America). Whether results would have been different had the analysis included all addresses delivered in all languages remains an empirical question,

but there is reason to believe differences would have been minimal. This study has shown that countries from every continent, with different levels of development, and with different political and social priorities deliver texts that are surprisingly similar. So while it is possible that the English translations of texts delivered in Arabic, Chinese, French, Russian and Spanish might exhibit different lexical and grammatical features, it is not particularly likely, and if they do, differences could potentially be due to the critical moderating variable of translation rather than due to characteristics of the country represented.

Methodological considerations such as operationalization of a collocate, distribution across texts, and corpus size have been discussed at some length in the literature (see Baker, 2004, 2006). One methodological innovation in this dissertation is rarely considered in corpus studies: the complementarity of register analysis and CADS.

Register analysis describes the linguistic features of a register using a bottom-up approach. All unusually frequent and sometimes infrequent lexical and grammatical features are investigated and their use is interpreted based on situational characteristics. In contrast, CADS investigates politically or socially charged words. The primary objective is to explore these concepts. However, CADS also reveals a great deal about language use in the texts analyzed. As a result, the CADS investigation of political and social ideology, identity, and relationships adds to the register analysis.

7.3. Implications for Political Discourse Analysis

The findings synthesized in Section 7.2 have important implications for political discourse analysis. Previous studies of international organization discourse have predominantly offered qualitative findings of written texts. Some research has been carried out on speeches, but to date no extensive corpus study has provided substantial quantitative data on a vast array of

481

linguistic features in speeches delivered at an international organization. The present study fills that gap by describing in detail the lexical and grammatical characteristics of addresses delivered in English at the United Nations General Assembly in 2015. The analysis has brought to light both similarities as well as differences across texts.

The present investigation of UNGA addresses supports with comprehensive quantitative data some findings that had been uncovered in previous studies. Like other scripted political texts (see Section 2.3.3), UNGA addresses use some modals and other features of overt persuasion but primarily less direct forms of persuasion. When modals are used, they are often in the passive voice or with subjects that include the speaker to mitigate their directive force and imply a shared objective rather than an order. Common goals and solidarity are also implied with the highly frequent use of *we* and *our,* collocating extensively with words related to cooperation and support. Other forms of indirect persuasion include expression of stance with the first person and a mental verb as well as with factive verbs, suggesting an objective reality rather than the speaker's personal opinion. In addition, the passive voice and nominalization are used throughout UNGA texts for greater neutrality and to maintain diplomacy.

Though the focus of this study has been on linguistic features, certain discourse conventions were also identified, such as congratulating the president of the proceedings. Donahue and Prosser (1997) also comment on this rhetorical pattern from their collection of UNGA addresses from 1945 to 1995. Interestingly, in the UNGA of 2015, many lexical signals of such formal discourse conventions (e.g., *Mr. President, Excellencies, Heads of state*) are significantly more frequent in texts from developing countries. Unfortunately, without any quantitative data on this information in Donahue and Prosser, it is unclear as to whether these

conventions are now disappearing or whether they have always been more frequent in the texts from particular groups of UN member states.

The importance of offering additional quantitative support for the findings from previous studies is not trivial, but the present study contributes much more to the current understanding of international organization discourse and UNGA addresses in particular. Specifically, the analyses have shown that member states tend to raise the same issues (e.g., development, peace and security, the environment) and say similar things about them. Any differences in the associations they make with key terms stem from countries relating issues of more general UN interest to a topic of vital national importance (e.g., how the environment in Ukraine has been impacted by the Russian occupation).

The present study has also yielded a great deal of information about discourse style and the grammatical structures preferred in UNGA addresses. Texts overall tend to be highly informational, with elaborated noun phrases that allow a great deal of information-packing. On average, UNGA speakers appear unconcerned about the difficulty for the listener of processing such dense text. Unlike other prepared monologues and even other scripted texts (see Sections 2.3.1 and 2.3.2), speakers at the UNGA rarely use signposting or other features to aid the listener in processing dense text in real time. Manual examination of several texts showed how these nominal structures often express opinion indirectly, with stance adjectives and nouns conveying evaluation. Stance verb complement clauses are also used as a persuasive device that is less overt than the use of modals. When modals are used, they are chosen carefully. As discussed in the UNGA keyword analysis (Section 4.4.5.2), *must* collocates most frequently with *we, United Nations,* and *international community.* This once again highlights the emphasis on shared commitment and joint efforts.

As well as offering specific details about the linguistic features of UNGA addresses, this study has highlighted some critical methodological considerations for undertaking corpus studies of political texts. It is hoped that future research in the field of political discourse analysis will continue to draw on corpus methods, report clearly the methods used, and make claims that are reflective of those methods and their inherent limitations.

7.4. Future Directions

The present study has raised a number of compelling avenues for future research, from methodological studies to diachronic studies to cross-linguistic and translation studies. This dissertation has also shown the UNGA to be a unique type of scripted register, revealing the need to further investigate scripted texts.

Methodological studies could examine in greater detail the importance of the choices researchers make when determining, for instance, what constitutes a collocate. Experimenting with different spans, from five words to fifteen, would reveal whether differences are minimal or, on the contrary, critical to the types of results that can be reported. Methodological studies could also continue to explore the ways in which register analysis and CADS complement each other.

Diachronic studies evaluate changes (and similarities) over time. A study of UNGA addresses every year from 2005-2015 could detect recurring patterns as well as variation. Of particular interest is the effect of the UNGA theme on the keywords for that year. An analysis of texts from a relatively short span of time such as a decade is likely to reveal more in terms of lexical features, but it could also help to support (or, on the contrary, repudiate) findings on grammatical structure preferences and dispreferences. In addition, a comparison of texts separated by a larger span of time (e.g., 1965 compared to 2015 or 1955-1965 compared to 2005-

2015) is likely to reveal not only lexical and grammatical differences (e.g., what associations were made with *women* in 1965), but rhetorical patterns as well. A study of the use of *Mr. President* or terms for complimenting the President of the Assembly (*the excellent credentials you bring, you will guide us judiciously*), for instance, could uncover changes or instead continuity in certain discourse conventions.

A translation or cross-linguistic study could investigate the effects of translation on lexico-grammatical features. For a translation study, it might be tempting to try breaking translated texts down into five separate categories, one for each of the five other official UN languages, in order to tease out differences based on the original language of the text, but this would be challenging given the relatively small number of texts delivered in Russian and Chinese. More judicious would be to combine all texts delivered in other languages and contrast their English translations with texts delivered in English.

In addition, a cross-linguistic study comparing the original French, for instance, with the translation in English could elucidate language-specific influence. Of particular interest would be expression of stance. Is the preference for less overt expression of persuasion in the English original respected in a translation from French? For instance, the French *il faut* uses an impersonal subject that has been translated by European and Canadian Parliament translators with great variation (*there is a need, I just want to make it clear though, we must, is required, there has to be, needs to, is to…,* Linguee, n.d.) Are there any patterns in the translations preferred for *il faut* based on topic, country, or collocates?

Once the effect of translation on the linguistic features of UNGA addresses is better understood, it would be possible to examine all UNGA texts regardless of language in which they were delivered. As discussed in Sections 3.3.2.1 and 3.6.2, a number of important countries

485

and sub-regions are not represented or are under-represented in the corpus used for this dissertation. Adding the countries that delivered addresses in other languages would make findings more robust and perhaps elucidate additional patterns of variation based on country group if the effect of translation can be fully accounted for.

Of substantial interest is the utility of corpus methods for translators and interpreters. The present study has revealed a great deal about preferred terms and collocations as well as preferred (and dispreferred) grammatical structures. The results help further our understanding of political discourse. Can this information also be useful for professionals who are not trained politicians but who are trained to write and sound like politicians? Interpreters, for instance, often prepare for conferences by pouring over dozens of speeches delivered previously either by the politicians they will be interpreting or at the venue at which they will be working. They compile glossaries of important terms based on their painstaking manual examination of the texts. Could automatized keyword and collocational analyses using software such as WordSmith (Scott, 2017) improve their efficiency and/or the usefulness of the terms identified? A study could, for instance, compare the glossaries amassed automatically with computerized software to glossaries compiled manually by interpreters. Do the glossaries differ greatly? Is it truly faster to compile a glossary from a computerized search given that results must then be sorted for ease of use during the conference? Finally, do interpretations differ significantly when interpreters use glossaries they have compiled after a manual reading from when they use glossaries compiled automatically? Such investigations are compelling because they are potential opportunities for corpus linguistics to become truly *applied,* conceivably improving the efficiency and perhaps even reliability of the work of translators and interpreters.

Finally, this study has shown a marked difference between UNGA addresses and the prepared speeches in Biber (1988). The differences between scripted international organization speeches and other scripted speeches merit further study. The extremely abstract, dense prose of UNGA texts suggests they were written to be read rather than heard. Interviews with speechwriters have been carried out in previous studies of international organization discourse (e.g., Waheed, Schuck, Neijens, & de Vreese, 2013) and could be useful in elucidating the reasons for dense information packing in UNGA addresses.

7.5. Conclusion

Chapter 7 has presented a summary of findings and their implications, including methodological considerations for future research. It has also shown just a few examples of additional research that can be carried out to add to the current understanding of political discourse and to apply that knowledge to other fields such as translation and interpretation. This study of UNGA addresses has furthered the current understanding of international organization speeches and shown the substantial interest of pursuing future corpus-assisted research in the field of political discourse analysis.

References

African Union. (2005). *The common African position on the proposed reform of the United Nations: "The Ezulwini consensus."* Retrieved from http://www.un.org/en/africa/osaa/pdf/au/cap_screform_2005.pdf.

Archer, D., Aijmer, K., & Wichmann, A. (2012). *Pragmatics: An advanced resource book for students*. Oxon, England: Routledge Applied Linguistics.

Baer, R. (2010). Foreword: An ex-CIA perspective. In R. Erlich (Ed.), *Conversations with terrorists* (pp. vii-xii). Sausalito, CA: PoliPointPress.

Baker, P. (2004). Querying keywords: Questions of difference, frequency and sense in keywords analysis. *Journal of English Linguistics, 32*(4), 346-359.

Baker, P. (2005). *Public discourses of gay men.* London, England: Routledge.

Baker, P. (2006). *Using corpora in discourse analysis.* London, England: Continuum.

Baker, P., & Levon, E. (2015). Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. *Discourse & Communication, 9*(2), 221-236.

Baker, P., & McEnery, T. (2005). A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of Language and Politics, 4*(2), 197-226.

Baturo, A., Dasandi, N., & Mikhaylov, S. (2017). Understanding state preferences with text as data: Introducing the UN General Debate corpus. *Research & Politics, 4*(2), 1-9.

Becker, H. S. (1967). Who's side are we on? *Social Problems, 14*(3), 239-247.

Bednarek, M. (2010). *The language of fictional television drama and identity*. New York, NY: Continuum International.

Bhatia, A. (2009). The discourses of terrorism. *Journal of Pragmatics*, *41*(2), 279-289.

Biber, D. (1988). *Variation across speech and writing*. Cambridge, England: Cambridge University Press.

Biber, D. (1994). An analytical framework for register studies. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 31-56). Oxford, England: Oxford University Press.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* Amsterdam, Netherlands: John Benjamins.

Biber (2009). Are there linguistic consequences of literacy? Comparing the potentials of language use in speech and writing. In D. R. Olson & N. Torrrance (Eds.), *The Cambridge handbook of literacy* (pp. 75-91). Cambridge, England: Cambridge University Press.

Biber, D., & Conrad, S. (2009). *Register, genre, and style.* Cambridge, England: Cambridge University Press.

Biber, D. & Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use.* Cambridge, England: Cambridge University Press.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multi-dimensional comparison. *TESOL Quarterly, 36*, 9-48.

Biber, D., and Egbert, J. (in press). *Register variation online.* Cambridge, England: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* London, England: Longman.

Biber et al. (2004). *Representing language use in the university: Analysis of the TOEFL 2000 spoken and written academic language corpus.* Princeton, NJ: ETS/TOEFL.

Bruner, M. L. (2002). Taming 'wild' capitalism. *Discourse & Society, 13*(2), 167-184.

Camiciottoli, B. (2007). *The language of business studies lectures a corpus-assisted analysis*. Amsterdam, Netherlands: John Benjamins.

Cap, P. (2015). Follow-ups in the US anti-terrorist discourse: Proposal for a macro-discursive approach to monologic follow-up sequences. *Discourse & Society 26*(5), 543-561.

Central Intelligence Agency (2016). *The World Factbook*. Retrieved from https://www.cia.gov/library/publications/resources/the-world-factbook.

Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-53). Norwood, NJ: Ablex.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York, NY: Academic Press.

Collet, T. (2009). Civilization and civilized in post-9/11 US presidential speeches. *Discourse & Society*, *20*(4), 455-475.

Cutting, J. (2011). Spoken discourse. In K. Hyland & B. Paltridge (Eds.), *Bloomsbury companion to discourse analysis* (pp. 155-170). London, England: Continuum.

D'Acquisto, G. (2017). *A linguistic analysis of diplomatic discourse: UN resolutions on the question of Palestine*. Cambridge, England: Cambridge Scholars Publishing.

Donahue, R. T. & Prosser, M. H. (1997). *Diplomatic discourse: International conflict at the United Nations - Addresses and analysis*. Greenwich, CT: Ablex Publishing Corporation.

Duchêne, A. (2008). *Ideologies across nations: The construction of linguistic minorities at the United Nations*. Berlin, Germany: Mouton de Gruyter.

Faini, R., Annez, P., & Taylor, L. (1984). Defense spending, economic structure, and growth: Evidence among countries and over time. *Economic Development and Cultural Change, 32*(3), 487-498.

Fairclough, N. (1989). *Language and power.* London, England: Longman.

Fairclough, N. (2002). Language in new capitalism. *Discourse & Society, 13*(2), 163-166.

Ferrarotti, L. (2009). The news presenter and the television audience: A comparative perspective of the use of *we* and *you.* In L. Haarman & L. Lombardo (Eds.), *Evaluation and stance in war news: A linguistic analysis of American, British and Italian television news reporting of the 2003 Iraqi war* (pp. 72-96). London, England: Continuum International.

Forchini, P. (2012). *Movie language revisited: Evidence from multi-dimensional analysis and corpora.* Bern, Switzerland: Peter Lang.

Fortanet, I. (2005). The use of 'we' in university lectures: Reference and function. *English for Specific Purposes, 23*(1), 45-66.

Friginal, E. (2009). *The language of outsourced call centers: A corpus-based study of cross-cultural interaction.* Amsterdam, Netherlands: John Benjamins.

Government Publishing Office. (2008). *H. Res. 1249.* Retrieved from https://www.congress.gov/110/bills/hres1249/BILLS-110hres1249ih.pdf.

Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. Amsterdam, Netherlands: John Benjamins.

Gregori-Signes, C. (2017). "Apparently, women don't know how to operate doors": A corpus-based analysis of women stereotypes in the TV series *3rd Rock from the Sun. International Journal of English Studies, 17*(2), 21-43.

Halmari, H. (2005). In search of "successful" political persuasion: A comparison of the styles of Bill Clinton and Ronald Reagan. In H. Halmari, & T. Virtanen (Eds.), *Persuasion across genres: A linguistic approach* (pp. 105-134)*. Amsterdam, Netherlands: John Benjamins.

Holmes, J. (2000). Politeness, power and provocation: How humour functions in the workplace. *Discourse studies*, *2*(2), 159-185.

Institute for Economics and Peace. (2016). *2016 global terrorism index*. Retrieved from http://economicsandpeace.org/wp-content/uploads/2016/11/Global-Terrorism-Index-2016.2.pdf.

Johansson, S., Leech, G. N., & Goodluck, H. (1978). *Manual of information to accompany the Lancaster-Oslo-Bergen Corpus of British English, for use with digital computers*. Department of English, University of Oslo.

Kim, K. H. (2014). Examining US news media discourses about North Korea: A corpus-based critical discourse analysis. *Discourse & Society, 25*(2), 221-244.

Kuczok, M. (2014). *The conceptualisation of the Christian life in John Henry Newman's parochial and plain sermons*. Newcastle upon Tyne, England: Cambridge Scholars.

Krijtenburg, F., & de Volder, E. (2015). How universal is UN 'peace'? A comparative linguistic analysis of the United Nations and Giryama (Kenya) concepts of 'peace'. *International Journal of Language and Culture, 2*(2), 194-218.

Kumaravadivelu, B. (1999). Critical classroom discourse analysis. *TESOL quarterly*, *33*(3), 453-484.

Lazar, M. M. (2005). *Feminist and critical discourse analysis: Gender, power and ideology in discourse*. Basingstoke, England: Palgrave Macmillan.

L'Hôte, E. (2009). New Labour and globalization: Globalist discourse with a twist? *Discourse & Society, 21*(4), 355-376.

Linguee. (n.d.). *Il faut.* Retrieved from https://www.linguee.com/french-english/translation/il+faut.html on January 9, 2018.

Lombardo, L. (2009). Positioning and stance in TV news reporting of the 2003 Iraq War: The anchor on CBS and the news presenter on BBC. In J. Morley & P. Bayley (Eds.), *Corpus-assisted discourse studies on the Iraq conflict: Wording the war* (pp. 141-164). New York, NY: Routledge.

McCarthy, P. A. (1998). *Hierarchy and flexibility in world politics: Adaptation to shifting power distributions in the United Nations Security Council and the International Monetary Fund.* Brookfield, VT: Ashgate Publishing Company.

McEntee-Atalianis, L. J. (2013). Stance and metaphor: Mapping changing representations of (organizational) identity. *Discourse & Communication, 7*(3), 319-340.

Miller, D. R., & Turci, M. (2006). Construing the 'social gospel' of Martin Luther King Jr.: A corpus assisted study of *free**. *Linguistics and the Human Sciences, 2*(3), 399-424.

Montgomery, M. (2007). *The discourse of broadcast news: A linguistic approach*. London, England: Routledge.

Morley, J., & Bayley, P. (Eds.). (2009). *Corpus-assisted discourse studies on the Iraq conflict*. New York, NY: Routledge.

Mulderrig, J. (2012). The hegemony of inclusion: A corpus-based critical discourse analysis of deixis in education policy. *Discourse & Society, 23*(6), 701-728.

Najarzadegan, S., Dabaghi, A., & Eslami-Rasekh, A. (2017). A critical discourse analysis of Iran and US presidential speeches at the UN: The sociopragmatic functions. *Theory and Practice in Language Studies, 7*(9), 764-774.

Oddo, J. (2011). War legitimation discourse: Representing 'us' and 'them' in four US presidential addresses. *Discourse & Society, 22*(3), 287-314.

Organisation for Economic Co-operation and Development. (2018). *Receipts for Sierra Leone*. Retrieved from https://public.tableau.com/views/OECDDACAidataglancebyrecipient_new/Recipients?:embed=y&:display_count=yes&:showTabs=y&:toolbar=no?&:showVizHome=no.

Otieno, T. M. (2015). Linguistic description of the rhetoric: Barack Obama's political speeches. *International Organization of Scientific Research Journal of Humanities and Social Science 20*(9), 80-92.

Pardo, M. L. (2001). Linguistic persuasion as an essential political factor in current democracies: Critical analysis of the globalization discourse in Argentina at the turn and at the end of the century. *Discourse & Society, 12*(1), 91-118.

Partington, A., Duguid, A., & Taylor, C. (2013) *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies*. Amsterdam, Netherlands: John Benjamins.

Partington, A., Morley, J., & Haarman, L. (2004). *Corpora and discourse*. Bern, Switzerland: Peter Lang.

Pew Research Center (2012). Table: Religious composition by country, in percentages. Retrieved from http://www.pewforum.org/2012/12/18/table-religious-composition-by-country-in-percentages/.

Puchala, D. J. (2005). World hegemony and the United Nations. *International Studies Review, 7*, 571-584.

Pujante, D. & Morales-López, E. (2008). A political action against popular opinion: Aznar's final speech before the Spanish Parliament justifying the war in Iraq (December 2003). *Journal of Language and Politics, 7*(1), 71-98.

Quaglio, P. (2009). *Television dialogue: The sitcom* Friends *vs. natural conversation*. Amsterdam, Netherlands: John Benjamins.

Reyes, A. (2011). Strategies of legitimization in political discourse: From words to actions. *Discourse & Society, 22*(6), 781-807.

Salama, A. H. Y. (2011). Ideological collocation and the recontextualization of Wahhabi-Saudi Islam post-9/11: A synergy of corpus linguistics and critical discourse analysis. *Discourse & Society, 22*(3), 315-342.

Sardinha, T. B., & Pinto, M. V. (2017). American television and off-screen registers: A corpus-based comparison. *Corpora*, *12*(1), 85-114.

Scott, M. (n.d.). *WordSmith tools manual*. Retrieved from http://www.lexically.net/downloads/version7/HTML/keywords_calculate_info.html.

Scott, M. (2017). WordSmith Tools version 7. Stroud, England: Lexical Analysis Software.

