

LINGUISTIC VARIATION ACROSS TEXTS:
THE ROLE OF REGISTER, INTRA-REGISTER COMMUNICATIVE DISTINCTIONS,
SOCIAL GROUP, AND INDIVIDUAL DIFFERENCES

By Marianna Gracheva

A Dissertation

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in Applied Linguistics

Northern Arizona University

May 2023

Approved:

Douglas Biber, Ph.D., Co-chair

Jesse Egbert, Ph.D., Co-chair

Naoko Taguchi, Ph.D.

Michaela Mahlberg, Ph.D.

ABSTRACT

LINGUISTIC VARIATION ACROSS TEXTS:

THE ROLE OF REGISTER, INTRA-REGISTER COMMUNICATIVE DISTINCTIONS, SOCIAL GROUP, AND INDIVIDUAL DIFFERENCES

MARIANNA GRACHEVA

Language variation is attributed to a variety of factors – linguistic and various nonlinguistic or external constraints determining language use. These external factors have been commonly classified as characteristics of language users and factors associated with the situation of use (e.g., Halliday, 1978). User characteristics have been the subject of sociolinguistic research, concerned with variation predicted by group membership on the basis of such factors as age, gender, region, ethnicity, educational or occupational background. Other fields of linguistics, such as idiolectal sociolinguistics, are also concerned with characteristics of the users, but these characteristics are individual rather than indicative of group affiliation.

On the other hand, in the research tradition focused on considerations of use, linguistic variation is explained by the fact that language is always produced in response to a particular communicative situation, and linguistic variation corresponds to situational differences. In the text-linguistic school of thought, linguistic variation has been traditionally attributed to the fact that texts belong to ‘registers’, culturally-recognized language varieties “associated with the situation of use” (Biber & Conrad, 2019, p. 6). On the other hand, it has been recognized that registers are not situationally uniform, but rather exhibit substantial internal variation across texts or even within texts. An additional consideration in linguistic variation according to the situation of use has therefore involved granular communicative differences within registers, shown to

explain additional linguistic variation, previously unaccounted for by existing register categories. (e.g., Biber, Egbert, & Keller, 2020; Egbert & Gracheva, 2023).

This dissertation examines the interactions that may exist between these distinct approaches to linguistic variation and investigates social group (age and gender), register, and individual speaker characteristics as predictors of functional linguistic variation in the ‘Corpus of 100 Idiolects’, comprised by 7 registers produced by 112 speakers from a range of social backgrounds. The goals of the study include (1) identifying underlying dimensions of functional linguistic variation in the corpus; (2) measuring the contribution of each predictor variable to functional linguistic variation; (3) providing accounts of language use by age and gender groups within and across registers; (4) accounting for possible communicative distinctions within registers; (5) analyzing individual authors across registers.

The study produces several key findings. First, social and individual characteristics do not contribute to explaining functional linguistic variation across texts, while the contribution of register is major. Second, as social groups are examined across registers, it is revealed that register is a significant predictor of variation for all groups, and age and gender groups generally follow the same cross-register patterns. When social groups are examined within registers, it is shown that age and gender do not predict register-internal variation, but each social group varies extensively within its scope. Third, as variation within registers is explored further, the study finds that this linguistic variation is systematically explained by highly granular communicative distinctions across texts. Finally, the study demonstrates how analysis of individual language use is enhanced through an account of authors’ approaches to a range of registers that allow varying degrees of internal linguistic freedom.

ACKNOWLEDGEMENTS

This study is important to me because it reflects and augments the work I have been inspired by the most over these years. I feel incredibly grateful that this is how this dissertation has turned out – I could not have asked for better luck and a more fitting ending to my time with NAU. But what has mattered in this journey extends far beyond it.

In many ways, I am thankful to Doug Biber for showing me the linguistics I truly loved. Considering what a fundamental part of me grammar has always been, it is hard to believe that I could have been unsure, but that was a discovery I made in your class four years ago when things just fell into place. I know now that I have never been more at home than doing grammar that semester or assisting you with that class a year later. I am grateful for register research – for all your work that shaped my thinking and defined me in yet another major way: When I hear that I am a “register researcher”, it now sounds right. Finally, I am sure I am just another one in thousands of people who are grateful for multidimensional analysis. To me, however, the MDA has become more than a method, but a new way of looking at language, so elegant and complex, yet so perfectly intuitive; the source of endless fascination and endless questions I can only hope to answer. Thank you for showing me who I am as a linguist – I cannot think of anything more valuable in this journey than that. When I reflect on my experience at NAU, I realize my first two years here were your last two years of teaching. I often think how lucky that was.

I am thankful to Jesse Egbert for guiding me through these years every step of the way. In these reflections on how it all started, I try to mentally record every big and little thing you have done for me. I always lose count. But what I know is this. Somehow you always knew what was good for me before I did and better than I did. Learning from you has always been transformative; statistics or programming, it was always something that intimidated me the most, something I was prepared to hate, but loved in spite of my bias. It always led to self-discovery

and change. Thank you for being patient with my stubbornness and listening to my strong opinions, for engaging with every thought I had, every tangent my mind went on, and helping me grow through all of them. You validated every small piece of work I did – there were few things that lifted me up along the way more than recognition that came from you. I am thankful for every question that you shared with me, all the conversations, and all the ideas that emerged from them. They were the highlight of my time with NAU and will be among my fondest memories. Thank you for making this experience this meaningful.

I would like to thank my committee members – Naoko Taguchi and Michaela Mahlberg, whose work has informed my thinking on questions of individual choice, and agency throughout this project. I am especially thankful for your keen interest and engagement with all facets of this topic, your enthusiasm, and your efforts to enrich and strengthen this work through thoughtful questions, insightful comments, and for the challenging and stimulating discussions at earlier stages of this research. Your insights in this project as well as your own work have given me the examples I need as a researcher and linguist.

I want to thank the person who over the past years has become my friend – Daniel Keller. I was learning from you every time I heard you talk, and every time I would discover something so gripping that it stayed with me and stretched my thoughts in ways I had not known existed. Thank you for your talent, for taking my mind to fascinating places, but most of all, for your friendship in the hardest of times. As I look ahead into this future in linguistics, one of my most inspiring thoughts is that of witnessing the extraordinary things I know you will do.

Linguistics for me began much earlier than this PhD program. To the people who first recognized and nurtured the linguist in me – the faculty of the Foreign Languages Department of Moscow State Pedagogical University: Elena Freydina, Olga Blokhina, Tatiana Nesterova,

Svetlana Stepanova, Olga Kleymenova, and Olga Silina – I want to say that everything you see me do I owe to your wisdom, your passion, and your dedication to linguistics that shaped me all those years ago. Thank you for giving me this language. For teaching me to love it.

Finally, I would like to thank my family for trying to understand why I wanted this, for being close while being so far, and for their faith and love. These four years have only been made possible by my mom, who has sacrificed the most and for whom they have been the hardest.

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements.....	iv
List of tables.....	x
List of figures.....	xii
Chapter 1. Introduction.....	1
1.1 Variation according to characteristics of the users.....	2
1.2 Variation according to the situation of use.....	3
1.3 Interplay between perspectives on variation according to users and use.....	5
1.4 Goals of the dissertation.....	6
1.5 Overview of the dissertation.....	7
Chapter 2. Variation according to use and characteristics of the users. Literature review.....	9
2.1 Linguistic variation according to use.....	9
2.1.1 Register in Systemic Functional Linguistics.....	10
2.1.2 Register in Text-Linguistics.....	14
Situational variation within registers – functional correspondence at new levels of analysis.....	19
2.2 Linguistic variation according to characteristics of the users.....	24
2.2.1 Social variation.....	24
Gender as a marker of social difference.....	26
Age as a marker of a life stage and skill development.....	33
2.2.2 Individual variation.....	36
2.3 Considerations of use and user characteristics – previous recognition of the mutual relationship.....	43
2.3.1 Social dialect and attention to speech.....	43
2.3.2 Social dialect and register.....	46
2.3.3 Social dialect and individual variation.....	48
2.3.4 Register and individual variation.....	50
2.4 Goals of the present study.....	55
Chapter 3. Method.....	59
3.1 Corpus of 100 Idiolects.....	59
3.1.1 Registers in the corpus.....	61
3.2 Linguistic features.....	63
3.3 Quantitative method.....	68
3.3.1 Patterns of functional linguistic variation.....	68
3.3.2 Variance explained by social group, individual differences, and register.....	72
3.3.3 Social group with relation to register.....	73
3.3.4 Communicative differences within registers.....	74
3.3.5 Individual variation across registers.....	80
Chapter 4. Dimensions of functional linguistic variation.....	85
4.1. Dimension 1. Oral elaboration vs. Information density.....	88
4.1.1 Registers along Dimension 1.....	91
4.1.2 Social group and individual variation on Dimension 1.....	102
4.2 Dimension 2. Abstract vs. Concrete discourse.....	108
4.2.1 Registers along Dimension 2.....	110

4.2.2 Social group and individual variation on Dimension 2.....	117
4.3 Dimension 3. Others-oriented descriptive vs. Self-oriented or interactive discourse.....	121
4.3.1 Registers along Dimension 3.....	122
4.3.2 Social group and individual variation on Dimension 3.....	130
4.4 Dimension 4. Evidence-based stance.....	133
4.4.1 Registers along Dimension 4.....	134
4.4.2 Social group and individual variation on Dimension 4.....	141
4.5 Conclusion.....	144
Chapter 5. Social group and register.....	146
5.1 Age and gender groups across register.....	146
5.1.1 Age Groups across registers on Dimension 1.....	147
5.1.2 Gender Groups across registers on Dimension 1.....	152
5.2 Age and gender groups within registers.....	155
5.2.1 Text analysis: Age within business memos on Dimension 1.....	157
5.2.2 Text analysis: Gender within business memos on Dimension 1.....	162
5.2.3 Text analysis: Age within text messages on Dimension 4.....	165
5.2.4 Text analysis: Gender within text messages on Dimension 4.....	168
5.3 Conclusion.....	170
Chapter 6. Communicative variation across texts.....	172
6.1 Communicative variation across texts: Oral elaboration vs. Information density.....	172
6.1.1 Application of a framework for situational analysis.....	176
6.1.2 Variation within essays.....	183
Text analysis: Central and peripheral essays above and below the register mean.....	196
6.1.2 Variation within emails.....	211
Text analysis: Central and peripheral emails above and below the register mean.....	221
6.1.3 Variation within image descriptions.....	228
Text analysis: Central and peripheral image descriptions above and below the register mean.....	231
6.1.4 Variation within business memos.....	235
Text analysis: Central and peripheral business memos above and below the register mean.....	238
6.2 Communicative variation across texts: Evidence-based stance.....	241
6.2.1 Variation within essays.....	244
Text analysis: Central and peripheral essays above and below the register mean.....	248
6.2.2 Variation within emails.....	253
Text analysis: Central and peripheral emails above and below the register mean.....	258
6.2.3 Variation within image descriptions.....	261
Text Analysis: Central and peripheral image descriptions above and below the register mean.....	263
6.2.4 Variation within business memos.....	266
Text analysis: Central and peripheral business memos above and below the register mean.....	267
6.2.5 Variation within evaluations.....	269
Text analysis: Central and peripheral evaluations above and below the register mean.....	271
6.3 Conclusion.....	273

Chapter 7. Individual variation across registers.....	275
7.1 Individual variation on Dimension 1: Oral elaboration vs. Information density.....	277
7.1.1 Case Study 1: Author ‘central’ across registers.....	280
7.1.2 Case Study 2: Authors ‘peripheral’ across registers	287
7.1.3 Case Study 3: Authors of different distance groups and directionality across registers.....	294
7.2 Individual variation on Dimension 4: Evidence-based stance.....	298
7.2.1 Case Study 1: Author ‘central’ across registers.....	301
7.2.2 Case Study 2: Authors ‘peripheral’ across registers.....	305
7.2.3 Case Study 3: Authors of different distance groups and directionality across registers.....	313
7.3 Conclusion.....	317
Chapter 8. Discussion and conclusions.....	318
8.1 Role of social group in linguistic variation across texts.....	319
8.2 Communicative variation across texts within registers: Functional correspondence between situation and language.....	326
8.2.1 The role of topic and purpose in variation across texts.....	327
8.2.2 The role of additional granular situational factors.....	336
8.2.3 Situational and linguistic granularity in broader context.....	338
8.3 Individual variation across registers: Individual freedom and ideas of appropriacy.....	344
8.4 Conclusion: Limitations and future directions.....	350
8.4.1 Limitations of the corpus.....	350
8.4.2 Limitations of the method.....	352
8.4.3 Future directions	354
Situation of use.....	354
Characteristics of the users.....	356
8.4.4 Conclusion.....	357
References.....	358
Appendix 1. Scree plot of eigenvalues.....	375
Appendix 2. Factor loadings.....	375
Appendix 3. Social groups across registers.....	378
Appendix 4. Social groups within registers.....	396
Appendix 5. Text scores, register means, measures of distance.....	427

LIST OF TABLES

Table 2.1. Framework for situational register analysis (Biber & Conrad, 2019, p. 40).....	16
Table 3.1. Corpus of the study: Variables, variable levels, authors, and texts.....	61
Table 3.2. Application of the framework for situational register analysis (Biber 1994; Biber & Conrad, 2019).....	63
Table 3.3. Features initially included in the MDA.....	66
Table 3.4. Summary of research questions and respective methods.....	83
Table 4.1. Dimension of functional linguistic variation.....	86
Table 4.2. Variance explained by each predictor variable (age, gender, author, register) in linguistic variation on each dimension.....	88
Table 4.3. Registers along Dimension 1. Descriptive statistics.....	92
Table 4.4. Age groups on Dimension 1. Descriptive statistics.....	103
Table 4.5. Gender groups on Dimension 1. Descriptive statistics.....	104
Table 4.6. Registers on Dimension 2. Descriptive statistics.....	110
Table 4.7. Age groups on Dimension 2. Descriptive statistics.....	118
Table 4.8. Gender groups on Dimension 2. Descriptive statistics.....	118
Table 4.9. Registers on Dimension 3. Descriptive statistics.....	123
Table 4.10. Age groups on Dimension 3. Descriptive statistics.....	131
Table 4.11. Gender groups on Dimension 3. Descriptive statistics.....	131
Table 4.12. Registers on Dimension 4. Descriptive statistics.....	135
Table 4.13. Age groups on Dimension 4. Descriptive statistics.....	143
Table 4.14. Gender groups on Dimension 4. Descriptive statistics.....	143
Table 5.1. Registers within age groups on Dimension 1. Descriptive statistics.....	148
Table 5.2. Variance explained by register within age groups on Dimension 1.....	148
Table 5.3. Age groups across registers on Dimension 1: Pairwise comparisons.....	151
Table 5.4. Gender groups across registers. Dimension 1.....	153
Table 5.5. Variance explained by register within gender group. Dimension 1.....	153
Table 5.6. Gender groups across registers on Dimension 1: Pairwise comparisons; Cohen's <i>d</i> ...	155
Table 5.7. Age groups within business memos. Descriptive statistics.....	158
Table 5.8. Variance explained by age within business memos.....	158
Table 5.9. Gender groups within business memos.....	164
Table 5.10. Variance explained by gender in business memos.....	164
Table 5.11. Age groups within text messages. Descriptive statistics.....	167
Table 5.12. Variance explained by age within text messages.....	167
Table 5.13. Gender groups within text messages. Descriptive statistics.....	170
Table 5.14. Variance explained by gender in text messages.....	170
Table 6.1 Number of texts per distance and direction group within each register. Dimension 1: Oral Elaboration vs. Information Density.....	174
Table 6.2. Application of the framework for situational register analysis (Biber 1994; Biber & Conrad, 2019).....	180
Table 6.3. Number of essays per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density.....	185
Table 6.4. Essay topics within each distance and direction group. Dimension 1: Oral Elaboration vs. Information Density	185
Table 6.5. Number of emails per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density.....	211

Table 6.6. Email topics within each distance and direction group. Dimension 1: Oral Elaboration vs. Information Density.....	215
Table 6.7. Number of descriptions per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density.....	229
Table 6.8. Number of business memos per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density.....	236
Table 6.9. Number of texts per distance and direction group within each register. Dimension 4: Evidence-based Stance.....	243
Table 6.10. Number of essays per distance and direction group. Dimension 4: Evidence-based Stance.....	244
Table 6.11. Essay purposes within each distance and direction group. Dimension 4: Evidence-based Stance.....	245
Table 6.12. Number of emails per distance and direction group. Dimension 4: Evidence-based Stance.....	253
Table 6.13. Email topics within each distance and direction group. Dimension 4: Evidence-based Stance.....	254
Table 6.14. Number of descriptions per distance and direction group. Dimension 4: Evidence-based Stance.....	261
Table 6.15. Number of business memos per distance group and direction of deviation. Dimension 4: Evidence-based Stance.....	266
Table 6.16. Number of evaluations per distance and direction group. Dimension 4: Evidence-based Stance.....	270

LIST OF FIGURES

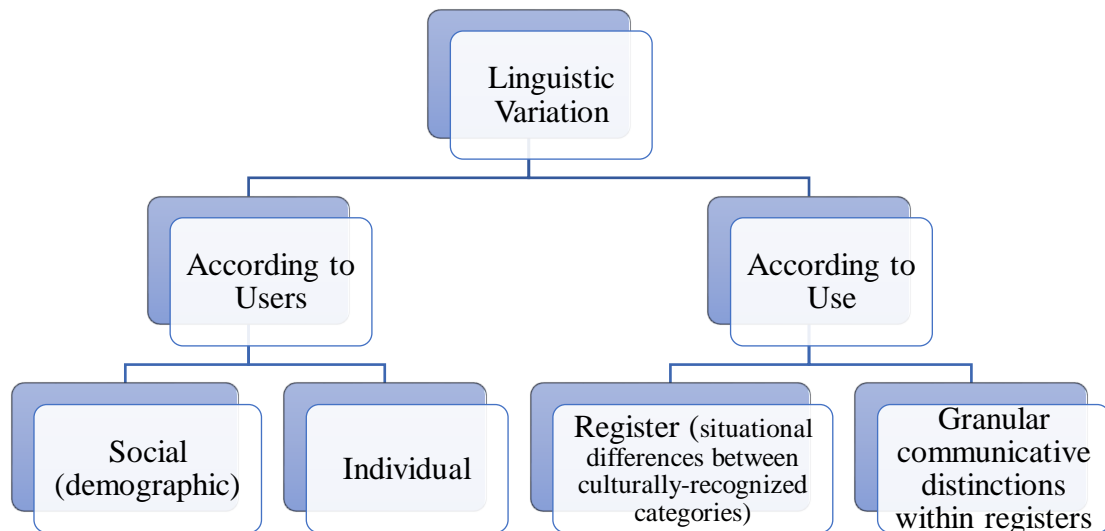
Figure 1.1. Approaches to investigating linguistic variation.....	2
Figure 2.1. Text-linguistic register framework: the relationship between situation, linguistic features, and function.....	15
Figure 2.2. Approaches to investigating linguistic variation	55
Figure 4.1. Registers along Dimension 1.....	93
Figure 4.2. Age groups on Dimension 1.....	104
Figure 4.3. Gender groups on Dimension 1.....	104
Figure 4.4. Registers on Dimension 2.....	111
Figure 4.5. Age groups on Dimension 2.....	118
Figure 4.6. Gender groups on Dimension 2.....	119
Figure 4.7. Registers on Dimension 3.....	123
Figure 4.8. Age groups on Dimension 3.....	131
Figure 4.9. Gender groups on Dimension 3.....	132
Figure 4.10. Registers on Dimension 4.....	136
Figure 4.11. Age groups on Dimension 4.....	143
Figure 4.12. Gender groups on Dimension 4.....	144
Figure 5.1. Register differences within the 19-year-old group. Dimension 1.....	149
Figure 5.2. Register differences within the 20-year-old group. Dimension 1.....	149
Figure 5.3. Register differences within the 21-year-old group. Dimension 1.....	150
Figure 5.4. Register differences within the 22-year-old group. Dimension 1.....	150
Figure 5.5. Register differences within the 23-year-old group. Dimension 1.....	151
Figure 5.6. Register differences within the male group. Dimension 1.....	154
Figure 5.7. Register differences within the female group. Dimension 1.....	154
Figure 5.8. Age groups within business memos: Dimension 1.....	159
Figure 5.9. Genders within business memos: Dimension 1.....	164
Figure 5.10. Age groups within text messages: Dimension 4.....	167
Figure 5.11. Gender groups within text messages: Dimension 4.....	170

CHAPTER 1. INTRODUCTION

Variation is an inherent attribute of language. While in many, especially earlier, linguistic traditions this statement could have been contested (e.g., de Saussure's (1916) structuralism, Bloomfield's (1933) descriptivism, or Chomsky's (1979) generative approach), for sub-fields concerned with the dynamic nature of language, explaining variation has been one of the central goals of the linguistic enterprise. In these linguistic fields, linguistic variation is pervasive and may be targeted at all linguistic levels – it may concern phonetic and phonological differences, morphology, lexis and semantics, syntax and grammatical categories, or occur on the level of discourse. Crucially, most of language variation on all these levels is not haphazard, but has a systematic basis and is attributed to a variety of factors – linguistic and various nonlinguistic or external constraints determining language use (e.g., Wolfram, 1995).