Sowinska, A. (2013). A critical discourse approach to the analysis of values in political discourse: The example of freedom in President Bush's State of the Union addresses (2001-2008). *Discourse & Society, 24*(6), 792-809.

Staples, S. (2015a). *The discourse of nurse-patient interactions: Contrasting the communicative styles of US and international nurses*. Amsterdam, Netherlands: John Benjamins.

Staples, S. (2015b). Spoken discourse. In D. Biber & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 271-291). Cambridge, England: Cambridge University Press.

Sullivan, M. P. (1990). *Power in contemporary international politics.* Columbia, SC: University of South Carolina Press.

Svartvik, J. & Quirk, R. (1980). *A corpus of English conversation.* Lund, England: CWK Gleerup.

UN Web TV. (2015). United Nations – Year in Review 2015. Retrieved from http://webtv.un.org/watch/united-nations-year-in-review-2015/4671606008001.

United Nations. (n.d.) *Global issues overview.* Retrieved from http://www.un.org/en/sections/issues-depth/global-issues-overview/.

United Nations Development Programme. (n.d.). *Human development statistical tables.* Retrieved from http://hdr.undp.org/en/data.

United Nations General Assembly. (n.d.). *General debate of the 70ᵗʰ session.* Retrieved from https://gadebate.un.org/.

United Nations General Assembly. (2014). *Resolution adopted by the General Assembly on 27 March 2014.* Retrieved from http://www.securitycouncilreport.org/atf/.

United Nations General Assembly. (2015). *Transforming our world: The 2030 agenda for sustainable development.* Retrieved from http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

United Nations High Commissioner for Refugees (n.d.). *Population statistics: Persons of concern.* Retrieved from http://popstats.unhcr.org/en/persons_of_concern.

United Nations High Commissioner for Refugees. (2016). *Global trends: Forced displacement in 2015.* Retrieved from http://www.unhcr.org/576408cd7.pdf.

United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States. (2018). *About the Landlocked Developing Countries (LLDCs).* Retrieved from http://unohrlls.org/about-lldcs/.

United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States. (2018). *About LDCs.* Retrieved from http://unohrlls.org/about-ldcs/.

United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States. (2018). *About the Small Island Developing States.* Retrieved from http://unohrlls.org/about-sids/.

United Nations Security Council. (n.d.) *The Security Council.* Retrieved from http://www.un.org/en/sc/.

United Nations Security Council. (2018). *Countries never elected members of the Security Council.* Retrieved from http://www.un.org/en/sc/members/notelected.asp.

United Nations Security Council Report. (2018). *UN documents for Eritrea.* Retrieved from http://www.securitycouncilreport.org/un-documents/eritrea/.

United Nations Statistics Division. (2016). *Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings.* Retrieved from http://unstats.un.org/unsd/methods/m49/m49regin.htm.

United States Agency for International Development. (2013). Violent extremism and insurgency in Nepal: A risk assessment. Retrieved from http://pdf.usaid.gov/pdf_docs/pa00jsjq.pdf.

Valentine, E. & Preston, D. (2002). Discourse, devices and diplomacy: Reporting in international settings. *Across Languages and Cultures, 3*(2), 227-238.

van Dijk, T. A. (1988). *News as discourse.* Hillsdale, NJ: Lawrence Erlbaum Associates.

van Dijk, T. A. (1993). *Elite discourses and racism.* Newbury Park, CA: Sage.

van Dijk, T. A. (1997). What is political discourse analysis? In J. Blommaert & C. Bulcaen (Eds.), *Political linguistics* (pp. 11-52). Amsterdam, Netherlands: John Benjamins.

van Dijk, T. A. (2006). Discourse, context and cognition. *Discourse & Society, 8*(1), 159-177.

Vital, D. (1967). *The inequality of states: A study of the small power in international relations.* London, England: Oxford University Press.

Waheed, M., Schuck, A., Neijens, P., & de Vreese, C. (2013). Values in UN speeches: Understanding the speechwriters' perspectives. *The International Communication Gazette, 75*(7), 597-618.

Webber, P. (2005). Interactive features in medical conference monologue. *English for Specific Purposes, 24*, 157-181.

Wharry, C. (2003). Amen and hallelujah preaching: Discourse functions in African American sermons. *Language in Society, 32*, 203-225.

Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied Linguistics, 21*(1), 3-25.

Widdowson, H. G. (2004). *Text, context, pretext: Critical issues in discourse analysis.* Oxford, England: Blackwell.

Wierzbicka, A. (1997). *Understanding cultures through their key words: English, Russian, Polish, German and Japanese*. Oxford, England: Oxford University Press.

Wodak, R. (Ed.). (1989). *Language, power and ideology: Studies in political discourse.* Amsterdam, Netherlands: John Benjamins.

Wodak, R. (1997). *Gender and discourse.* London, England: Sage.

Woods, N. (2006). *Describing discourse: A practical guide to discourse analysis*. London, England: Hodder Education.

World Bank. (2016a). *GDP (current US$).* Retrieved from
http://data.worldbank.org/indicator/NY.GDP.MKTP.CD.

World Bank. (2016b). *Military expenditure (% of GDP).* Retrieved from
http://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS?name_desc=false.

World Bank. (2018). *Net official development assistance and official aid received (current US$).*
Retrieved from https://data.worldbank.org/indicator/DT.ODA.ALLD.CD.

APPENDIX A: TAGGED FEATURES FOR THE UNGA

Feature	UNGA Mean	UNGA SD
Verb (not including auxiliary)	113.00	13.70
Present tense verb	77.31	13.20
Past tense verb	10.71	4.00
Progressive aspect verb	7.28	2.79
Perfect aspect verb	8.14	2.50
Infinitive verb	17.62	3.51
Pro-verb 'do'	0.43	0.57
Verb 'be'	2.10	1.52
Verb 'have'	1.43	1.14
Private verb	7.40	2.65
Public verb	2.31	1.28
Suasive verb	1.45	1.03
Activity verb	16.13	4.71
Communication verb	8.06	2.64
Mental verb	12.67	4.40
Causative verb	3.98	1.88
Occurrence verb	2.67	1.60
Existence verb	7.44	2.95
Aspectual verb	2.62	1.65
All modals	14.39	5.03
Modal of prediction	5.83	2.58
Modal of possibility	3.59	2.33
Modal of necessity	4.95	2.77
Split auxiliary-verb	4.01	1.78
All passives	8.84	2.80
Agentless passive verb	5.91	2.16
Passive verb + by	1.24	0.94
Passive postnominal modifier	1.70	1.07
All personal pronouns	45.39	14.27
1st person pronoun	35.58	12.35
2nd person pronoun	3.01	2.46
3rd person pronoun	6.80	3.49
Pronoun 'it'	9.45	3.31
Demonstrative pronoun	3.49	2.16
Nominal pronoun	1.61	1.29
All nouns	303.03	25.34
Common noun	250.37	19.07

Feature	UNGA Mean	UNGA SD
Proper noun	39.95	12.85
Animate noun	5.06	2.43
Process noun	15.24	5.00
Cognitive noun	3.74	1.87
Abstract noun	23.57	6.13
Concrete noun	4.35	2.36
Technical noun	1.71	1.31
Quantity noun	5.86	2.86
Place noun	6.46	2.99
Group noun	2.28	1.59
Noun-noun sequence	28.93	9.21
Stance noun + preposition	1.84	1.46
Determiner + stance noun	0.54	0.69
All indefinite articles	17.70	5.25
All definite articles	64.76	15.86
Preposition	124.88	13.44
Stranded preposition	0.53	0.65
All adjectives	88.93	12.35
Attributive adjective	68.46	12.01
Attributive adjective – Size	1.73	1.59
Attributive adjective – Time	1.63	1.40
Attributive adjective – Color	0.05	0.18
Attributive adjective – Evaluative	1.30	1.05
Attributive adjective – Relational	2.29	1.37
Attributive adjective – Topical	9.22	4.25
Predicative adjective	4.68	1.84
Predicative attitudinal adjective	1.33	1.07
Predicative epistemic adjective	0.45	0.58
All adverbs	30.22	6.99
Adverb – time	4.03	2.01
Adverb – place	4.19	2.15
All stance adverbs	1.81	1.57
Non-factive adverb	0.12	0.28
Attitudinal adverb	0.27	0.44
Factive adverb	1.25	1.28
Likelihood adverb	0.16	0.39
Adverb – downtoner	1.79	1.56
Amplifier	1.77	1.39

Feature	UNGA Mean	UNGA SD
Emphatic	2.11	1.37
Hedge	0.18	0.34
Discourse particle	0.26	0.51
'That' deletion	0.93	0.84
Contraction	1.20	1.98
All conjunctions	17.69	6.19
Subordinating conjunction – causative	0.48	0.78
Subordinating conjunction – conditional	1.17	1.17
Coordinating conjunction – clausal	7.80	5.34
Coordinating conjunction – phrasal	4.81	2.27
Adverbial – conjuncts	4.64	2.55
All wh- words	0.42	0.63
Wh question	0.28	0.58
Wh clause	0.15	0.29
All wh- relative clauses	3.65	2.10
Wh pronoun – object relative clause	0.84	0.80
Wh pronoun – subject relative clause	2.24	1.44
Wh pro. – rel. cls. with preposition fronting	0.56	0.57
Wh CCC communicaton verb	0.05	0.16
Wh CCC attitudinal verb	0.03	0.14
Wh CCC factive verb	0.12	0.26
Wh CCC likelihood verb	0.04	0.14
'That' relative clause	3.40	2.15
'That' CCC verb	3.23	1.64
'That' CCC communication verb	1.29	0.97
'That' CCC attitudinal verb	1.67	1.04
'That' CCC factive verb	1.11	0.88
'That' CCC likelihood verb	0.91	0.89
'That' CCC attitudinal adjective	0.30	0.43
'That' CCC factive adjective	0.30	0.45
'That' CCC attitudinal noun	0.11	0.31
'That' CCC factive noun	0.32	0.53
'That' CCC likelihood noun	0.14	0.29
All stance 'that' CCC verbs	4.97	2.07
All stance 'that' CCC adjectives	0.59	0.69
All stance 'that' CCC nouns	0.57	0.74
All stance 'that' complement clauses	6.14	2.73
'To' CCC speech act verb	0.17	0.35

Feature	UNGA Mean	UNGA SD
'To' CCC verb of cognition	0.27	0.34
'To' CCC verb of desire	2.39	1.41
'To' CCC verb of effort	1.66	1.09
'To' CCC probability verb	0.08	0.21
'To' CCC adjective of certainty	0.00	0.03
'To' CCC adjective of ability	0.14	0.31
'To' CCC adjective of affect	0.01	0.06
'To' CCC adjective of ease	0.02	0.11
'To' CCC evaluative adjective	0.11	0.28
'To' CCC stance noun	3.28	1.77
All stance 'to' CCC verb	4.57	1.82
All stance 'to' CCC adjective	0.28	0.43
All stance 'to' complement clauses	8.12	2.81
Dimension 1	-1403.65	5.90
Dimension 2	-283.34	0.64
Dimension 3	72.03	1.78
Dimension 4	-23.99	1.98
Dimension 5	153.83	1.78

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX B: TAGGED FEATURES BY REGION

Feature	Africa		Americas		Asia		Europe		Oceania	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Verb (not including auxiliary)	111.51	14.69	107.01	15.01	113.22	14.42	114.86	13.21	117.33	9.65
Present tense verb	75.89	12.93	71.05	16.00	76.68	13.38	80.52	12.97	80.33	9.39
Past tense verb	11.40	4.38	9.76	4.53	11.25	5.03	10.48	3.18	10.19	2.63
Progressive aspect verb	7.57	2.54	5.54	2.11	8.22	3.46	7.16	2.72	7.38	2.36
Perfect aspect verb	8.55	2.60	7.15	2.50	8.54	2.05	8.14	2.49	7.88	3.03
Infinitive verb	16.68	4.35	17.44	3.00	17.22	2.59	18.21	3.28	18.76	4.06
Pro-verb 'do'	0.21	0.43	0.48	0.59	0.52	0.71	0.35	0.46	0.75	0.60
Verb 'be'	1.63	1.08	2.08	1.19	1.97	1.57	2.50	1.95	2.31	1.37
Verb 'have'	1.34	1.27	1.19	1.02	1.12	1.10	1.60	1.06	1.92	1.19
Private verb	6.53	2.05	7.04	3.79	8.18	2.60	7.48	2.44	7.84	2.56
Public verb	1.84	1.33	2.45	0.78	2.66	1.47	2.27	1.41	2.53	0.87
Suasive verb	1.52	0.96	1.55	0.90	1.48	1.19	1.28	1.13	1.49	0.89
Activity verb	15.41	5.22	15.52	3.62	16.32	4.31	16.22	4.13	17.48	6.53
Communication verb	8.11	3.13	7.76	1.62	8.63	2.53	7.74	2.04	8.12	3.86
Mental verb	11.47	3.70	10.55	4.26	13.46	3.71	14.80	4.51	11.34	4.75
Causative verb	3.39	1.57	3.88	1.63	4.16	2.02	4.37	1.75	3.98	2.54
Occurrence verb	2.60	2.22	2.29	1.08	2.43	1.18	2.91	1.40	2.99	1.83
Existence verb	7.21	3.26	7.18	1.89	8.30	2.82	7.73	3.19	6.25	2.95
Aspectual verb	2.39	1.70	2.58	1.13	2.69	2.00	3.02	1.81	2.14	1.10
All modals	11.95	4.12	13.63	5.11	13.88	3.61	16.57	4.99	15.47	6.53
Modal of prediction	4.83	2.04	5.39	2.09	6.51	3.21	6.50	2.71	5.55	2.21
Modal of possibility	3.34	2.15	3.92	3.37	2.94	2.32	3.80	1.81	4.22	2.39
Modal of necessity	3.77	1.80	4.28	2.45	4.44	2.68	6.25	2.59	5.68	3.78
Split auxiliary-V	4.08	2.12	3.19	1.22	4.43	1.59	4.18	1.80	3.74	1.83
All passives	9.51	3.01	10.68	2.05	7.66	2.59	8.31	2.81	8.73	2.50
Agentless passive	6.42	2.19	6.55	1.99	4.81	1.77	5.78	2.32	6.30	2.13

Feature	Africa		Americas		Asia		Europe		Oceania	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Passive + by	1.35	1.22	1.78	0.94	1.05	0.76	1.14	0.77	0.99	0.89
Passive postnominal modifier	1.74	1.11	2.33	0.93	1.78	1.24	1.41	1.04	1.45	0.65
All personal pronouns	47.97	14.93	46.18	12.74	46.22	17.57	40.01	12.10	49.98	11.97
1st person pronoun	36.10	13.00	37.36	11.96	36.68	15.73	30.94	9.53	40.62	9.55
2nd person pronoun	4.71	2.67	2.05	1.04	3.27	3.17	2.08	1.74	2.68	1.83
3rd person pronoun	7.16	3.50	6.75	2.83	6.26	3.70	7.00	3.60	6.68	3.96
Pronoun 'it'	9.32	3.42	8.53	2.70	9.25	3.62	10.30	2.85	9.18	4.16
Demonstrative pronoun	3.62	2.00	2.95	2.19	3.53	2.73	3.62	2.10	3.52	1.79
Nominal pronoun	1.50	1.33	1.13	1.02	1.67	1.58	1.91	1.02	1.58	1.49
All nouns	304.00	28.78	311.32	23.27	304.74	25.33	296.08	19.57	304.55	31.60
Common noun	252.86	21.15	254.41	16.65	249.77	20.08	249.48	16.78	244.98	21.96
Proper noun	37.87	14.71	43.45	9.04	41.40	15.60	36.28	8.91	44.99	14.10
Animate noun	4.95	2.81	5.29	2.97	5.62	2.28	5.01	1.96	4.28	2.42
Process noun	14.99	5.36	14.64	4.34	13.72	3.99	15.92	4.22	17.15	7.25
Cognitive noun	3.40	1.86	4.75	1.13	3.14	1.92	4.22	1.81	3.17	2.09
Abstract noun	20.64	6.12	23.70	4.25	23.11	5.49	26.78	5.07	22.42	8.18
Concrete noun	3.30	2.24	3.70	1.30	4.61	1.92	5.06	2.14	4.87	3.64
Technical noun	1.43	1.21	1.56	0.71	1.89	1.10	1.72	1.70	2.00	1.43
Quantity noun	4.57	2.97	5.11	2.11	5.94	2.42	6.54	2.99	7.18	3.02
Place noun	6.86	2.83	7.52	3.24	6.63	3.50	5.87	2.82	5.72	2.50
Group noun	2.15	2.36	2.62	1.25	2.56	1.16	2.05	1.34	2.20	1.44
Noun-noun sequence	29.87	7.91	27.75	8.62	26.98	9.68	27.63	8.85	34.01	11.00
Stance noun + preposition	1.83	1.20	1.94	1.86	2.08	1.77	1.76	1.24	1.55	1.48
Determiner + stance noun	0.54	0.58	0.59	0.62	0.47	0.80	0.63	0.76	0.43	0.65
All indefinite articles	15.33	5.41	17.18	4.00	17.43	4.14	20.35	5.14	17.16	6.12
All definite articles	64.04	19.87	68.04	10.98	63.95	14.36	68.30	13.23	56.73	18.52
Preposition	129.86	11.03	129.60	11.59	122.77	14.06	123.50	14.81	117.97	12.22
Stranded preposition	0.66	0.82	0.37	0.53	0.44	0.51	0.58	0.66	0.55	0.61

502

Feature	Africa		Americas		Asia		Europe		Oceania	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All adjectives	86.26	10.56	92.69	10.33	91.39	12.81	87.78	13.40	88.23	14.24
Attributive adjective	66.23	9.65	71.86	11.39	70.53	11.83	68.24	13.21	66.08	14.18
Attributive adjective – Size	1.45	0.97	2.92	3.01	1.35	1.21	1.64	0.97	1.70	1.55
Attributive adjective – Time	1.51	1.76	1.77	1.25	1.58	1.16	1.81	1.28	1.39	1.59
Attributive adjective – Color	0.04	0.12	0.10	0.21	0.02	0.07	0.05	0.25	0.05	0.19
Attributive adjective – Evaluative	1.09	1.21	0.61	0.46	1.23	0.69	1.80	1.11	1.45	1.12
Attributive. adjective – Relational	1.93	1.13	2.11	1.05	2.45	1.36	2.72	1.66	1.97	1.32
Attributive adjective – Topical	8.42	5.14	9.95	4.24	9.30	3.97	10.12	3.92	7.88	3.74
Predicative adjective	5.17	1.72	4.02	1.17	4.43	2.01	4.52	2.00	5.24	1.87
Predicative attitudinal adjective	1.25	0.70	1.26	0.78	1.14	1.09	1.53	1.35	1.39	1.25
Predicative epistemic adjective	0.31	0.43	0.33	0.45	0.41	0.55	0.70	0.76	0.35	0.38
All adverbs	29.43	6.62	28.52	7.37	31.31	7.12	31.23	7.27	29.55	6.93
Adverb – time	3.70	1.88	3.40	1.46	4.19	1.85	4.18	1.82	4.61	3.09
Adverb – place	4.36	2.67	3.45	2.00	4.77	2.36	4.27	1.89	3.66	1.31
All stance adverbs	1.84	1.97	1.20	0.84	1.93	1.76	2.40	1.49	1.02	0.62
Non-factive adverb	0.20	0.36	0.09	0.19	0.17	0.36	0.08	0.22	0.05	0.17
Attitudinal adverb	0.31	0.48	0.15	0.33	0.11	0.27	0.44	0.55	0.18	0.30
Factive adverb	1.31	1.66	0.87	0.74	1.49	1.49	1.52	1.19	0.66	0.45
Likelihood adverb	0.01	0.07	0.08	0.19	0.16	0.41	0.35	0.56	0.12	0.24
Adverb – downtoner	1.92	1.98	1.29	0.88	1.50	1.06	2.15	1.80	1.81	1.38
Amplifier	1.76	1.63	1.13	0.74	1.69	1.15	2.22	1.48	1.60	1.47
Emphatic	1.71	1.03	1.58	0.83	1.83	1.42	2.73	1.51	2.47	1.50
Hedge	0.18	0.28	0.17	0.38	0.24	0.40	0.20	0.36	0.09	0.23
Discourse particle	0.33	0.83	0.28	0.49	0.21	0.28	0.26	0.40	0.24	0.37
'That' deletion	1.11	0.73	0.98	1.01	1.08	1.04	0.64	0.66	0.94	0.79
Contraction	0.72	0.92	1.34	2.37	2.09	3.39	0.90	1.03	1.13	1.15
All conjunctions	16.11	5.59	18.66	6.73	18.02	6.33	17.57	6.55	19.04	6.00
Subordinating conjunction – causative	0.43	0.71	0.42	0.63	0.51	0.98	0.42	0.69	0.68	0.92

Feature	Africa		Americas		Asia		Europe		Oceania	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Subordinating conjunction – conditional	1.01	1.18	1.03	0.82	0.94	1.02	1.28	1.42	1.65	1.07
Coordinating conjunction – clausal	5.99	5.26	7.79	6.07	8.45	5.66	8.18	5.20	9.01	4.57
Coordinating conjunction – phrasal	5.00	2.63	5.08	2.48	4.77	2.28	4.67	2.20	4.57	1.85
Adverbial – conjuncts	5.80	2.65	4.95	2.10	4.75	2.37	3.92	2.49	3.76	2.71
All wh- words	0.27	0.37	0.30	0.54	0.55	0.95	0.55	0.66	0.35	0.35
Wh question	0.12	0.27	0.28	0.54	0.30	0.92	0.38	0.57	0.26	0.36
Wh clause	0.15	0.30	0.02	0.06	0.25	0.35	0.17	0.34	0.08	0.21
All wh- relative clauses	4.40	1.90	4.68	2.18	2.95	1.68	3.30	2.44	3.14	1.62
Wh pronoun – object relative clause	1.02	0.86	1.19	0.97	0.76	0.84	0.71	0.72	0.57	0.47
Wh pronoun – subject relative clause	2.57	1.30	2.68	1.43	1.83	1.00	2.10	1.78	2.18	1.41
Wh pro. – rel. cls. with preposition fronting	0.81	0.62	0.79	0.64	0.37	0.48	0.48	0.57	0.38	0.35
Wh CCC communication verb	0.00	0.00	0.09	0.19	0.06	0.19	0.07	0.19	0.03	0.11
Wh CCC attitudinal verb	0.00	0.00	0.00	0.00	0.03	0.11	0.08	0.24	0.00	0.00
Wh CCC factive verb	0.13	0.30	0.02	0.06	0.17	0.27	0.13	0.28	0.12	0.22
Wh CCC likelihood verb	0.05	0.15	0.00	0.00	0.04	0.13	0.04	0.16	0.05	0.17
'That' relative clause	3.90	3.11	3.77	2.22	2.88	1.69	3.59	1.81	2.62	1.10
'That' CCC verb	2.96	1.44	3.66	2.19	3.35	1.73	2.77	1.53	4.01	1.12
'That' CCC communication verb	1.47	1.01	1.43	0.87	1.54	1.15	0.97	0.83	1.18	0.94
'That' CCC attitudinal verb	1.74	1.17	2.10	1.09	1.73	1.09	1.38	0.81	1.61	1.10
'That' CCC factive verb	1.23	0.97	1.27	1.10	0.88	0.92	1.00	0.76	1.28	0.68
'That' CCC likelihood verb	0.82	0.67	0.68	0.91	1.52	1.10	0.82	0.85	0.55	0.49
'That' CCC attitudinal adjective	0.42	0.52	0.38	0.48	0.22	0.35	0.20	0.40	0.31	0.39
'That' CCC factive adjective	0.44	0.64	0.18	0.24	0.37	0.48	0.26	0.38	0.15	0.30
'That' CCC attitudinal noun	0.10	0.22	0.20	0.44	0.19	0.37	0.01	0.06	0.12	0.44
'That' CCC factive noun	0.36	0.43	0.36	0.46	0.35	0.82	0.33	0.48	0.15	0.32
'That' CCC likelihood noun	0.08	0.18	0.32	0.47	0.04	0.16	0.12	0.26	0.26	0.35
All stance 'that' CCC verbs	5.24	2.32	5.50	2.70	5.64	1.91	4.16	1.46	4.62	1.89
All stance 'that' CCC adjectives	0.87	1.02	0.56	0.62	0.58	0.53	0.47	0.58	0.46	0.44

Feature	Africa		Americas		Asia		Europe		Oceania	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All stance 'that' CCC nouns	0.53	0.54	0.88	0.81	0.58	0.98	0.46	0.63	0.53	0.76
All stance 'that' complement clauses	6.65	2.98	6.97	3.38	6.80	2.70	5.09	2.07	5.63	2.38
'To' CCC speech act verb	0.16	0.29	0.10	0.19	0.27	0.56	0.12	0.26	0.19	0.31
'To' CCC verb of cognition	0.30	0.33	0.31	0.39	0.19	0.29	0.18	0.28	0.48	0.43
'To' CCC verb of desire	2.52	1.50	1.42	1.11	2.85	1.21	2.65	1.45	1.95	1.28
'To' CCC verb of effort	1.95	1.17	1.85	1.06	1.35	1.08	1.73	1.01	1.35	1.15
'To' CCC probability verb	0.10	0.22	0.15	0.23	0.04	0.16	0.11	0.27	0.00	0.00
'To' CCC adjective of certainty	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06	0.00	0.00
'To' CCC adjective of ability	0.13	0.29	0.00	0.00	0.07	0.22	0.23	0.42	0.21	0.28
'To' CCC adjective of affect	0.00	0.00	0.00	0.00	0.03	0.14	0.00	0.00	0.00	0.00
'To' CCC adjective of ease	0.00	0.00	0.04	0.14	0.03	0.11	0.02	0.10	0.05	0.17
'To' CCC evaluative adjective	0.13	0.31	0.04	0.14	0.09	0.31	0.10	0.29	0.17	0.27
'To' CCC stance noun	2.75	1.70	3.70	1.63	3.13	1.26	3.55	2.09	3.39	2.01
All stance 'to' CCC verb	5.05	1.89	3.84	1.97	4.68	1.47	4.77	1.58	3.98	2.33
All stance 'to' CCC adjective	0.26	0.38	0.08	0.19	0.22	0.48	0.37	0.46	0.42	0.52
All stance 'to' complement clauses	8.05	3.34	7.58	2.60	8.03	2.11	8.67	2.47	7.80	3.77
Dimension 1	-16.30	5.32	-17.12	6.28	-15.26	7.84	-14.09	4.58	-14.04	5.60
Dimension 2	-3.05	0.68	-3.28	0.69	-2.95	0.52	-3.09	0.62	-3.11	0.76
Dimension 3	1.43	1.65	2.02	1.39	0.05	1.72	0.38	1.97	0.36	1.12
Dimension 4	-1.08	1.71	-1.03	1.45	-0.27	1.26	0.59	2.09	0.14	2.75
Dimension 5	2.57	1.55	2.57	1.54	1.46	1.50	1.03	1.64	0.92	2.21

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX C: TAGGED FEATURES BY AREA

Feature	Small Area		Medium Area		Large Area		Very Large Area	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Verb (not including auxiliary)	112.82	14.46	115.49	13.36	110.55	13.75	112.12	13.37
Present tense verb	78.15	12.19	79.82	11.93	75.05	14.21	74.86	13.57
Past tense verb	9.45	3.56	10.33	3.71	11.39	4.66	11.92	3.65
Progressive aspect verb	5.65	1.89	7.11	1.92	7.97	2.72	7.91	3.16
Perfect aspect verb	7.52	2.58	8.70	2.53	7.90	2.22	8.41	2.73
Infinitive verb	17.98	3.78	18.65	3.58	16.40	3.30	17.36	3.26
Pro-verb 'do'	0.50	0.44	0.46	0.64	0.47	0.60	0.31	0.58
Verb 'be'	2.63	1.35	1.88	1.27	2.27	1.60	1.38	1.09
Verb 'have'	1.60	1.06	1.51	1.15	1.20	1.40	1.41	0.95
Private verb	7.03	2.77	7.43	2.55	7.48	2.41	7.70	3.02
Public verb	2.39	1.04	2.15	1.08	2.47	1.58	2.31	1.40
Suasive verb	1.36	0.94	1.62	1.28	1.66	0.98	1.19	0.77
Activity verb	15.64	5.57	17.37	3.89	15.32	4.18	16.24	5.20
Communication verb	7.86	3.22	7.92	1.94	8.36	2.95	8.03	2.51
Mental verb	12.09	4.37	12.35	3.54	13.65	4.68	12.20	4.79
Causative verb	3.60	2.12	4.56	1.89	3.91	1.53	3.75	1.94
Occurrence verb	2.50	1.56	3.35	1.55	2.66	1.94	2.04	1.01
Existence verb	7.18	2.80	8.07	3.03	6.61	2.35	7.82	3.55
Aspectual verb	2.37	1.47	3.11	1.87	2.31	1.58	2.70	1.67
All modals	15.82	5.14	13.68	4.26	14.43	5.26	13.30	5.16
Modal of prediction	5.38	1.97	6.14	2.54	5.88	3.34	5.70	2.27
Modal of possibility	4.21	2.05	3.33	1.86	3.36	2.37	3.40	2.97
Modal of necessity	6.18	3.24	4.20	2.70	5.20	2.60	4.19	2.12
Split auxiliary-verb	4.05	1.83	4.12	1.60	3.72	1.83	4.10	1.97
All passives	9.44	2.75	9.06	2.96	8.39	2.95	8.61	2.56
Agentless passive verb	6.39	2.41	6.09	2.62	5.37	1.73	5.80	1.79

Feature	Small Area		Medium Area		Large Area		Very Large Area	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Passive verb + by	1.42	0.99	1.20	0.86	1.24	1.08	1.15	0.85
Passive postnominal modifier	1.62	0.82	1.78	0.98	1.79	1.33	1.66	1.09
All personal pronouns	49.80	14.17	44.78	13.49	40.98	14.63	46.31	14.65
1st person pronoun	40.37	13.16	35.55	11.17	31.10	12.22	35.66	12.15
2nd person pronoun	1.92	1.46	2.91	2.77	3.07	2.41	4.14	2.65
3rd person pronoun	7.50	3.75	6.32	3.25	6.83	3.52	6.50	3.62
Pronoun 'it'	8.99	2.85	9.87	2.74	10.49	4.16	8.40	3.21
Demonstrative pronoun	3.70	2.19	3.66	1.98	3.42	2.29	2.97	2.08
Nominal pronoun	1.18	1.13	1.62	1.25	1.93	1.50	1.57	1.13
All nouns	296.29	21.40	303.60	22.92	306.06	28.45	307.94	26.99
Common noun	244.27	15.35	253.96	19.04	252.63	20.67	251.82	19.31
Proper noun	38.69	12.44	38.06	14.70	40.98	12.49	42.26	12.18
Animate noun	4.52	2.36	5.44	1.93	4.73	2.70	5.58	2.73
Process noun	16.20	5.58	15.95	4.22	14.97	5.15	14.00	5.07
Cognitive noun	4.31	2.02	3.58	1.60	3.87	1.72	3.13	2.08
Abstract noun	22.53	7.32	25.33	5.91	24.27	4.53	21.98	6.46
Concrete noun	4.40	2.67	4.21	2.05	4.95	2.59	3.92	2.11
Technical noun	1.51	0.93	1.23	0.79	2.00	1.63	1.87	1.18
Quantity noun	5.60	2.43	6.49	3.08	5.09	2.90	6.22	3.01
Place noun	5.28	2.59	7.05	3.24	6.86	3.21	6.61	2.78
Group noun	2.53	1.34	1.93	1.30	1.98	1.33	2.70	2.22
Noun-noun sequence	27.57	8.41	31.03	9.01	26.42	9.09	31.38	9.36
Stance noun + preposition	1.56	1.33	1.96	1.14	1.91	1.94	1.98	1.35
Determiner + stance noun	0.36	0.44	0.55	0.67	0.67	0.94	0.61	0.60
All indefinite articles	17.99	5.76	18.67	5.00	18.03	4.63	15.90	5.61
All definite articles	62.38	17.32	64.51	13.14	69.14	12.15	63.25	20.30
Preposition	124.83	13.90	121.03	11.59	128.57	15.37	125.12	12.75
Stranded preposition	0.47	0.46	0.58	0.82	0.59	0.73	0.52	0.53

Feature	Small Area		Medium Area		Large Area		Very Large Area	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All adjectives	86.67	12.81	90.46	12.14	88.23	12.93	91.08	11.48
Attributive adjective	65.71	12.26	69.30	11.93	69.77	13.05	69.85	10.39
Attributive adjective – Size	2.63	2.03	1.38	0.94	1.70	1.83	1.26	1.05
Attributive adjective – Time	1.49	1.37	1.87	1.63	1.48	1.27	1.69	1.39
Attributive adjective – Color	0.03	0.13	0.12	0.31	0.02	0.08	0.03	0.09
Attributive adjective – Evaluative	1.27	0.92	1.57	1.36	1.35	1.05	1.00	0.73
Attributive adjective – Relational	2.29	1.08	1.98	1.62	2.59	1.55	2.24	1.14
Attributive adjective – Topical	8.39	3.90	9.48	4.51	10.88	4.37	8.24	3.89
Predicative adjective	4.76	1.69	4.99	1.98	4.30	1.99	4.64	1.75
Predicative attitudinal adjective	1.16	0.80	1.62	1.39	1.15	1.02	1.36	0.98
Predicative epistemic adjective	0.34	0.44	0.54	0.61	0.41	0.57	0.38	0.39
All adverbs	32.62	7.14	29.38	6.32	28.85	6.97	30.16	7.55
Adverb – time	4.09	2.29	4.23	2.14	3.78	1.75	3.90	1.94
Adverb – place	3.15	1.86	4.08	1.62	4.32	1.97	5.05	2.63
All stance adverbs	1.94	1.28	2.00	1.90	1.73	1.45	1.36	1.33
Non-factive adverb	0.05	0.15	0.09	0.23	0.17	0.38	0.15	0.26
Attitudinal adverb	0.25	0.46	0.38	0.49	0.21	0.46	0.18	0.30
Factive adverb	1.43	1.06	1.30	1.51	1.17	1.11	1.00	1.31
Likelihood adverb	0.21	0.41	0.22	0.50	0.18	0.39	0.04	0.14
Adverb – downtoner	1.86	1.33	1.81	1.88	1.70	1.51	1.57	1.04
Amplifier	1.82	1.38	1.95	1.32	1.47	1.42	1.85	1.51
Emphatic	2.28	1.71	2.06	1.39	2.17	1.27	1.95	1.15
Hedge	0.10	0.21	0.15	0.26	0.34	0.48	0.16	0.29
Discourse particle	0.25	0.39	0.37	0.82	0.31	0.40	0.13	0.24
'That' deletion	1.05	0.82	0.88	0.84	0.72	0.75	1.12	0.93
Contraction	0.96	0.96	1.23	2.88	1.23	1.60	1.38	2.02
All conjunctions	20.43	5.11	17.03	6.24	15.57	5.87	17.50	6.67
Subordinating conjunction – causative	0.66	0.87	0.43	0.62	0.42	0.83	0.30	0.59

Feature	Small Area		Medium Area		Large Area		Very Large Area	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Subordinating conjunction – conditional	1.51	1.18	1.15	1.11	0.96	1.39	1.02	0.96
Coordinating conjunction – clausal	9.58	5.25	7.26	4.71	6.81	5.39	7.40	5.94
Coordinating conjunction – phrasal	5.16	1.77	4.74	2.02	4.23	2.88	5.31	2.16
Adverbial – conjuncts	5.71	2.51	3.54	1.88	4.36	3.05	5.16	2.27
All wh- words	0.59	0.88	0.47	0.65	0.41	0.48	0.25	0.44
Wh question	0.50	0.86	0.32	0.53	0.19	0.44	0.11	0.37
Wh clause	0.09	0.23	0.15	0.31	0.23	0.33	0.13	0.29
All wh- relative clauses	4.36	2.51	3.05	1.75	3.56	2.05	3.68	2.01
Wh pronoun – object relative clause	1.00	0.90	0.76	0.67	0.83	0.93	0.78	0.73
Wh pronoun – subject relative clause	2.81	1.81	1.73	1.15	2.26	1.21	2.23	1.44
Wh pro. – rel. cls. with preposition fronting	0.52	0.57	0.56	0.64	0.48	0.55	0.67	0.55
Wh CCC communicaton verb	0.05	0.14	0.03	0.12	0.11	0.24	0.01	0.04
Wh CCC attitudinal verb	0.00	0.00	0.05	0.18	0.06	0.21	0.00	0.00
Wh CCC factive verb	0.11	0.24	0.10	0.21	0.14	0.25	0.09	0.26
Wh CCC likelihood verb	0.06	0.19	0.03	0.10	0.04	0.13	0.03	0.13
'That' relative clause	2.86	1.55	3.88	2.41	3.70	2.45	3.14	2.06
'That' CCC verb	3.73	1.47	2.93	1.43	2.99	1.89	3.29	1.74
'That' CCC communication verb	1.14	0.86	1.31	0.94	1.31	1.08	1.37	1.05
'That' CCC attitudinal verb	2.07	1.18	1.80	1.10	1.20	0.88	1.60	0.87
'That' CCC factive verb	1.05	0.87	1.10	0.81	1.16	0.81	1.12	1.10
'That' CCC likelihood verb	0.89	0.96	0.61	0.67	1.08	1.07	1.07	0.80
'That' CCC attitudinal adjective	0.30	0.39	0.43	0.52	0.17	0.36	0.29	0.42
'That' CCC factive adjective	0.22	0.33	0.32	0.40	0.30	0.64	0.31	0.38
'That' CCC attitudinal noun	0.07	0.19	0.08	0.19	0.22	0.52	0.07	0.20
'That' CCC factive noun	0.20	0.25	0.44	0.76	0.27	0.39	0.30	0.45
'That' CCC likelihood noun	0.18	0.34	0.10	0.27	0.10	0.23	0.14	0.29
All stance 'that' CCC verbs	5.13	2.23	4.82	1.81	4.75	1.75	5.14	2.56
All stance 'that' CCC adjectives	0.54	0.57	0.75	0.67	0.47	0.92	0.60	0.55

Feature	Small Area		Medium Area		Large Area		Very Large Area	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
All stance 'that' CCC nouns	0.45	0.49	0.62	0.90	0.59	0.68	0.50	0.66
All stance 'that' complement clauses	6.13	2.68	6.19	2.58	5.83	2.59	6.25	3.14
'To' CCC speech act verb	0.11	0.24	0.20	0.28	0.25	0.55	0.10	0.20
'To' CCC verb of cognition	0.45	0.43	0.20	0.28	0.19	0.30	0.26	0.30
'To' CCC verb of desire	2.14	1.42	2.07	1.23	2.73	1.41	2.60	1.56
'To' CCC verb of effort	1.40	1.16	1.97	1.12	1.53	1.17	1.73	0.89
'To' CCC probability verb	0.12	0.24	0.03	0.12	0.10	0.22	0.05	0.15
'To' CCC adjective of certainty	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06
'To' CCC adjective of ability	0.19	0.32	0.06	0.17	0.16	0.39	0.12	0.26
'To' CCC adjective of affect	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.13
'To' CCC adjective of ease	0.03	0.13	0.02	0.10	0.04	0.14	0.00	0.00
'To' CCC evaluative adjective	0.15	0.32	0.12	0.31	0.08	0.27	0.08	0.18
'To' CCC stance noun	3.38	1.82	3.89	2.07	2.98	1.54	2.85	1.55
All stance 'to' CCC verb	4.20	2.31	4.46	1.67	4.82	1.60	4.75	1.71
All stance 'to' CCC adjective	0.36	0.57	0.20	0.35	0.28	0.45	0.24	0.30
All stance 'to' complement clauses	7.94	3.40	8.53	2.63	8.06	2.23	7.83	3.06
Dimension 1	-13.90	6.06	-15.11	5.50	-15.81	6.11	-16.56	5.91
Dimension 2	-3.19	0.68	-3.04	0.62	-3.08	0.71	-3.00	0.59
Dimension 3	1.42	1.60	0.47	1.75	0.60	2.05	0.80	1.56
Dimension 4	0.37	2.41	-0.45	1.86	-0.27	1.96	-0.78	1.51
Dimension 5	2.53	2.02	1.00	1.32	1.44	1.82	1.90	1.61

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX D: TAGGED FEATURES BY POPULATION