These external nonlinguistic factors have been commonly classified as characteristics of language users and factors associated with the situation of use – a conception of variation that dates back to at least Halliday's distinction between diatypic and dialectal variation. While dialectal variation is based on characteristics of the user, such as social class or regional affiliation, gender, age, or ethnicity, diatypic variation reflects the kind of social activity performed by the speakers (Halliday, 1978). This classification, however, involves further distinct approaches to variation based on the specific nature of the predictors as illustrated in Figure 1.1

Figure 1.1. Approaches to investigating linguistic variation



1.1 Variation according to Characteristics of the Users

Figure 1.1 shows that one set of nonlinguistic factors involves social and individual characteristics of the users. The study of language in social context has concerned itself with questions regarding the relationship between social characteristics of the speakers and their language use, and the goal of sociolinguistics has been defined as the attempt to find correlations between social structure and linguistic use (Gumperz, 1971, p. 223). Social structure in turn has been viewed as geographic, social class, occupational or educational backgrounds, as well as racial and ethnic, gender, or age group identities (e.g., Francis, 1983; Kroch, 1978). The norms of linguistic use are thus determined by these cultural and social group affiliations, and it is belonging to a social group or a community, sharing in its established norms, participating in common activities and tasks, and interacting with the other members that shapes behavior (e.g., Eckert, 2000). Since in the sociolinguistic perspective language use reveals group affiliation,

conformity to group norms is often seen as subconscious linguistic behavior. Such linguistic behavior is thus seen as indexical, i.e., indicating belonging to the group (e.g., Wolfson, 1989).

Figure 1.1 further shows that linguistic variation according to users may additionally be due to preferences of individual speakers, i.e., speaker idiolects, and on an even finer-grained level, reflect further variation within a single speaker's language use (Francis, 1983). The individual is central to this view (e.g., Johnstone, 1996), and a social or regional dialect is seen to be composed of idiolects. While individuals adhere to the social expectations of the communities that they belong to, within the norms of these communities, the individual has the creative freedom to express themselves, choosing from the available linguistic repertoire. In this view, it is this individual linguistic freedom, rather than group membership, that is treated as the locus of variation and is believed to provide the most detailed and complete view of language use. This language use, not attributed to social affiliation, is typically interpreted as a matter of individual linguistic preference for a particular way of self-expression.

1.2 Variation according to the Situation of Use

Another set of nonlinguistic factors contributing to linguistic variation involves considerations of use (Figure 1.1). A user from any region, social network, racial or ethnic background, age group, or gender navigates a variety of communicative contexts and makes linguistic choices in accordance with their purpose in communication, the relationship with their interlocutors, amount of shared knowledge between them, the nature of that knowledge (such as specialist or general), channel of communication, amount of time available for planning and preparation, or the production circumstances in which the text is created (Biber, 1994; Biber & Conrad, 2019). In this tradition, linguistic variation is explained through the fact that texts are always

produced in response to the demands of a particular communicative situation and can be characterized in terms of multiple situational factors.

Figure 1.1 further illustrates that the effect of the situation of use has also been approached from different perspectives. On the one hand, an extensive body of research points to the paramount role of ‘register’, a language variety “associated with the situation of use” (Biber & Conrad, 2019, p. 6) in predicting linguistic variation (e.g., Biber, 1988). In a traditional conceptualization of register, texts unified by their situational characteristics belong to a register category and differ from texts that are united by a different set of situational parameters and thus belong to a different register. It is therefore register categories that are seen as predictors of linguistic variation. In this research paradigm, general accounts of language without attention to register are considered meaningless (Biber, 2012; Seoane & Biber, 2021, p. 3), as the differences between situational varieties are so great that no such general account of language use can be true for any particular variety.

Recently, however, an additional consideration in variation according to the situation of use has been that of granular communicative differences within registers (Figure 1.1). This research has shown that registers are not always as situationally uniform as they were previously accepted to be, but rather exhibit substantial internal variation across texts or even within texts. This register-internal variation has been associated with highly specific communicative distinctions, such as differences in communicative purpose (e.g., Biber, Egbert & Keller, 2020; Egbert & Gracheva, 2023). Importantly, these specific communicative differences within registers have been shown to explain additional linguistic variation, previously unaccounted for by existing register categories. This increasingly granular approach to variation by use has led to a reconceptualization of registers as named, culturally recognized varieties (true in the older

approach to register as well), but not defined in their entirety through a set of situational characteristics, as these characteristics vary from one text of a register to another.

1.3 Interplay between Perspectives on Variation according to Users and Use

The research traditions focusing on variation according to use and according to characteristics of the users have each produced a wealth of knowledge on language variation explained by their respective groups of non-linguistic factors. However, it becomes apparent from this overview of their agenda and goals that variation is viewed through a different lens by each. While these perspectives are in no way mutually exclusive – i.e., a speaker who belongs to a certain age cohort, gender, ethnic group, and social class produces language in a range of diverse situations of use, and their language is at the same time distinct from the language produced by other speakers – so far, little research has been done to bring together these distinct approaches to linguistic variation. Research paradigms focusing on social and situational predictors, for example, have been said to follow what looks like a path of mutual avoidance: Macaulay (1991, p. 6), for instance, states that despite their interest in social context, sociolinguists seem to have ignored considerations of context, while quantitative discourse analysts, who have investigated “textual distribution of variables” based on situational considerations, are typically not concerned with social characteristics of the authors of the texts. It will be shown in the next chapter that these statements made over three decades ago are almost equally true today.

Similarly, investigations of situationally determined and individual variation have been treated as distinct analytical perspectives. The register perspective is concerned with investigating features that are frequent and pervasive in a particular situation of use and perform a particular function associated with that situation. This kind of situationally determined variation has been on the sidelines of research into individual expression, which has either been

done within a single domain (e.g., Johnstone's, 1996 personal narratives), or has included a variety of registers without accounting for the effect of that choice (e.g., Pinker, 2014). The goal of this dissertation is to investigate the interactions that may exist between these distinct approaches to linguistic variation.

1.4 Goals of the Dissertation

While non-linguistic factors associated with users and use have been shown to predict linguistic variation, the interaction between these factors is under-researched or, in some cases, not researched at all. The overarching goal of this dissertation is to bring together these independent strands of language research which have so far focused on one set of possible predictors of language variation, namely social group, individual characteristics, register, or more recently, communicative differences within registers, in order to investigate the relative importance of these factors in accounting for functional linguistic variation. This single overarching goal is comprised by a combination of specific quantitative and qualitative goals:

1. To investigate the relative importance of non-linguistic factors associated with the users – authors' gender, age, and individual speaker characteristics – and the non-linguistic factors associated with use – register – in explaining functional linguistic variation across texts

The first goal of the study is quantitative: To address the question of the relative importance of the variables of age, gender, individual characteristics, and register, the study measures the extent to which these nonlinguistic factors affect linguistic variation across texts cumulatively and independently, thus identifying the unique contribution of each factor once the others have been accounted for.

As noted above, despite the acknowledgement that linguistic variation is predicted by a combination of interrelated non-linguistic factors, the interaction among them has not been investigated. Specifically, the role of register has been understated in sociolinguistic research, centered primarily around demographic characteristics of speakers, as well as in research on individual differences, the goal of which is to examine individual patterns of use irrespective of the situational distinctions. The study thus aims to reconcile the research paradigms investigating demographic factors, individual expressions, and register through the following qualitative goals:

2. To provide accounts of language use by age and gender groups within and across registers
3. To account for possible communicative distinctions within registers
4. To analyze authors' individual language use across registers

To accomplish this, the study makes use of the corpus of '100 Idiolects', which consists of speech and writing samples of 112 individuals and is annotated for the speakers' gender and age – the demographic variables used in the study. Each speaker in the corpus has produced texts in 7 registers: interviews on a single topic, image descriptions, emails, text messages, academic essays, written responses to a task (positive and negative evaluations of a place), and business memos.

1.5 Overview of the Dissertation

Chapter 2 outlines the scope, main focus, and key findings in the three research traditions outlined in this introduction. Specifically, Chapter 2 presents the common non-linguistic variables and the insights gained from them in 1) the sociolinguistic tradition aiming to explain linguistic variation associated with the social characteristics of the users, with a focus on age and

gender variation; 2) individual linguistic variation, and 3) situational variation defined as variation due to “situation of use” (Biber & Conrad, 2019, p. 6) at different levels of analysis – between and within registers.

Chapter 3 presents the corpus used in the study, the linguistic features investigated, and the methodology. It introduces the corpus of ‘100 Idiolects’ and discusses its design with a focus on the characteristics of the users. The chapter then provides a description of the seven registers featured in the corpus – interviews, emails, text messages, academic essays, image descriptions, written responses to a task, and business memos. The chapter then focuses on the dependent variables of the study and motivates the linguistic features associated with demographic, individual, and register variation selected for the analysis and discusses the quantitative method used to (1) identify the underlying patterns of functional linguistic variation in the corpus; (2) investigate the cumulative and unique contribution of the nonlinguistic factors to linguistic variation; (3) examine social groups across and within registers; (4) analyze communicative distinctions within registers; (5) investigate individual authors across registers.

Chapters 4-7 present the results of the study. Chapter 4 describes the identified dimensions of functional linguistic variation, reports the contribution of each nonlinguistic factor to explaining linguistic variation on each of the dimensions, and presents the analyses demonstrating the relevant group differences. Chapter 5 examines social group with respect to register; Chapter 6 is devoted to communicative differences within registers; and Chapter 7 presents results on individual variation across registers. Chapter 8 synthesizes these findings and makes conclusions regarding the interrelationship between the demographic variables, individual differences, and register and their respective roles in functional linguistic variation.

CHAPTER 2. VARIATION ACCORDING TO USE AND CHARACTERISTICS OF THE USERS. LITERATURE REVIEW

The present chapter aims to i. review the current state of knowledge within the linguistic sub-fields investigating variation according to the situation of use (register and granular communicative variation within registers) and characteristics of the users (social and individual factors); ii. discuss the previous attempts to integrate approaches to linguistic variation associated with users and use, and iii. state the research gap and research questions addressed by the present study.

2.1 LINGUISTIC VARIATION ACCORDING TO USE

Situation of use has long been of interest to linguists from several research traditions. Reid (1956, p. 32) investigates registers of speech and writing and refers to them as varieties associated with “different social situations”; Joos (1962, p. 10) emphasizes the importance of recognizing that “a community has a complex structure, with various different needs and occasions”, which leads him to distinguish between five varieties associated with use from ‘frozen’ to ‘familiar’; Crystal and Davy (1969, p. 11) write about language users’ ability to relate features of their language to their non-linguistic experiences or events, which they term ‘situations’ or ‘extra-linguistic contexts’; Hymes (1974, p. 440) discusses language variation in relation to “types of situation”; Halliday and Hasan (1976, p. 22) point out the relationship between linguistic features and “configurations of situational features”; Ure (1982, p. 5) writes that registers are “situations for which appropriate patterns are available” to speakers within certain social norms; Brown and Yule (1983) contemplate the role of context for discourse analysis, and in communication studies, researchers observe a transition in research practices

beginning in mid-20th century from examining abstract language models to “forms of situated language use within social practices” and a focus on language users’ tendency to adapt their linguistic resources “within specific varieties of communicative conduct” (Asif & Frog, 2015, p. 13). While there is a considerable inconsistency in terminology, with the terms ‘style’ (e.g., Crystal & Davy, 1969), ‘register’ (Halliday, 1978), and ‘genre’ (Halliday & Hasan, 1976) used by these scholars to refer to situational variation, the general consensus is that any account of language as a monolithic, unvaried phenomenon will be inaccurate. Language users are adept at recognizing situational distinctions in their environment as they process and produce language.

The following sub-sections are devoted to register as a predictor of linguistic variation and present two major schools of thought that have shaped register research – Systemic Functional Linguistics (SFL) (Section 2.1.1) and Text-Linguistics (Section 2.1.2). Section 2.1.2 additionally discusses granular communicative distinctions within registers as sources of linguistic variation, recently identified and investigated within the Text-Linguistic research tradition.

2.1.1 Register in Systemic Functional Linguistics

The role of situation has been central in Systemic Functional Linguistics (SFL), a school of thought pioneered by M.A.K. Halliday and his collaborators, but its foundation is thought to have been laid even earlier by Halliday’s mentor, the linguist J.R. Firth, and the anthropologist B. Malinowsky, a collaborator of Firth’s. In SFL, language is seen as a meaning-constructing (semogenetic) system, and specific instances of language produced in particular situations (spoken or written texts) are concrete actualizations of the meaning potential of a language as a whole. Registers therefore are construed as the intermediaries between the level of language and the specific instantiations of its meaning potential in each spoken or written text. That is, SFL

scholars see the major role of situational context in the construction of meaning, and from this point of view, registers represent sub-potentials of a language in conveying meaning (Shore, 2015).

In the teachings of the British linguist J.R. Firth, this line of thought came into being in large part in response to American structuralism – specifically, L. Bloomfield’s (1933) notion of describing language without attention to meaning as well as Saussure’s statement that it is *langue*, language as an abstract system, rather than *parole*, its manifestation in discourse, that should be the object of scientific study. In contrast, Firth’s (1957) conception of linguistic analysis crucially involved interpretations of the meaning emergent in communication and reflected the views of Malinowsky (1946), who maintained that language performs certain functions in particular situations of use and that way acquires meaning. Firth (1957, p. 182; 1968, pp. 177-178) further develops Malinowsky’s notion ‘the context of situation’, which for Firth includes such considerations as the characteristics of the participants and their personalities, their verbal and non-verbal actions (i.e., what is said and done), the objects in the participants’ environment and events associated with the communicative act, and the effect achieved by the communicative event. Based on these characteristics, contexts of situation can be classified, and as a result, language typical of certain situational contexts – the ‘restricted language’ (Firth, 1968, pp. 29-30, 87, 106, 112) – can be analyzed. Restricted language may be found both in speech and writing and necessarily involves “specialized vocabulary, grammar and style.” It is apparent that this view is in stark contrast with the structuralist holistic approach to language as a single system. Conversely, Firth (1968, p. 186) conceptualizes language as a “polysystemic” phenomenon with multiple varied structures, adapted to situational contexts. The meaning

potential of language is realized in concrete instances of communication on all linguistic levels from phonology to syntax.

Firth's view of language as a network of systems is thus at the foundation of Systemic Functional Linguistics and persists in M.A.K. Halliday's work. Halliday advanced the notion of language variation in response to the situation of use and the functional nature of this variation:

A functional approach to language means, first of all, investigating how language is used: trying to find out what are the purposes that language serves for us, and how we are able to achieve these purposes through speaking and listening, reading and writing. But it also means more than this. It means seeking to explain the nature of language in functional terms: seeing whether language itself has been shaped by use, and if so, in what ways – how the form of language has been determined by the functions it has evolved to serve (Halliday, 1973, p. 7).

Similar to Firth's view, the importance of registers for Halliday lies in their semogenetic potential – i.e., the construction (or 'realization') of meaning in different situations through lexicogrammatical resources. Halliday proposes that linguistic variation according to the situation of use can be described based on three situational variables: 'field', 'tenor', and 'mode' of discourse (Halliday, 1978), where 'field' encompasses characteristics of the setting, 'tenor' refers to the relationship between participants, and 'mode' relates to the role played by language. Each of the parameters roughly corresponds to the three metafunctions of language (i.e., general superordinate language functions encompassing more specific ones) identified by Halliday: ideational (the function of processing, interpreting, and organizing the knowledge about the events of the world), interpersonal (the function of forging relationships, making judgments,

expressing attitudes and feelings), and textual (the function that relates to cohesion and organization of texts in discourse).

The parameters of ‘field’, ‘tenor’, and ‘mode’ each encompass a vast number of more specific situational characteristics (e.g., specific information about the participants and the relationships between them within ‘tenor’). Gregory (1967), for example, notes that communicative purpose is a sub-parameter of the field of discourse, and ‘purposive roles’ may be further subcategorized. Nevertheless, the framework has been criticized as vague and imprecise even by scholars working within the SFL tradition. Hasan (2009a), for example, notes that the framework lacks “checkable” criteria to analyze the situation of the register, while Shore (2015, p. 63) states that although the three parameters are “umbrella categories” and are intuitive, it is not immediately clear what specific situational characteristics each of them comprises. This concern is partially addressed by Halliday through ‘first-order’ and ‘second-order’ field, tenor, and mode, which allow more specificity of situational description. For example, if first-order field denotes the setting of an academic lecture, the second-order field specifies the discipline and/or the subject matter of that lecture (ethics, quantum physics, etc.); while first-order tenor relates to participants’ social roles, such as buyer and seller or interviewer and interviewee, second-order tenor specifies their linguistic roles, such as questioner and answerer; finally, mode at its basic level encapsulates the distinction in medium between speech and writing, but speech and writing in their turn are comprised of numerous distinctions based on immediacy of production (e.g., shared physical setting or mediated by technology), time (language produced in real-time production or prerecorded), frequency of turns, among many others (Shore, 2015, p. 63). It is obvious that this abstract nature of the three main situational parameters leaves

considerable room for additional analysis, and some SFL scholars have shared a vision for a more concrete and comprehensive approach. Matthiessen (1993, p. 274), for example, wrote:

It would be an important contribution to describe the overall semiotic space in which these ‘registers’ are located relative to one another – to provide a general account of field, tenor and mode and to specify the values for each variety listed above. This would introduce greater precision in register analysis and might very well invite us to re-interpret some of the varieties that have been identified in the past.

It is this vagueness of criteria for situational analysis that is considered a major weakness of the SFL framework by researchers who adopt the Text-Linguistic approach to register studies, discussed in the next section.

2.1.2 Register in Text-Linguistics

Similar to the SFL conceptualization, in the Text-Linguistic tradition, registers are named culturally recognized language varieties associated with the situation of use (Biber, Egbert & Keller, 2020). Unlike the SFL approach, however, the Text-Linguistic approach (Biber, 1986; Biber, 1988) is associated with the rise of corpus linguistic methods and sets the goal of comprehensive linguistic descriptions of registers as well as cross-register comparisons on the basis of a wide range of specific situational parameters. The use of corpora, large collections of texts from a register, is conducive to the goal of a comprehensive linguistic description generalizable to other texts from that register. This goal is achieved through a description of specific situational characteristics outlined by a framework for situational register analysis, an application of quantitative methods to identify the linguistic composition of each text in a corpus,

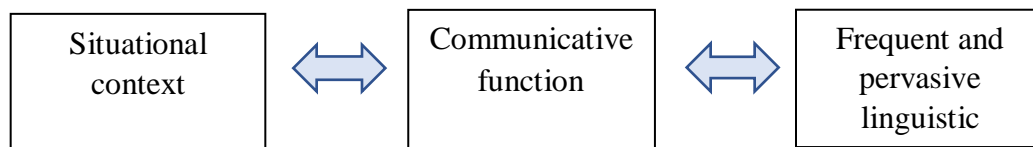
i.e., the rates of occurrence of lexico-grammatical features in each text, and an interpretation of the observed linguistic patterns in functional terms.

The Text-Linguistic approach to analysis of register variation thus hinges on the following core considerations:

1. Register categories are described in terms of their situational and linguistic characteristics
2. There is a systematic functional relationship between the situational characteristics of a register and the prominent, pervasive linguistic features of that register

The links between the situation, the linguistic features associated with the situation, and the function performed by these features in context has been referred to as ‘the register triangle’ and are graphically represented in Figure 2.1.

Figure 2.1. Text-linguistic register framework: the relationship between situation, linguistic features, and function (e.g., Biber & Conrad, 2019, p. 6)



Since it is situation that determines language use, situation is seen as primary to language. Analyzing the situational context is therefore the starting point for register analysis in the Text-Linguistic approach. The situational context is defined through specific situational parameters, which include characteristics of participants and the relations among them, the channel of communication, processing circumstances, setting, communicative purposes, and topic (Biber, 1994; Biber & Conrad, 2019, p. 40). Specifically, participants are characterized in terms of their number and possibility to concretely identify them (single/ plural/ unenumerated/ unidentified/ institutional), their social characteristics (age, education, profession), their relative social status

and power, the type of relationship between them (friendly, professional, etc.), the amount and type of shared knowledge between them (specialist or general), and the presence or absence of onlookers. The channel of communication includes consideration of mode (spoken or written) and medium – permanent, i.e., printed or transcribed, or transient, such as a face-to-face or a phone conversation. Processing circumstances are defined in terms of language production and comprehension – occurring in real time or planned, revised and edited (production) and quick reading/ skimming or careful reading (comprehension). The parameter of setting involves considerations such as whether the time and physical space are shared by the participants; setting can be further defined as private or public, with the exact setting specified. The framework further defines communicative purpose, type of factuality, and expression of stance. It is possible to identify a general communicative purpose (to inform, to edify, etc.) and a range of specific purposes (summarize information, describe methods, etc.). The message is further characterized as factual, based on opinion, speculative, or imaginative, expressing epistemic, attitudinal, or no overt stance. Finally, the general topical domain is identified (business, sport, etc.) and the specific topics within it. Table 2.1 summarized the framework for situational analysis in the Text-Linguistic school of thought (Biber & Conrad, 2019, p. 40).

Table 2.1. Framework for situational register analysis (Biber & Conrad, 2019, p. 40)

I. Participants

A. Addressor(s) (i.e., speaker or author)

1. single / plural / institutional / unidentified
2. social characteristics: age, education, profession, etc.

B. Addressees

1. single / plural / un-enumerated
2. self / other

C. Are there on-lookers?

II. Relations among participants

A. Interactiveness

B. Social roles: relative status or power

C. Personal relationship: e.g., friends, colleagues, strangers

D. Shared knowledge: personal and specialist

III. Channel

A. Mode: speech / writing / signing

B. Specific Medium:

permanent: taped / transcribed / printed / handwritten / e-mail / etc.

transient speech: face-to-face / telephone / radio / TV / etc.

IV. Processing circumstances

A. Production: real time / planned / scripted / revised and edited

B. Comprehension: real time / skimming / careful reading

V. Setting

A. Is the time and place of communication shared by participants?

B. Place of communication

1. Private / public

2. Specific setting

C. Time: contemporary, historical time period

VI. Communicative purposes

A. General purposes: narrate / report, describe, inform / explain / interpret, persuade, how-to / procedural, entertain, edify, reveal self

B. Specific purposes: e.g., summarize information from numerous sources, describe methods, present new research findings, teach moral through personal story

C. Factuality: factual, opinion, speculative, imaginative

D. Expression of stance: epistemic, attitudinal, no overt stance

VII. Topic

A. General topical “domain”: e.g., domestic, daily activities, business / workplace, science, education / academic, government / legal / politics, religion, sports, art / entertainment, etc.

B. Specific topic

C. Social status of person being referred to

On the other end of Figure 2.1 (last component of the ‘triangle’) are the linguistic features that occur frequently in the register as a result of its situational characteristics. The second step in register analysis in Text-Linguistics is therefore computing the rates of occurrence of lexicogrammatical features in each text of a register, thus measuring the linguistic composition of the register in continuous terms. According to Biber and Conrad (2019), the register approach to linguistic variation is concerned with frequent and pervasive features of a given register. A feature is not characteristic for a register if it occurs once; rather, particular situational configurations result in relatively high rates of occurrence of features associated with a particular

situation of use. Pervasiveness of linguistic features in a register is thus key to the Text-Linguistic conceptualization of the register approach to language variation.

Finally, the third, crucial component of the ‘register triangle’ and a link between situation and language is communicative function (Figure 2.1). Linguistic features occur frequently in a register because they serve a particular communicative function in a particular situation of use. The pervasive linguistic patterns identified by register studies are always interpreted in functional terms, thus consistently demonstrating the existence of functional correspondence between situation and language.

Register analyses produced within the Text-Linguistic research tradition include numerous descriptions of particular registers or comparative studies examining register variation in a range of registers (see, e.g., overview by Barbieri & Wizner, Appendix 1, Biber & Conrad, 2019). Comparative studies, in their turn, have been carried out between more general situational varieties, such as the registers of speech and writing (Biber, 1988), and more specific ones, such as a comparison of academic registers (Biber, 2006) or, more specifically, journal articles in published academic journals (Gray, 2015). Apart from register descriptions and comparisons, studies have addressed numerous specific questions about the linguistic features characteristic for particular registers on the basis of their situational distinctions (e.g., an investigation of complexity patterns characteristic for conversation and academic writing (Biber & Gray, 2010)). Regardless of their scope and the exact research goals, these studies have repeatedly observed that register is a powerful predictor of variation on all linguistic levels (Biber, 2012) and have all contributed to the state of the art on register variation – the systematic relationship between language and the situation of use.

Situational Variation within Registers – Functional Correspondence at New Levels of Analysis

While between-register comparisons have been the most prominent in register research, the notion of functional correspondence between situation and language is not restricted to differences between existing culturally recognized varieties. More recent work in Text-Linguistics has tackled new questions, many of which were uncharted territory in traditional register research – variation across texts within a single register, texts without an *a priori* established register category, alternative textual units, and the continuous rather than categorical nature of situational parameters – and observed trends pointing to a much more pervasive nature of the relationship between situation and language.

Substantial variation within registers has consistently been observed even in studies whose main focus was on differences between culturally recognized categories (e.g., Biber, 1988, Chapter 8) or has been targeted by studies focusing more closely on a particular register. Gray (2015), for example, investigates variation in academic journals and distinguishes between empirical and nonempirical studies, further identifying specific article types, such as qualitative, quantitative, and mixed methods empirical studies, and theoretical and evaluative articles within the nonempirical group. Further, theoretical articles include ‘author interpretation’ and ‘general theoretical’ papers, and the evaluative group includes reviews, ‘commentaries/ forums’, and research syntheses. Thus, the goal of the study was to examine linguistic variation across these publication types. As the situational distinctions between the publication types are highly specific (unlike considerations such as mode or production circumstances in the framework in Table 2.1, which do not vary across journal articles), Gray proceeds to adapt the framework and identify specific situational parameters, such as the type of journal, nature of evidence and data,

object of study, the method used, explicitness of research design, and others. Gray then shows that these specific situational factors explain linguistic variation across texts, adding substantially to the effect of discipline.

While Gray (2015) shows that variation *across* texts of a register has a situational basis, situational variation within registers has also been observed *within* texts. For example, several studies have recognized that linguistic differences exist between the sections of research articles and correspond to their IMRD structure (Introduction, Methods, Results, Discussion), which Biber and Finegan (1994, p. 221) associate with “micro-purpose variation”. That is, systematic linguistic differences among sections correspond to their distinct communicative purposes – to motivate the study, to describe the methods used, to present findings, and to situate the results in the existing body of work.

While the sections of a research article are marked with existing external boundaries (i.e., headings signaling a clear beginning and end of a section), commonly used in register research to identify textual units for analysis, variation among texts may go beyond such conveniently identifiable textual units. An example is a large-scale investigation of the searchable web (Biber & Egbert, 2018), in which many of the texts did not have an initially assigned register category or any external cues suggesting a category. As a result, texts were placed into categories through the process of coding – identification of the mode, interactivity, and purpose for each document. However, a recurring combination of communicative purposes rather than a clear purpose for each text resulted in hybrid registers, where texts were classified as how to/ opinion, opinion/ informational description, narrative/ opinion, narrative/ informational description, and others. The existence of such hybrid texts suggests that communicative purpose is the situational parameter that can account for additional linguistic variation within them.

Situational complexity within registers is further explored as situation is conceptualized in continuous rather than categorical terms (Biber, Egbert & Keller, 2020). Building on the insights gained from the work on online registers, this study does not attempt to assign texts into register categories on the basis of certain situational parameters, but codes texts from online registers for the extent to which they can be defined through a number of situational parameters. This continuous approach (i.e., measuring the extent to which a particular situational parameter is present in a text) enables the study to identify dimensions of situational variation in online registers. The study shows that not all registers are situationally homogenous. For example, some registers may be more situationally uniform (song lyrics, encyclopedia articles), while others, such as opinion blogs or interactive discussion forums, contain texts quite diverse situationally. The study thus demonstrates that while registers are culturally recognized categories, they may vary with regard to their situational characteristics, and measuring these situational characteristics continuously captures the reality of this situational variation more accurately than defining the situational characteristics for the entire register. The study proceeds to identify situational text types – groups of texts unified in their situational characteristics regardless of their register category – and shows that some especially varied registers (such as opinion blogs) can include multiple text types. While this study did not carry out a linguistic analysis on the basis of the identified situational differences, a follow-up investigation shows that dimensions of situational variation in online registers correlate with dimensions of linguistic variation (Egbert, Biber & Keller, in preparation).

Situational distinctions within texts and the existence of textual units not delineated by existing textual boundaries or any other external cues were further explored in the register of conversation as distinct discourse units, “sequences of coherent discourse”, were identified

within conversations in the British National Corpus (BNC Spoken) (Biber, Egbert, Keller & Wizner, 2021, p. 22). Each discourse unit was characterized by the purposes it fulfilled and the extent to which the purpose was present. The study then identifies conversational text types – groups of texts united by the same purpose, again providing scope for further linguistic investigations of cohesive textual units within the register.

Goulart, Biber and Reppen (2022) develop a framework for analysis of purpose in undergraduate student writing and investigate variability by purpose in essays in a range of genres (authors' term) of the BAWE corpus, essays, critiques, case studies, methodology recounts, and explanations, coding texts for their major and minor communicative purposes. The main findings of that study include substantial variation by purpose among texts of each genre, the possibility of combinations of purposes in a single text, and the fact that different genres allow a different degree of this internal variation. Goulart (2022) again focuses on purpose in undergraduate student writing, using the developed framework, and investigates variation attributable to communicative text type, identified on the basis of purpose, and discipline.

Building on this work, whose results clearly demonstrated in a variety of ways that registers are internally varied, Egbert and Gracheva (2023) show that accounting for granular situational distinctions within registers consistently results in linguistic gains and explain this pattern through the notion of functional correspondence. The four case studies presented in that paper show variation across as well as within texts of a register. Variation among texts is illustrated in a case study of political memoirs, in which it is shown that an author's books can vary with regard to their focus on narration; in a case study on political speeches, which demonstrates that texts representing different speech types, such as rallies, commencement addresses, addresses to the nation, and remarks on a specific occasion, vary with respect to self-

or addressee focus; and in a case study on university textbooks, coded for presence or absence of communicative purpose, which then is shown to explain variation on Biber (1988) dimensions. Situational and linguistic variation within texts was observed when discourse types, such as dialogue and narration, were identified within 19th century fiction (originally studied by Egbert & Mahlberg, 2020). In all these cases, the specific situational parameters considered by the study yielded corresponding linguistic differences.

It is apparent that while different specific situational parameters have been investigated by studies of intra-register variation, such as publication type in journal articles or speech type in political speeches, the situational factor that has received the most attention is communicative purpose. This focus on communicative purpose is understandable as variation among texts of a register is examined closely. Communicative purpose has long stood out as the basis for intra-register textual distinctions (e.g., the IMRD-related linguistic differences associated with differences in “micro-purpose” discussed above – Biber & Finegan, 1994), and more recently purpose has been targeted directly as this section has shown. However, communicative purpose is not the only *textual* situational parameter in Table 2.1, and while the relations between participants, the setting, the mode, or the production circumstances are text-external situational characteristics, topic is another situational factor that pertains to the text and may therefore be worth studying as the source of linguistic variation among texts of a register. However, so far, topic has not been the subject of such investigations. The role of topic in variation among texts of a register is among the key outcomes of this study and is discussed further in Chapter 8.

Taken together, this body of research has led to a reconceptualization of the phenomenon of register and an emerging theory of register (Biber & Egbert, 2023). While variation among texts has been traditionally explained through the fact that texts belong to registers, categories

unified by their situational characteristics, register has not explained all linguistic variation among texts. Studies presented in this section have shown that this unexplained variance is due (at least in part) to variation among texts within registers or even to variation within texts of a register and is associated with highly specific situational distinctions.

2.2 LINGUISTIC VARIATION ACCORDING TO CHARACTERISTICS OF THE USERS

The previous section discussed the situation of use at different levels of analysis – between and within registers – as the source of linguistic variation. The focus of this section is on the linguistic fields explaining linguistic variation through characteristics of the users. Section 2.2.1 discusses social variation, i.e., linguistic variation accounted for by the group affiliation on the basis of the demographic characteristics of the language users. Specifically, the focus of the section is on the two social characteristics investigated by this study – gender and age of the speakers. Section 2.2.2 then focuses on the individual user and overviews approaches to speaker individual characteristics as a predictor of linguistic variation.

2.2.1 Social Variation

One core feature of sociolinguistic research, whose focus is often on ethnographic or dialectal variation (e.g., Alim, 2004; Auer, 2007, Rickford & Blake, 1990), is its strong interest in the group rather than the individual. This emphasis is apparent as the speech community is defined as the “fundamental unit of analysis” and a “sociolinguistic entity” (Gumperz, 1968, cited in Romaine, 1982); the object of study is defined as a way of speaking observed within a certain social group (Coupland, 2007, pp. 2-3), speech adhering to the norms of a certain community of practice, showing conformity to the norms of the population from which features are sampled, or

linguistic usage that “consistently distinguishes different classes of text or different populations” (Sandell, 1977, p.9).

Linguistic features explaining such variation are thus viewed as indexical markers of group membership rather than interpreted in functional terms (as noted, e.g., by Finegan & Biber, 1994). It is not common for studies concerned with social dialect variation to provide an explanation for the observed linguistic patterns that would extend beyond language serving indexical purposes – that is, marking the speakers’ group affiliation. Some scholars note that to an extent, this may be due to the fact that acquisition of norms by a member of a speech community is seen as natural – members are not taught the norms of speech, and there is presumably little awareness about these norms among community members. Their use is therefore perceived as unconscious rather than as intentional choices within the speakers’ control (Wolfson, 1989, p. 37). It is then to be expected that language use investigated from this perspective is not seen to have a functional basis.

A social group or a community in this view is formed on the basis of such uniting characteristics as the place of residence, ethnicity, age, gender, occupational or educational background or practices, or rather a complex interrelationship between these factors. Each of these factors often does not present as much interest in itself as it does as a contributor to larger questions of social importance. Regional differences, for example, are examined as sources of language change, with both time and space seen as the reasons for dialect development as migrations occur and time passes. The American and British English varieties, for example, or even varieties within each are separated by distance and centuries of political independence (Wardhaugh, 1993, p. 133). Regional differences are thus of interest not only to dialect geography but also to historical linguistics. Alternatively, dialect distinctions are relevant to

studies of rural and urban communities, their differences in lifestyle, and their respective effects on language use, each associated not as much with location as with certain outlooks, reflected in more ‘conservative’ or innovative use of language.

Similarly, ethnicity in and of itself is not the focus of sociolinguistics as it is not assumed that speakers’ ethnic origins directly affect language. Rather, it is the social practices and experiences associated with ethnic affiliation that are thought to predict linguistic patterns. An investigation of ethnic groups, for instance, may involve related questions of education and considerations of the ‘cultural force’ that inform conclusions regarding the influence of ‘family’ versus ‘peer group’ in the development of the linguistic system (Laferriere, 1979).

Socioeconomic class is perhaps central to sociolinguistic research, with the other variables often associated with class distinctions. Social stratification is studied as a reflection of “social differentiation and social evaluation” (Barber, 1957, cited in Labov, 1986), a phenomenon seen as natural and pervasive in society, as “the normal workings of society [produce] systematic differences between certain institutions or people, and [...] these differentiated forms have been ranked in status or prestige by general agreement” (Labov, 1986, p. 305). Occupational groups, in turn, are closely associated with socioeconomic class and, according to Labov, are one of the most important indexes of social stratification.

Important user characteristics are age and gender of the speaker, although the two seem to have received different degrees of focus in sociolinguistic research. The next sub-sections focus in more detail on each factor and overview the key findings from research on these predictors.

Gender as a Marker of Social Difference

According to Trudgill (1974b, cited in Trudgill, 1986, p. 396), while regional, ethnic group, and social-class varieties reflect social *distance*, linguistic patterns predicted by gender groups are the result of social *difference*. Research on gender has occupied a central place in sociolinguistics, examining gender in binary and nonbinary sense, although a considerably larger body of work exists on the distinct socialization practices between men and women (Wardhaugh, 1993). The focus of sociolinguistics is on gender as a socially constructed and performed phenomenon rather than sex as biological differences. Language used by gender groups reflects preferential differences (Friginal & Hardy, 2014, p. 109), i.e., certain features favored by representatives of either group, rather than an absolute difference between the groups (i.e., features only used by one of them and never used by the other).

This body of work has resulted in some persistent generalizations. Representative early studies include Lakoff's (1975) observation that women use features of politeness more than men as well as tag questions and question forms with affirmative functions, interpreted as markers of hesitancy and uncertainty, discourse markers as exclamations (e.g., *goodness*), hedges and modal verbs, and 'empty' evaluative adjectives (e.g., *adorable*, *charming*). Labov (1990, 2001) noticed women's greater likelihood of choosing a standard over a nonstandard form if a choice was possible. At the same time, when nonstandard forms were used by women, women were found to use novel features and thus contribute to change. Such trends are then commonly interpreted through the lens of social power – for example, increased politeness as well as the tendency towards the standard or 'correct' usage may be associated with a lack of assertive stance, while a tendency to incorporate linguistic innovation associated with prestige may be a tactic to gain power. Biber and Burges (2000, 2001, cited in Friginal & Hardy, 2014, p. 110) summarize some other generalizations made by sociolinguistic gender studies. For example,

conversations in mixed gender settings are said to feature less women talk (Crawford, 1995; James & Drakich, 1993); women are thought to focus on interpersonal and interactional matters, and men tend to favor information presentation (Holmes, 1995; Lakoff, 1990); studies observe women's increased tentativeness (Coates, 1996, 2004; Rubin & Greene, 1992) and women's predilection for narration and men's tendency towards persuasion (Fleishman, 1998; Rubin & Greene, 1992). A book-length treatment of gender by Coates (2004) elaborates on these and other established distinctions in gender language, such as the idea that women gravitate toward parataxis as a more involved and emotional mode of expression with less complex organization of ideas, while men use hypotaxis to a greater extent (Jespersen, 1922, cited in Coates, 2004, p. 18). Other differences include women's verbosity and a tendency towards euphemisms, and women's tendency towards a cooperative rather than competitive style of talk in turn-taking, use of questions, minimal responses, and other conversational strategies.

Further insight into gender differences is offered by corpus-based studies, which highlight the use of politeness features: Ishikawa (2011) observes their prevalence in male speech in BNC dialogues contrary to previous findings by Lakoff (1975) and Holmes (1995), while Friginal (2009b), conversely, confirms Lakoff's (1975) and Holmes's (1995) results (i.e., more prominent use of politeness features by women) in a study of American speakers' phone interactions. Stance markers – affect, evidentials, and quantifiers – are examined by Precht (2008) in the Longman Corpus of Spoken and Written English, and the findings confirm the trend of men using more expletives, insults, explicit language, and an overall higher numbers of negative stance expressions than women.

It is important to note, however, that these studies focus on individual linguistic markers they then find to distinguish between men's and women's speech. This approach has been

criticized by some researchers for the inherent premise that a given linguistic feature will always have the same function across contexts. Aries (1996), for example, writes that since individual features are multifunctional, there should not be a tacit assumption that the features identified by Lakoff (1975) or other studies as “women’s language” will always distinguish between genders. *I think*, for instance, can serve an assertive function as well as tentative. Likewise, a tag question may not express uncertainty, but conversely, may serve to encourage the addressee. To illustrate, Aries cites studies (e.g., Dubois & Crouch, 1975 and Hartman, 1976) that have reached contradictory conclusions with regard to gender differences in the use of tag questions, among other features, and attributes these discrepancies to differences in contexts (professional meeting following a presentation and interviews in the participants’ homes) as well as the status of the participants. Aries claims that studying gender without attention to context as well as participants’ social roles and power differences but simply assuming that certain features will always be distinct in all cases may lead to meaningless conclusions.

An alternative approach to gender has been to examine a combination of linguistic features in aggregate thought of as indices of “women’s language” (Aries, 1996, p. 131) and compare men’s and women’s speech on such indices. Aries notes that this approach is also not without its faults as functionally unrelated features are combined into a single index simply because the features are expected to reflect gender differences, but “a high frequency of occurrence of one speech feature in the index [is] not correlated with that of another” (O’Barr & Atkins, 1980, cited in Aries, 1996, p. 131) and the functional meaning of that index is not clear. Aries gives an example of an index of tentativeness that combines qualifiers, hedges, and tag questions, in which case while women were found to use more of these features, each of them may have served functions that did not contribute to tentativeness.

Using such indices, Preisler (1986) and Crosby and Nyquist (1977) (cited in Aries, 1996, p. 131) examine the effect of social role and status alongside gender, and their results indicate that genders do not systematically differ with regard to such holistic measures, but the differences depend on the participants' roles and the situational context. Specifically, gender differences were more likely to occur in informal contexts and were not found in formal, more informational ones. Johnson (1994), Soskin and John (1963), Moore, Shaffer, Goodsell and Baringoldz (1983), and Sagrestano (1992) (cited in Aries, 1996, p. 136) likewise find that authority is a better predictor of language differences and observe no gender differences in the features commonly seen as distinctive (e.g., tag questions, personal pronouns, disclaimers, qualifiers, etc.) when the status is equal. Such findings have led researchers to suggest that differences in "women's language" actually reflect differences in social roles rather than gender (O'Barr & Atkins, 1980, cited in Aries, 1996) and that factors such as the setting and topic of communication mediate gender differences.

Mulac, Incontro and James (1985) use discriminant function analysis studying recordings produced in a communications studies class to investigate whether the frequency of use of a combination of features can distinguish between the speech of men and women. As a result, Mulac et al. identified several features whose weighted combinations predicted gender with 100% accuracy. Interestingly, however, the 'male' features that emerged from this analysis did not align with previous research (including, e.g., personal pronouns, typically seen as a 'female' feature). The 'female' features were more in line with other studies and included speech fillers, intensifying adverbs, and references to emotions. Mulac and Lundell (1986) then replicate Mulac et al.'s study, this time asking participants to describe landscape photographs, and identify a different set of distinctive features. While these features predicted gender with 87.5% accuracy,

again the emerging ‘male’ features had not been previously associated with men’s discourse, and some of the features associated with women’s speech in the previous study (e.g., fillers) were in this study associated with men’s speech. Women’s speech in Mulac and Lundell’s study was marked by three features consistent with previous research – tag questions, intensifying adverbs, and personal pronouns. Mulac, Lundell and Bradac (1986) replicate the study again in conversation to test the approach in an interactive setting. Their results in this case were the most closely aligned with the findings of other research: the features characteristic for men’s discourse included interruptions, directives, and sentence-initial conjuncts/ fillers, while the features found in women’s speech were personal pronouns, sentence-initial adverbials, questions, intensifying adverbs, and justifications for previous statements (cited in Aries, 1996, p. 134). The results of all their studies lead the authors to the conclusion that the features that distinguish between men’s and women’s language depend on the situational context, that different contexts allow different degrees of overlap in their speech, and that very few features (3 in total: intensifying adverbs and personal pronouns for women and directives for men) consistently distinguish between genders across contexts, while all other features (out of the 35 initially tested) either do not predict gender, predict gender only in some contexts, or predict different genders across contexts. In all their studies, human subjects could never differentiate between male and female language, but differences were detected only through the automatically identified combination of features.

Some researchers have identified men’s and women’s ‘speech styles’ and labeled them as restrictive and enabling (Maccoby, 1990) or agency and communion (Bakan, 1966), respectively. These labels appear in line with the sets of features commonly associated with men’s and women’s discourse – discourse that manifests agency and in a way restricts communication uses

features such as directives and no facilitating strategies, while the kind of discourse that enables cooperation and communion uses backchannels and minimal responses, facilitative tags, and politeness markers. However, these researchers too acknowledge that both men and women use both types of discourse depending on the situation and the communicative goals (cited in Aries, 1996, p. 143).

This overview shows several key trends. First, research on gender in sociolinguistics tends to be restricted to the register of conversation or speech-based registers. Friginal and Hardy, for example, state that investigations of written registers for gender differences are relatively rare due to the assumption that the editing process and adherence to style requirements may even out the existing gender differences in the language and the fact that the unspecified audience of many written texts could make the effect of gender less prominent. As reported by several studies discussed above, it is this limitation that appears to have resulted in some persistent gender differences. In other contexts, however, only a few features proved distinctive, and in some situations of use (public setting, informational contexts) gender differences were not observed. Another major finding that emerges from this body of work then is that gender differences are mediated by the situation of use, and there seem to be no invariant ways of expression characteristic for men or women (Aries, 1996, p. 144). Finally, researchers have identified methodological problems with targeting specific, functionally disparate linguistic features, and thus setting up studies in such a way as to target and eventually find gender differences. Instead, they propose that gender be accounted for in discourse, but the focus be shifted to the study of language produced in different roles and interactional contexts.