Feature	Small Pop		Medium Pop		Large Pop		Very Large Pop	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Verb (not including auxiliary)	112.98	13.33	111.66	13.85	115.60	13.17	111.87	15.00
Present tense verb	78.16	11.24	74.43	12.76	81.37	13.87	75.45	14.64
Past tense verb	9.75	3.34	10.48	3.61	10.63	4.09	12.09	4.79
Progressive aspect verb	5.97	1.91	7.37	2.12	8.00	2.40	7.88	4.02
Perfect aspect verb	7.56	2.48	8.25	2.92	7.97	2.66	8.82	1.77
Infinitive verb	17.97	3.72	18.60	3.51	16.66	3.51	17.15	3.16
Pro-verb 'do'	0.51	0.47	0.34	0.51	0.57	0.70	0.30	0.56
Verb 'be'	2.55	1.33	1.73	1.26	2.03	1.76	2.10	1.70
Verb 'have'	1.67	1.08	1.55	1.32	1.21	1.01	1.25	1.13
Private verb	6.86	2.86	7.49	1.99	7.94	2.70	7.33	3.02
Public verb	2.39	1.13	2.02	0.91	2.26	1.37	2.61	1.63
Suasive verb	1.34	0.92	1.92	1.23	1.39	1.04	1.11	0.72
Activity verb	15.90	5.30	16.73	4.15	15.91	5.42	15.97	4.04
Communication verb	7.73	3.16	8.08	2.25	7.76	2.95	8.72	2.09
Mental verb	12.45	4.30	11.25	2.93	13.91	5.75	13.24	4.09
Causative verb	3.46	2.01	4.54	1.86	4.20	1.84	3.71	1.70
Occurrence verb	2.73	1.37	2.65	1.71	2.54	1.63	2.74	1.77
Existence verb	7.21	2.74	7.54	2.90	7.19	3.23	7.85	3.11
Aspectual verb	2.50	1.48	2.73	1.95	2.45	1.43	2.81	1.77
All modals	15.65	4.97	12.70	4.01	14.76	5.90	14.48	4.98
Modal of prediction	5.57	1.86	5.42	2.55	5.99	3.23	6.41	2.63
Modal of possibility	4.17	2.00	3.18	1.57	3.70	2.70	3.30	2.91
Modal of necessity	5.87	3.30	4.09	2.41	5.08	2.91	4.75	2.14
Split auxiliary-verb	3.82	1.71	3.82	1.64	5.00	1.84	3.42	1.65
All passives	9.47	2.67	8.65	2.79	9.11	3.26	8.10	2.40
Agentless passive verb	6.51	2.34	5.50	2.19	6.19	2.19	5.43	1.80

Feature	Small Pop		Medium Pop		Large Pop		Very Large Pop	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Passive verb + by	1.30	1.00	1.35	0.92	1.41	1.05	0.87	0.72
Passive postnominal modifier	1.65	0.86	1.82	0.96	1.50	1.32	1.80	1.13
All personal pronouns	49.94	13.28	43.32	13.51	46.95	16.94	41.13	12.34
1st person pronoun	40.53	12.33	34.95	12.10	35.72	14.10	30.73	9.05
2nd person pronoun	1.97	1.41	2.38	1.51	4.48	3.18	3.37	2.71
3rd person pronoun	7.44	3.60	6.00	3.02	6.76	2.88	7.03	4.37
Pronoun 'it'	9.02	3.04	9.71	3.39	9.59	2.76	9.51	4.11
Demonstrative pronoun	3.80	2.16	2.67	1.51	3.62	2.12	3.91	2.65
Nominal pronoun	1.33	1.09	1.49	1.12	1.92	1.58	1.73	1.34
All nouns	298.24	22.10	307.39	24.22	306.38	25.98	300.14	29.34
Common noun	246.22	15.80	254.41	17.96	253.65	21.28	247.21	20.90
Proper noun	38.55	12.04	40.83	11.77	40.30	16.30	40.15	11.68
Animate noun	4.85	2.45	4.43	1.81	5.15	2.87	5.88	2.45
Process noun	16.08	5.49	15.21	4.43	14.90	6.13	14.70	3.85
Cognitive noun	4.07	2.02	3.85	1.53	3.65	2.27	3.32	1.62
Abstract noun	23.12	7.25	23.84	5.32	23.84	7.19	23.49	4.72
Concrete noun	4.70	2.55	3.99	2.00	4.46	2.91	4.24	1.96
Technical noun	1.63	1.08	1.47	1.15	1.66	1.37	2.10	1.62
Quantity noun	5.92	2.58	6.22	3.28	5.14	2.70	6.11	2.89
Place noun	5.74	2.79	7.75	3.36	5.89	3.33	6.42	2.00
Group noun	2.30	1.33	1.97	1.40	2.51	2.16	2.36	1.42
Noun-noun sequence	28.95	8.57	29.24	8.43	30.00	10.20	27.49	10.09
Stance noun + preposition	1.46	1.31	2.11	1.58	1.96	1.36	1.83	1.57
Determiner + stance noun	0.38	0.43	0.63	0.76	0.59	0.89	0.58	0.62
All indefinite articles	17.69	5.68	18.78	4.92	17.36	6.30	16.88	3.95
All definite articles	62.39	16.58	68.23	12.48	58.32	17.61	69.98	14.68
Preposition	123.81	12.68	126.79	12.14	120.16	15.20	128.69	13.07
Stranded preposition	0.57	0.61	0.49	0.59	0.50	0.80	0.58	0.61

Feature	Small Pop		Medium Pop		Large Pop		Very Large Pop	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All adjectives	86.85	12.89	92.08	12.66	88.15	12.84	88.56	11.03
Attributive adjective	65.54	12.38	71.17	11.90	68.84	12.56	68.32	11.20
Attributive adjective – Size	2.45	1.97	1.59	1.81	1.54	0.94	1.27	1.14
Attributive adjective – Time	1.49	1.36	1.88	1.33	1.69	1.73	1.45	1.19
Attributive adjective – Color	0.08	0.20	0.07	0.27	0.01	0.06	0.03	0.09
Attributive adjective – Evaluative	1.33	1.01	1.26	0.94	1.35	1.40	1.27	0.87
Attributive adjective – Relational	2.28	1.14	2.14	1.42	2.36	1.61	2.40	1.38
Attributive adjective – Topical	7.82	3.37	10.58	4.15	9.83	4.77	8.67	4.38
Predicative adjective	4.77	1.71	4.76	2.02	4.52	1.84	4.66	1.87
Predicative attitudinal adjective	1.27	0.82	1.47	1.34	1.40	1.10	1.15	1.01
Predicative epistemic adjective	0.33	0.43	0.48	0.59	0.45	0.57	0.55	0.71
All adverbs	32.06	7.15	28.08	6.56	31.43	6.32	29.32	7.55
Adverb – time	4.13	2.20	3.53	2.18	4.24	1.86	4.25	1.78
Adverb – place	3.47	1.95	3.84	1.69	5.03	2.61	4.53	2.06
All stance adverbs	1.92	1.24	1.17	0.96	2.33	2.02	1.87	1.78
Non-factive adverb	0.04	0.14	0.09	0.23	0.17	0.37	0.21	0.33
Attitudinal adverb	0.30	0.47	0.18	0.31	0.39	0.57	0.21	0.37
Factive adverb	1.34	1.05	0.79	0.86	1.54	1.66	1.38	1.41
Likelihood adverb	0.24	0.40	0.12	0.42	0.22	0.44	0.07	0.24
Adverb – downtoner	1.87	1.40	1.68	1.00	1.76	2.06	1.87	1.72
Amplifier	1.73	1.34	1.59	1.22	1.97	1.72	1.79	1.31
Emphatic	2.40	1.65	1.66	1.19	2.50	1.43	1.91	1.01
Hedge	0.11	0.22	0.14	0.32	0.24	0.39	0.26	0.39
Discourse particle	0.31	0.48	0.08	0.17	0.49	0.82	0.19	0.30
'That' deletion	0.98	0.83	0.89	0.78	0.98	0.81	0.87	0.98
Contraction	1.07	0.91	0.79	1.12	1.85	3.21	1.13	1.91
All conjunctions	19.84	5.72	15.34	5.48	17.75	6.57	17.84	6.53
Subordinating conjunction – causative	0.63	0.84	0.40	0.67	0.28	0.63	0.59	0.94

Feature	Small Pop		Medium Pop		Large Pop		Very Large Pop	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Subordinating conjunction – conditional	1.48	1.13	1.07	1.14	1.03	1.30	1.08	1.11
Coordinating conjunction – clausal	9.28	5.17	5.62	3.23	8.59	5.60	7.77	6.57
Coordinating conjunction – phrasal	4.90	2.05	4.99	1.83	4.73	3.04	4.60	2.20
Adverbial – conjuncts	5.33	2.53	4.25	2.39	4.01	2.28	4.95	2.90
All wh- words	0.54	0.86	0.33	0.60	0.49	0.55	0.33	0.43
Wh question	0.45	0.83	0.27	0.59	0.23	0.36	0.13	0.37
Wh clause	0.08	0.23	0.07	0.18	0.26	0.41	0.20	0.30
All wh- relative clauses	4.30	2.37	3.38	2.33	3.10	1.66	3.79	1.84
Wh pronoun – object relative clause	1.08	0.88	0.81	0.86	0.60	0.65	0.85	0.76
Wh pronoun – subject relative clause	2.71	1.72	2.05	1.45	1.93	1.20	2.27	1.25
Wh pro. – rel. cls. with preposition fronting	0.49	0.57	0.53	0.56	0.57	0.62	0.67	0.55
Wh CCC communicaton verb	0.05	0.14	0.03	0.12	0.09	0.23	0.03	0.11
Wh CCC attitudinal verb	0.00	0.00	0.03	0.12	0.07	0.24	0.02	0.11
Wh CCC factive verb	0.10	0.23	0.07	0.18	0.17	0.31	0.15	0.29
Wh CCC likelihood verb	0.05	0.18	0.02	0.08	0.04	0.14	0.04	0.13
'That' relative clause	2.66	1.56	3.71	2.42	3.41	1.71	3.87	2.66
'That' CCC verb	3.48	1.50	3.08	1.47	2.79	1.86	3.59	1.70
'That' CCC communication verb	1.04	0.81	1.55	0.87	1.10	1.11	1.49	1.03
'That' CCC attitudinal verb	1.86	1.15	1.87	1.20	1.25	0.74	1.65	0.95
'That' CCC factive verb	1.05	0.84	1.05	0.76	1.17	0.95	1.15	1.03
'That' CCC likelihood verb	0.80	0.91	0.70	0.72	1.03	1.03	1.13	0.87
'That' CCC attitudinal adjective	0.33	0.38	0.37	0.54	0.25	0.42	0.23	0.37
'That' CCC factive adjective	0.19	0.32	0.33	0.38	0.24	0.52	0.43	0.56
'That' CCC attitudinal noun	0.08	0.20	0.25	0.46	0.06	0.30	0.05	0.15
'That' CCC factive noun	0.20	0.28	0.53	0.75	0.15	0.30	0.40	0.58
'That' CCC likelihood noun	0.19	0.37	0.11	0.21	0.11	0.25	0.15	0.33
All stance 'that' CCC verbs	4.75	2.09	5.15	1.73	4.55	2.24	5.41	2.23
All stance 'that' CCC adjectives	0.53	0.54	0.70	0.69	0.49	0.78	0.65	0.76

Feature	Small Pop		Medium Pop		Large Pop		Very Large Pop	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
All stance 'that' CCC nouns	0.47	0.53	0.88	0.94	0.34	0.47	0.59	0.81
All stance 'that' complement clauses	5.75	2.43	6.75	2.58	5.38	3.00	6.67	2.81
'To' CCC speech act verb	0.13	0.27	0.11	0.20	0.09	0.24	0.34	0.54
'To' CCC verb of cognition	0.43	0.44	0.26	0.30	0.16	0.27	0.22	0.30
'To' CCC verb of desire	2.18	1.49	1.96	1.31	2.63	1.49	2.85	1.21
'To' CCC verb of effort	1.36	1.08	2.16	1.14	1.45	1.10	1.67	0.89
'To' CCC probability verb	0.12	0.25	0.02	0.08	0.06	0.19	0.13	0.27
'To' CCC adjective of certainty	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06
'To' CCC adjective of ability	0.17	0.31	0.08	0.23	0.16	0.38	0.15	0.30
'To' CCC adjective of affect	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.13
'To' CCC adjective of ease	0.07	0.18	0.00	0.00	0.00	0.00	0.02	0.11
'To' CCC evaluative adjective	0.17	0.33	0.04	0.13	0.17	0.39	0.05	0.14
'To' CCC stance noun	3.25	1.86	4.18	2.02	2.66	1.60	2.96	1.19
All stance 'to' CCC verb	4.20	2.30	4.50	1.76	4.39	1.77	5.23	1.18
All stance 'to' CCC adjective	0.40	0.54	0.11	0.27	0.33	0.49	0.27	0.32
All stance 'to' complement clauses	7.84	3.30	8.78	2.72	7.38	2.93	8.44	2.09
Dimension 1	-13.93	5.70	-17.14	4.98	-14.26	6.29	-15.65	6.43
Dimension 2	-3.18	0.65	-3.16	0.70	-3.13	0.65	-2.84	0.53
Dimension 3	1.21	1.56	1.06	1.64	0.09	2.21	0.71	1.54
Dimension 4	0.16	2.42	-0.74	1.62	0.15	2.20	-0.61	1.43
Dimension 5	2.23	2.03	1.49	1.54	1.40	1.73	1.54	1.79

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX E: TAGGED FEATURES BY HUMAN DEVELOPMENT INDEX

Feature	Low HDI		Medium HDI		High HDI		Very High HDI	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Verb (not including auxiliary)	110.01	11.08	105.59	15.00	112.49	14.11	120.76	10.45
Present tense verb	73.56	10.19	70.07	12.71	75.09	13.21	85.84	10.82
Past tense verb	11.36	3.95	9.93	4.93	11.11	4.35	10.75	3.47
Progressive aspect verb	7.54	3.24	7.47	3.00	6.85	2.31	7.45	2.91
Perfect aspect verb	8.15	2.52	8.24	2.23	7.69	2.33	8.70	2.57
Infinitive verb	16.99	4.16	16.63	3.27	17.53	3.26	19.41	2.83
Pro-verb 'do'	0.13	0.34	0.33	0.47	0.51	0.53	0.60	0.71
Verb 'be'	1.36	1.02	2.01	1.64	2.24	1.20	2.48	1.99
Verb 'have'	1.29	1.29	1.01	0.99	1.21	1.18	1.79	0.95
Private verb	5.96	1.73	7.81	2.33	7.32	2.71	8.18	3.06
Public verb	2.15	1.59	1.97	0.92	2.61	1.04	2.56	1.37
Suasive verb	1.25	0.99	1.51	0.99	1.43	0.78	1.51	1.30
Activity verb	15.37	3.39	12.97	6.06	16.48	4.18	17.67	3.71
Communication verb	8.78	2.65	7.81	4.21	8.35	1.85	7.69	2.22
Mental verb	11.92	2.99	9.77	5.01	12.29	4.00	15.13	4.58
Causative verb	3.30	1.48	3.78	2.09	3.77	1.73	5.00	1.88
Occurrence verb	2.45	1.74	2.39	1.97	2.82	1.60	2.76	1.46
Existence verb	7.95	3.27	5.97	3.37	8.30	2.75	7.38	2.45
Aspectual verb	2.26	1.50	2.39	1.97	2.60	1.33	3.19	1.89
All modals	10.76	3.59	11.97	2.77	14.06	3.91	17.80	4.91
Modal of prediction	4.67	2.08	5.91	3.06	5.71	2.04	6.74	2.99
Modal of possibility	2.74	2.04	2.45	1.99	3.43	1.90	4.68	2.59
Modal of necessity	3.34	1.87	3.61	2.12	4.90	2.57	6.36	2.48
Split auxiliary-verb	4.05	2.11	3.63	1.63	3.86	1.49	4.40	1.86
All passives	9.75	3.06	8.31	2.88	9.43	2.29	7.85	2.97
Agentless passive verb	6.44	2.24	5.22	1.97	5.95	2.07	5.70	2.42

Feature	Low HDI		Medium HDI		High HDI		Very High HDI	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Passive verb + by	1.13	1.25	1.33	0.91	1.48	0.89	1.11	0.77
Passive postnominal modifier	2.18	1.12	1.77	1.08	2.00	1.03	1.05	0.72
All personal pronouns	43.82	10.93	39.98	14.03	49.97	15.45	43.75	11.97
1st person pronoun	33.23	9.51	30.99	12.37	40.28	13.95	34.00	9.10
2nd person pronoun	4.01	2.57	3.65	1.97	2.54	1.80	2.42	2.91
3rd person pronoun	6.57	3.11	5.35	4.24	7.14	3.32	7.34	3.55
Pronoun 'it'	9.11	3.50	7.75	2.50	9.28	3.27	11.10	3.38
Demonstrative pronoun	3.59	2.10	2.43	2.12	3.11	1.75	4.59	2.42
Nominal pronoun	1.37	1.28	1.00	1.06	1.80	1.50	2.03	1.19
All nouns	314.15	31.69	313.06	23.07	303.39	21.93	290.96	20.10
Common noun	260.35	20.45	252.79	21.36	249.74	15.69	243.98	19.00
Proper noun	39.18	15.67	45.36	13.25	41.14	11.21	37.13	11.03
Animate noun	5.53	2.15	4.24	2.91	5.04	2.46	5.20	2.11
Process noun	15.85	4.10	15.15	7.77	14.56	4.67	15.62	4.29
Cognitive noun	3.21	1.71	3.01	2.05	4.19	1.60	4.04	1.64
Abstract noun	23.27	4.35	20.45	8.91	24.30	5.51	25.10	5.84
Concrete noun	3.71	2.31	3.89	2.43	4.27	1.56	5.08	2.42
Technical noun	1.66	1.19	1.67	1.20	1.67	1.04	1.86	1.80
Quantity noun	5.41	3.25	4.31	2.33	5.96	2.79	6.60	2.78
Place noun	7.12	2.34	5.75	3.77	7.55	3.44	5.52	2.18
Group noun	2.52	2.41	1.61	1.33	2.54	1.33	2.05	1.19
Noun-noun sequence	32.71	10.37	30.57	9.09	26.97	8.71	27.97	8.78
Stance noun + preposition	2.19	1.39	2.14	1.97	1.71	1.34	1.62	1.28
Determiner + stance noun	0.61	0.61	0.29	0.48	0.46	0.71	0.69	0.84
All indefinite articles	15.57	4.24	14.55	6.30	18.17	5.30	19.90	3.96
All definite articles	66.05	9.79	64.87	30.41	65.76	11.57	63.21	10.81
Preposition	127.59	10.27	132.79	9.58	126.46	15.84	117.85	10.16
Stranded preposition	0.68	0.89	0.38	0.39	0.47	0.58	0.62	0.68

Feature	Low HDI		Medium HDI		High HDI		Very High HDI	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All adjectives	91.03	11.64	89.63	12.16	89.72	12.38	84.56	12.63
Attributive adjective	71.15	11.32	70.37	11.52	68.51	13.65	64.26	10.60
Attributive adjective – Size	1.31	0.78	1.83	2.30	2.01	1.87	1.38	1.00
Attributive adjective – Time	1.58	1.76	1.44	1.54	1.89	1.38	1.52	1.17
Attributive adjective – Color	0.00	0.00	0.10	0.22	0.04	0.15	0.07	0.27
Attributive adjective – Evaluative	1.15	1.27	1.21	0.87	1.19	1.00	1.59	1.05
Attributive adjective – Relational	2.12	1.34	2.03	1.53	2.48	1.18	2.33	1.61
Attributive adjective – Topical	10.48	5.06	7.60	3.81	9.35	4.38	8.97	3.71
Predicative adjective	4.82	1.81	4.17	1.60	4.49	1.62	4.90	2.10
Predicative attitudinal adjective	1.33	0.62	1.11	0.97	1.17	0.90	1.56	1.52
Predicative epistemic adjective	0.25	0.35	0.23	0.36	0.47	0.59	0.70	0.74
All adverbs	28.08	7.30	28.71	8.04	30.37	5.61	31.99	7.31
Adverb – time	3.78	1.82	3.96	2.59	3.90	1.75	4.32	1.99
Adverb – place	4.52	2.14	4.03	2.40	3.82	2.28	4.68	1.98
All stance adverbs	1.66	2.07	1.31	1.21	1.63	1.02	2.48	1.82
Non-factive adverb	0.21	0.38	0.23	0.40	0.06	0.16	0.09	0.23
Attitudinal adverb	0.21	0.47	0.20	0.35	0.17	0.33	0.40	0.47
Factive adverb	1.20	1.65	0.77	1.08	1.22	1.00	1.70	1.38
Likelihood adverb	0.03	0.14	0.14	0.37	0.17	0.44	0.29	0.46
Adverb – downtoner	1.78	2.09	1.08	0.74	1.49	1.03	2.33	1.75
Amplifier	1.80	1.63	1.49	1.08	1.28	0.94	2.30	1.52
Emphatic	1.58	1.17	1.87	1.47	1.94	1.05	2.76	1.59
Hedge	0.16	0.29	0.31	0.48	0.14	0.29	0.19	0.34
Discourse particle	0.41	0.87	0.17	0.34	0.19	0.37	0.32	0.38
'That' deletion	1.09	0.95	0.67	0.77	0.98	0.82	0.97	0.90
Contraction	0.55	0.77	0.89	1.46	1.20	1.19	1.76	3.26
All conjunctions	16.34	5.19	14.19	4.69	19.55	6.54	17.95	6.77
Subordinating conjunction – causative	0.28	0.53	0.14	0.26	0.72	0.99	0.59	0.86