Age as a Marker of a Life Stage and Skill Development

Sociolinguists recognize the importance of age as well as its complexity. While the years after birth may be the most straightforward way of measuring age, the construct of age and its sociolinguistic significance do not necessarily correspond to biological age (Friginal & Hardy, 2014, p. 107). Eckert (2017) writes about age and the process of aging as fundamental to human experience; age represents human achievement in skills development as a result of the individual's participation in social activities and a construction of a personal identity and history. Age is therefore a marker shared by a group united by a life stage and a set of characteristic experiences. Llamas (2006) also stresses the social rather than biological meaning of age, pointing out that age serves as the critical threshold that determines the allowed or prohibited activities in a society and their consequences to a greater extent than any other social marker. Age further determines social behavior, presentations of a particular social identity (e.g., through clothes, language, and other choices that conform to societal norms associated with a life stage), obligations, and responsibilities. Age is thus a performed construct, and expectations of a particular performance corresponding to a certain age determine social perceptions and attitudes.

Young adults and adolescents have received the most attention in this line of work, with studies primarily focused on early developmental stages, such as acquisition of communicative competence in childhood and adolescence (e.g., Stenström & Jørgensen, 2009). Friginal and Hardy (2014) explain this exclusive focus on adolescents through the interest in language innovation and change, believed to be carried forward by that demographic. Slang, sound change, expletives, quotative markers, expressive pronunciation, syntactic deviations, new coinages, intensifiers, and discourse markers are among the features studied extensively in the language of adolescents (Friginal & Hardy, 2014, p. 120). Eckert (2004) explains this strong

interest saying that adolescents' status in the social arrangement is unique due to their isolated position – they are a group expected to mainly interact with peers, and in that their position is akin to that of the elderly. Thus, although it has been shown that changes in a person's language in adult life, associated with major life events, do happen and development is not limited to early life (Cheshire, 1987; Seifert, Hoffnung & Hoffnung, 2000, cited in Murphy, 2010), young language users are of primary interest to sociolinguistics.

With regard to the linguistic features under analysis, Cheshire (2005) notes that the focus of sociolinguistic research on age differences has been morphophonological rather than syntactic, lexical, and pragmatic variation. Lexico-grammatical studies include investigations in social psychology, such as Pennebaker and Stone (2003) studying linguistic as well as psychological processes, focusing on the use of past tense, parts of speech, word count, speech fillers, and assent, among other features. Some of the findings included a change from more negative to more positive emotion words with age, fewer references to other entities or themselves, more expressions of modal meaning, fewer past references, and more features of cognitive complexity, such as words of causation or inner processes.

A corpus-based comparative study was conducted by Barbieri (2008), who identified keywords in conversations of two age groups: 15-25-year-olds and 35-60-year-olds and found differences in intensifiers, personal pronouns, modal verbs, quotative verbs, attitudinal adjectives, stance adverbs, inserts, discourse markers, and slang words, noting specifically the marked prevalence of overt stance in the speech of the younger group (cited in Friginal & Hardy, 2014, pp. 121-122). Murphy (2010) investigates hedges, vague language, boosters, and amplifiers across age groups (20-year-olds, 40-year-olds, and 70/80-year-olds) and observes an increased use of hedges in the youngest group compared to the other groups and the lowest use

of the feature in the oldest group. However, the author proceeds to attribute the observed differences to the relationship between the participants, the degree of familiarity, different in all groups, and the nature of the topics discussed (e.g., personal or sensitive). The author thus acknowledges the role of context in the use of the feature; it is not entirely clear, however, whether the language differences are due to the social or the situational variable, and the interpretation suggests that the role of the situational factors is seen by the author as primary (e.g., the effect is attributed to topic as the topics typically discussed by 20- and 70-year-olds differ due to differences in upbringing). Similarly, the use of vague category markers, which prevails in the oldest group, is attributed to differences in the amount of shared knowledge and the strength of the bonds between the participants. The difference in amplifiers and boosters is again explained through “the type of talk” (p. 133) the groups engage in. This study illustrates that while age is the main target of this investigation and the basis for participant grouping, the interpretation of age differences is tied to considerations of context, suggesting that variation according to user characteristics is enhanced in important ways by considerations of use.

Although indirectly, age has been investigated by studies of grammatical complexity in the academic context from a register-functional perspective. Academic writing, in particular, has been associated with phrasal complexity features (Biber, 2006; Biber & Gray, 2016), such as nominalizations, prepositional phrases, and nominal sequences. Building on this work, Staples, Egbert, Biber and Gray (2016) investigate whether academic writers develop phrasal complexity over the course of their university careers. The groups of the study include students in the first, second, and final years of their undergraduate programs and graduate students, and complexity features are examined in the papers written in the disciplines of Arts and Humanities, Life Sciences, Physical Sciences, and Social Sciences in four registers – essays, critiques, case

studies, and explanations. While the increase in the use of different phrasal features (pre-modifying nouns, attributive adjectives, nominalizations, and of-genitives) happens at different rates across levels, the overall trend observed by Staples et al. is that of university students progressing from clausal structures (i.e., finite and nonfinite clauses) at lower levels of study to phrasal complexity (embedded phrases) at later stages. A particularly distinct difference between entry-level and advanced academic writing lay in the marked increase in the use of pre-modifying nouns – a feature more strongly associated with academic writing than with any other informational variety (e.g., Biber & Gray, 2016, cited in Staples et al., 2016, p. 163). The use of clausal structures, on the other hand, decreases from lower to higher levels. The study thus shows a clear trajectory of complexity development across levels.

2.2.2 Individual Variation

The previous section discussed linguistic variation on the basis of speakers' group membership. It was emphasized that sociolinguistic research views group affiliation as the ground for linguistic differences among members of social communities, and language users are grouped on the basis of region, ethnicity, age, gender, social class, occupational or educational backgrounds. This section is devoted to a different approach to linguistic variation according to characteristics of the users – namely, the idiolectal perspective, in which variation is attributed to individual differences. It is important to distinguish the idiolectal approach from the theoretical frameworks in which the unique linguistic characteristics of each speaker are viewed in isolation from the social setting, unaffected by the linguistic environment of the social group or the situation of use (e.g., Chomskian view of language as the product of the 'language faculty'). Linguists working in the idiolectal tradition acknowledge the indisputable role of society and context, where the individual is perceived as an agent in convention-building and as the driving force of linguistic

innovation and change occurring through small-scale contacts. Hocket (1987), for instance, states that language is a “social system to the extent that individual systems overlap” (cited in Johnston, 1996, p. 19), and LePage and Tabouret-Keller (1985) perceive individual choices as a repertoire of language available to the individual within certain social conventions, to which some speakers may choose to adhere more consistently, while others may choose to be more idiosyncratic. That is, individuals select the linguistic resources available in the system to express their individual selves (Johnstone, 1996, p. 93).

This research tradition hinges on the view that “no two people have the same knowledge of language or the same way of speaking” (Johnstone, 1996, p. 3). A similar view is expressed by discourse analysts, whose methods of close text examinations, according to Johnstone, result in major contributions to understanding the individual. Tannen (1989, p. 80) writes that “cultural patterns do not prescribe the form that a speaker’s discourse will take, but provide a range from which individuals choose strategies that they habitually use in expressing their [individuality].” While Johnstone admits that as we focus on the linguistic individual, we inevitably think about ways such factors as gender, ethnicity, region, vocational identification, community, and social roles “constrain and facilitate speakers’ choices” (p. 178), she finds the view that speakers are “linguistic creatures of [their] social environments” rather than “agents of [their] speech” (p. 179) limiting and incomplete and calls for investigations of the unique self in discourse. Johnstone maintains that self-expression is necessary regardless of how prominent conformity to social norms may be. This results in idiosyncratic expression, which is then understood by other speakers through cognitive processes of contextualization (Gumperz, 1982, cited in Johnstone, 1996, p. 187) or interpretations based on the relevance maxim (Sperber & Wilson, 1986, cited in Johnstone, 1996, p. 187) rather than on fixed rules.

Johnstone studies personal narratives as the most straightforward expression of individuality, which reveal different ancestries and personal histories and their manifestation in language. Specifically, Johnstone observes differences in discourse organization and syntax, which she ascribes to different personal backgrounds and circumstances of the participants. Johnstone's goal in such investigations is not to generalize beyond the speakers under analysis; rather, her interest lies in the very study of these people's language use, which in her view, is much more informative than the focus on group trends and simultaneously reveals normative and creative language use. That is, while a linguistic individual may be in some ways representative of "larger sets of voices" (pp. 24-25), close text analysis reveals how individuals diverge from group trends.

Johnstone claims that people use language creatively in order to express individuality not only in registers like personal narratives, but often in situations when creative use is not expected, such as phone opinion surveys with given response options. Johnstone notes that in such communicative events neither interviewers nor interviewees exhibit socially conditioned patterns of linguistic behavior, although language is often scripted, but they systematically deviate from these scripts, revealing their own unique motivations. A similar claim is made in an analysis of articulateness and clarity of two of the eleven speakers (graduate students and professors) discussing a topic at a scholarly conference. Johnstone maintains that while diverging too far from the norm of the academic context is not common as it may be misconstrued as 'academic ego' (p. 62) and be subject to criticism, individual differences are still apparent. To illustrate, she discusses the use of various types of clauses among other features by the two speakers and ways they contribute to personal and interactional fluency, clarity, and effectiveness. Johnstone notes that while narratives are meant to naturally evoke individual

expression and the results pointing to individual differences are less surprising, academic discussions and scripted phone opinion surveys are settings where scope for individual variation is typically not expected and yet consistently illustrated.

In line with this perspective, Dąbrowska (2012) argues that contrary to the common view of speakers of a language sharing a grammar of that language, adult monolingual speakers in fact have different mental representations of grammar. Dąbrowska illustrates this on different linguistic levels (morphology and syntax) and with regard to several features – passive voice, quantifiers, and different types of subordination. She presents evidence that not all adult monolingual speakers have an equal grasp of these structures and attributes these differences to a number of factors. One possibility is that speakers pay attention to different cues in the input they receive. Alternatively, in the case of certain complex constructions, as language users encounter them, some speakers may make only ‘local’ generalizations, i.e., extract linguistic material as it applies to only some linguistic items. This ‘local’ processing of linguistic information thus does not result in higher-level generalizations to a construction; other speakers, in contrast, may deduce abstract rules from concrete instances of use, which then leads to higher-level schematization. These individual differences in processing then lead to different mental grammars formed on the basis of these experiences with language. Dąbrowska admits that at least some of these individual differences are due to the effect of education, with more educated speakers acquiring more general schemas, which may be the consequence of a more frequent exposure to a more diverse linguistic repertoire. Nevertheless, Dąbrowska’s major claim is that individual speakers of the same language have different linguistic knowledge and do not eventually converge on the same grammar in the process of language acquisition.

Likewise, Schmid, Würschinger, Fischer and Küchenhoff (2021) claim that individual variation should not be treated as residual variance, but if it is brought into focus, can inform our understanding of grammatical, lexical, and pragmatic variation as well as of language change. Their study investigates the ‘THAT’S + ADJ’ construction (e.g., *that’s great*) in BNC2014 and emphasizes individual variation in the use of the construction overall, the semantic classes of the adjectives in it, and the individual adjectives speakers tend to favor. Their results show that while some patterns of use of the construction can be attributed to social and situational factors, they find low effect of the variables of age, gender, and social class, but substantial individual differences. Not only do individuals use the pattern to different extents, but different speakers’ results show a prevalence of epistemic, evaluative, emotive, or ‘uptake’ (i.e., of what was said by the previous speaker) adjectives in the construction, and further variation is seen among the speakers when the most frequent adjectives in the variable slot are inspected (e.g., the most frequent adjective ‘right’ is used extremely frequently by some speakers, but others do not contribute to this pattern). The study then identifies speakers with ‘extreme habits’ (i.e. a marked preference for a particular adjective) with other competing adjectives used as well and speakers with ‘extreme habits’ without any other adjectives in the variable slot. These trends are then discussed in terms of frequency-driven entrenchment, and the authors reach the conclusion that speakers differ in the extent to which a particular pattern is mentally represented as a holistic chunk or the extent of how strongly a particular higher-level (THAT’S + ADJ) or lower-level (THAT’S + epistemic/evaluative/emotive/etc. adjective) is represented in their mental lexicon. That is, speakers who clearly prefer the use of a particular adjective in the structure, which does not have competitors, are unlikely to assemble the structure each time from its components, while speakers who allow other possibilities in the slot are more likely to process the structure

compositionally (p.13). Similarly, individual differences in representation of multiword sequences and individual judgments of such sequences are studied by Verhagen and Mos (2016), who also conclude that individual variation is likely a reflection of different linguistic and metalinguistic cognitive representations. In line with these conclusions, Sabino (2018, p. 59) maintains that speaker idiolects, which may change over time, expand in different ways depending on “what has been previously entrenched.”

Individual language use has been of interest to applied research, such as comparative studies of literary language, comparing authors of a particular time period. Egbert (2012) carries out a multidimensional analysis of 19th century authors’ fiction and distinguishes between the language use of ten 19th century fiction writers along three dimensions of variation. The study observes considerable between-author variation as well as several unique cases of wide within-author variation with regard to their reliance on the functional patterns of ‘Thought Presentation vs. Description’; high-scoring authors on the dimension of ‘Abstract Exposition vs. Concrete Action’ were characterized by an intrusive explicit expression, which allowed less freedom of interpretation to the reader, while authors on the opposite end of the dimension showed a tendency towards concreteness; finally, the dimension of ‘Dialogue vs. Narration’ identified authors with interactive tendencies, relying on dialogue, and authors favoring informational prose, prioritizing narration. Gracheva (2022) conducts the same type of comparative analysis among modern nonfiction authors of literary essays and analyzes authors’ language along dimensions of ‘Interactive vs. Informational’, ‘Abstract Expository vs. Concrete Descriptive’, ‘Immediate vs. Removed’, and ‘Hypothetical’ expression.

Language of political leaders has also been a prominent avenue for research of individual variation. Notable examples include Hart’s (1984) book-length treatment of American

presidential discourse, and more recently, the language of presidential and vice-presidential candidates in the latest elections has become the focus of several investigations. Slatcher, Chung, Pennebaker and Stone (2006) investigate the relationship between personality characteristics and psychological states and word use as they compare the language of John Kerry, John Edwards, George W. Bush, and Dick Cheney in their 2004 interviews, press conferences, and campaign debates. As a result, they identify individual differences with regard to linguistic features they associate with aging, cognitive complexity, depression, femininity, honesty, and presidentiality and characterize each individual with respect to their linguistic performance. In a similar vein, individual language use has been the subject of interdisciplinary studies at the intersection of linguistics and psychology, whose goal is to explore the links between language and personality types or particular personality traits and the contribution understanding individual language can make into personality research (e.g., Pennebaker & King, 1999; Oberlander & Gill, 2004).

While this overview is far from complete, its goal here is to illustrate the main agenda undertaken by the research tradition whose focus is on the linguistic individual. The studies presented here have demonstrated that linguistic evidence of individual differences has often contributed to our understanding of larger linguistic or cognitive phenomena. The linguistic features under analysis are attributed to the author's individual linguistic choices and are seen as a representation of one's unique self rather than of group membership. Importantly, however, these linguistic choices in the idiolectal perspective are not the result of a 'language faculty' at work as in the generative tradition, but are the product of the individual's interaction with the environment – language exposure and use.

2.3 CONSIDERATIONS OF USE AND USER CHARACTERISTICS – PREVIOUS RECOGNITION OF THE MUTUAL RELATIONSHIP

Sections 2.1-2.2 of this chapter aimed to overview the research agenda of the sub-fields of linguistics examining variation according to the situation of use and according to characteristics of the users. Variation according to use included register variation, i.e., variation due to differences between culturally recognized register categories, and granular communicative variation within registers. Variation according to characteristics of the users encompassed demographic variation, in which case speaker group affiliation is treated as the basis for linguistic differences, and individual variation, i.e., linguistic differences reflecting individual use. Sections 2.1-2.2 thus demonstrated the different research foci of these distinct approaches to linguistic variation. Yet, several attempts to explore the relationship between these approaches have been made. The following sub-sections discuss the most seminal work uniting considerations of user characteristics and use. Specifically, Sections 2.3.1 and 2.3.2 illustrate the relationship between social dialectal and situational variation, but represent steps taken towards unifying the two perspectives from different directions: namely, the sections distinguish between the approach adopted in sociolinguistics, where attention to situational variation has mostly meant “attention to speech” (Section 2.3.1), and the approach undertaken by register researchers (Section 2.3.2). Section 2.3.3 discusses the role of the individual in sociolinguistic research, and Section 2.3.4 then presents studies of individual variation and register.

2.3.1 Social Dialect and Attention to Speech

Situational considerations have featured in sociolinguistic research, with some scholars acknowledging the critical role of “the speech situation” (Wolfson, 1989, p. 189). It is accepted that speakers do not use language the same way across contexts, and while speaker

characteristics like education level and social class will have an effect on their speech, they may provide only a partial view of linguistic variation; the ability to vary speech in response to characteristics of the listener, social distance between the speaker and the addressee, or the setting of the interaction is, on the other hand, a major factor in communicative competence (Hymes, 1974; Wolfson, 1989). “Speech situation”, however, does not acquire prominence in sociolinguistics and is largely synonymous with degree of formality or attention paid to speech, dating back to Labov’s (1966) seminal work on variation in a New York department store, in which attention paid to speech was measured in his sociolinguistic interviews on a single dimension from casual to careful speech, the latter represented through a range of reading tasks (a reading passage, a word list, and a minimal pair list).

Since then, several limitations of Labov’s approach have been pointed out, some of the major ones being the linear, unidimensional approach to situation, where formality or the different degrees of attention to speech is the only situational parameter (Macaulay, 1991). Additionally, Labov’s framework has been criticized for its limited and rather contrived nature and the fact that speakers varying their speech in response to different tasks or topics during a sociolinguistic interview (e.g., answering the interviewer’s questions, reading a text, reading wordlists) does not represent real communication (Rickford & McNair-Knox, 1994).

These criticisms led to new approaches to investigating the ‘speech situation’, such as the convergent accommodation theory (Giles & Powesland, 1975) and later audience design (Bell, 1984), which postulate that speakers accommodate their speech to that of their interlocutors. Alternatively, speakers may demonstrate ‘initiative’ language use, in which case they set the tone for communication rather than accommodating their language to the audience.

Bell draws on Giles's (1984) accommodation theory and his own research on the language of the same radio news broadcasters who addressed different audiences. While the theory of audience design (Bell, 1984) may appear to have taken a step towards addressing some of the limitations of previous work by introducing a new situational parameter, the role of situation in this approach is in fact subordinate to considerations of status and social distance between the speakers. In Bell's view, for example, variation that occurs due to audience distinctions derives from differences in social standing that may exist between the interlocutors, in essence echoing variation on the basis of social class (Bell's study on radio news broadcasts features audiences of different socioeconomic classes). It is thus these differences in social positions and power that lead to convergent accommodation and guide the speakers in their choice of language (Bell, 1984) rather than specific audience considerations related to a particular communicative event, such as the nature and amount of shared knowledge between the speakers. While for Preston (1991) linguistic variation is first attributed to internal linguistic constraints, like Bell, he views situational variation as embedded within considerations of social status and speakers' tendency to accommodate their speech to that of their audience following an assessment of the interlocutor's social dialect. Finegan and Biber (2001) note, however, that neither Bell, nor Preston consider it necessary to explain patterns of social variation.

Investigations of speaker familiarity with the addressee relying on Giles's (1984) convergent accommodation theory and Bell's (1984) audience design gained traction (e.g., Blom & Gumperz, 1972; Coupland, 1980; Trudgill, 1986), but while Rickford and McNair-Knox (1994) illustrate the benefits of this approach by two interviews, in which the relationships between the subject and the two interviewers were different, they admit that other situational considerations that may have played a role include the setting, topics, and the tempo of the

interview. The study also does not control for other interviewer characteristics such as race, which makes it not immediately clear what exactly caused speech accommodation.

Overall, despite the acknowledgement of the importance of the speech situation, which “intersects with social class” (Wolfson, 1989, p. 189), sociolinguistic studies, having come from the tradition of dialectology with its emphasis on region, age, and education as the main predictors, have not paid due attention to situational context. While efforts have been made to unify the two perspectives, these efforts do not seem to have resulted in a reevaluation of the role of situation, and a “pattern of underplaying communicative goals” and “the functional complexity of language in use” still persists in sociolinguistics (Coupland, 2001, p. 188).

2.3.2 Social Dialect and Register

A notable attempt to bring register considerations to the forefront of sociolinguistic research and treat functionally-determined linguistic variation as the basis for social dialect variation is a study by Finegan and Biber (1994), who suggested that linguistic choices are determined by the social environment in combination with the situational characteristics of the text. Sociolinguists, such as Labov (1969), Romaine (1980), or Kroch (1978) (cited in Finegan & Biber, 1994) also comment on the parallel trends in the use of a feature by social groups and its use in certain situational contexts, but they mostly do not discuss the reasons for this overlap. In contrast, Finegan and Biber emphasize the close systematic relationship between the socioeconomic and situational perspectives and the fact that socioeconomic group differences in fact reflect their members’ differential access to certain situational contexts. It is therefore consistent communicative constraints that determine language used by these social groups. To illustrate, the authors show that language characteristic for less formal communicative situations exhibits a tendency for economy and is found in groups of lower social standing, while linguistic choices

that mark more formal interactions tend to include explicit, elaborate structures and are common for groups higher in the social hierarchy. Finegan and Biber conclude that social group variation appears secondary to situational variation, which is functional in nature, while social variation has been perceived to be “arbitrary or conventional” (Finegan & Biber, 1994, p. 317). The authors reference the works of several sociolinguists (Gumperz, 1964; Lindenfeld, 1969; Hymes, 1974; Heath, 1986; Kay, 1977; Van den Broeck, 1977) who have acknowledged the fact that groups of higher social standing acquire a broader repertoire of communicative contexts in which they are commonly involved. As a result, these groups’ consistent exposure to a particular range of contexts will shape their language use. Based on this, Finegan and Biber propose that studies of sociolinguistic variation should address this parallel, systematic socially and situationally determined variation.