Feature	Low HDI		Medium HDI		High HDI		Very High HDI	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Subordinating conjunction – conditional	0.76	0.99	0.74	0.68	1.41	1.15	1.37	1.37
Coordinating conjunction – clausal	5.77	5.09	5.98	3.14	8.33	5.40	9.06	6.00
Coordinating conjunction – phrasal	5.72	2.57	4.29	1.99	5.20	1.93	3.86	2.37
Adverbial – conjuncts	5.26	3.01	4.40	2.19	4.89	2.40	3.85	2.38
All wh- words	0.19	0.30	0.36	0.55	0.40	0.81	0.65	0.65
Wh question	0.08	0.24	0.29	0.53	0.34	0.81	0.36	0.55
Wh clause	0.12	0.23	0.07	0.20	0.06	0.20	0.29	0.36
All wh- relative clauses	3.48	1.90	4.02	2.07	3.81	2.31	3.46	2.05
Wh pronoun – object relative clause	0.79	0.81	0.87	0.66	1.00	0.96	0.64	0.68
Wh pronoun – subject relative clause	2.02	1.25	2.65	1.46	2.29	1.41	2.23	1.47
Wh pro. – rel. cls. with preposition fronting	0.68	0.63	0.52	0.57	0.53	0.58	0.56	0.55
Wh CCC communicaton verb	0.00	0.00	0.07	0.18	0.06	0.18	0.08	0.19
Wh CCC attitudinal verb	0.03	0.11	0.00	0.00	0.00	0.00	0.09	0.25
Wh CCC factive verb	0.06	0.19	0.07	0.20	0.06	0.16	0.22	0.31
Wh CCC likelihood verb	0.00	0.00	0.03	0.10	0.02	0.10	0.06	0.17
'That' relative clause	3.29	2.78	3.19	2.69	3.31	1.44	3.86	2.02
'That' CCC verb	2.99	1.29	3.80	1.60	2.97	1.45	3.36	2.14
'That' CCC communication verb	1.66	1.16	1.23	0.83	1.37	0.97	1.07	0.92
'That' CCC attitudinal verb	1.76	1.17	1.56	1.14	1.68	1.07	1.55	0.98
'That' CCC factive verb	1.20	0.96	0.86	0.80	1.05	0.71	1.26	1.07
'That' CCC likelihood verb	0.79	0.69	1.03	1.20	0.86	0.80	0.98	0.90
'That' CCC attitudinal adjective	0.48	0.53	0.10	0.28	0.26	0.39	0.25	0.40
'That' CCC factive adjective	0.51	0.68	0.25	0.44	0.21	0.30	0.26	0.37
'That' CCC attitudinal noun	0.05	0.16	0.17	0.40	0.13	0.33	0.12	0.35
'That' CCC factive noun	0.34	0.44	0.15	0.29	0.31	0.35	0.50	0.82
'That' CCC likelihood noun	0.06	0.19	0.05	0.12	0.25	0.37	0.14	0.32
All stance 'that' CCC verbs	5.42	2.00	4.66	2.57	4.94	1.86	4.87	2.25
All stance 'that' CCC adjectives	0.98	1.04	0.36	0.53	0.47	0.51	0.52	0.53

Feature	Low HDI		Medium HDI		High HDI		Very High HDI	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All stance 'that' CCC nouns	0.45	0.44	0.37	0.54	0.70	0.59	0.76	1.10
All stance 'that' complement clauses	6.88	2.46	5.37	3.23	6.12	2.21	6.17	3.28
'To' CCC speech act verb	0.14	0.29	0.14	0.25	0.28	0.50	0.10	0.23
'To' CCC verb of cognition	0.32	0.34	0.19	0.31	0.27	0.33	0.21	0.31
'To' CCC verb of desire	2.84	1.49	2.07	1.44	2.04	1.14	2.73	1.52
'To' CCC verb of effort	1.92	1.19	1.13	0.94	1.57	0.93	2.03	1.16
'To' CCC probability verb	0.03	0.14	0.03	0.10	0.14	0.25	0.11	0.27
'To' CCC adjective of certainty	0.00	0.00	0.00	0.00	0.01	0.06	0.00	0.00
'To' CCC adjective of ability	0.11	0.27	0.11	0.25	0.07	0.20	0.20	0.40
'To' CCC adjective of affect	0.03	0.14	0.00	0.00	0.00	0.00	0.00	0.00
'To' CCC adjective of ease	0.00	0.00	0.03	0.13	0.02	0.10	0.02	0.10
'To' CCC evaluative adjective	0.14	0.32	0.08	0.21	0.11	0.29	0.07	0.27
'To' CCC stance noun	2.85	1.53	2.53	1.89	3.51	1.50	3.76	1.87
All stance 'to' CCC verb	5.25	1.59	3.57	1.87	4.30	1.80	5.17	1.60
All stance 'to' CCC adjective	0.27	0.40	0.23	0.30	0.21	0.42	0.29	0.45
All stance 'to' complement clauses	8.35	2.89	6.31	3.14	8.02	2.46	9.23	2.50
Dimension 1	-17.80	4.70	-18.44	4.34	-15.28	6.60	-11.79	5.25
Dimension 2	-3.10	0.65	-3.21	0.62	-3.09	0.65	-2.90	0.60
Dimension 3	1.18	1.76	0.91	1.02	1.15	1.93	-0.02	1.86
Dimension 4	-1.54	1.23	-1.21	1.36	-0.26	1.60	0.90	1.97
Dimension 5	2.21	1.81	1.51	1.62	2.11	1.86	0.84	1.37

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX F: TAGGED FEATURES BY GROSS DOMESTIC PRODUCT

Feature	Low GDP		Medium GDP		High GDP		Very High GDP	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Verb (not including auxiliary)	110.85	13.71	114.42	9.13	112.65	16.47	124.76	9.77
Present tense verb	74.72	12.75	78.67	10.99	76.40	11.21	89.52	11.41
Past tense verb	11.02	4.13	9.36	3.80	10.90	4.12	11.31	4.09
Progressive aspect verb	7.34	2.89	6.27	1.52	7.00	3.19	8.17	3.21
Perfect aspect verb	8.20	2.41	8.08	2.80	8.58	3.29	8.64	2.53
Infinitive verb	16.78	3.49	19.96	2.68	19.05	3.28	19.59	2.78
Pro-verb 'do'	0.35	0.50	0.45	0.42	0.83	0.60	0.70	0.82
Verb 'be'	1.93	1.33	2.27	1.49	2.18	2.02	2.74	2.25
Verb 'have'	1.35	1.22	1.73	0.88	1.00	0.97	1.85	1.09
Private verb	7.23	2.51	6.77	1.75	7.88	2.58	9.19	3.15
Public verb	2.22	1.23	2.52	1.17	2.88	1.20	2.49	1.48
Suasive verb	1.42	0.94	1.59	0.84	2.70	1.75	1.24	1.20
Activity verb	15.52	5.08	15.73	3.56	15.93	2.25	19.37	3.19
Communication verb	8.35	2.90	8.00	2.18	7.75	2.57	7.51	2.20
Mental verb	11.86	4.11	11.97	3.18	12.73	4.06	16.84	5.03
Causative verb	3.61	1.77	4.44	1.70	5.30	2.77	5.06	1.82
Occurrence verb	2.67	1.73	3.22	1.40	2.08	0.70	2.65	1.49
Existence verb	7.43	3.26	7.82	1.54	7.43	2.84	7.51	2.91
Aspectual verb	2.37	1.45	3.93	2.28	2.18	1.92	3.14	1.26
All modals	13.25	4.37	16.13	4.20	12.98	4.70	18.79	5.53
Modal of prediction	5.55	2.37	6.38	2.70	5.20	1.97	7.42	3.06
Modal of possibility	3.26	2.15	3.09	1.47	3.58	1.39	5.56	3.17
Modal of necessity	4.44	2.78	6.61	2.44	4.20	2.77	5.81	2.44
Split auxiliary-verb	3.90	1.70	3.72	2.34	4.13	1.65	4.56	1.61
All passives	8.95	2.64	9.78	3.27	7.45	3.06	7.76	2.40
Agentless passive verb	5.77	2.08	6.90	2.70	5.48	2.26	5.72	1.85

Feature	Low GDP		Medium GDP		High GDP		Very High GDP	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Passive verb + by	1.27	1.00	1.37	0.82	0.73	0.56	1.09	0.72
Passive postnominal modifier	1.91	1.11	1.52	0.74	1.28	0.51	0.94	0.84
All personal pronouns	46.47	14.64	42.00	17.70	48.43	14.45	47.09	8.75
1st person pronoun	36.53	12.75	33.53	15.39	38.68	11.04	36.63	6.90
2nd person pronoun	3.39	2.35	1.73	1.02	1.58	1.23	3.24	3.64
3rd person pronoun	6.54	3.63	6.78	3.97	8.18	4.30	7.23	2.64
Pronoun 'it'	8.64	3.16	9.95	2.85	11.75	2.90	11.43	3.62
Demonstrative pronoun	3.14	2.00	2.90	1.57	5.25	3.24	5.04	2.26
Nominal pronoun	1.54	1.35	1.63	0.98	1.48	0.80	2.24	1.39
All nouns	306.83	25.81	300.25	21.67	293.05	26.38	289.09	20.96
Common noun	252.44	18.99	253.90	16.46	243.78	9.08	237.26	18.87
Proper noun	40.66	13.27	36.12	9.70	35.78	15.54	41.37	11.59
Animate noun	5.03	2.56	4.08	2.14	6.55	2.04	5.56	2.40
Process noun	14.65	5.43	15.51	4.49	15.98	2.52	15.94	4.20
Cognitive noun	3.49	1.90	3.92	1.32	4.45	1.50	4.19	1.84
Abstract noun	22.75	6.27	27.06	6.43	21.83	5.42	23.80	4.92
Concrete noun	4.21	2.35	4.05	1.97	4.10	1.13	5.58	2.60
Technical noun	1.66	1.13	1.56	1.32	1.05	0.45	2.36	2.02
Quantity noun	5.66	2.99	6.33	3.18	4.55	1.30	7.04	2.57
Place noun	6.96	3.16	6.11	2.87	6.20	2.62	5.39	2.42
Group noun	2.42	1.74	1.92	1.41	1.85	1.57	2.31	1.12
Noun-noun sequence	29.46	9.15	31.11	7.88	26.18	4.50	26.27	9.42
Stance noun + preposition	1.94	1.55	1.83	1.07	1.95	1.72	1.10	1.10
Determiner + stance noun	0.48	0.62	0.59	0.83	0.33	0.38	0.77	0.87
All indefinite articles	17.05	5.70	17.67	2.86	20.53	8.30	19.37	3.42
All definite articles	64.24	18.11	67.17	11.35	65.85	15.95	62.10	10.29
Preposition	127.60	13.70	122.38	11.04	123.50	6.83	115.35	11.25
Stranded preposition	0.54	0.66	0.60	0.59	0.23	0.26	0.57	0.78

Feature	Low GDP		Medium GDP		High GDP		Very High GDP	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All adjectives	89.93	11.79	89.15	11.84	88.58	9.92	81.22	9.23
Attributive adjective	69.58	12.15	68.66	10.31	66.58	6.18	60.85	8.36
Attributive adjective – Size	1.74	1.65	2.01	2.27	1.08	0.25	1.66	1.06
Attributive adjective – Time	1.72	1.55	1.60	1.09	1.63	0.84	1.51	1.25
Attributive adjective – Color	0.05	0.15	0.11	0.38	0.00	0.00	0.04	0.09
Attributive adjective – Evaluative	1.25	1.06	1.26	1.10	0.98	0.92	1.69	1.05
Attributive adjective – Relational	2.31	1.33	2.21	1.53	2.38	0.62	2.29	1.73
Attributive adjective – Topical	8.96	4.27	10.61	4.72	8.60	2.99	7.75	2.96
Predicative adjective	4.70	1.72	4.23	2.08	4.38	2.34	5.12	1.84
Predicative attitudinal adjective	1.20	0.83	1.63	1.71	1.43	1.26	1.65	1.32
Predicative epistemic adjective	0.37	0.49	0.41	0.46	0.35	0.70	0.92	0.81
Adverb – time	4.02	2.12	3.29	2.16	4.28	1.18	4.78	1.73
Adverb – place	4.20	2.25	3.55	1.77	3.35	1.69	5.18	2.27
All stance adverbs	1.57	1.46	1.51	1.08	2.38	1.53	2.78	2.03
Non-factive adverb	0.14	0.31	0.06	0.14	0.00	0.00	0.14	0.29
Attitudinal adverb	0.20	0.38	0.27	0.37	0.60	0.47	0.34	0.55
Factive adverb	1.10	1.24	1.02	0.89	1.45	1.60	2.08	1.49
Likelihood adverb	0.13	0.37	0.14	0.35	0.33	0.65	0.21	0.36
Adverb – downtoner	1.70	1.49	1.65	1.19	1.78	1.68	2.52	2.12
Amplifier	1.62	1.33	1.80	0.75	2.60	1.53	2.19	1.90
Emphatic	1.86	1.24	2.27	1.54	1.85	1.47	2.61	1.00
Hedge	0.21	0.36	0.16	0.29	0.13	0.25	0.17	0.37
Discourse particle	0.21	0.55	0.09	0.17	0.80	0.70	0.40	0.41
'That' deletion	0.93	0.85	0.82	0.67	1.03	0.92	1.19	0.97
Contraction	1.04	1.23	0.54	0.57	0.93	0.74	2.67	4.17
All conjunctions	17.40	6.39	17.24	3.77	16.15	2.87	18.96	8.46
Subordinating conjunction – causative	0.38	0.72	0.58	0.86	0.25	0.30	0.86	1.02
Subordinating conjunction – conditional	1.07	1.06	1.33	1.43	1.20	0.90	1.74	1.40

Feature	Low GDP		Medium GDP		High GDP		Very High GDP	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Coordinating conjunction – clausal	7.46	5.36	6.15	2.37	6.15	2.02	10.85	6.73
Coordinating conjunction – phrasal	4.99	1.96	5.18	1.62	4.50	1.26	2.87	2.37
Adverbial – conjuncts	4.98	2.61	4.18	2.43	5.15	2.33	3.14	1.23
All wh- words	0.38	0.65	0.54	0.63	0.40	0.29	0.52	0.62
Wh question	0.26	0.64	0.44	0.64	0.10	0.20	0.23	0.37
Wh clause	0.11	0.26	0.10	0.23	0.30	0.36	0.29	0.40
All wh- relative clauses	3.55	1.97	3.41	2.90	5.05	1.75	3.63	1.72
Wh pronoun – object relative clause	0.89	0.78	0.59	0.91	1.65	0.64	0.59	0.63
Wh pronoun – subject relative clause	2.10	1.31	2.16	1.82	2.90	0.91	2.40	1.33
Wh pro. – rel. cls. with preposition fronting	0.56	0.59	0.63	0.70	0.45	0.53	0.63	0.43
Wh CCC communicaton verb	0.04	0.14	0.08	0.20	0.10	0.20	0.06	0.19
Wh CCC attitudinal verb	0.01	0.07	0.05	0.17	0.00	0.00	0.11	0.29
Wh CCC factive verb	0.10	0.24	0.10	0.23	0.30	0.36	0.18	0.30
Wh CCC likelihood verb	0.04	0.14	0.00	0.00	0.00	0.00	0.05	0.13
'That' relative clause	3.28	2.25	3.77	1.43	3.45	2.85	3.91	2.16
'That' CCC verb	3.18	1.47	3.02	1.25	4.73	1.16	3.39	2.52
'That' CCC communication verb	1.45	1.00	1.19	0.76	1.43	1.37	0.96	0.89
'That' CCC attitudinal verb	1.62	1.08	2.13	1.04	2.73	1.23	1.21	0.72
'That' CCC factive verb	0.99	0.80	1.23	0.87	0.73	0.59	1.69	1.13
'That' CCC likelihood verb	0.92	0.89	0.53	0.70	1.40	1.15	0.99	0.87
'That' CCC attitudinal adjective	0.31	0.45	0.33	0.51	0.43	0.31	0.21	0.31
'That' CCC factive adjective	0.30	0.51	0.33	0.30	0.45	0.53	0.21	0.36
'That' CCC attitudinal noun	0.10	0.31	0.07	0.16	0.45	0.31	0.11	0.43
'That' CCC factive noun	0.28	0.38	0.25	0.28	0.98	1.69	0.45	0.66
'That' CCC likelihood noun	0.15	0.30	0.11	0.21	0.00	0.00	0.19	0.40
All stance 'that' CCC verbs	4.96	2.08	5.08	1.57	6.23	1.32	4.88	2.70
All stance 'that' CCC adjectives	0.61	0.77	0.67	0.57	0.90	0.32	0.42	0.47
All stance 'that' CCC nouns	0.54	0.57	0.43	0.30	1.43	1.72	0.76	1.12

Feature	Low GDP		Medium GDP		High GDP		Very High GDP	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
All stance 'that' complement clauses	6.11	2.60	6.21	1.84	8.53	2.85	6.09	3.73
'To' CCC speech act verb	0.23	0.40	0.10	0.23	0.00	0.00	0.09	0.23
'To' CCC verb of cognition	0.28	0.35	0.32	0.40	0.50	0.36	0.14	0.26
'To' CCC verb of desire	2.26	1.30	1.99	1.49	2.58	1.53	2.96	1.45
'To' CCC verb of effort	1.46	1.03	2.34	0.85	2.65	1.42	1.91	1.07
'To' CCC probability verb	0.06	0.17	0.11	0.25	0.15	0.30	0.10	0.28
'To' CCC adjective of certainty	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00
'To' CCC adjective of ability	0.11	0.24	0.11	0.25	0.00	0.00	0.32	0.49
'To' CCC adjective of affect	0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00
'To' CCC adjective of ease	0.03	0.12	0.00	0.00	0.00	0.00	0.04	0.13
'To' CCC evaluative adjective	0.11	0.26	0.04	0.14	0.10	0.20	0.09	0.35
'To' CCC stance noun	2.94	1.69	4.44	1.79	3.48	1.08	3.51	1.81
All stance 'to' CCC verb	4.29	1.87	4.86	1.74	5.85	1.68	5.21	1.66
All stance 'to' CCC adjective	0.26	0.42	0.15	0.27	0.10	0.20	0.45	0.53
All stance 'to' complement clauses	7.48	2.85	9.43	2.95	9.38	2.06	9.17	2.37
Dimension 1	-16.31	5.81	-15.65	4.67	-14.26	3.55	-9.97	5.38
Dimension 2	-3.09	0.62	-3.10	0.70	-2.83	0.61	-2.91	0.61
Dimension 3	0.83	1.55	1.24	2.21	1.35	0.99	-0.50	1.76
Dimension 4	-0.67	1.76	0.68	2.10	-0.29	2.35	0.95	2.15
Dimension 5	1.95	1.83	1.58	1.49	1.36	1.47	0.35	0.99

Note. CCC = complement clause controlled by.