The notion of situational variation being directly related to socioeconomic status is not exclusive to Finegan and Biber’s study and can be traced back to Bernstein’s (1971) investigation of restricted and elaborated expression among children from upper and working class backgrounds as well Halliday’s (1978, p. 34) discussion of dialectal and diatypic (situational) variation. Halliday too makes the claim that dialectal and diatypic variation are related and follow a pattern determined by the differential access members of any society have to different registers. Similar to Finegan and Biber, Halliday associates ‘the standard’ with formal situations and a more colloquial variety with more informal ones. In the same vein, Bernstein discovers that children from upper social classes tend to have a more varied, versatile repertoire, whereas those from working class families a more restricted one.

Finegan and Biber (2001, p. 235) revisit the “classic sociolinguistic finding”, which postulates that features of formal contexts tend to be associated with higher social class, while

features common in informal contexts are characteristic for groups of lower social standing (e.g., Romaine, 1980). In the 2001 study, the model proposed by Finegan and Biber (1994), in which differential access of social groups to literate or oral registers was seen as the basis for dialect differences, gains empirical support from a case study of three social groups in the British National Corpus (BNC). The study finds that the same pattern in the use of economy vs. elaboration features observed earlier in situational and social variation bears out in the case of three features used by BNC speakers. However, the fact that not all register features also reflect dialect differences leads the authors to refute Bell's (1984) claim that situational variation derives from (and in that is subsidiary to) social variation. As a result, Finegan and Biber propose the 'register axiom', which posits that the greater use of variable linguistic features by a social group is determined by this group's greater exposure to and experience with the registers in which these linguistic features are frequent. Thus, the authors call for investigations of a broader range of registers in sociolinguistic studies than the sociolinguistic interview has traditionally allowed – an approach that has been unable to adequately represent a full range of speakers' experience with language. Finegan and Biber (2001, p. 236) see an integrated approach to language variation accounting for social and situational factors as essential, as “enough is now known about language variation to confirm that no single dimension can adequately explain it,” but note that no previous model has been successful at conceptualizing the interrelationship of these factors.

2.3.3 Social Dialect and Individual Variation

As was noted earlier, the linguistic individual is typically not of interest to sociolinguistics with its overwhelming focus on the demographic characteristics of a group. However, some sociolinguists recognize the oversimplification of linguistic reality that occurs as a result of

exclusive group focus. Coupland (2010), for example, reflects on Eckert's (2003) ideas regarding 'the authentic speaker', i.e., the speaker who is assumed to be a perfect representation of a group and, importantly, nearly identical in their characteristics to the other members of that group. Eckert states that such an ideal representative has been at the center of sociolinguistic research, most of which has operated on the tacit assumption of group homogeneity. Such 'authenticity', in Eckert's words, implies a static nature of a group, which makes it a convenient object of analysis. As a result, the normative assumptions formed on the basis of sociolinguistic findings, such as adults' language being relatively fixed, resistant to change or innovation, are questionable (Eckert, 2003, cited in Coupland, 2010, p. 99-100). Eckert refers to this problem as a 'structure vs. agency debate' as typical sociolinguistic studies aim to capture a reflection of social structure in language, disregarding the construction of personal meanings in discourse within that structure through speaker agency. The role of social structure, according to Eckert, is still obvious in that it provides individual speakers the resources for personal expression so that they "creatively re-enact or rework established indexical relationship in their talk", structure and agency thus depending on each other (Eckert, 2003, cited in Coupland, 2010, p. 100). Coupland (2010, pp. 101-102) adds that it is striking that the concept of a 'speech community' seems completely devoid of "aspects of subjectivity" and states that research practices should at least account for the "subjective experience of communal participation". Coupland illustrates the role of the individual in the construction of a community with a case study of a TV show host, whose real as well as media identity and personal narrative are shown to be instrumental for the community built around him as other show participants are involved.

Stuart-Smith and Timmins (2010) foreground the role of the individual in language change as they investigate a particular phonological variant in Glasgow's inner city districts. The

authors emphasize that the observed change cannot be generalized to a group of adolescent speakers who are commonly assumed to be agents of change. Rather, the observed variant appears to exhibit individual engagement with a TV show that may have influenced the speaker's pronunciation. Similarly, in a study of individual voice recognition, Watt (2010, p. 76) states that variation is not only accounted for by belonging to a social group, but by people's perceptions of themselves as "self-contained individuals distinct from all others" and notes that in view of a vast number of possibilities (phonological in the context of Watt's study, although the potential role of lexico-grammar is also acknowledged), it is to be expected that variation in speech will be explained on the idiolectal level.

2.3.4 Register and Individual Variation

Just as the effect of register has not been fully explored or sometimes even acknowledged in research on social variation (Preston, 2001, for example, critiques and rejects Finegan and Biber's approach to social dialect variation described above), its role in research on individual variation has also been marginal. On the one hand, the focus on the individual has led researchers to investigate individual expression within a single domain (e.g., political speeches). This research goal does not entail attention to situational factors, which in these single-register studies are presumably the same across texts. On the other hand, analysis of individual language use may cross register boundaries. Pinker (2014:13-26), for example, analyzes examples of "good writing" from works of different authors, united by the same topic rather than register. These works span texts from popular science, biography, obituaries, and historical nonfiction. As in the case of single-register studies, it is again impossible to know if the same author's language would remain unchanged if they were responding to the situational characteristics of popular science versus an obituary – texts vastly different situationally. Additionally, since the texts are

all examples of different registers, the linguistic features under analysis may be attributes of register distinctions rather than individual choices. Finally, it is also not acknowledged that the definition of “good writing” may vary substantially depending on what is appropriate and effective within particular contexts of use.

However, an alternative perspective and a need for a nuanced conceptualization of individual language use with attention to register has been recognized. Bakhtin (1986) treats the role of situation in individual variation as obvious and expected. In fact, Bakhtin does not treat individual variation as separate from functional use of language at all and states that language is never produced with the sole purpose of expressing the author’s personality, while it is always produced in response to some communicative need (pp. 63-65). The expression of individuality is thus inextricably linked to this functional use. While Bakhtin acknowledges the value of studies whose sole focus is individual expression, he stresses that the success of such efforts depends on the level of awareness of communicative distinctions and a “a preliminary study of the subcategories of speech genres” (Bakhtin’s term) (p. 64), which is not the typical basis for research of individual variation and hence, according to Bakhtin, its weakness. Importantly, Bakhtin observes that not all situational varieties are equally conducive to individual expression, with creative genres (Bakhtin’s term) allowing the most individual freedom. While he does not directly relate this observation to the situational characteristics of registers (genres), his statement implies that registers vary in the extent to which their situational characteristics are defined (Biber, 1994), and more narrow, situationally better-defined registers, which require “a standard form” (e.g., business documents) (Bakhtin, 1986, p. 63) may not give language users as much linguistic latitude as broad, less situationally fixed registers will allow.

Similarly, Carter (2007, p. 598) advocates a reevaluation of linguistic creativity as contextually determined use of language, emergent in communication, and proposes examining its relationship to the type of interaction and the social roles of the participants. In Carter (2004), creativity is analyzed along a cline representing differences in social contexts and illustrating degrees to which speakers can be prone to creative expression across the following contexts: transactional (public with no prior relationship between participants), professional (shared professional context and a professional relationship, although not necessarily of equal status), socializing (more relaxed context with a greater equality of roles assumed), and intimate (private linguistically 'off-guard' contexts). Additionally, interaction type was considered, and distinctions were made on a continuum from collaborative to non-collaborative activities. Collaborative interactions were further categorized into task-oriented ones (communication is geared towards completing a task) and those that were not (collaborative ideas, i.e., sharing thoughts rather than completing a task). Carter investigates different speakers in a range of registers in the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) along this cline for evidence of creative expression. He concludes that creativity (pattern-forming and pattern-reforming features: covert or subconscious repetitions, parallelism, echoes, and style-matching in the case of the former and overt uses of figures of speech, puns, metaphors, idioms, or decomposition of set phrases in the case of the latter) varies depending on the speakers' relationships and nature of the tasks. It is claimed that contexts that are more interactive, collaborative, and engaging (e.g., joking among friends, friends cooking together, reminiscing with friends), or intimate (e.g., siblings discussing their childhood, couple decorating a room or discussing a film they saw) are more conducive to creative expression, while one-way information presentation (e.g., transactional contexts like choosing and purchasing a product,

conversation at a service encounter or professional contexts such as reports at a meeting) allows less scope for linguistic creativity.

Even more recently, Cvrček et al. (2020) quantify the amount of variation explained by the variables of the individual and register and examine different social contexts and a range of communicative purposes in elicited letters from the same authors (a cover letter representing the corporate context with the goal to persuade, a letter from vacation as friendly communication with the goal to convince, a letter of complaint – official communication with the goal to complain, and a letter of apology – friendly communication with the goal to apologize), different power dynamics (submissive in the case of cover letters and letters of apology vs. dominant in the case of letters of complaint and letters from vacation), and degree of formality (formal – cover letters and letters of complaint – and informal – letters of apology and letters from vacation). The study finds that the distance in multidimensional space between texts of the same author is greater than that between texts on the same scenario and concludes that the role of register is major, cautioning researchers against including both variables in their design.

Other studies take a different approach, however, and rather than viewing register as a confounding variable that will obscure individual patterns, analyze individual expression and register with relation to one another. Sheffler, Kern and Seeman (2022) study individual variation in German blogs and tweets written by the same 44 authors. While the study's main focus is on the distinction between register (texts coded as informative, narrative, and persuasive) and medium, i.e., the platform used (webblog or tweet), the study observes an interplay between individual use and the media, thus illustrating the benefits of such analyses. Marko, Reitbauer and Pickl (2022) investigate the same authors on Instagram and Twitter, approaching the relationship between individual use and register from a forensic linguistic

perspective and aiming to identify features that are resistant to the effect of register. As a result, the authors distinguish between stable and variable features in the language of three authors based on accounts on both platforms. Gracheva (under review) is a study of presidential discourse across registers (Biden's, Trump's, Obama's, and Bush's language is investigated in political memoirs, political speeches, and official correspondence). That study shows that while a large number of linguistic features vary as a result of register, the effect of individual differences is much lower (as indicated by the number of distinctive linguistic features and differences in effect sizes). The study then carries out a cross register analysis of the presidents' language and demonstrates consistent cross-register differences with regard to functional linguistic patterns of oral expression, information density, situation-dependent discourse, and narration versus immediacy. While the study observes some clear individual trends (e.g., Bush is more nominal than the other presidents across registers), even in these cases of well-defined personal trends individuals showed consistent register patterns.

While such studies are still in the minority, the emerging state-of-the-art appears to include two main findings. First, the language of the individual has been consistently shown to vary depending on the situation of use. This proves that individual expression is not uniform; rather, individuals are adept at adjusting their language to the demands of the situation of use. Second, statements by Bakhtin and Carter suggest that situational contexts may vary with regard to the amount of linguistic latitude they allow individual authors. These statements are in line with Biber's (1994) view that registers are defined to varying extents by their situational characteristics and thus may be expected to impose varying degrees of situational constraints. This emerging understanding thus calls for new investigations of individual language use across

a range of registers not only distinct in their situational characteristics, but diverse in the amount of linguistic freedom they may allow.

2.4 GOALS OF THE PRESENT STUDY

The goal of this chapter was to overview the main agenda, questions, and findings in research traditions concerned with variation according to users and use. The chapter has also aimed to demonstrate that while previous studies at the interface of these different sub-fields have each made important contributions, the interactions between approaches to variation according to users and use are still under-investigated. The Figure below once again shows these distinct approaches and in this section serves to illustrate the goals of the present study.

Figure 2.2. Approaches to investigating linguistic variation

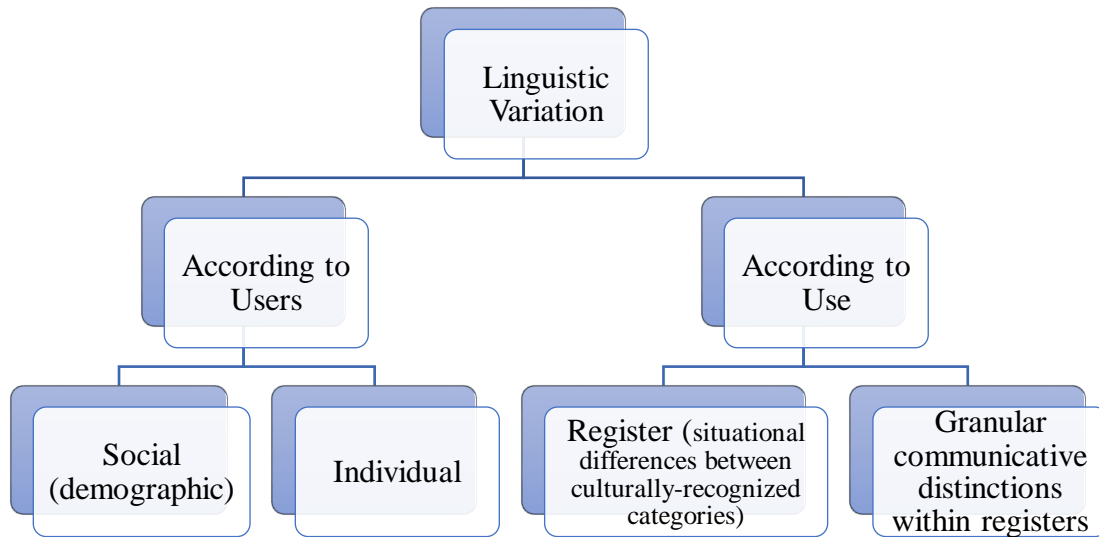


Figure 2.2 shows that these research traditions, despite their dramatic differences, are united by the common goal of explaining linguistic variation. Linguistic variation, however, is often operationalized differently within these fields (e.g., variation in phonetic and phonological

features seen as indexical markers, lexical features seen as markers of individual expression, etc.). The present study is concerned with functional linguistic variation across texts. Functional linguistic variation is typically not investigated in sociolinguistic research, where variation is interpreted as indexical rather than in functional terms.

The first goal of this study, therefore, is to identify general underlying patterns of functional linguistic variation in the corpus. To this end, the study uses a macroscopic analytical approach (Biber, 1985) – a multidimensional analysis – which identifies dimensions of linguistic variation among texts. These dimensions are based on patterns of feature co-occurrence in texts and in that differ from a microscopic analysis focused on the functions of individual linguistic features. Microscopic analyses provide crucial insight into the specific communicative functions of particular features, and thus lay the necessary foundation for and inform macroscopic analyses (Biber, 1985), which integrate and rely on the functional interpretations of individual features. However, due to their focus on a particular feature, microscopic analyses have a narrow focus and cannot identify the “overall parameters of linguistic variation within a set of texts” (Biber, 1985, p. 339). The central claim in favor of a macro approach to investigations of functional variation is that there is not an infinite number of separate communicative functions served by individual linguistic features; rather, there is a limited number of underlying communicative functions shared by features frequently co-occurring in discourse (Biber, 1985, p. 340). The macroscopic approach thus allows an investigation of these systematic relationships existing between individual features and an identification of underlying communicative functions that are “conceptually clearer than the many linguistic measures considered individually” (p. 340). This study uses this macroscopic analytical approach to address the following question:

RQ1: What underlying functional dimensions of linguistic variation exist in the corpus?

Figure 2.2 further shows two major distinctions in approaches to linguistic variation – investigating variation according to characteristics of the users and according to use, with each major branch further divided into the specific sub-fields based on the type of predictors of interest. Variation according to users encompasses social (demographic) and individual characteristics, while variation according to use is explained through differences between culturally-recognized register categories and further granular communicative distinctions within registers. The present study examines the variables of register, individual author, age, and gender and aims to quantify the relative contribution of these social, individual, and situational predictors to explaining functional linguistic variation across texts. The study thus poses the following research question:

RQ2: To what extent is linguistic variation along each of the identified dimensions predicted by social group (age and gender), individual differences, and register?

The study then focuses on the social groups' approach to register and investigates age and gender groups across registers and within each register. An examination of each age and gender group across registers is intended to illustrate a) the extent of register effect for each b) whether age and gender groups reveal the same cross-register patterns. An examination of age and gender within each register is expected to show a) whether social characteristics of the authors predict linguistic variation within registers; b) whether age and gender groups show the same patterns of variation within different registers. The study addresses the following questions:

RQ3: How do age and gender groups vary *across* registers?

- Within each age group, to what extent does register predict linguistic variation?
- Within each gender group, to what extent does register predict linguistic variation?

RQ4: How do age and gender groups vary *within* registers?

- Within each register, to what extent does age predict linguistic variation?
- Within each register, to what extent does gender predict linguistic variation?

The study then proceeds to investigate reasons other than social characteristics of the authors for variation across texts within registers. Specifically, the study explores granular communicative differences that may exist among texts of a single register and investigates each register with regard to that possibility. Drawing on previous research that has established varying degrees of situational uniformity in different registers, this study aims to characterize the registers in the corpus in terms of the degree to which they are situationally well-defined. To this end, the study identifies constant and variable situational parameters in each register and explores the effect of the variable situational parameters on linguistic variation among texts. In this process, the study examines the roles of previously established situational parameters that account for variation among texts of a register, such as communicative purpose and topic, as well as the possibility that in some quite specific registers, additional, highly granular situational distinctions may explain linguistic variation. The study poses the following research question:

RQ5: To what extent do communicative differences across texts of a register explain linguistic variation within registers?

Finally, the study turns its attention to individual linguistic variation and examines individual authors across registers. The study addresses the following research question:

RQ6: How does the language of individuals vary across registers?

Each question is discussed at length in the following chapter.

CHAPTER 3. METHOD

This chapter presents the corpus of the study, the included variables and their levels – the characteristics of the participants and the situational characteristics of the registers (Section 3.1), the linguistic features selected for the analysis (Section 3.2), and the methods used to address each research question (Section 3.3).

3.1 Corpus of 100 Idiolects

The study makes use of the corpus of ‘100 Idiolects’ (Heini, Kredens & Pezik, 2021), collected and annotated by a research team in the Aston Institute for Forensic Linguistics of Aston University, UK. The corpus consists of speech and writing samples of 112 individuals, undergraduate students at Aston University, aged 19-25, from a variety of ethnic backgrounds and geographic regions of the United Kingdom. Each participant provided the following background information: name, age, gender, place(s) of residence, ethnicity, and languages spoken. Identifying gender, the participants were asked to choose from four options – ‘male’, ‘female’, ‘other’, and ‘prefer not to say’. All identifying information, such as participant names as well as any proper names used by the participants in their texts, was removed by corpus creators in the process of data anonymization.

Each speaker in the corpus produced texts in a range of registers: oral interviews on a single topic (food preferences and eating habits), image descriptions, emails, text messages, academic essays/reports, written responses to a task (positive and negative evaluations of Aston facilities), and business memos elicited in response to a scenario. Four registers of the study – business memos, evaluations, interviews, image descriptions – were elicited in the process of data collection, and each participant produced one text within each register. Three registers – essays, emails, and text messages – contain naturally occurring texts, supplied by the same 112

participants for research purposes. Thus, the provided essays were written by these students as course projects, and the text messages and emails represent real instances of personal and professional communication.

The present study uses a portion of the corpus, consistently including one text from an author in each register, which amounts to 112 texts per register. An exception is the register of evaluations, with 64 texts due to not all respondents participating in data collection. In the cases of essays, emails, and text messages – registers in which more than one text was provided – the longest text submitted by each participant was selected. Length was used as a criterion for text selection due to the overall shortness of emails and text messages and problems this limitation posed for quantitative corpus analysis.

The corpus contains six age groups, of which five are investigated by the study (19-23). The only 25-year-old respondent was excluded from the analyses. All participants are undergraduate students or recent graduates of Aston University. The age groups may therefore be aligned with years in their undergraduate programs, although there is no evidence in the metadata of the one-to-one correspondence between age and year in the program¹. These age groups also do not represent sufficient age differences in order to shed light on questions that are of interest to sociolinguistics, such as age-grading (e.g., Murphy's (2010) study cited above with 20- and 30-year increments). These groups do, however, represent the demographic commonly investigated in sociolinguistic research, namely, young adults, and thus insights this study aims to offer regarding this group's language use with relation to register may add to the existing descriptions of this group in the sociolinguistic literature.

¹ Thus, while by including the variable of age the study tests the possibility that age groups may represent the length of immersion in the academic context, it is hard to assert that this is the case. The results therefore need to be interpreted with caution, and this issue will be revisited in the discussion of findings.

With regard to gender, the present study investigates only the ‘male’ and ‘female’ gender groups, as the other two labels – ‘other’ and ‘prefer not to say’ – do not specify the gender category. Table 3.1 presents the corpus composition of the present study, and the following subsections provide more detail on each of the analyzed registers.

Table 3.1. Corpus of the study: Variables, variable levels, number of authors, and texts

Variable	Level	Authors (N)	Texts (N)
Gender	Male	20	132
	Female	90	590
	Prefer not to say	2	14
Total analyzed:		110	722
Age	19	6	38
	20	28	181
	21	37	250
	22	32	210
	23	8	50
	25	1	7
Total analyzed:		111	729
Register	Written	Speakers (N)	Texts (N)
	Business memo	112	112
	Email	112	112
	Essay	112	112
	Evaluation	64	64
	Text Message	112	112
	Oral		
	Description	112	112
	Interview	112	112
Total analyzed:		112	736

3.1.1 Registers in the Corpus

This section presents general descriptions of each register in the corpus of the study. The section distinguishes between elicited registers, produced for the purposes of corpus compilation, and the

naturally occurring registers, in which case authors contributed texts they produced for real-life purposes of authentic communication. This section briefly introduces the tasks the authors were presented with in the case of elicited registers and the typical characteristics of the naturally-occurring registers. Table 3.2 summarizes the situational characteristics of each register, relying on the Text-Linguistic framework for situational analysis described in Chapter 2 (Table 2.1). A detailed situational analysis of the texts of these registers and an extensive discussion of Table 3.2 will be presented in Chapter 6 as the framework is applied to the analysis of communicative distinctions within registers.

Table 3.2. Application of the framework for situational register analysis (Biber 1994; Biber & Conrad, 2019)

Register	Addressee	Relationship	Relative status/power	Interactiveness	Shared knowledge	Domain	Setting; Shared time/ place	Mode	Specific medium	Production & comprehension	General Purpose	Specific purposes	Topic
Essays	Sg (course instructor)	Academic	Lower>higher	Low	High (specialist)	School	(Semi)public /no	Written	Permanent electronic	Revised & edited/ careful reading	–	–	–
Emails	Sg	1.– (undefined for non-academic emails) 2.academic	1.– (undefined for non-academic emails) 2.Lower>higher	High	1.— (undefined for non-academic emails) 2. med/high	1. – (undefined for non-academic emails) 2. school	Private/ no	Written	Permanent electronic	Revised & edited/ careful reading	–	–	–
Interviews	Sg	– (prof/academic or 1 st time encounter)	– (lower>higher or equal)	High	Low (general)	Research project	Public/ yes	Spoken	Transient face-to-face	Online	Answer questions/ inform	Share personal experience, give examples	Food, eating habits
Descriptions	Sg/pl	Participant-researcher	– (lower>higher or equal)	Low	Tied to the visual	Research project	Public/ no	Spoken	Permanent recorded	Online production/ offline comprehension	Describe	Other more specific purposes are unlikely	Content of the visual
Evaluations	Sg/pl	Participant-researcher	– (lower>higher or equal)	Low	– (general)	School, research project	(Semi)public /no	Written	Permanent electronic	Revised & edited/ careful reading	Describe Evaluate	–	Aston campus facilities
Business memos	Sg	Professional	Lower>higher	Low	High (general)	Work	Private/ no	Written	Permanent electronic	Revised & edited/ careful reading	Inform	–	CEO's trip to Helsinki

Elicited Registers

Business Memos

Business memos in the corpus are written texts elicited in response to a scenario. The participants were asked to do research and provide an imagined CEO of a company they work for with an itinerary and specific suggestions for leisure time activities during the CEO's business trip to Helsinki. As part of this task, the participants were expected to research options that satisfied the criteria specified by the scenario, such as accommodation budget and preferences, dietary restrictions, shopping needs, entertainment preferences, scenery, food, and nightlife. The exact directions of the task are unclear; the overall goal of the memo is to provide the CEO with details about their itinerary and stay.

Evaluations

Evaluations are written responses the participants provided to the task of sharing their positive and negative impressions of the Aston University campus. The texts are directed at an unspecified audience or the researcher(s) administering the task. The texts describe the strong and weak points of building locations, the location of the university campus with relation to transportation, student facilities, student life, campus environment, services provided by the university, etc. Texts may support evaluations with the students' personal past experiences in the form of anecdotes. The mini-essays are typically divided into two parts focused on the 'likes' and 'dislikes', thus addressing the two main points of the prompt.

Image Descriptions

Image description are oral texts produced by the participants in a public research setting with the help of speech-to-text software for an unspecified audience. Alternatively, the audience includes

the researcher, who was likely to be present during data collection. The audience, however, is not addressed directly and does not participate in communication, but only witnesses text production. Texts are monologic, noninteractive. Participants were asked to describe three visuals depicting scenes of university life. Descriptions include discussions of appearance, relationships, university life and activities, and actions performed by the people in the visuals. The same visuals were used for all participants.

Interviews

Interviews are conducted in a research setting in a face-to-face interaction with the interviewer and represent online oral dialogic communication. The interviewer's lines were removed, and the texts contain only interviewees' responses. All interviews are on the topic of food preferences and eating habits, with the same questions covered by all interviews.

Naturally-occurring Registers

Emails

Emails represent instances of real typically professional communication. While these emails are not restricted to any particular setting, addressees, or topics, most texts are addressed to a university professor, administrator, or another authority, and classmates. The topics tend to be coursework and other university-related matters. Other emails are addressed at external addressees from other contexts, although such texts are sparse.

Essays

Academic essays or reports are works of academic writing produced by the participants in an academic setting as part of university coursework. Disciplines range from psychology (e.g.,

biological and psychoanalytical perspectives in the field, parenting styles) to social studies (e.g., paid family leave and gender equality, effects of immigration), law (e.g., international sales laws, Birk's legal taxonomy), history (e.g., democracy in WWII), and cognitive science (e.g., application of Stroop test).

Text Messages

Similar to emails and essays, text messages are instances of naturally-occurring communication. Texts represent personal communication and tend to be addressed to friends, family, significant others. Texts display an extremely broad array of communicative purposes and are written on a wide range of personal topics, including school, family, relationships, work, and others.

3.2 Linguistic Features

In feature selection, the study draws on Biber's (1988) comprehensive study of functional variation in speech and writing. Feature selection is a critical first step in macro analyses of functional linguistic variation and involves analysis of "the range of communicative situations and purposes within a given domain" (Biber, 1985, p. 341). While "the range of communicative situations and purposes" in the corpus of this study, outlined in the previous section, is not as broad as that of the study of the full range of variation in speech and writing (Biber, 1988), the registers of the present study, both the naturally-occurring and elicited ones, include spoken and written, interactive and non-interactive, monologic and dialogic varieties, with close personal and socially distant professional relationship between interactants, communication occurring in formal and informal settings, in shared physical space and time and mediated by technology, online spontaneous interactions and prepared and edited speech, and perhaps most importantly, includes a wide range of communicative goals. This range thus appears to warrant an inclusion

of linguistic features from a variety of functional domains featured in Biber (1988). These domains include pronouns and pro-verbs, nominal forms, adjectives and adjective classes, prepositions, lexical classes of verbs, modal verbs, tense and aspect markers, passive voice, stative forms, lexical classes of adverbs, place and time adverbials, coordination, subordination, questions, lexical specificity, reduced forms and dispreferred structures, and text length. Table 3.3 below contains a complete list of features within each of the general domains of the study.

While these features have been investigated by register studies and shown to constitute linguistic dimensions of situationally-determined variation (e.g., Biber, 2006; Gray, 2015; Biber & Egbert, 2018, among others), lexico-grammatical variation does not only reflect register differences. In research on individual variation, lexico-grammatical features are viewed as “formal characteristics”, “independent of their specific content or meaning” (Moerk, 1970, p. 225), “nonsemantic linguistic variables” (Balasubramanian, 1982, p. 11), “linguistic form” rather than “semantic units” (Esser, 1993, p. 298). Lexico-grammatical features have been shown to distinguish between and capture important patterns in individual language use by several studies, using both a microscopic approach, focusing on a particular feature (as the studies of individual differences overviewed above) and studies employing the macroscopic approach revealing the underlying patterns of variation among the authors’ works and examining the authors’ language with respect to these patterns (e.g., Biber & Finegan, 1994; Egbert, 2012; Gracheva, 2022).

A number of these lexico-grammatical features have been shown to be predicted by social variables, although gender variation has received overwhelmingly more prominence in such investigations. Coates (2004), for example, discusses the effect of gender on clausal coordination and adverbial subordination, passive voice (as a feature of ‘impersonal discourse’), degree adverbs, and text length; Eckert and McConnell-Ginet (2003) and Coates (2004) note that

hedges are a characteristic of female speech, while ‘boosters’ are viewed as a feature common for female speech by Eckert and McConnell-Ginet (2003), Coates (2004), Aijmer (2018), and Brezina, Love and Aijmer (2018). The relative scarcity of research and the importance of investigating lexico-grammatical variation with relation to age are highlighted by Barbieri (2008), who stresses that, compared to the insights gained into variation due to other social variables, such as gender or social class, little is known about the effect of age on the use of syntactic and lexical features. Some exceptions include Aijmer (2018), who investigates intensifiers with relation to age; hedges, vague category markers, amplifiers, and boosters feature in Murphy’s (2010) book-length treatment of linguistic variation by age in female speech; in addition to these features, Barbieri (2008) investigates modal verbs, adjectives, pronouns, stance adverbs, and discourse markers, all included by the present study.

This brief overview shows that it has been proposed previously that lexico-grammatical features vary not only by register, but may be equally informative for accounts of individual and social variation. The study begins with a large pool of linguistic features (Table 3.3), which have been shown to play an important role in functional linguistic variation (the exact functions of each are described in detail by Biber, 1988, pp. 223-245) as well as have been shown to serve as indexical markers of group membership and individual expression. The linguistic features included in the study are subsequently reduced to factors or dimensions through factor analysis based on frequent co-occurrence patterns of these features in texts. The next section describes the statistical procedure of factor analysis and the features that comprise the factors, which are used for all further analyses in the study.

Table 3.3. Features initially included in the MDA

General Domain	Linguistic Feature
1. Pronouns & Proverbs	<p><i>Personal (including possessive)</i></p> <ol style="list-style-type: none"> 1. 1st person (Sg and Pl forms of personal, possessive, objective, and reflexive pronouns: e.g., I, my, me, myself) 2. 2nd person (personal, possessive, objective, and reflexive pronouns: e.g., you, your, yourself) 3. 3rd person (Sg and Pl forms of personal, possessive, objective, and reflexive pronouns: e.g., he, his, him, himself) <p><i>Impersonal</i></p> <ol style="list-style-type: none"> 4. Impersonal ‘it’ 5. Demonstrative (Sg and Pl forms: this, these, that, those) 6. Indefinite (someone, anyone, everyone, etc.) <p><i>Proverbs</i></p> <ol style="list-style-type: none"> 7. Proverb ‘do’ (e.g., She asked me to call, but I never did.)
2. Nominal Forms	<ol style="list-style-type: none"> 8. Nominalizations (management, organization, worker) 9. Abstract nouns (art, history) 10. Concrete nouns (car, vase) 11. Pre-modifying nouns (working conditions) 12. ‘S genitive (Anna’s story)
3. Adjectives	<ol style="list-style-type: none"> 13. Attributive adjectives (nice hotel) 14. Predicative adjectives (it was nice)
4. Adjective classes	<ol style="list-style-type: none"> 15. Attitudinal adjectives (best option, this is ideal)
5. Prepositions	<ol style="list-style-type: none"> 16. Epistemic adjectives (evident, obvious) 17. Of-genitive: Noun + of (value of this approach) 18. Prepositional phrases: Noun + prepositions (other than ‘of’: e.g., relationship with) 19. Prepositional phrases: Adjective + preposition (e.g., important for) 20. Other prepositions
6. Lexical Classes of Verbs	<ol style="list-style-type: none"> 21. Prepositional complement: preposition + Wh-clause (talk about what may be useful) 22. Public verbs (e.g., say, speak) 23. Private verbs (e.g., think, wonder) 24. Suasive verbs (e.g., instruct, demand) 25. Activity verbs (e.g., go, leave) 26. Seem/ appear 27. Verb ‘have’

General Domain	Linguistic Feature
7. Modal Verbs	28. Possibility modals (can, may, might, could) 29. Necessity modals (should, must) 30. Predictive modals (will, would, shall)
8. Tense & Aspect	31. Past tense (V2) 32. Perfect aspect (have/has/had + V3) 33. Present tense 34. Progressive aspect (be + ing form)
9. Passive Voice	35. Agentive passive (passive + by phrase) 36. Agentless passive (passive without by phrase)
10. Stative Forms	37. 'Be' as main verb 38. Existential 'there'
11. Lexical Classes of Adverbs	39. Downtoners (e.g., merely, nearly) 40. Hedges (e.g., almost, maybe) 41. Amplifiers (e.g., completely, absolutely) 42. Emphatics (e.g., really, most) 43. Conjuncts (e.g., however, nevertheless) 44. Discourse particles (e.g., well, anyway) 45. Other adverbs
12. Place & Time Adverbials	46. Place adverbials (e.g., here, there) 47. Time adverbials (e.g., now, later)
13. Coordination	48. Phrasal coordination 49. Clausal coordination
14. Subordination	<i>Complementation</i> 50. Verb complements (that & to) controlled by public verbs (He said that...) 51. Verb complements (that & to) controlled by private verbs (I thought that...) 52. 'That' stance verb complements (She admitted that...) 53. 'To' stance verb complements (I refuse to...) 54. 'That' deletion controlled by public verbs (He said (that del) he'd come) 55. 'That' deletion controlled by private verbs (I thought (that del) you finished it) 56. 'Wh' complements (He told me where to go) 57. 'Ing' verb complement clauses (I'd avoid doing this) 58. Adjective complement clauses (It's impossible to achieve...) 59. 'That' noun complement clauses (the fact that this is true) 60. 'To' noun complement clauses (ability to read)

General Domain	Linguistic Feature
	<i>Relative Clauses</i>
	61. ‘That’ relative clauses on subject position (the question that worries me)
	62. ‘That’ relative clauses on object position (the question that you are asking)
	63. ‘Wh’ relative clauses on subject position (the person who promised to...)
	64. ‘Wh’ relative clauses on object position (the book which I’m reading now...)
	65. Pied-piping (the person about whom she asked)
	66. ‘Ing’ relative clauses (a person admiring ...)
	67. ‘Ed’ relative clauses (a person admired by)
	68. ‘To’ relative clauses (a person to admire)
	<i>Adverbial Clauses</i>
	69. Causative (e.g., because, as)
	70. Concessive (e.g., though, although)
	71. Conditional (e.g., if, unless)
	72. Other (e.g., after, before)
	73. ‘To’ adverbial clauses (To achieve this, one should...)
	74. ‘Ing’ adverbial clauses (Looking out of the window, he said...)
15. Questions	75. ‘Wh’ questions (What do you think of...?)
16. Lexical Specificity	76. Type-token ratio
	77. Word length
17. Reduced forms and dispreferred structures	78. Contractions (e.g., I’d; isn’t; She’s, etc.)
	79. Stranded prepositions (e.g., something I was dealing with)
	80. Split auxiliaries (to honestly say ...)
18. Text length	81. Text length

3.3 Quantitative Method

3.3.1 Patterns of Functional Linguistic Variation

RQ1: What underlying functional dimensions of linguistic variation exist in the corpus?

The overarching goal of the study is to explore the relationship between social, individual, and situationally determined linguistic variation. As discussed in Section 2.4, to this end, the study

uses a macroscopic approach and first identifies the underlying patterns (dimensions) of functional linguistic variation among texts of the corpus. The study thus aims to describe and compare the registers, discourse patterns characteristic for the age groups, gender groups, and individual speakers of the study along these underlying dimensions and measure the amount of linguistic variation explained by each of the independent variables (age, gender, individual author, register) on each of the dimensions. Therefore, to address the first research question regarding the patterns of functional linguistic variation that exist in the corpus, the study conducts exploratory factor analysis, a multivariate statistical technique that reduces the large pool of initially included linguistic variables to a smaller set of latent variables or underlying dimensions, based on patterns of feature co-occurrence. As noted earlier, this macroscopic approach to linguistic variation is based on the assumption that linguistic features co-occur to jointly contribute to a particular underlying function (Biber, 1985). Feature co-occurrence is therefore not random, but always reflects functional patterns of variation in a domain and lends itself to interpretation (Biber, 1988). Factors, or dimensions, identified through factor analysis, thus represent these general underlying functional patterns of linguistic variation among texts.

Patterns of linguistic co-occurrence are revealed through a correlation matrix, which is at the core of factor analysis and shows the extent to which individual linguistic features are associated with each other and co-vary in texts. Linguistic features correlate positively with each other if they tend to frequently occur together and negatively if they occur in complementary distribution – that is, if one set of features are common in texts, it is unlikely that the other set of features will be present. Based on these correlation patterns, the features load on the poles of the factor: the features that show positive correlations with each other group on one end of the factor, while features that correlate negatively with the first set load on its opposite end.

The statistical assumptions for factor analysis include linear correlations among the variables, factorability of the sample (Kaiser-Meyer-Olkin factor adequacy), and a sufficient sample size. No curvilinear relationships were observed between the variables; therefore, the assumption of linearity was met. With the sample size of $N = 736$ and 81 linguistic variables, the sample is robust enough for factor analysis, as the study meets the established sample size criterion of five observations (i.e., texts) per variable (i.e., linguistic feature) (Gorsuch, 1983). Factorability of the sample, or sampling adequacy, was checked through examining the Kaiser-Meyer-Olkin (KMO) values – indicators of bivariate and partial correlations between the variables. The KMO statistic for the sample in the study is 0.72, which suggests a ‘middling’ sampling adequacy, namely overall sufficiently strong correlations between the variables and, therefore, a factorable sample (Kaiser, 1974).

The study determines the optimal number of factors that best capture variation in the corpus through a scree plot of eigenvalues (Appendix 1) which shows the total amount of shared variance explained by the factors (Biber, 1985). A break in the eigenvalues indicates the number of factors that best account for the variance in the data and a point beyond which new factors do not (or minimally) contribute to explaining additional variance in the data. In this study, a four- and five-factor solutions were examined, and due to a more diluted, less clearly interpretable factor structure produced by the latter, a four-factor solution was used in the final analysis.

Since in certain factor extraction techniques, such as ‘factor extraction solution’, used for exploratory factor analysis (Gorsuch, 1983) the first factor accounts for the most variance and has the most feature loadings, this results in less insight into the other factors. The rotation of the factors is a way of addressing this issue by ensuring that the variables load on the smallest number of factors and a limited number of only large feature loadings that best account for and

characterize some aspect of shared variance appear on each factor rather than a large number of small ones (Biber, 1985, p. 350). The present study uses the Promax factor rotation method, which assumes that the factors are correlated (Gorsuch, 1983; Brown, 2009). While the correlations between the four identified factors are relatively low (ranging between ± 0.12 to ± 0.25), which suggests that the factors explain unique variance in the data, some correlation is to be expected, as linguistic dimensions are not independent of each other just as the linguistic features at their core correlate with each other to some degree. The four identified factors cumulatively account for 21% of variance in the data (6%, 6%, 5%, and 4%, respectively).

To identify the features most strongly associated with each of the four factors, communalities were inspected, and an accepted cut-off of .20 was used for feature inclusion on the factors, beyond which features were considered for removal. Communalities indicate the degree to which a given feature correlates with the other features on the factor and how much it contributes to that factor. Next, the factor loadings of each linguistic feature were examined – a measure of how representative the linguistic feature is of the factor it loads on or the extent to which the interpretation of the factor can be generalized to the function of a particular linguistic feature (Biber, 1985, p. 350). A high factor loading of a feature thus indicates that this feature is strongly associated with the factor and its function contributes considerably to the interpretation of this factor. The study follows the commonly accepted guideline of factor loadings of $\pm .30$ and above for feature retention. Following the practice of some factor analyses (e.g., Biber & Egbert, 2018), however, the study also records high factor loadings approaching this cut-off in parentheses, as these features are likely to be informative in the interpretation of the factor. It is important to note that factor loadings of linguistic features may load on and exceed this cut-off on several factors, as all linguistic features correlate to some degree. Each linguistic feature,

however, is included on only one factor, where this correlation, and by extension, the contribution the feature makes to the underlying communicative function identified by the factor is largest. Appendix 2 contains a complete list of the linguistic features included in factor analysis with factor loadings of each feature on each factor.

The factors are then interpreted with attention to the functions that the individual features comprising them perform in the texts, and dimension labels that best reflect the underlying pattern of variation captured by the factor are proposed. These factors or dimensions thus represent underlying general patterns of functional linguistic variation among the texts in the corpus. These patterns were identified by the present study with the goal to then explore the extent to which social and individual characteristics of the users (age, gender, individual author identity) and the situation of use (register) account for linguistic variation. Therefore, these identified linguistic dimensions are used as the dependent variables of the study in all subsequent analyses. That is, the following chapters will examine the role of each predictor variable in linguistic variation on each of the identified dimensions. The functional interpretations of these dimensions are presented in Chapter 4.

3.3.2 Variance Explained by Social Group, Individual Differences, and Register

RQ2: To what extent is linguistic variation along each of the identified dimensions predicted by social group (age and gender), individual differences, and register?

The study measures the amount of variance accounted for by the four predictor variables (age, gender, individual author, and register) cumulatively and the proportion of unique variance explained by each of the independent variables separately on each of the identified dimensions of variation in the corpus. To this end, the study makes use of a random effects model and treats each of the independent variables as random effects. At this point in the study, the specific levels

of the variables are not the focus of analysis; rather, the study is concerned with the overall variance explained by the four predictors and thus computes the overall conditional R^2 and the proportion of unique variance explained by each of the variables. The study therefore does not distinguish between fixed and random effects. The random effects model allows measuring the amount of variance accounted for by each of the variables independently after all the other variables have been accounted for. Additionally, the random effects model allows for possibilities of some levels of some variables not being represented in the data (in the case of this study, only half of the authors are represented in the register of evaluations) (see Egbert & Gracheva, 2023 for more detail on this use of random effects model). Four random effects models were run on each of the identified dimensions. The results of this analysis are presented in Chapter 4.

3.3.3 Social Group with Relation to Register

RQ3: How do the social groups of the study vary *across* registers?

- Within each age group, to what extent does register predict linguistic variation?
- Within each gender group, to what extent does register predict linguistic variation?

The study further investigates social groups with relation to register on each of the identified dimensions of variation. Research Question 3 focuses on the language use of each social group – five age groups (19-23) and two gender groups of the study (men and women) – *across* registers. To address the question of the extent of variation explained by register for each of the social groups, four one-way analysis of variance procedures (ANOVA) were run within each of the social groups, with register as the independent variable and each of the four dimensions of variation in the corpus as the dependent variable (a total of 28 ANOVAs for the

five age groups and two gender groups; alpha adjusted for multiple comparisons). This analysis sheds light on ways register space is navigated by men and women and whether register awareness manifests itself differently across age groups or whether it develops with time spent in the university context.