APPENDIX G: TAGGED FEATURES BY MILITARY SPENDING

Feature	Low Military		High Military	
	Mean	**SD**	**Mean**	**SD**
Verb (not including auxiliary)	113.26	13.77	111.91	13.75
Present tense verb	77.75	13.05	75.49	14.03
Past tense verb	10.58	4.01	11.26	4.03
Progressive aspect verb	7.01	2.59	8.39	3.36
Perfect aspect verb	7.96	2.53	8.90	2.33
Infinitive verb	17.87	3.51	16.59	3.42
Pro-verb 'do'	0.40	0.53	0.56	0.72
Verb 'be'	2.20	1.61	1.69	1.07
Verb 'have'	1.51	1.21	1.08	0.69
Private verb	7.31	2.51	7.77	3.19
Public verb	2.22	1.31	2.68	1.07
Suasive verb	1.45	1.07	1.43	0.84
Activity verb	15.88	4.78	17.17	4.38
Communication verb	8.16	2.81	7.68	1.80
Mental verb	12.77	4.55	12.29	3.78
Causative verb	3.91	1.79	4.25	2.23
Occurrence verb	2.76	1.68	2.29	1.14
Existence verb	7.48	3.10	7.28	2.34
Aspectual verb	2.58	1.67	2.78	1.64
All modals	14.53	5.12	13.83	4.72
Modal of prediction	5.93	2.54	5.41	2.77
Modal of possibility	3.49	2.13	4.03	3.05
Modal of necessity	5.09	2.83	4.39	2.50
Split auxiliary-verb	3.93	1.73	4.33	1.99
All passives	9.00	2.76	8.20	2.94
Agentless passive verb	6.03	2.24	5.42	1.75
Passive verb + by	1.28	0.96	1.06	0.87
Passive postnominal modifier	1.70	1.05	1.69	1.16
All personal pronouns	45.20	13.85	46.16	16.28
1st person pronoun	35.63	11.92	35.37	14.36
2nd person pronoun	2.86	2.17	3.64	3.41
3rd person pronoun	6.72	3.48	7.16	3.62
Pronoun 'it'	9.46	3.44	9.42	2.82
Demonstrative pronoun	3.46	2.02	3.59	2.74
Nominal pronoun	1.55	1.22	1.86	1.55
All nouns	302.67	26.70	304.49	19.32
Common noun	250.42	19.31	250.16	18.61
Proper noun	39.26	12.52	42.75	14.17

Feature	Low Military		High Military	
	Mean	**SD**	**Mean**	**SD**
Animate noun	4.87	2.40	5.84	2.46
Process noun	15.68	5.10	13.47	4.21
Cognitive noun	3.85	1.92	3.28	1.63
Abstract noun	23.74	6.36	22.87	5.19
Concrete noun	4.32	2.42	4.44	2.17
Technical noun	1.60	1.30	2.16	1.30
Quantity noun	5.84	2.98	5.91	2.40
Place noun	6.12	2.95	7.89	2.84
Group noun	2.16	1.42	2.79	2.12
Noun-noun sequence	29.45	9.41	26.77	8.25
Stance noun + preposition	1.82	1.50	1.91	1.31
Determiner + stance noun	0.49	0.65	0.78	0.79
All indefinite articles	17.79	5.38	17.33	4.79
All definite articles	64.93	16.62	64.03	12.58
Preposition	124.46	13.24	126.63	14.51
Stranded preposition	0.58	0.67	0.37	0.52
All adjectives	88.92	12.70	88.99	11.11
Attributive adjective	68.13	12.39	69.83	10.48
Attributive adjective – Size	1.89	1.68	1.05	0.90
Attributive adjective – Time	1.61	1.48	1.73	1.07
Attributive adjective – Color	0.05	0.20	0.05	0.12
Attributive adjective – Evaluative	1.33	1.12	1.21	0.73
Attributive adjective – Relational	2.18	1.41	2.74	1.14
Attributive adjective – Topical	9.37	4.45	8.63	3.35
Predicative adjective	4.66	1.84	4.78	1.89
Predicative attitudinal adjective	1.32	1.01	1.36	1.32
Predicative epistemic adjective	0.43	0.58	0.53	0.57
All adverbs	30.26	7.30	30.02	5.67
Adverb – time	3.95	2.01	4.34	2.04
Adverb – place	3.81	1.84	5.76	2.63
All stance adverbs	1.81	1.51	1.82	1.86
Non-factive adverb	0.12	0.29	0.16	0.27
Attitudinal adverb	0.28	0.46	0.19	0.33
Factive adverb	1.22	1.22	1.39	1.54
Likelihood adverb	0.19	0.41	0.08	0.26
Adverb – downtoner	1.81	1.68	1.73	0.94
Amplifier	1.77	1.44	1.73	1.20
Emphatic	2.02	1.31	2.48	1.58
Hedge	0.19	0.34	0.18	0.32

Feature	Low Military		High Military	
	Mean	**SD**	**Mean**	**SD**
Discourse particle	0.29	0.55	0.17	0.29
'That' deletion	0.86	0.82	1.19	0.88
Contraction	1.00	1.19	2.04	3.72
All conjunctions	17.88	6.37	16.90	5.47
Subordinating conjunction – causative	0.53	0.81	0.28	0.60
Subordinating conjunction – conditional	1.20	1.20	1.03	1.04
Coordinating conjunction – clausal	7.85	5.32	7.58	5.59
Coordinating conjunction – phrasal	4.87	2.21	4.55	2.56
Adverbial – conjuncts	4.61	2.70	4.77	1.86
All wh- words	0.42	0.65	0.43	0.57
Wh question	0.29	0.61	0.23	0.47
Wh clause	0.14	0.29	0.20	0.31
All wh- relative clauses	3.67	2.08	3.58	2.23
Wh pronoun – object relative clause	0.83	0.80	0.88	0.82
Wh pronoun – subject relative clause	2.23	1.45	2.30	1.43
Wh pro. – rel. cls. with preposition fronting	0.61	0.58	0.39	0.48
Wh CCC communicaton verb	0.05	0.16	0.05	0.17
Wh CCC attitudinal verb	0.04	0.16	0.00	0.00
Wh CCC factive verb	0.12	0.26	0.14	0.26
Wh CCC likelihood verb	0.04	0.15	0.02	0.07
'That' relative clause	3.53	2.22	2.87	1.80
'That' CCC verb	3.16	1.49	3.56	2.16
'That' CCC communication verb	1.30	1.00	1.27	0.89
'That' CCC attitudinal verb	1.71	1.04	1.51	1.06
'That' CCC factive verb	1.11	0.78	1.08	1.25
'That' CCC likelihood verb	0.76	0.81	1.53	0.93
'That' CCC attitudinal adjective	0.27	0.40	0.41	0.53
'That' CCC factive adjective	0.31	0.49	0.24	0.29
'That' CCC attitudinal noun	0.11	0.32	0.13	0.26
'That' CCC factive noun	0.32	0.55	0.32	0.44
'That' CCC likelihood noun	0.12	0.28	0.21	0.34
All stance 'that' CCC verbs	4.87	2.00	5.37	2.32
All stance 'that' CCC adjectives	0.58	0.72	0.65	0.55
All stance 'that' CCC nouns	0.55	0.75	0.66	0.71
All stance 'that' complement clauses	6.01	2.69	6.69	2.89
'To' CCC speech act verb	0.17	0.36	0.17	0.26
'To' CCC verb of cognition	0.28	0.36	0.24	0.29
'To' CCC verb of desire	2.45	1.46	2.16	1.16
'To' CCC verb of effort	1.64	1.15	1.78	0.80

Feature	Low Military		High Military	
	Mean	**SD**	**Mean**	**SD**
'To' CCC probability verb	0.08	0.21	0.09	0.22
'To' CCC adjective of certainty	0.00	0.00	0.02	0.07
'To' CCC adjective of ability	0.17	0.33	0.03	0.12
'To' CCC adjective of affect	0.01	0.07	0.00	0.00
'To' CCC adjective of ease	0.03	0.12	0.00	0.00
'To' CCC evaluative adjective	0.11	0.30	0.07	0.17
'To' CCC stance noun	3.29	1.85	3.23	1.48
All stance 'to' CCC verb	4.60	1.96	4.43	1.12
All stance 'to' CCC adjective	0.32	0.46	0.12	0.23
All stance 'to' complement clauses	8.20	2.97	7.79	2.08
Dimension 1	-15.23	5.70	-15.38	6.86
Dimension 2	-3.14	0.65	-2.83	0.53
Dimension 3	0.96	1.67	0.04	2.04
Dimension 4	-0.19	2.11	-0.57	1.28
Dimension 5	1.70	1.84	1.56	1.55

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

Feature	Least Developed Country	
	Mean	**SD**
Verb (not including auxiliary)	111.19	10.61
Present tense verb	76.20	10.56
Past tense verb	10.95	4.11
Progressive aspect verb	7.88	3.57
Perfect aspect verb	8.47	2.87
Infinitive verb	15.91	3.54
Pro-verb 'do'	0.23	0.50
Verb 'be'	1.91	1.43
Verb 'have'	1.36	1.26
Private verb	6.68	2.31
Public verb	1.91	1.51
Suasive verb	1.26	0.84
Activity verb	15.52	5.07
Communication verb	7.94	3.57
Mental verb	12.16	4.33
Causative verb	3.27	1.61
Occurrence verb	2.52	1.69
Existence verb	6.21	3.19
Aspectual verb	2.41	1.27
All modals	12.62	4.44
Modal of prediction	5.75	2.51
Modal of possibility	2.77	2.34
Modal of necessity	4.09	2.66
Split auxiliary-verb	3.84	1.77
All passives	9.53	3.33
Agentless passive verb	6.34	2.40
Passive verb + by	1.08	1.13
Passive postnominal modifier	2.09	1.29
All personal pronouns	46.81	15.32
1st person pronoun	35.57	12.83
2nd person pronoun	4.68	2.79
3rd person pronoun	6.56	2.83
Pronoun 'it'	8.96	3.51
Demonstrative pronoun	3.29	2.18
Nominal pronoun	1.36	1.28
All nouns	308.63	30.24
Common noun	255.48	20.71
Proper noun	38.56	14.67

Feature	Least Developed Country	
	Mean	**SD**
Animate noun	5.69	2.71
Process noun	14.96	5.75
Cognitive noun	3.09	2.28
Abstract noun	22.77	6.71
Concrete noun	3.66	2.78
Technical noun	1.41	1.27
Quantity noun	6.09	3.16
Place noun	6.31	2.84
Group noun	2.71	2.36
Noun-noun sequence	31.24	9.13
Stance noun + preposition	1.79	1.41
Determiner + stance noun	0.45	0.56
All indefinite articles	15.91	5.62
All definite articles	58.94	18.68
Preposition	126.59	11.88
Stranded preposition	0.59	0.87
All adjectives	90.53	11.48
Attributive adjective	69.22	11.38
Attributive adjective – Size	1.52	1.24
Attributive adjective – Time	1.58	1.91
Attributive adjective – Color	0.04	0.16
Attributive adjective – Evaluative	1.24	1.38
Attributive adjective – Relational	1.78	1.31
Attributive adjective – Topical	9.49	5.69
Predicative adjective	5.57	1.61
Predicative attitudinal adjective	1.29	0.77
Predicative epistemic adjective	0.29	0.41
All adverbs	29.98	6.90
Adverb – time	4.30	1.83
Adverb – place	3.67	2.23
All stance adverbs	1.94	2.07
Non-factive adverb	0.21	0.44
Attitudinal adverb	0.30	0.51
Factive adverb	1.32	1.66
Likelihood adverb	0.10	0.34
Adverb – downtoner	1.64	2.17
Amplifier	1.86	1.65
Emphatic	1.92	1.49
Hedge	0.21	0.40

Feature	Least Developed Country	
	Mean	**SD**
Discourse particle	0.38	0.89
'That' deletion	0.93	0.85
Contraction	0.98	1.48
All conjunctions	16.78	5.93
Subordinating conjunction – causative	0.36	0.59
Subordinating conjunction – conditional	0.98	1.05
Coordinating conjunction – clausal	6.56	5.68
Coordinating conjunction – phrasal	5.28	2.80
Adverbial – conjuncts	5.07	3.05
All wh- words	0.26	0.39
Wh question	0.16	0.32
Wh clause	0.09	0.29
All wh- relative clauses	3.07	1.58
Wh pronoun – object relative clause	0.67	0.56
Wh pronoun – subject relative clause	1.84	1.21
Wh pro. – rel. cls. with preposition fronting	0.56	0.63
Wh CCC communicaton verb	0.02	0.09
Wh CCC attitudinal verb	0.03	0.12
Wh CCC factive verb	0.09	0.29
Wh CCC likelihood verb	0.03	0.14
'That' relative clause	2.99	2.22
'That' CCC verb	3.19	1.56
'That' CCC communication verb	1.54	1.30
'That' CCC attitudinal verb	1.73	1.14
'That' CCC factive verb	1.00	0.75
'That' CCC likelihood verb	0.86	1.08
'That' CCC attitudinal adjective	0.29	0.42
'That' CCC factive adjective	0.58	0.71
'That' CCC attitudinal noun	0.13	0.35
'That' CCC factive noun	0.29	0.42
'That' CCC likelihood noun	0.00	0.00
All stance 'that' CCC verbs	5.12	2.41
All stance 'that' CCC adjectives	0.86	1.00
All stance 'that' CCC nouns	0.42	0.50
All stance 'that' complement clauses	6.42	3.12
'To' CCC speech act verb	0.12	0.29
'To' CCC verb of cognition	0.27	0.33
'To' CCC verb of desire	2.86	1.48
'To' CCC verb of effort	1.08	0.99

Feature	Least Developed Country	
	Mean	**SD**
'To' CCC probability verb	0.00	0.00
'To' CCC adjective of certainty	0.00	0.00
'To' CCC adjective of ability	0.02	0.09
'To' CCC adjective of affect	0.03	0.14
'To' CCC adjective of ease	0.00	0.00
'To' CCC evaluative adjective	0.16	0.33
'To' CCC stance noun	2.33	1.71
All stance 'to' CCC verb	4.33	2.00
All stance 'to' CCC adjective	0.21	0.37
All stance 'to' complement clauses	6.86	3.18
Dimension 1	-16.70	5.22
Dimension 2	-3.09	0.69
Dimension 3	0.73	1.71
Dimension 4	-0.91	1.67
Dimension 5	2.02	1.93

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX I: TAGGED FEATURES BY LANDLOCKED DEVELOPING COUNTRY

Feature	Landlocked developing country	
	Mean	**SD**
Verb (not including auxiliary)	110.19	14.86
Present tense verb	74.60	14.17
Past tense verb	11.17	3.89
Progressive aspect verb	7.72	2.68
Perfect aspect verb	8.23	2.39
Infinitive verb	16.46	4.71
Pro-verb 'do'	0.10	0.22
Verb 'be'	1.35	1.23
Verb 'have'	1.61	1.57
Private verb	6.81	1.93
Public verb	2.13	1.45
Suasive verb	1.55	1.16
Activity verb	15.06	5.17
Communication verb	7.27	2.97
Mental verb	11.35	4.72
Causative verb	3.39	2.22
Occurrence verb	2.56	2.01
Existence verb	6.91	3.22
Aspectual verb	2.09	1.47
All modals	11.45	3.00
Modal of prediction	5.44	2.30
Modal of possibility	2.73	2.46
Modal of necessity	3.29	1.94
Split auxiliary-verb	4.50	1.72
All passives	9.49	3.22
Agentless passive verb	6.05	2.48
Passive verb + by	1.18	1.11
Passive postnominal modifier	2.24	1.40
All personal pronouns	44.57	13.26
1st person pronoun	34.67	13.02
2nd person pronoun	3.59	2.64
3rd person pronoun	6.31	1.92
Pronoun 'it'	9.48	3.36
Demonstrative pronoun	3.03	2.15
Nominal pronoun	1.59	1.43
All nouns	302.47	26.38
Common noun	256.82	16.98
Proper noun	33.89	15.77

Feature	Landlocked developing country	
	Mean	**SD**
Animate noun	4.60	2.60
Process noun	15.01	5.20
Cognitive noun	3.72	1.92
Abstract noun	21.50	6.72
Concrete noun	3.94	2.59
Technical noun	2.03	1.15
Quantity noun	5.37	3.46
Place noun	7.52	3.21
Group noun	2.17	2.45
Noun-noun sequence	30.70	7.61
Stance noun + preposition	2.50	1.67
Determiner + stance noun	0.56	0.85
All indefinite articles	15.31	5.67
All definite articles	66.05	24.24
Preposition	131.87	12.95
Stranded preposition	0.68	0.95
All adjectives	90.05	10.03
Attributive adjective	70.56	9.69
Attributive adjective – Size	1.27	0.68
Attributive adjective – Time	1.80	1.87
Attributive adjective – Color	0.00	0.00
Attributive adjective – Evaluative	1.47	1.31
Attributive adjective – Relational	1.76	1.23
Attributive adjective – Topical	9.73	4.75
Predicative adjective	5.28	1.80
Predicative attitudinal adjective	1.81	0.96
Predicative epistemic adjective	0.41	0.37
All adverbs	31.48	6.76
Adverb – time	3.47	2.03
Adverb – place	3.64	2.52
All stance adverbs	2.32	2.16
Non-factive adverb	0.25	0.40
Attitudinal adverb	0.23	0.52
Factive adverb	1.81	1.77
Likelihood adverb	0.02	0.08
Adverb – downtoner	1.93	2.28
Amplifier	1.96	1.73
Emphatic	1.78	1.24
Hedge	0.16	0.30

Feature	Landlocked developing country	
	Mean	**SD**
Discourse particle	0.50	0.96
'That' deletion	1.03	0.99
Contraction	0.68	1.16
All conjunctions	16.93	6.01
Subordinating conjunction – causative	0.37	0.71
Subordinating conjunction – conditional	1.13	1.12
Coordinating conjunction – clausal	6.49	5.31
Coordinating conjunction – phrasal	5.09	2.19
Adverbial – conjuncts	6.03	3.27
All wh- words	0.16	0.28
Wh question	0.08	0.21
Wh clause	0.08	0.21
All wh- relative clauses	3.12	1.68
Wh pronoun – object relative clause	0.66	0.63
Wh pronoun – subject relative clause	1.86	1.08
Wh pro. – rel. cls. with preposition fronting	0.61	0.72
Wh CCC communicaton verb	0.00	0.00
Wh CCC attitudinal verb	0.03	0.13
Wh CCC factive verb	0.08	0.21
Wh CCC likelihood verb	0.00	0.00
'That' relative clause	3.75	2.92
'That' CCC verb	3.21	1.09
'That' CCC communication verb	1.79	1.25
'That' CCC attitudinal verb	1.57	1.34
'That' CCC factive verb	0.91	0.77
'That' CCC likelihood verb	0.86	0.74
'That' CCC attitudinal adjective	0.45	0.59
'That' CCC factive adjective	0.73	0.69
'That' CCC attitudinal noun	0.05	0.15
'That' CCC factive noun	0.32	0.49
'That' CCC likelihood noun	0.06	0.17
All stance 'that' CCC verbs	5.11	2.20
All stance 'that' CCC adjectives	1.18	1.09
All stance 'that' CCC nouns	0.44	0.47
All stance 'that' complement clauses	6.73	2.86
'To' CCC speech act verb	0.13	0.30
'To' CCC verb of cognition	0.18	0.28
'To' CCC verb of desire	2.25	1.24
'To' CCC verb of effort	1.78	1.14

Feature	Landlocked developing country	
	Mean	**SD**
'To' CCC probability verb	0.00	0.00
'To' CCC adjective of certainty	0.00	0.00
'To' CCC adjective of ability	0.03	0.13
'To' CCC adjective of affect	0.04	0.15
'To' CCC adjective of ease	0.00	0.00
'To' CCC evaluative adjective	0.07	0.19
'To' CCC stance noun	2.55	1.89
All stance 'to' CCC verb	4.36	2.04
All stance 'to' CCC adjective	0.15	0.25
All stance 'to' complement clauses	7.04	3.21
Dimension 1	-17.13	5.16
Dimension 2	-3.10	0.59
Dimension 3	0.87	1.40
Dimension 4	-0.94	1.13
Dimension 5	2.69	1.83

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX J: TAGGED FEATURES BY SMALL ISLAND DEVELOPING STATE

Feature	SIDS	
	Mean	**SD**
Verb (not including auxiliary)	111.13	13.02
Present tense verb	74.78	12.78
Past tense verb	9.82	3.73
Progressive aspect verb	6.22	2.12
Perfect aspect verb	7.23	2.60
Infinitive verb	17.95	3.57
Pro-verb 'do'	0.55	0.49
Verb 'be'	2.31	1.31
Verb 'have'	1.45	1.17
Private verb	6.87	2.68
Public verb	2.48	0.84
Suasive verb	1.49	0.88
Activity verb	15.76	5.02
Communication verb	8.23	2.91
Mental verb	10.93	4.19
Causative verb	3.52	1.69
Occurrence verb	2.85	1.65
Existence verb	7.18	2.65
Aspectual verb	2.45	1.39
All modals	14.12	5.46
Modal of prediction	5.30	2.09
Modal of possibility	3.70	2.17
Modal of necessity	5.09	3.26
Split auxiliary-verb	3.82	1.82
All passives	9.78	2.34
Agentless passive verb	6.50	2.04
Passive verb + by	1.43	0.94
Passive postnominal modifier	1.84	0.93
All personal pronouns	49.31	13.03
1st person pronoun	40.38	11.93
2nd person pronoun	2.42	1.44
3rd person pronoun	6.50	3.16
Pronoun 'it'	8.66	2.80
Demonstrative pronoun	3.02	1.69
Nominal pronoun	1.25	1.07
All nouns	308.32	27.42
Common noun	249.86	19.09
Proper noun	43.96	12.59

Feature	SIDS	
	Mean	**SD**
Animate noun	4.32	2.48
Process noun	15.82	5.93
Cognitive noun	3.93	1.86
Abstract noun	23.15	6.49
Concrete noun	4.05	2.45
Technical noun	1.58	0.89
Quantity noun	5.81	2.60
Place noun	6.28	3.17
Group noun	2.34	1.35
Noun-noun sequence	30.37	9.97
Stance noun + preposition	1.73	1.66
Determiner + stance noun	0.37	0.46
All indefinite articles	16.63	4.77
All definite articles	63.27	15.48
Preposition	124.35	13.77
Stranded preposition	0.47	0.46
All adjectives	90.17	12.28
Attributive adjective	69.00	13.05
Attributive adjective – Size	2.55	2.39
Attributive adjective – Time	1.56	1.40
Attributive adjective – Color	0.07	0.20
Attributive adjective – Evaluative	1.05	0.96
Attributive adjective – Relational	2.13	1.12
Attributive adjective – Topical	8.92	4.12
Predicative adjective	4.48	1.56
Predicative attitudinal adjective	1.16	0.77
Predicative epistemic adjective	0.24	0.35
All adverbs	29.38	6.72
Adverb – time	3.87	2.42
Adverb – place	3.31	1.74
All stance adverbs	1.35	0.96
Non-factive adverb	0.06	0.18
Attitudinal adverb	0.17	0.31
Factive adverb	0.98	0.90
Likelihood adverb	0.12	0.23
Adverb – downtoner	1.58	1.29
Amplifier	1.45	1.15
Emphatic	1.90	1.21
Hedge	0.13	0.31