RQ4: How do the social groups of the study vary *within* registers?

- Within each register, to what extent does age predict linguistic variation?
- Within each register, to what extent does gender predict linguistic variation?

The social groups of the study are then examined within each register. To address the question regarding the extent to which age and gender predict linguistic variation within registers, four one-way analysis of variance procedures (ANOVA) were run within each of the seven registers, with age as the independent variable and each of the four dimensions of variation as the dependent variable. The procedure was then repeated with gender as the independent variable and each of the dimensions as the dependent variable (resulting in a total of 56 ANOVAs, alpha adjusted for multiple comparisons). Each register is thus analyzed with regard to the internal variation by social group allowed within its scope: which groups differ within each register and what constitutes these differences in functional terms.

3.3.4 Communicative Differences within Registers

RQ 5: To what extent do communicative differences across texts of a register explain linguistic variation within registers?

To address the question regarding variation across texts within a register, the study examines each text with relation to the other texts in the register on each of the identified dimensions. Specifically, each text is examined with relation to the register mean. That is, the

study gains an insight into the place each text occupies in their respective register with relation to the central tendency of that register. To this end, the study takes the following steps:

1. The study introduces a **measure of distance** of the text in a given register from the register mean, which is taken to represent the register central tendency. This measure is computed as the difference between the register mean and the individual text score on a given dimension ($\text{Distance} = \text{Individual Text Score} - \text{Register Mean}$). The resultant measure of distance reveals whether the text is close to the central tendency (i.e., register mean) within a given register or, conversely, deviates from the central tendency in some way, thus representing a ‘peripheral’ rather than ‘central’ instantiation of a register. The terms ‘central’ and ‘peripheral’ (originally used in psychology literature on processes of categorization – e.g., Rosch (1976)) were applied to texts by Biber, Egbert and Keller (2020) in a study on situational text types. As that study identifies clusters of texts unified by their situational characteristics, it observes some texts that are ‘central exemplars’ (p. 22) of their clusters (or situational text types) and others that are not quite representative of their text type or ‘peripheral to [its] situational characteristics’ (p. 29).

The present study investigates central and peripheral texts within a single register (whereas in Biber et al.’s study these texts represented a situational text type which encompassed texts from a variety of registers) and groups the texts on the basis of linguistic rather than situational dimension scores (applying different methods), but the terms ‘central’ and ‘peripheral’ are used in the same sense as in the original study. That is, ‘central’ texts are representative of their group – in the case of the present study, the register they belong to. ‘Peripheral’ texts stand out from the trend and demonstrate some deviation from the central tendency. The approach adopted here is thus intended to shed light on the way in which each

text responds to the demands of a register and the possible reasons why a text may deviate from the central tendency, causing variation within a register.

Importantly, the ‘peripheral’ or ‘central’ place of a text in a register is not a dichotomy. Rather, texts can occupy various positions on this spectrum, being more or less removed from the register mean. Texts are therefore classified in how far or how close they are in relative terms from the register mean. Distance between an individual text and the register mean is operationalized through five categories that represent a cline from central to peripheral texts. These categories include:

- 1) **Central texts:** The text is considered ‘central’ if the text score is extremely close to the mean: the distance between the register mean and the individual text score is under 0.1
- 2) **Minimally removed:** The text is considered minimally removed from the mean if the distance between the register mean and the individual text score is between 0.1 and 0.3
- 3) **Moderately removed:** The text is considered moderately removed from the mean if the distance between the register mean and the individual text score is between 0.3 and 0.6
- 4) **Peripheral 1:** The text is considered ‘peripheral’ if the text score is quite far from the mean: the distance between the register mean and the individual text score is between 0.6 and 0.9
- 5) **Peripheral 2:** The text is considered ‘peripheral’ if the text score is extremely far from the mean: the distance between the register mean and the individual text score is over 0.9²

² The procedure described here (i.e., computing measures of distance between each individual text score and the register mean) was performed separately on each dimension. The identified dimensions have different scales (resultant from the different number of features that load on each dimension). It follows from this difference in scales that the ‘distance groups’, operationalized through the established cut-off criteria, will in fact have a different meaning on each dimension. That is, a distance from the register mean of <0.1 on Dimension 1 with a scale from 2 to –3 is measured in units different from that on Dimension 4 with a scale from 6 to –2.5. While the study does not compare distance groups across dimensions, these differences in scales are important to acknowledge as the results on different dimensions are discussed with reference to the same cut-off criteria.

The cut-off points in the five groups are arbitrary in the sense that they do not rely on any established cut-off criteria; however, they are established based on the measures of distance between individual texts and the register means observed in the data and are meant to represent common contrasts among sets of values. For example, a text with a value of distance 0.03 was deemed to be extremely close to the mean, as was a text with a measure of distance 0.06. The data contained a large number of texts with such values of distance. These extremely low numbers stood in contrast from measures around 0.15 or 0.25 (the group between 0.1-0.3), which in their turn were closer to the mean than the measures around 0.40 (the group between 0.3-0.6), etc. In each case, there were substantial numbers of texts that could be classified in the group, which was seen as an indicator of the stability of the group. On the other end of this spectrum, texts whose distance from the mean exceeded 0.6 and reached 0.9 stood in stark contrast from the rest of the data, while texts whose measure of distance exceeded 0.9 (e.g., 1, 1.15, or 1.25) were seen as extreme deviations from the central tendency. Again, such texts were not singular occurrences, and several texts in several registers could be placed in these groups, which suggested that the selected cut-off criteria reflected the tendency in the data (Appendix 5 contains the text scores with relation to register means on each dimension and the measures of distance which were then classified into the five groups described here).

2. The study further identifies the **direction of deviation**, i.e., the position of each text in each register with relation to the register mean: above or below the mean on the dimension. The position of each text with relation to the mean reveals a tendency towards the particular functional pattern on each dimension relative to the other texts, which constitute the central tendency. For example, if a text in a given register is above the mean on Dimension 1, a cline

from oral elaboration to informationally dense literate discourse (described in detail in Chapter 4), this author's text in this case is more oral than the typical representatives of the register³. The study then identifies the basis for this deviation.

Direction of deviation is possible to identify in each of the five distance groups. That is, texts that are very close to the mean and thus in the classification adopted in the study are considered 'central' are also positioned higher or lower relative to the mean. This tendency, however, is so weak (distance from the mean <0.1) that texts in this group were not examined with regard to the direction of deviation, but were all viewed as representative of the register. Deviation is the most prominent (and thus much more clearly interpretable) in texts that are considerably removed from the mean – texts labeled 'Peripheral 1' and 'Peripheral 2'. It is these groups of texts that represent the most stark contrast with the central tendency of a register and thus are among the most informative for this study. The group of texts labeled 'Moderately Removed' is inspected for confirmatory purposes: for example, if a pattern emerges from a systematic comparison of the peripheral groups to the central group, the texts in the 'Moderately Removed' group are used as additional evidence for the observed trend. Alternatively, the 'Moderately Removed' group is inspected if there are no texts in the 'Peripheral 1' and 'Peripheral 2' groups of the register. The group of texts labeled 'Minimally Removed' from the register mean is considered to be quite close to the central tendency, i.e., highly similar to the 'Central' distance group and at the same time form less of a contrast with the 'Peripheral' groups. This group was therefore not deemed informative for

³ Here it is important to make the distinction between the position of the texts relative to the mean (above or below) and the positive and negative range on the dimension and interpret the place of a text on the dimension in relative terms. If the text in this example, for instance, is more oral than the mean of the register, it does not necessarily appear in the positive (oral) range since the register mean itself may fall in the negative range. The texts that are above the mean may thus also fall in the negative range (e.g., if the register mean is in the negative range and the text is not far removed from it). While in this example this author's text would be generally informational rather than oral, what is important for the analysis here is that the text is *more* oral than the central tendency.

the analyzed contrasts. That said, identifying the ‘Minimally Removed’ group and including it in the five-group classification was still important, as first, it characterizes a substantial number of texts in each register (i.e., a substantial number of texts are that far from the mean), and second, it illustrates the continuous nature of the cline from central to peripheral texts in a register, with multiple possibilities between the two extremes.

3. After each text in each register is classified with regard to a) the distance group it belongs to (i.e., how far this text is positioned from the register mean) and b) the direction of deviation (i.e., whether this text shows a tendency towards one or the other functional possibility of the dimension), a qualitative analysis of texts within the distance groups is conducted to identify the basis for the observed patterns. Three types of texts are the focus of this analysis:

- Texts that represent the central tendency (Distance from $M < 0.1$)

As noted earlier, while direction of deviation was identified for this group, the tendency towards one or the other dimension function is not prominent in texts in this distance group due to their extreme closeness to the mean. More often, ‘Central’ texts are a rather homogenous representation of the central tendency.

- Peripheral texts *below the mean*

Texts that show a considerable deviation from the mean (mostly Distance from $M = 0.6-0.9$ and Distance from $M > 0.9$) in the direction of the functional pattern represented by the negative (or lower) pole of the dimension. As mentioned earlier, the tendency towards the negative pole, i.e., the position of a text *below* the mean, does not necessarily suggest that the text is in the negative range of the dimension. Rather, this direction of deviation suggests that

the text is *closer* to the negative pole and thus gravitates towards the function represented by that end and uses relatively more of its linguistic features.

- Peripheral texts *above the mean*

Texts that show a considerable deviation from the mean (mostly Distance from $M = 0.6-0.9$ and Distance from $M > 0.9$) in the direction of the functional pattern represented by the positive (or upper) pole of the dimension. Likewise, this direction of deviation does not indicate that the text is positioned in the positive range. Instead, it shows an inclination towards that pole and that functional pattern of the dimension, which results in relatively higher rates of occurrence of the features associated with that functional pattern.

3.3.5 Individual Variation across Registers

RQ 6: How do authors vary across registers?

Addressing RQ 5 above, the study explores variation across texts within a register. Each text has been characterized by a measure of distance from the register mean in each register and by a tendency towards a particular functional pattern of the dimension. While the analysis in RQ 5 solely focused on differences among texts of a register, it is now possible to apply the findings of that stage to an examination of individual language use across registers. That is, each author in the corpus has produced a text within each register⁴. Each register has now been characterized in terms of the amount of internal variation it allows across its texts (RQ 5). The study thus gained insight into the place each author's text occupies within each register with relation to the register mean. The study further uses these data to reveal individual cross-register patterns by comparing

⁴ As discussed above, evaluations are an exception, with 64 out of 112 authors supplying texts in that register.

authors with regard to these tendencies across registers. For example, an author may consistently use language in ways characteristic for a given register – that is, consistently produce texts that represent the central tendency in all registers, which could suggest that this author shows sensitivity to register, is highly aware of the situational demands of each register, and consistently adheres to these demands; alternatively, an author may consistently be the ‘outlier’ on the periphery of the register, using language in some new, uncommon, idiosyncratic ways. This author may show this trend for a number of reasons – they may not be clear on what constitutes the situational constraints of a given register, i.e., what goals are typically set, what topics are discussed, in what setting, etc.; they may have a highly persistent individual way of expression across registers and this individual language use may be so uncharacteristic for some (or all) registers in the corpus that the speaker systematically occurs on the periphery; finally, this author may be introducing some additional situational characteristic that is not commonly associated with a particular register based on their unique communicative need. Another group of authors may exhibit elements of both trends, namely represent the central tendency in some registers while occurring on the periphery of others. At the same time, in all these possibilities, the authors may consistently show a preference towards a particular dimension function across registers (e.g., be noticeably more informational than the central tendency on Dimension 1), or vary with regard to their preferred functional pattern across registers (i.e., be more informational than the central tendency in some registers, but more oral and elaborated in others). Thus, at this stage of the study, individual variation is explored in terms of authors’ approaches to register-specific language use, and several contrasting examples are presented as case studies.

Analyses of communicative differences within registers (RQ 5) and the analyses of individual variation across registers that builds on that stage (RQ 6) are carried out in six

registers of the study and on two dimensions. The register of text messages is excluded from this stage, where each text is analyzed in terms of its relative preference for a particular dimension function. This analysis is problematic in text messages, whose dimension scores are inflated due to the extreme shortness of the texts. That is, often low negative scores of some text messages are the result of most, if not all, dimension features being absent in these extremely short texts. Alternatively, some text messages with low negative scores do not contain negative features, but still contain some positive features, or positive features outnumber the negative ones. The reason for this seemingly counterintuitive ratio is the considerably larger number of positive features on the dimension, which fail to occur in text messages due to their shortness. Therefore, when only a few of the positive features occur in the texts, in the absence of the majority of the other positive features their weight is still not sufficient to result in a positive score. These texts are then found on the extreme negative end even if they do not contain any negative features. This makes it impossible to analyze the tendency of the text towards a particular dimension function relying on differences in scores between the register mean and that text, as the linguistic features in that text may not reflect that tendency. Therefore, while it is possible to observe the functional patterns of the dimensions in many text messages, this problem observed in other text messages warranted the exclusion of the register from the analysis at this stage.

All analyses are carried out on two of the identified dimensions – Dimension 1 (‘Oral Elaboration vs. Information Density’) and Dimension 4 (‘Evidence-based Stance’), discussed at length in the following chapter. These dimensions are selected as two of the three dimensions that have repeatedly been shown to emerge in multidimensional studies of numerous domains and, as a result, have been referred to as ‘universal dimensions’ of linguistic variation – the oral/literate divide, the narrative/ nonnarrative dimension, and stance (Biber, 2004). The dimensions

of this study do not include the distinction between narrative and nonnarrative mode of expression, as none of the registers in the corpus contain substantial amounts of narration. However, the other two of the ‘universal’ dimensions – the oral/literate divide and stance – are present, and while the exact features comprising them may vary across studies, the analyses carried out by this study contribute to the body of knowledge about these universal discourse patterns. These dimensions also differ substantially with regard to their linguistic composition (discussed in detail in Chapter 4), with Dimension 1 comprised by nominal, pronominal, clausal, and adverbial features and Dimension 4 exclusively comprised by clausal features. Basing the analyses on these quite different linguistic dimensions can thus offer different insight into variation across texts of a register.

The methods used in the study to address each research question are summarized in Table 3.4 below.

Table 3.4. Summary of research questions and respective methods

Research question	Method	Outcome
1. What underlying functional dimensions of linguistic variation exist in the corpus?	Multidimensional analysis	Dimensions of functional linguistic variation in the corpus
2. To what extent is linguistic variation along each of the identified dimensions predicted by social group (age and gender), individual differences, and register?	Random effects model	Total variance explained by all predictors (conditional R^2) and proportion of variance (R^2) explained by each predictor
3. How do the social groups of the study vary <i>across</i> registers?	One-way ANOVAs within each social group	Significant register pairs within each social group

4. How do the social groups of the study vary <i>within</i> registers?	One-way ANOVAs within each register	Significant age and gender pairs within each register
5. To what extent do communicative differences across texts of a register explain linguistic variation within registers?	Analysis of each text with relation to register means: -Measure of distance between register mean and individual text score -Direction of deviation: above/ below the mean	1) texts representing central tendency 2) peripheral texts above the mean 3) peripheral texts below the mean
6. How do individual speakers vary across registers?	For each author: -Measure of distance between register mean and individual text score across registers -Direction of deviation: above/ below the mean across registers	Author's tendency to: 1) Be close to or far from the register mean across registers 2) Stay above or below the mean across registers

CHAPTER 4. DIMENSIONS OF FUNCTIONAL LINGUISTIC VARIATION

This chapter first presents the functional interpretation of the four factors identified by the study – the general underlying patterns of linguistic variation in the corpus. Table 4.1 contains the factors (hereafter, dimensions) with the proposed dimension labels and factor loadings of each linguistic variable associated with each dimension organized from the highest feature correlation with the dimension to lowest. The interpretations that follow are based on the functions of these individual linguistic features comprising a dimension, the manifestation of these functions in the texts of the corpus found in the positive or negative range on a given dimension, and similar dimensions observed and interpreted by previous macroscopic research studies. Table 4.2 shows the amount of variance each of the predictor variables of the study (age, gender, individual author, register) explains on each dimension. These results are then discussed in detail in the following sections.

Sections 4.1, 4.2, 4.3, and 4.4 present the dimension structure of each dimension and explain their functional configuration. Sections 4.1.1, 4.2.1, 4.3.1, and 4.4.1 place the registers of the study along each dimension and discuss the features of the dimensions in context, illustrating how these features co-construct the discourse of each register. It is important to note that many registers exhibit quite extensive internal variation along the dimensions, thus not necessarily representing a particular place on the dimension (e.g., highly informational or highly elaborated), but include texts that incorporate features of both extremes to varying degrees. This internal variation and its possible sources will be the subject of Chapters 5-6, where social group and communicative variation within registers is analyzed. The analyzed text excerpts in this chapter (sections 4.1.1, 4.2.1, 4.3.1, and 4.4.1) thus represent only one possibility of how the dimension features are realized in a given register and only serve the goal of providing a general register

comparison on the dimensions. These text samples should not, however, be viewed as representative of the entire register on the dimension.

Similarly, sections 4.1.2, 4.2.2, 4.3.2, and 4.4.2 place the social groups of the study, age groups and gender groups, on the identified dimensions. Each section (sections devoted to registers as well as social groups) reports the relative importance of each variable in predicting linguistic variation on each dimension.

Table 4.1. Dimensions of functional linguistic variation

Dimension	Associated linguistic features
Dimension 1:	Positive features:
<i>Online oral elaboration vs. Information density</i>	Contractions (0.73)
	Conjuncts (0.63)
	Word count (0.59)
	Present tense (0.58)
	Clausal coordination (0.54)
	Demonstrative pronouns (0.54)
	Stranded prepositions (0.45)
	Hedges (0.43)
	Emphatics (0.38)
	Pronoun 'it' (0.37)
	Split auxiliaries (0.32)
	Causative adverbial clauses (0.31)
	Adverbs (0.30)
	Negative features:
Pre-modifying nouns (-0.50)	
Word length (-0.46)	
Dimension 2:	Positive features:
<i>Abstract vs. Concrete discourse</i>	Nominalizations (0.56)
	Type-token ratio (0.51)
	'Of' genitive (0.48)
	Prepositional phrases (noun + preposition) (0.47)
	'To' noun complement clauses (0.41)
	Predicative adjectives (0.41)
	'That' noun complement clauses (0.35)
	Adjective complement clauses (0.35)
	Agentive passive voice (0.35)
	Agentless passive voice (0.34)
	Abstract nouns (0.34)

	'Ed' relative clauses (0.30) (Pied relative clauses (0.28)) Negative features: Concrete nouns (-0.45) Nominal/ indefinite pronouns (-0.40)
Dimension 3:	Positive features:
<i>Others-oriented descriptive vs. Self-oriented/ Interactive discourse</i>	Progressive aspect (0.53) 3 rd person pronouns (0.52) 'Seem'/'appear' (0.52) Attributive adjectives (0.38) Existential 'there' (0.36) 'Ing' relative clauses (0.34) ('Wh' relative clauses on subject position (0.28)) Negative features: 1 st person pronouns (-0.61) 2 nd person pronouns (-0.37) Verb 'have' (-0.38) Modal verbs of prediction (-0.31) Time adverbials (-0.31)
Dimension 4:	Positive features:
<i>Evidence-based stance</i>	'That' verb complement clauses controlled by stance verbs (0.76) Public verbs (0.56) 'That' deletion controlled by public verbs (0.48) 'To' verb complement clauses controlled by stance verbs (0.44) Private verbs (0.38) Past tense (0.37) Verb complement clauses controlled by public verbs (0.32) (Modal verbs of possibility (0.28)) ('That' deletion controlled by private verbs (0.28)) (Verb complement clauses controlled by private verbs (0.26)) Negative features: -

Table 4.2. Variance explained by each predictor variable (age, gender, author, register) in linguistic variation on each dimension

Dimension	Age	Gender	Register	Author	Total
Dim 1	0	0	76%	1%	77%
Dim 2	0	0.05%	79.4%	0.3%	80%
Dim 3	0	0	77%	0.5%	77.5%
Dim 4	0.1%	0	38.3%	0.5%	39%

4.1. DIMENSION 1. ORAL ELABORATION VS. INFORMATION DENSITY

The contrast observed between the features loading on the positive and negative poles of Dimension 1 resembles the commonly identified continuum of oral vs. literate discourse consistently discovered in multidimensional studies (e.g., Biber, 1988; Biber, 2006) focused on both oral and literate registers as well as exclusively oral or exclusively written ones. There is thus no one-to-one correspondence between the oral pole of this dimension and spoken registers or between the literate pole and the written ones. Rather, the oral/ literate dimension is a continuum from texts relying on features characteristic for oral interactive and involved discourse to texts that prioritize information transmission over interpersonal communication. On this dimension, oral discourse is constructed through verbal, clausal features, while literate discourse is associated with nominal features, which contribute to efficiency and concision in the task of information packaging (e.g., Biber, 1988).

Dimension 1 identified in this study exhibits the same verbal vs. nominal contrast. Clausal features associated with the positive end of the dimension include present tense (0.58), clausal coordination (0.54), split auxiliaries (0.32), causative adverbial clauses (0.31), and a feature most strongly associated with the dimension and which implies verbal structures – contractions (0.73). These features, all pointing to a focus on actions rather than information presentation, also indicate present-orientedness and immediacy, apparent through the high factor loading of present tense (0.58) and a tendency to elaboration and reasoning, conveyed through

causative adverbial clauses (0.31). This tendency to elaboration also appears to be confirmed through the high loading of clausal coordination (0.54) on the positive pole.

Alongside these verbal structures, a group of adverbial features, such as conjuncts (0.63), hedges (0.43), emphatics (0.38), and other adverbs (0.30), qualifying propositions and expressing speaker stance, are an important component of the dimension. Adverbs are viewed as markers of interpersonal and situated discourse (Biber, Johansson, Leech, Conrad & Finegan, 1999/2021, pp. 771-775) – functions which complement immediacy of communication, argumentation and reasoning, and elaboration, conveyed through the clausal coordination, causative adverbial clauses, and present tense, discussed above. As will be illustrated by the text samples below, these adverbial features contribute to the involved nature of the registers occurring on the positive end.