Feature	SIDS	
	Mean	**SD**
Discourse particle	0.22	0.42
'That' deletion	0.95	0.83
Contraction	0.94	0.93
All conjunctions	18.85	6.48
Subordinating conjunction – causative	0.57	0.83
Subordinating conjunction – conditional	1.46	1.15
Coordinating conjunction – clausal	8.12	5.30
Coordinating conjunction – phrasal	5.04	1.89
Adverbial – conjuncts	4.49	2.37
All wh- words	0.47	0.83
Wh question	0.44	0.84
Wh clause	0.02	0.12
All wh- relative clauses	3.89	2.15
Wh pronoun – object relative clause	0.91	0.81
Wh pronoun – subject relative clause	2.40	1.48
Wh pro. – rel. cls. with preposition fronting	0.57	0.55
Wh CCC communicaton verb	0.05	0.16
Wh CCC attitudinal verb	0.00	0.00
Wh CCC factive verb	0.04	0.14
Wh CCC likelihood verb	0.02	0.12
'That' relative clause	3.08	1.38
'That' CCC verb	3.57	1.37
'That' CCC communication verb	1.28	0.87
'That' CCC attitudinal verb	1.98	1.09
'That' CCC factive verb	1.15	0.77
'That' CCC likelihood verb	0.61	0.66
'That' CCC attitudinal adjective	0.32	0.44
'That' CCC factive adjective	0.18	0.28
'That' CCC attitudinal noun	0.12	0.33
'That' CCC factive noun	0.20	0.29
'That' CCC likelihood noun	0.26	0.37
All stance 'that' CCC verbs	5.02	1.98
All stance 'that' CCC adjectives	0.50	0.54
All stance 'that' CCC nouns	0.58	0.58
All stance 'that' complement clauses	6.12	2.39
'To' CCC speech act verb	0.17	0.27
'To' CCC verb of cognition	0.41	0.41
'To' CCC verb of desire	1.74	1.18
'To' CCC verb of effort	1.45	1.02

Feature	SIDS	
	Mean	**SD**
'To' CCC probability verb	0.10	0.22
'To' CCC adjective of certainty	0.00	0.00
'To' CCC adjective of ability	0.13	0.25
'To' CCC adjective of affect	0.00	0.00
'To' CCC adjective of ease	0.04	0.15
'To' CCC evaluative adjective	0.15	0.32
'To' CCC stance noun	3.52	1.77
All stance 'to' CCC verb	3.88	2.02
All stance 'to' CCC adjective	0.33	0.53
All stance 'to' complement clauses	7.71	3.08
Dimension 1	-15.71	6.37
Dimension 2	-3.26	0.68
Dimension 3	1.35	1.59
Dimension 4	-0.24	2.40
Dimension 5	1.86	1.98

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX K: TAGGED FEATURES BY RELIGIOUS AFFILIATION

Feature	Low Religious Non-Affiliation		High Religious Non-Affiliation	
	Mean	SD	Mean	SD
Verb (not including auxiliary)	110.91	13.52	113.96	13.79
Present tense verb	75.17	12.85	78.29	13.34
Past tense verb	10.78	3.79	10.67	4.12
Progressive aspect verb	7.38	2.78	7.23	2.81
Perfect aspect verb	8.47	2.95	7.99	2.28
Infinitive verb	16.47	3.25	18.16	3.53
Pro-verb 'do'	0.50	0.64	0.40	0.53
Verb 'be'	1.80	1.42	2.25	1.56
Verb 'have'	1.25	1.16	1.51	1.13
Private verb	7.44	2.62	7.38	2.68
Public verb	2.48	1.53	2.24	1.15
Suasive verb	1.25	0.83	1.54	1.10
Activity verb	14.79	5.08	16.75	4.43
Communication verb	8.80	3.19	7.72	2.30
Mental verb	12.37	4.32	12.81	4.46
Causative verb	3.64	2.00	4.13	1.81
Occurrence verb	2.37	1.77	2.80	1.51
Existence verb	7.46	3.02	7.43	2.95
Aspectual verb	2.18	1.64	2.82	1.64
All modals	12.31	4.19	15.34	5.12
Modal of prediction	5.55	2.79	5.96	2.50
Modal of possibility	2.69	2.19	4.01	2.29
Modal of necessity	4.06	2.33	5.36	2.87
Split auxiliary-verb	3.97	1.75	4.02	1.81
All passives	8.52	2.88	8.99	2.77
Agentless passive verb	5.18	1.79	6.24	2.25

Feature	Low Religious Non-Affiliation		High Religious Non-Affiliation	
	Mean	SD	Mean	SD
Passive verb + by	1.38	1.01	1.18	0.91
Passive postnominal modifier	1.96	1.27	1.57	0.94
All personal pronouns	47.68	18.50	44.33	11.86
1st person pronoun	37.60	16.06	34.65	10.22
2nd person pronoun	3.55	2.48	2.76	2.43
3rd person pronoun	6.51	3.64	6.94	3.44
Pronoun 'it'	9.26	3.85	9.54	3.07
Demonstrative pronoun	3.64	2.41	3.42	2.06
Nominal pronoun	1.32	1.34	1.74	1.25
All nouns	307.37	30.25	301.03	22.72
Common noun	252.09	19.86	249.58	18.81
Proper noun	41.81	15.41	39.09	11.53
Animate noun	5.07	2.46	5.05	2.44
Process noun	14.60	5.64	15.54	4.69
Cognitive noun	3.23	2.02	3.97	1.76
Abstract noun	23.28	6.44	23.70	6.03
Concrete noun	3.91	2.29	4.55	2.38
Technical noun	1.54	1.05	1.78	1.42
Quantity noun	5.70	2.91	5.93	2.86
Place noun	6.47	3.35	6.46	2.84
Group noun	2.61	2.05	2.13	1.31
Noun-noun sequence	29.11	10.66	28.84	8.56
Stance noun + preposition	2.01	1.74	1.76	1.32
Determiner + stance noun	0.51	0.74	0.56	0.66
All indefinite articles	16.68	5.54	18.17	5.09
All definite articles	63.91	18.05	65.14	14.88
Preposition	127.42	16.00	123.71	12.06
Stranded preposition	0.54	0.57	0.53	0.68

Feature	Low Religious Non-Affiliation		High Religious Non-Affiliation	
	Mean	SD	Mean	SD
All adjectives	89.27	11.61	88.78	12.76
Attributive adjective	69.47	12.56	68.00	11.82
Attributive adjective – Size	1.54	1.08	1.81	1.77
Attributive adjective – Time	1.53	1.40	1.67	1.41
Attributive adjective – Color	0.00	0.00	0.07	0.22
Attributive adjective – Evaluative	1.18	1.01	1.36	1.07
Attributive adjective – Relational	2.17	1.21	2.35	1.45
Attributive adjective – Topical	9.24	4.32	9.21	4.25
Predicative adjective	4.88	1.82	4.59	1.85
Predicative attitudinal adjective	1.26	0.92	1.36	1.14
Predicative epistemic adjective	0.41	0.55	0.47	0.59
All adverbs	29.63	7.60	30.48	6.73
Adverb – time	4.13	1.88	3.98	2.08
Adverb – place	4.00	2.33	4.28	2.07
All stance adverbs	1.80	1.73	1.81	1.51
Non-factive adverb	0.15	0.36	0.11	0.24
Attitudinal adverb	0.16	0.41	0.31	0.45
Factive adverb	1.34	1.51	1.21	1.17
Likelihood adverb	0.14	0.34	0.18	0.41
Adverb – downtoner	1.40	1.10	1.97	1.70
Amplifier	1.88	1.47	1.71	1.36
Emphatic	1.69	1.25	2.31	1.38
Hedge	0.21	0.38	0.17	0.32
Discourse particle	0.19	0.30	0.30	0.59
'That' deletion	0.96	0.84	0.92	0.84
Contraction	0.99	1.37	1.30	2.20
All conjunctions	17.53	6.23	17.76	6.22
Subordinating conjunction – causative	0.39	0.71	0.52	0.81

Feature	Low Religious Non-Affiliation		High Religious Non-Affiliation	
	Mean	SD	Mean	SD
Subordinating conjunction – conditional	0.92	1.04	1.28	1.21
Coordinating conjunction – clausal	7.74	5.67	7.83	5.24
Coordinating conjunction – phrasal	4.99	2.39	4.73	2.24
Adverbial – conjuncts	5.01	2.96	4.48	2.34
All wh- words	0.40	0.77	0.44	0.57
Wh question	0.24	0.75	0.29	0.49
Wh clause	0.16	0.32	0.15	0.28
All wh- relative clauses	3.41	1.74	3.76	2.25
Wh pronoun – object relative clause	0.93	0.81	0.80	0.80
Wh pronoun – subject relative clause	2.05	1.23	2.33	1.53
Wh pro. – rel. cls. with preposition fronting	0.43	0.50	0.63	0.59
Wh CCC communicaton verb	0.06	0.17	0.05	0.15
Wh CCC attitudinal verb	0.02	0.09	0.03	0.16
Wh CCC factive verb	0.12	0.28	0.12	0.24
Wh CCC likelihood verb	0.04	0.14	0.04	0.14
'That' relative clause	3.28	1.86	3.46	2.28
'That' CCC verb	3.32	1.52	3.20	1.70
'That' CCC communication verb	1.33	1.06	1.28	0.94
'That' CCC attitudinal verb	1.64	1.15	1.68	1.00
'That' CCC factive verb	0.90	0.74	1.20	0.93
'That' CCC likelihood verb	1.07	0.99	0.83	0.83
'That' CCC attitudinal adjective	0.25	0.35	0.32	0.47
'That' CCC factive adjective	0.32	0.56	0.29	0.40
'That' CCC attitudinal noun	0.12	0.31	0.10	0.31
'That' CCC factive noun	0.26	0.40	0.35	0.58
'That' CCC likelihood noun	0.12	0.25	0.15	0.31
All stance 'that' CCC verbs	4.92	2.01	4.98	2.11
All stance 'that' CCC adjectives	0.57	0.71	0.61	0.69

Feature	Low Religious Non-Affiliation		High Religious Non-Affiliation	
	Mean	SD	Mean	SD
All stance 'that' CCC nouns	0.51	0.53	0.60	0.82
All stance 'that' complement clauses	6.01	2.45	6.20	2.86
'To' CCC speech act verb	0.25	0.50	0.13	0.24
'To' CCC verb of cognition	0.26	0.29	0.27	0.37
'To' CCC verb of desire	2.76	1.48	2.22	1.35
'To' CCC verb of effort	1.03	0.84	1.96	1.07
'To' CCC probability verb	0.08	0.20	0.08	0.22
'To' CCC adjective of certainty	0.00	0.00	0.00	0.04
'To' CCC adjective of ability	0.10	0.24	0.16	0.33
'To' CCC adjective of affect	0.02	0.11	0.00	0.00
'To' CCC adjective of ease	0.02	0.09	0.03	0.11
'To' CCC evaluative adjective	0.10	0.33	0.11	0.25
'To' CCC stance noun	3.33	1.83	3.26	1.76
All stance 'to' CCC verb	4.37	1.97	4.67	1.76
All stance 'to' CCC adjective	0.24	0.46	0.30	0.42
All stance 'to' complement clauses	7.94	3.06	8.20	2.71
Dimension 1	-16.18	6.73	-14.83	5.49
Dimension 2	-3.00	0.68	-3.12	0.62
Dimension 3	0.74	1.86	0.80	1.75
Dimension 4	-0.95	1.44	0.06	2.11
Dimension 5	1.98	1.97	1.53	1.69

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX L: TAGGED FEATURES FOR THE UNITED STATES

Feature	US frequency	US Cohen's *d*
Verb (not including auxiliary)	143.5	2.30
Present tense verb	111.3	2.69
Past tense verb	11.2	0.12
Progressive aspect verb	4.7	-0.93
Perfect aspect verb	7.5	-0.26
Infinitive verb	17.4	-0.06
Pro-verb 'do'	1.9	2.71
Verb 'be'	3.7	1.06
Verb 'have'	1.7	0.24
Private verb	15.5	3.25
Public verb	2.4	0.07
Suasive verb	1.1	-0.34
Activity verb	21.7	1.20
Communication verb	7.7	-0.14
Mental verb	19.2	1.51
Causative verb	6.2	1.20
Occurrence verb	2.2	-0.29
Existence verb	8.6	0.39
Aspectual verb	2.6	-0.01
All modals	27.3	2.68
Modal of prediction	8.0	0.85
Modal of possibility	14.2	5.22
Modal of necessity	5.2	0.09
Split auxiliary-verb	3.7	-0.17
All passives	9.5	0.24
Agentless passive verb	7.7	0.84
Passive verb + by	1.3	0.06
Passive postnominal modifier	0.4	-1.23
All personal pronouns	62.4	1.21
1st person pronoun	46.9	0.93
2nd person pronoun	2.8	-0.09
3rd person pronoun	12.7	1.73
Pronoun 'it'	10.1	0.20
Demonstrative pronoun	9.5	2.92
Nominal pronoun	3.4	1.41
All nouns	262.1	-1.65
Common noun	227.9	-1.19

Feature	US frequency	US Cohen's d
Proper noun	29.9	-0.79
Animate noun	10.8	2.45
Process noun	9.9	-1.08
Cognitive noun	4.9	0.63
Abstract noun	23.9	0.05
Concrete noun	2.4	-0.83
Technical noun	0.9	-0.62
Quantity noun	6.9	0.37
Place noun	6.0	-0.16
Group noun	3.9	1.03
Noun-noun sequence	12.7	-1.80
Stance noun + preposition	1.9	0.04
Determiner + stance noun	1.5	1.42
All indefinite articles	25.0	1.41
All definite articles	45.4	-1.24
Preposition	107.2	-1.34
Stranded preposition	0.0	-0.84
All adjectives	74.5	-1.18
Attributive adjective	52.7	-1.33
Attributive adjective – Size	0.4	-0.84
Attributive adjective – Time	2.8	0.84
Attributive adjective – Color	0.2	0.84
Attributive adjective – Evaluative	0.6	-0.67
Attributive adjective – Relational	2.8	0.37
Attributive adjective – Topical	4.9	-1.03
Predicative adjective	6.7	1.11
Predicative attitudinal adjective	0.9	-0.40
Predicative epistemic adjective	1.1	1.15
All adverbs	38.7	1.23
Adverb – time	3.7	-0.16
Adverb – place	5.2	0.47
All stance adverbs	2.2	0.25
Non-factive adverb	0.2	0.27
Attitudinal adverb	0.0	-0.61
Factive adverb	1.9	0.51
Likelihood adverb	0.0	-0.43
Adverb – downtoner	1.7	-0.06
Amplifier	0.9	-0.63

Feature	US frequency	US Cohen's d
Emphatic	3.4	0.95
Hedge	0.0	-0.55
Discourse particle	0.6	0.66
'That' deletion	3.0	2.57
Contraction	9.0	4.37
All conjunctions	33.4	2.65
Subordinating conjunction – causative	1.7	1.60
Subordinating conjunction – conditional	2.4	1.07
Coordinating conjunction – clausal	24.7	3.37
Coordinating conjunction – phrasal	0.2	-2.09
Adverbial – conjuncts	5.4	0.30
All wh- words	0.4	-0.04
Wh question	0.2	-0.13
Wh clause	0.2	0.17
All wh- relative clauses	5.2	0.74
Wh pronoun – object relative clause	1.7	1.09
Wh pronoun – subject relative clause	3.2	0.67
Wh pro. – rel. cls. with preposition fronting	0.2	-0.64
Wh CCC communicaton verb	0.2	0.97
Wh CCC attitudinal verb	0.0	-0.21
Wh CCC factive verb	0.2	0.31
Wh CCC likelihood verb	0.0	-0.28
'That' relative clause	9.0	2.72
'That' CCC verb	9.3	4.05
'That' CCC communication verb	2.6	1.37
'That' CCC attitudinal verb	1.9	0.22
'That' CCC factive verb	4.1	3.65
'That' CCC likelihood verb	3.2	2.70
'That' CCC attitudinal adjective	0.0	-0.69
'That' CCC factive adjective	0.6	0.67
'That' CCC attitudinal noun	0.0	-0.36
'That' CCC factive noun	1.5	2.30
'That' CCC likelihood noun	1.1	3.50
All stance 'that' CCC verbs	11.8	3.55
All stance 'that' CCC adjectives	0.6	0.01
All stance 'that' CCC nouns	2.6	2.89
All stance 'that' complement clauses	15.1	3.52
'To' CCC speech act verb	0.4	0.68

Feature	US frequency	US Cohen's d
'To' CCC verb of cognition	0.2	-0.21
'To' CCC verb of desire	0.9	-1.07
'To' CCC verb of effort	3.0	1.24
'To' CCC probability verb	0.4	1.53
'To' CCC adjective of certainty	0.0	-0.10
'To' CCC adjective of ability	0.0	-0.46
'To' CCC adjective of affect	0.0	-0.10
'To' CCC adjective of ease	0.0	-0.21
'To' CCC evaluative adjective	0.0	-0.39
'To' CCC stance noun	2.6	-0.39
All stance 'to' CCC verb	4.9	0.18
All stance 'to' CCC adjective	0.0	-0.65
All stance 'to' complement clauses	7.5	-0.22
Dimension 1	1.1	
Dimension 2	-2.9	
Dimension 3	-0.9	
Dimension 4	0.6	
Dimension 5	2.0	

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX M: TAGGED FEATURES FOR UKRAINE

Feature	Ukraine frequency	Ukraine Cohen's d
Verb (not including auxiliary)	101.8	-0.82
Present tense verb	67.4	-0.76
Past tense verb	14.5	0.96
Progressive aspect verb	5.8	-0.53
Perfect aspect verb	9.0	0.35
Infinitive verb	15.4	-0.64
Pro-verb 'do'	0.3	-0.23
Verb 'be'	1.9	-0.13
Verb 'have'	1.3	-0.11
Private verb	6.1	-0.49
Public verb	1.3	-0.80
Suasive verb	1.6	0.15
Activity verb	16.1	-0.01
Communication verb	8.7	0.24
Mental verb	11.9	-0.18
Causative verb	1.6	-1.29
Occurrence verb	3.2	0.34
Existence verb	3.9	-1.22
Aspectual verb	2.9	0.17
All modals	15.1	0.14
Modal of prediction	7.7	0.73
Modal of possibility	3.9	0.13
Modal of necessity	3.5	-0.53
Split auxiliary-verb	3.9	-0.06
All passives	9.3	0.16
Agentless passive verb	6.7	0.37
Passive verb + by	1.0	-0.26
Passive postnominal modifier	1.6	-0.09
All personal pronouns	30.8	-1.03
1st person pronoun	21.2	-1.18
2nd person pronoun	6.1	1.27
3rd person pronoun	3.5	-0.96
Pronoun 'it'	13.2	1.15
Demonstrative pronoun	2.9	-0.27
Nominal pronoun	3.5	1.49
All nouns	310.9	0.31
Common noun	251.1	0.04

Feature	Ukraine frequency	Ukraine Cohen's d
Proper noun	51.7	0.92
Animate noun	4.8	-0.11
Process noun	12.8	-0.49
Cognitive noun	5.5	0.95
Abstract noun	27.0	0.56
Concrete noun	6.1	0.75
Technical noun	2.9	0.92
Quantity noun	6.7	0.30
Place noun	10.6	1.40
Group noun	1.9	-0.24
Noun-noun sequence	25.4	-0.39
Stance noun + preposition	2.6	0.53
Determiner + stance noun	0.6	0.08
All indefinite articles	13.5	-0.81
All definite articles	81.9	1.09
Preposition	144.8	1.51
Stranded preposition	0.3	-0.37
All adjectives	89.0	0.01
Attributive adjective	70.6	0.18
Attributive adjective – Size	1.0	-0.46
Attributive adjective – Time	0.6	-0.74
Attributive adjective – Color	0.0	-0.26
Attributive adjective – Evaluative	2.2	0.86
Attributive adjective – Relational	2.2	-0.07
Attributive adjective – Topical	8.3	-0.22
Predicative adjective	4.5	-0.10
Predicative attitudinal adjective	0.6	-0.68
Predicative epistemic adjective	0.6	0.27
All adverbs	25.0	-0.75
Adverb – time	2.9	-0.56
Adverb – place	1.6	-1.22
All stance adverbs	1.6	-0.14
Non-factive adverb	0.3	0.63
Attitudinal adverb	0.0	-0.61
Factive adverb	1.3	0.04
Likelihood adverb	0.0	-0.43
Adverb – downtoner	2.6	0.52
Amplifier	2.9	0.82

Feature	Ukraine frequency	Ukraine Cohen's d
Emphatic	3.5	1.03
Hedge	0.0	-0.55
Discourse particle	0.0	-0.52
'That' deletion	1.3	0.45
Contraction	1.9	0.36
All conjunctions	17.3	-0.06
Subordinating conjunction – causative	0.6	0.16
Subordinating conjunction – conditional	3.2	1.78
Coordinating conjunction – clausal	6.1	-0.32
Coordinating conjunction – phrasal	4.2	-0.27
Adverbial – conjuncts	3.5	-0.45
All wh- words	1.6	1.90
Wh question	1.6	2.36
Wh clause	0.0	-0.51
All wh- relative clauses	7.4	1.83
Wh pronoun – object relative clause	1.3	0.58
Wh pronoun – subject relative clause	5.1	2.04
Wh pro. – rel. cls. with preposition fronting	1.0	0.77
Wh CCC communicaton verb	0.0	-0.31
Wh CCC attitudinal verb	0.0	-0.21
Wh CCC factive verb	0.0	-0.47
Wh CCC likelihood verb	0.0	-0.28
'That' relative clause	2.6	-0.38
'That' CCC verb	4.2	0.59
'That' CCC communication verb	2.6	1.37
'That' CCC attitudinal verb	1.0	-0.65
'That' CCC factive verb	1.9	0.91
'That' CCC likelihood verb	1.3	0.45
'That' CCC attitudinal adjective	0.0	-0.69
'That' CCC factive adjective	0.6	0.67
'That' CCC attitudinal noun	0.0	-0.36
'That' CCC factive noun	0.3	-0.04
'That' CCC likelihood noun	0.0	-0.48
All stance 'that' CCC verbs	6.7	0.85
All stance 'that' CCC adjectives	0.6	0.01
All stance 'that' CCC nouns	0.3	-0.37
All stance 'that' complement clauses	7.7	0.58
'To' CCC speech act verb	0.3	0.39

Feature	Ukraine frequency	Ukraine Cohen's d
'To' CCC verb of cognition	0.3	0.09
'To' CCC verb of desire	1.3	-0.78
'To' CCC verb of effort	1.9	0.22
'To' CCC probability verb	0.0	-0.39
'To' CCC adjective of certainty	0.3	
'To' CCC adjective of ability	0.0	-0.46
'To' CCC adjective of affect	0.0	-0.10
'To' CCC adjective of ease	0.0	-0.21
'To' CCC evaluative adjective	0.3	0.71
'To' CCC stance noun	2.6	-0.39
All stance 'to' CCC verb	3.9	-0.37
All stance 'to' CCC adjective	0.6	0.75
All stance 'to' complement clauses	7.1	-0.36
Dimension 1	-13.5	
Dimension 2	-3.1	
Dimension 3	4.0	
Dimension 4	0.4	
Dimension 5	0.9	

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

Feature	Eritrea frequency	Eritrea Cohen's d
Verb (not including auxiliary)	91.4	-1.61
Present tense verb	64.6	-0.97
Past tense verb	6.1	-1.17
Progressive aspect verb	8.5	0.44
Perfect aspect verb	4.9	-1.31
Infinitive verb	12.2	-1.57
Pro-verb 'do'	0.0	-0.77
Verb 'be'	2.4	0.20
Verb 'have'	0.0	-1.27
Private verb	4.9	-0.95
Public verb	0.0	-1.86
Suasive verb	2.4	0.94
Activity verb	13.4	-0.59
Communication verb	4.9	-1.21
Mental verb	12.2	-0.11
Causative verb	2.4	-0.85
Occurrence verb	1.2	-0.93
Existence verb	3.7	-1.29
Aspectual verb	2.4	-0.13
All modals	7.3	-1.43
Modal of prediction	2.4	-1.35
Modal of possibility	2.4	-0.52
Modal of necessity	2.4	-0.93
Split auxiliary-verb	3.7	-0.17
All passives	14.6	2.12
Agentless passive verb	7.3	0.65
Passive verb + by	3.7	2.73
Passive postnominal modifier	3.7	1.93
All personal pronouns	32.9	-0.88
1st person pronoun	21.9	-1.12
2nd person pronoun	4.9	0.78
3rd person pronoun	6.1	-0.20
Pronoun 'it'	14.6	1.58
Demonstrative pronoun	4.9	0.66
Nominal pronoun	0.0	-1.27
All nouns	322.8	0.79
Common noun	285.0	1.86

Feature	Eritrea frequency	Eritrea Cohen's d
Proper noun	30.5	-0.74
Animate noun	4.9	-0.07
Process noun	21.9	1.35
Cognitive noun	1.2	-1.38
Abstract noun	28.0	0.73
Concrete noun	0.0	-1.89
Technical noun	0.0	-1.32
Quantity noun	3.7	-0.76
Place noun	8.5	0.69
Group noun	0.0	-1.46
Noun-noun sequence	21.9	-0.77
Stance noun + preposition	1.2	-0.44
Determiner + stance noun	1.2	0.97
All indefinite articles	23.1	1.04
All definite articles	60.9	-0.24
Preposition	126.7	0.14
Stranded preposition	0.0	-0.84
All adjectives	112.1	1.92
Attributive adjective	87.7	1.63
Attributive adjective – Size	1.2	-0.33
Attributive adjective – Time	0.0	-1.18
Attributive adjective – Color	0.0	-0.26
Attributive adjective – Evaluative	0.0	-1.26
Attributive adjective – Relational	0.0	-1.70
Attributive adjective – Topical	20.7	2.83
Predicative adjective	4.9	0.12
Predicative attitudinal adjective	0.0	-1.25
Predicative epistemic adjective	0.0	-0.78
All adverbs	18.3	-1.74
Adverb – time	3.7	-0.16
Adverb – place	3.7	-0.23
All stance adverbs	0.0	-1.17
Non-factive adverb	0.0	-0.44
Attitudinal adverb	0.0	-0.61
Factive adverb	0.0	-0.99
Likelihood adverb	0.0	-0.43
Adverb – downtoner	0.0	-1.17
Amplifier	0.0	-1.29

Feature	Eritrea frequency	Eritrea Cohen's d
Emphatic	2.4	0.21
Hedge	0.0	-0.55
Discourse particle	0.0	-0.52
'That' deletion	1.2	0.32
Contraction	0.0	-0.61
All conjunctions	17.1	-0.10
Subordinating conjunction – causative	1.2	0.94
Subordinating conjunction – conditional	0.0	-1.01
Coordinating conjunction – clausal	1.2	-1.25
Coordinating conjunction – phrasal	13.4	4.14
Adverbial – conjuncts	2.4	-0.89
All wh- words	0.0	-0.67
Wh question	0.0	-0.48
Wh clause	0.0	-0.51
All wh- relative clauses	4.9	0.60
Wh pronoun – object relative clause	1.2	0.45
Wh pronoun – subject relative clause	3.7	1.02
Wh pro. – rel. cls. with preposition fronting	0.0	-1.00
Wh CCC communicaton verb	0.0	-0.31
Wh CCC attitudinal verb	0.0	-0.21
Wh CCC factive verb	0.0	-0.47
Wh CCC likelihood verb	0.0	-0.28
'That' relative clause	6.1	1.27
'That' CCC verb	1.2	-1.26
'That' CCC communication verb	0.0	-1.35
'That' CCC attitudinal verb	1.2	-0.45
'That' CCC factive verb	1.2	0.11
'That' CCC likelihood verb	1.2	0.33
'That' CCC attitudinal adjective	0.0	-0.69
'That' CCC factive adjective	0.0	-0.66
'That' CCC attitudinal noun	0.0	-0.36
'That' CCC factive noun	0.0	-0.61
'That' CCC likelihood noun	0.0	-0.48
All stance 'that' CCC verbs	3.7	-0.62
All stance 'that' CCC adjectives	0.0	-0.87
All stance 'that' CCC nouns	0.0	-0.78
All stance 'that' complement clauses	3.7	-0.90
'To' CCC speech act verb	0.0	-0.48

557

Feature	Eritrea frequency	Eritrea Cohen's d
'To' CCC verb of cognition	0.0	-0.79
'To' CCC verb of desire	6.1	2.76
'To' CCC verb of effort	0.0	-1.56
'To' CCC probability verb	0.0	-0.39
'To' CCC adjective of certainty	0.0	-0.10
'To' CCC adjective of ability	0.0	-0.46
'To' CCC adjective of affect	0.0	-0.10
'To' CCC adjective of ease	0.0	-0.21
'To' CCC evaluative adjective	1.2	4.39
'To' CCC stance noun	3.7	0.24
All stance 'to' CCC verb	6.1	0.85
All stance 'to' CCC adjective	1.2	2.21
All stance 'to' complement clauses	11.0	1.04
Dimension 1	-21.2	
Dimension 2	-4.3	
Dimension 3	5.3	
Dimension 4	-2.6	
Dimension 5	3.0	

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX O: TAGGED FEATURES FOR TUVALU

Feature	Tuvalu frequency	Tuvalu Cohen's d
Verb (not including auxiliary)	105.9	-0.52
Present tense verb	78.0	0.05
Past tense verb	7.9	-0.71
Progressive aspect verb	7.0	-0.10
Perfect aspect verb	6.0	-0.86
Infinitive verb	11.1	-1.90
Pro-verb 'do'	0.5	0.12
Verb 'be'	4.6	1.67
Verb 'have'	3.3	1.68
Private verb	5.1	-0.88
Public verb	1.4	-0.72
Suasive verb	2.3	0.84
Activity verb	17.6	0.31
Communication verb	3.7	-1.69
Mental verb	9.8	-0.66
Causative verb	4.2	0.12
Occurrence verb	2.8	0.08
Existence verb	1.9	-1.93
Aspectual verb	1.4	-0.74
All modals	18.6	0.85
Modal of prediction	5.6	-0.09
Modal of possibility	1.9	-0.73
Modal of necessity	11.1	2.30
Split auxiliary-verb	1.9	-1.20
All passives	9.8	0.34
Agentless passive verb	7.9	0.93
Passive verb + by	0.0	-1.33
Passive postnominal modifier	1.9	0.19
All personal pronouns	49.2	0.27
1st person pronoun	40.9	0.43
2nd person pronoun	1.9	-0.45
3rd person pronoun	6.5	-0.09
Pronoun 'it'	6.5	-0.90
Demonstrative pronoun	1.9	-0.74
Nominal pronoun	0.9	-0.55
All nouns	320.5	0.70
Common noun	249.9	-0.02

Feature	Tuvalu frequency	Tuvalu Cohen's d
Proper noun	52.9	1.02
Animate noun	5.6	0.22
Process noun	13.0	-0.45
Cognitive noun	1.4	-1.27
Abstract noun	21.8	-0.29
Concrete noun	2.3	-0.88
Technical noun	2.3	0.46
Quantity noun	6.5	0.23
Place noun	7.4	0.31
Group noun	1.4	-0.56
Noun-noun sequence	35.3	0.70
Stance noun + preposition	0.9	-0.65
Determiner + stance noun	0.0	-0.80
All indefinite articles	20.4	0.52
All definite articles	49.2	-0.99
Preposition	106.8	-1.37
Stranded preposition	0.5	-0.05
All adjectives	98.9	0.81
Attributive adjective	68.7	0.02
Attributive adjective – Size	5.6	2.54
Attributive adjective – Time	1.9	0.19
Attributive adjective – Color	0.0	-0.26
Attributive adjective – Evaluative	0.9	-0.39
Attributive adjective – Relational	1.4	-0.65
Attributive adjective – Topical	4.2	-1.20
Predicative adjective	7.0	1.28
Predicative attitudinal adjective	2.8	1.40
Predicative epistemic adjective	0.0	-0.78
All adverbs	22.8	-1.07
Adverb – time	0.9	-1.58
Adverb – place	1.4	-1.32
All stance adverbs	1.4	-0.26
Non-factive adverb	0.0	-0.44
Attitudinal adverb	0.0	-0.61
Factive adverb	1.4	0.11
Likelihood adverb	0.0	-0.43
Adverb – downtoner	2.3	0.33
Amplifier	1.9	0.10

Feature	Tuvalu frequency	Tuvalu Cohen's d
Emphatic	3.3	0.88
Hedge	0.0	-0.55
Discourse particle	0.0	-0.52
'That' deletion	0.5	-0.52
Contraction	1.9	0.36
All conjunctions	16.3	-0.23
Subordinating conjunction – causative	1.4	1.20
Subordinating conjunction – conditional	1.9	0.63
Coordinating conjunction – clausal	4.6	-0.60
Coordinating conjunction – phrasal	7.0	0.97
Adverbial – conjuncts	0.9	-1.50
All wh- words	0.5	0.12
Wh question	0.5	0.39
Wh clause	0.0	-0.51
All wh- relative clauses	1.9	-0.84
Wh pronoun – object relative clause	0.9	0.08
Wh pronoun – subject relative clause	0.9	-0.94
Wh pro. – rel. cls. with preposition fronting	0.0	-1.00
Wh CCC communicaton verb	0.0	-0.31
Wh CCC attitudinal verb	0.0	-0.21
Wh CCC factive verb	0.0	-0.47
Wh CCC likelihood verb	0.0	-0.28
'That' relative clause	2.8	-0.28
'That' CCC verb	3.7	0.29
'That' CCC communication verb	0.0	-1.35
'That' CCC attitudinal verb	2.8	1.10
'That' CCC factive verb	1.9	0.91
'That' CCC likelihood verb	0.0	-1.04
'That' CCC attitudinal adjective	0.0	-0.69
'That' CCC factive adjective	0.0	-0.66
'That' CCC attitudinal noun	0.0	-0.36
'That' CCC factive noun	0.0	-0.61
'That' CCC likelihood noun	0.0	-0.48
All stance 'that' CCC verbs	4.6	-0.18
All stance 'that' CCC adjectives	0.0	-0.87
All stance 'that' CCC nouns	0.0	-0.78
All stance 'that' complement clauses	4.6	-0.57
'To' CCC speech act verb	0.0	-0.48

Feature	Tuvalu frequency	Tuvalu Cohen's d
'To' CCC verb of cognition	0.0	-0.79
'To' CCC verb of desire	0.9	-1.07
'To' CCC verb of effort	0.0	-1.56
'To' CCC probability verb	0.0	-0.39
'To' CCC adjective of certainty	0.0	-0.10
'To' CCC adjective of ability	0.0	-0.46
'To' CCC adjective of affect	0.0	-0.10
'To' CCC adjective of ease	0.0	-0.21
'To' CCC evaluative adjective	0.5	1.45
'To' CCC stance noun	0.9	-1.36
All stance 'to' CCC verb	0.9	-2.07
All stance 'to' CCC adjective	0.5	0.52
All stance 'to' complement clauses	2.3	-2.13
Dimension 1	-16.6	
Dimension 2	-3.8	
Dimension 3	2.6	
Dimension 4	2.4	
Dimension 5	-1.3	

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.

APPENDIX P: TAGGED FEATURES FOR BHUTAN

Feature	Bhutan frequency	Bhutan Cohen's d
Verb (not including auxiliary)	121.2	0.60
Present tense verb	85.3	0.61
Past tense verb	5.2	-1.40
Progressive aspect verb	10.4	1.13
Perfect aspect verb	9.9	0.71
Infinitive verb	19.1	0.42
Pro-verb 'do'	0.6	0.30
Verb 'be'	4.1	1.33
Verb 'have'	1.7	0.24
Private verb	8.1	0.27
Public verb	1.2	-0.88
Suasive verb	2.9	1.44
Activity verb	20.3	0.89
Communication verb	7.5	-0.21
Mental verb	9.9	-0.64
Causative verb	8.1	2.27
Occurrence verb	1.2	-0.93
Existence verb	9.9	0.84
Aspectual verb	1.2	-0.87
All modals	13.3	-0.22
Modal of prediction	10.4	1.81
Modal of possibility	0.0	-1.57
Modal of necessity	2.9	-0.75
Split auxiliary-verb	3.5	-0.29
All passives	9.9	0.38
Agentless passive verb	4.6	-0.61
Passive verb + by	2.9	1.80
Passive postnominal modifier	2.3	0.57
All personal pronouns	58.0	0.89
1st person pronoun	51.6	1.32
2nd person pronoun	2.9	-0.04
3rd person pronoun	3.5	-0.96
Pronoun 'it'	11.6	0.65
Demonstrative pronoun	3.5	0.00
Nominal pronoun	0.0	-1.27
All nouns	299.9	-0.12
Common noun	258.7	0.44

Feature	Bhutan frequency	Bhutan Cohen's d
Proper noun	26.1	-1.09
Animate noun	2.9	-0.90
Process noun	22.0	1.37
Cognitive noun	4.6	0.47
Abstract noun	22.0	-0.26
Concrete noun	3.5	-0.36
Technical noun	2.9	0.92
Quantity noun	6.4	0.19
Place noun	10.4	1.34
Group noun	0.6	-1.07
Noun-noun sequence	29.6	0.07
Stance noun + preposition	3.5	1.15
Determiner + stance noun	0.0	-0.80
All indefinite articles	19.1	0.27
All definite articles	58.6	-0.39
Preposition	127.0	0.16
Stranded preposition	0.0	-0.84
All adjectives	83.5	-0.44
Attributive adjective	62.6	-0.49
Attributive adjective – Size	1.7	-0.02
Attributive adjective – Time	0.6	-0.74
Attributive adjective – Color	0.0	-0.26
Attributive adjective – Evaluative	1.7	0.38
Attributive adjective – Relational	0.0	-1.70
Attributive adjective – Topical	9.3	0.02
Predicative adjective	5.2	0.29
Predicative attitudinal adjective	2.9	1.49
Predicative epistemic adjective	0.6	0.27
All adverbs	22.6	-1.10
Adverb – time	1.7	-1.17
Adverb – place	4.1	-0.04
All stance adverbs	3.5	1.09
Non-factive adverb	0.0	-0.44
Attitudinal adverb	0.0	-0.61
Factive adverb	3.5	1.80
Likelihood adverb	0.0	-0.43
Adverb – downtoner	1.7	-0.06
Amplifier	1.2	-0.41

Feature	Bhutan frequency	Bhutan Cohen's d
Emphatic	1.7	-0.30
Hedge	0.0	-0.55
Discourse particle	0.0	-0.52
'That' deletion	1.7	0.93
Contraction	0.6	-0.31
All conjunctions	15.1	-0.42
Subordinating conjunction – causative	0.0	-0.62
Subordinating conjunction – conditional	1.2	0.03
Coordinating conjunction – clausal	5.2	-0.49
Coordinating conjunction – phrasal	6.4	0.71
Adverbial – conjuncts	6.4	0.69
All wh- words	0.0	-0.67
Wh question	0.0	-0.48
Wh clause	0.0	-0.51
All wh- relative clauses	2.3	-0.65
Wh pronoun – object relative clause	0.6	-0.30
Wh pronoun – subject relative clause	1.7	-0.38
Wh pro. – rel. cls. with preposition fronting	0.0	-1.00
Wh CCC communicaton verb	0.0	-0.31
Wh CCC attitudinal verb	0.0	-0.21
Wh CCC factive verb	0.0	-0.47
Wh CCC likelihood verb	0.0	-0.28
'That' relative clause	5.8	1.13
'That' CCC verb	4.6	0.84
'That' CCC communication verb	0.6	-0.72
'That' CCC attitudinal verb	2.9	1.20
'That' CCC factive verb	1.7	0.68
'That' CCC likelihood verb	0.6	-0.35
'That' CCC attitudinal adjective	0.0	-0.69
'That' CCC factive adjective	0.0	-0.66
'That' CCC attitudinal noun	0.0	-0.36
'That' CCC factive noun	0.0	-0.61
'That' CCC likelihood noun	0.0	-0.48
All stance 'that' CCC verbs	5.8	0.41
All stance 'that' CCC adjectives	0.0	-0.87
All stance 'that' CCC nouns	0.0	-0.78
All stance 'that' complement clauses	5.8	-0.13
'To' CCC speech act verb	0.0	-0.48

Feature	Bhutan frequency	Bhutan Cohen's d
'To' CCC verb of cognition	0.0	-0.79
'To' CCC verb of desire	1.7	-0.49
'To' CCC verb of effort	0.6	-0.99
'To' CCC probability verb	0.0	-0.39
'To' CCC adjective of certainty	0.0	-0.10
'To' CCC adjective of ability	0.0	-0.46
'To' CCC adjective of affect	0.0	-0.10
'To' CCC adjective of ease	0.0	-0.21
'To' CCC evaluative adjective	0.0	-0.39
'To' CCC stance noun	5.2	1.10
All stance 'to' CCC verb	2.3	-1.26
All stance 'to' CCC adjective	0.0	-0.65
All stance 'to' complement clauses	7.5	-0.22
Dimension 1	-17.5	
Dimension 2	-3.3	
Dimension 3	1.6	
Dimension 4	0.1	
Dimension 5	4.0	

Note. Mean frequencies normed per 1,000 words. CCC = complement clause controlled by.