Another notable pattern on the positive pole of the dimension involves the features of reduced surface form (Biber, 1988), such as contractions (0.74), demonstrative pronouns (0.54), the pronoun ‘it’ (0.37), common in oral communication typically due to the shared physical context (space and time) between the participants which eliminates the need to always name the referenced entities, the time constraints of online speech production, and the increased cognitive load associated with it. Demonstrative pronouns – linguistic tools that allow quick references to objects known to the participants in a shared communicative context – are thus to be expected in such fragmentary, inexplicit oral discourse. Similarly, the pronoun ‘it’ functions to refer to known entities in a vague general manner and is made possible by the shared physical context, in which specifying each referent is typically not necessary. Alternatively, the pronoun functions as an exophoric reference enabled by the shared knowledge between the interactants. Finally,

contractions, the result of time constraints introduced by online production, serve as a time-saving device in fluent, fast, immediate communication (Biber, 1988).

Another set of features characteristic for oral discourse found on the positive pole of Dimension 1 is dispreferred structures, such as stranded prepositions (0.45) and split auxiliaries (0.32), the latter mentioned above as a clausal feature. Alongside contractions, stranded preposition and split auxiliaries are proscribed in written communication, at least in the prescriptive school of thought. These proscriptions, however, have been shown to be systematic rather than arbitrary (Finegan, 1980, cited in Biber, 1988, p. 243). Specifically, these structures represent a “mismatch between surface form and underlying representation, resulting either in a reduced surface form [...] or a weakened isomorphism.” What makes these structures common in oral discourse is thus likely to relate to its situational constraints, specifically the immediate nature of communication which calls for reduced forms, easier and faster to produce, rather than simply the fact that they are allowed in oral communication.

Overall, the positive features overviewed here all represent a particular aspect of oral communication, such as the immediacy of such communication, occurring in the present moment; the situated nature of such communication and the speaker’s position in it; its online production not conducive to planning and thus resulting in fragmented, vague, inexplicit discourse. Importantly, the speakers’ tendency to extensively elaborate on, explain, and develop their ideas appears to be a major characteristic feature of the oral end of this dimension, as indicated by clausal coordination and causative adverbial clauses. This oral elaboration is enhanced by the variable of ‘word count’ (0.53) loading on the positive end of the dimension, which suggests that oral elaboration in the corpus occurs in longer texts.

In contrast, pre-modifying nouns (-0.50), forming noun-noun sequences, and long words (-0.46), characterized as a feature of “specific, specialized meaning” (Zipf, 1949; Osgood, 1960; Drieman, 1962; Blankenship, 1974, cited in Biber, 1988, p. 238), typical for literate rather than oral production due to the lack of production constraints and ample opportunities for revision, appear on the negative end. Nominal structures have been discussed by studies investigating phrasal complexity and its evolution as features of efficiency and concision in information presentation, introduced and advanced into use by the needs of modern information-driven society (e.g., Biber & Gray, 2012; Biber et al., 2016; Staples et al., 2016). Noun-noun sequences, in particular, have received attention as a highly functionally diverse linguistic phenomenon, capable of conveying manifold complex relationships between its components (e.g., Egbert & Davies, 2019). Communicating “specific, specialized meaning” thus appears to be a function shared by the two features on the negative pole of this dimension, which represents specific, concise, and efficient information presentation. The discussed features are further illustrated in context as the registers of the study are analyzed along the dimension.

4.1.1 Registers along Dimension 1

Register is found to be a major predictor of variation on the dimension, explaining 76% of variance ($F(6, 729) = 363.5, p < 0.001$), all register pairs significantly different from each other with the exceptions of text messages and business memos ($p > 0.003$, Cohen’s $d < 0.2$) and evaluations and emails ($p > 0.003$, Cohen’s $d < 0.5$). Figure 4.1 shows the registers in the corpus along Dimension 1, and Table 4.3 contains the descriptive statistics (means and standard deviations) for each. The Figure shows that the registers that tend to occupy the positive end of the dimension, i.e., prioritize online oral elaboration over information density, are interviews, descriptions, and evaluations. As was described in Section 3.1, interviews and image descriptions

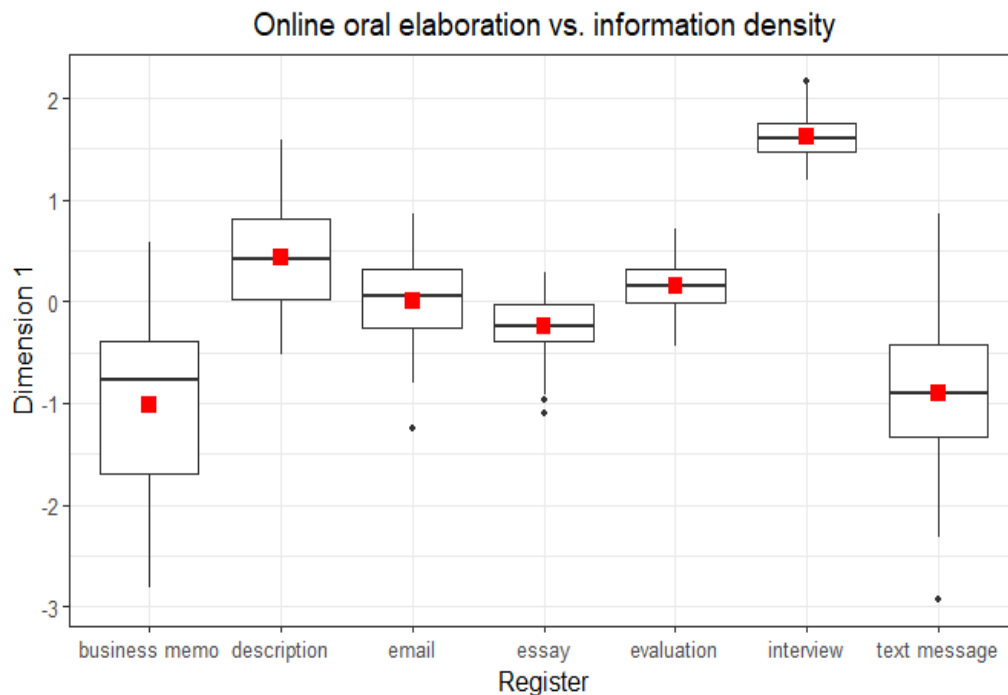
are spoken registers, marked by such situational characteristics of spoken discourse as real time production and, consequently, lack of opportunities for planning, editing, or revision.

Evaluations, mini-essays on the students' likes and dislikes of Aston campus, on the other hand, are written, the mode thus not imposing the same constraints. These results are not surprising: as noted earlier, the oral/ literate divide, repeatedly observed in multidimensional studies of both spoken and written registers, does not only reflect differences in mode. In fact, an important outcome of comprehensive investigations of speech and writing has been the fact that no clear distinction between spoken and written registers of English has been uncovered (Biber, 1988). Rather, registers vary along a continuum between highly oral, interactive, and involved text varieties on one end and highly informational ones on the other, regardless of the mode. Thus, evaluations in the present study, while being a written register, occupy an intermediate position on this continuum, combining features of information transmission, typical of literate discourse, and oral elaboration, interactivity and involvement, more characteristic for oral discourse, the latter outweighing the former. Text Samples 1, 2, and 3 illustrate features of oral elaboration (bolded) in these registers, contrasted with rather rare negative features of information density (underscored).

Table 4.3. Registers along Dimension 1. Descriptive statistics

Register	Dimension 1	
	<i>M</i>	<i>SD</i>
Business memos	-1.01	0.79
Description	0.44	0.51
Emails	0	0.38
Essays	-0.24	0.27
Evaluations	0.15	0.24
Interviews	1.62	0.21
Text messages	-0.9	0.66

Figure 4.1. Registers along Dimension 1



Text Sample 1 is an example of an interview, the highest scoring register on the dimension ($M = 1.62$; $SD = 0.21$). The text shows how the positive features of the dimension co-construct oral elaboration: specifically, frequent clausal coordination in this passage serves to string ideas together, most often adding to the previous ones (*and*) and occasionally indicating more complex idea relations (the disjunctive *but*), and the three causative adverbial clauses (*because*) elaborate on propositions, adding detail. This causal relationship between ideas is further enhanced by conjuncts (*so*), indicating consequence. Discourse markers serve as a minimal response to the interlocutor's previous statement and mark the beginning of the speaker's turn (*yeah*). The passage also illustrates the importance of adverbs for oral elaboration, apparent not only through frequency of use, but also through the variety of functions performed by adverbs. Stance adverbs are the first prominent class in the speaker's description of their habits. These stance adverbs (*obviously* x2, *actually*) express certainty and strengthen the speaker's statements; the use of these adverbs also appears to assume a common ground with the listener – the speaker offers an

explanation or a justification for their choices and by qualifying them as obvious assumes solidarity on the listener's part. Speaker stance is also expressed through the emphatics in the passage (e.g., *quite a bit, very, still, even, just*) or hedges (*quite*). Adverbs like *normally, then* or *anymore* situate the speaker's experience in time. Since this experience is habitual, it is presented entirely in the present tense. Finally, the passage demonstrates several instances of reduced surface form structures common in online production, such as contractions, and structures of generalized discourse such as demonstrative pronouns (*I normally eat **that***) and the pronoun 'it' (*it's (=the diet) **changed a bit***). The single occurrence of the negative feature, noun-noun sequence, is akin to a compound noun denoting a kind of food, as evidenced by the stress on the first element as well as the holistic meaning of the sequence (spring rolls). It is thus most likely processed as a single word and hardly contributes to information density.

Text Sample 1. Author 82, Interview (text id: 179), Dimension 1 score: 2.17

Erm. **Yeah, so ob-obviously it's** changed **quite a bit** during [...]. Erm, **so it's** not **very** healthy **anymore, because** no- we **normally eat**, er, at sunset, and **it's normally** fried food.

Like not **even** in the air-fryer, **it's just** fried food. Erm, samosas and spring rolls, and pastries and stuff like **that**. **And yeah, so I normally eat that** in the evening. **And just because** I've been fasting all day, I **don't actually eat** a lot. **So I get** full er, full **pretty** quick.

Erm, **but it's still quite** an unhealthy option. **And then obviously** during and sunrise, **that's** when we **have** the next meal. **And n-normally** towards then I **have** my fruit, **because** I can't have **it** in the morning **anymore. So I have** my fruit **then. And** erm, I **have** cereal or tea and biscuits. **So it's** like healthy and unhealthy at the same time.

Text Sample 2 is an example of image descriptions ($M = 0.44$; $SD = 0.51$). Due to the nature of the task, the tense of the text is present. Clausal coordination is a means of building on previously expressed ideas and moving forward in the description. The impersonal pronoun 'it', another feature associated with the positive pole of the dimension, serves as a formal subject and introduces the idea units (*it looks like/ it doesn't look like*). Adverbs again play a central role, expressing stance (*actually*) and organizing the discourse (*then*), while conjuncts (*for example*,

so) express the logical relationship between ideas. Among the expressions of stance, hedges (*maybe* x3, *almost*, *a bit of*, *quite*) are particularly prominent in this excerpt and seem to be common in this register generally. In descriptions, the speaker hypothesizes about the nature of the event in the picture, the relationship between the participants of this event, the task the participants are working on – the kind of specifics, in which the speaker cannot have confidence; in Text Sample 2, even emphatics, such as *really*, are used to express doubt (*you can't really see what's going on*). Discourse markers (*okay*) do not occur throughout descriptions, most likely due to their monologic nature in contrast to interviews, but seem to consistently mark the beginnings of descriptions, signaling the speaker's readiness to start the task. Similar to interviews, image descriptions also contain features indicative of the online production of these texts, such as contractions.

Text Sample 2. Author 62, description (text id: 3551), Dimension 1 score: 1.58

Okay. So, er, the first image **is** two women, erm, they **look** like they're good friends. Erm, she's having **a bit of a** laugh with her, **a bit of a** giggle about something. **It looks** like, **maybe**, the, the lady who **is** standing up next to the lady who's sat down **is** showing her something? Erm, **it, it doesn't look** like a formal setting, I wouldn't say, like, **for example**, in a lecture or in a meeting. Erm, **it looks** more like, **maybe**, they're sat down doing some revision together **and** they're having **a bit of a** laugh about something **or, maybe**, they're talking about something informal. **But it does look** like she's showing her something or grabbing something, she's **almost (split aux)** reaching her arm out in front of her. **And then** the other girl **has** her hand on her chest, laughing. Erm, the background, **actually, is quite** plain, you can't **really** see what's going on.

Text Sample 3 illustrates oral elaboration in the register of evaluations ($M = 0.15$; $SD = 0.24$). Evaluations dwell on 'general truths', i.e., positives and negatives of the Aston campus, which warrants the prominent use of the present tense. Speaker stance is central to evaluations, whose communicative goal is to describe the advantages and drawbacks of the university facilities, and in this passage is expressed through frequent emphatics and adverbs of time and manner, which always strengthen the propositions (*very easily*, *temporarily fixed constantly*,

regularly did not work, especially on hotter days). The passage abounds in conjuncts, which serve to organize and reinforce ideas, particularly the author's reasoning. Such reinforcement of reasoning appears to be the function of *again* throughout the passage. Demonstrative pronouns, a feature of inexplicit, vague discourse allowing references to entities known to the participants, are used in the passage to refer back to the issues listed, as the author elaborates on the root causes of the problems or why these problems deserve attention (*This may link to the first point; This is a big issue*). A similar function is performed in the passage by the pronoun 'it' (*this either was not the case or it did not appear to be*).

While there is a focus on oral elaboration, relative scarcity of some positive features may reflect the impact of the written mode. The absence of contractions, for example, a feature of compression necessitated by the time constraints of online production, is one indicator of evaluations being somewhat more literate on the continuum than interviews and image descriptions (the single contraction in this passage is a rare exception in evaluations). The passage also contains two instances of nominal sequences that convey its slightly higher informational focus.

Text Sample 3. Author 25, evaluation (text id: 80), Dimension 1 score: 0.63

Again, this could **very easily** be amended **just** by using paint and making **it** appear less dull. **Secondly**, one thing I found with Aston Campus which was a reoccurring problem **is** the issues with the lifts. **This** may link to the first point due to the buildings being old **but** they **regularly** did not work. **This is** a big issue when you **have** to go to a high level floor, and **especially** on hotter days. When you **are** paying such a large amount of money, **this is** not an issue that should keep reoccurring. **Again, I feel this links** to be poor upkeep of the buildings and them **being (split aux) temporarily** fixed **constantly** rather than having something done that's permanent and will stop the issue. **Lastly**, though **it states** that there **is** twenty-four hour security, I often found that **this either** was not the case **or it** did not appear to be. **Again**, when you **are** paying large amounts of money to attend university, **it is** important to feel safe.

Emails ($M = 0$; $SD = 0.38$), only slightly lower on the dimension than evaluations, form a perfect split between the positive and negative ends of the dimension (Figure 4.1). The

situational characteristics of the register are conducive to both oral elaboration and information transmission, and texts within the register may vary with respect to their more specific situational characteristics, which lead them to prioritize oral expression or information density. The text excerpt presented here (Text Sample 4) does not represent either of the extremes but shows a balance of oral and informational features possible in emails. The writer of this email describes their current interests, experience, and difficulties, which makes the texts present-oriented. Conjunctions (*however*) and other adverbs (*also*) serve to show the logical relations between the propositions, while the causative adverbial clause (*as I haven't yet launched*) explains the reasons for the described difficulty and motivates the questions for the addressee that follow. Generalized references, common for oral communication, are present (pronoun *it*, demonstrative *these*) but minimal: part of the writer's goal in writing this email is to provide the addressee with information which is new for them and pertains solely to the writer's professional life. Vague references in this communicative situation could therefore negatively impact clarity. The email is written by a student to their supervisor, and in view of the nature of the relationship, the goal of the email, and the professional context, the text does not feature overt expressions of stance through emphatics or other adverbs.

On the other hand, information presentation is an important goal of the email, as the writer informs the addressee about their experience and interests before asking for guidance. It becomes apparent that it is the writer's past experience with particular platforms or tasks (*network marketing and forex trading*), the subjects of their interests (*media marketing, media platforms*), and the related aspects of the job the writer is applying to (*time management, day-to-day operations*) that are conveyed through the noun-noun sequences in the text. The email thus combines the goals of presenting this information in a concise and efficient manner, dictated by

the conventions of professional communication, with elaboration necessary to make this information accessible to the reader. It is this balance that is reflected in its equidistant position from the positive and negative poles of the dimension (text score of 0.01) and the combination of their features.

Text Sample 4. Author 89, email (text id: 10172), Dimension 1 score: 0.01

I **am** mid launch for my own and **have** a strong interest in Social Media Marketing with multiple profiles on several social media platforms. In my spare time, I **also learn** about digital marketing from platforms like YouTube or other Digital Marketers. I **am also** a part of a company that **does** Forex Trading and Network Marketing. **However, as I haven't yet launched, I am** finding **it** difficult in applying experience to the job description. **So**, my questions **are** as **follows**:

I **understand** that I cannot say 'learnt digital marketing from YouTube', so how would you suggest I tackle **this**? For the requirement of 'Support in day to day operations', **is it** possible that I can mention 'being responsible for the time management and completion of tasks' pertaining to my students? In regard to my network marketing and forex trading, how could I implement **these**? would I have to prove how many connections I **have made** over the world in relation to other forex traders? Thank you for your help **and I hope** my email **isn't too** long/complicated!

The remaining three registers, essays ($M = -0.24$; $SD = 0.27$), text messages ($M = -0.90$; $SD = 0.66$), and business memos ($M = -1.01$; $SD = 0.79$), are found predominantly in the negative range of the dimension (Figure 4.1). Communication is often set in professional contexts, and the relationship between participants tends to be professional, unequal in status rather than personal (although text messages are in many ways an exception). Features of online oral elaboration, many of which express speaker stance and involvement, are thus uncommon in this situational context. The registers are written; their production circumstances therefore allow opportunities for planning, editing, and revision and do not require features of reduction and compression, such as contractions. At the same time, these allowances of the mode enable the writers to make use of complex nominal structures, which are cognitively demanding (e.g., Coleman & Blumenfeld, 1963; Cheung, 2017) and therefore rare or impossible in the spoken mode. In writing, on the other hand, such phrasal complexity is an important attribute of

information packaging (e.g., Biber et al., 2016; Staples et al., 2016). The efficiency and concision of this kind of information transmission may serve to satisfy the space constraints often imposed by the medium (e.g., word count in university essays; character restrictions in text messages). At the same time, the use of nominal structures, which are harder to process, is justified by the opportunity for careful reading. The text excerpts below demonstrate how information presentation is achieved through noun-noun sequences.

Text Sample 5 is an excerpt from an essay from a psychology class. It is apparent that the subject matter of the essay is specialist, and a substantial amount of shared specialist knowledge is expected of the addressee. This characteristic of the audience, namely the fact that the essays target specialist readers and represent scholarly discourse, as well as opportunities for careful reading compensate for the cognitive load imposed by the complex nominal structures. The text reports the results of an experiment investigating a relationship between constructs. This goal warrants a heavy informational focus, reflected in the variety of noun-noun sequences in the passage. Nominal sequences denote notions central to the goal of the essay and include the phenomena investigated by the study (*eating behavior, eating habits, food consumption, food categories*), the instrument used in the study (*Food Frequency Questionnaire*), the authority that developed the instrument (*Swiss Food Panel*), construct measurements (*neuroticism score; conscientiousness, openness, agreeableness score*), and the results (*study result*). If the pre-modifying nouns of these sequences are analyzed in more detail, we can see that pre-modification conveys a range of semantic relations denoting an institution (*Food Panel*), content (*food categories* = categories of food; *Food Frequency Questionnaire* = questionnaire about food frequency), objective relations, with the pre-modifying noun being the object of the process described in the second noun (*food consumption* = consumption of food), or relations of location

(*study results* = results of the study/ observed in the study), where the second noun is found at the “location” of the pre-modifying noun (Biber et al., 2021, pp. 584-585). Some noun-noun sequences seem to express more nuanced or complex relations: *eating habits* or *eating behavior* denote habits or behavior formed as a result of eating a certain way; *neuroticism scores*, *conscientiousness score*, *openness score*, and *agreeableness score* denote scores that measure their respective construct. It is apparent from this complexity that noun-noun sequences contain information vital to the goal and the topic of the passage while serving as a means of economy and precision. Features of elaboration are rare and are mostly limited to the present tense.

Text Sample 5. Author 27, essay (text id 10862), Dimension 1 score: – 0.85

The purpose of this study **is** to see **is** there a relation between personality and unhealthy eating behavior based on the Big Five Inventory and Swiss Food Panel Food Frequency Questionnaire. Researcher **predicts** neuroticism score will have a positive relationship with unhealthy eating habit; there **is** a negative correlation between conscientiousness, openness, agreeableness score and unhealthy eating habits. Finding from correlation and regression **shows** that there **is** a significant negative relationship of conscientiousness and unhealthy eating habit.

To look at the relationship between personality and eating behaviour, Keller & Siegrist (2015) study result **show** that neuroticism had a significant positive indirect effect on sweet and savoury food consumption, whereas conscientiousness had a significant negative indirect effect on consumption of food categories that **are** indicators of an unbalanced diet. **This shows** that there might be link between personality and food consumption and eating habits.

Another register occurring rather low on the informational end ($M = -0.90$; $SD = 0.66$) is text messages. Text Sample 6 presents an example of a text message with a clear informational focus. Surprisingly, the noun-noun sequences in the text also vary widely in the meaning relations they express, which include purpose (*shopping items*), content (*sandwich bags*), composition (*cranberry juice, beef hamburgers, custard cakes, chocolate biscuits, masala curry*), subjective relations (*Hollander sauce* = Hollander is the brand that makes the sauce), and what may be considered partative (*cross buns* = the image of a cross is on the buns). The individual coinages of the noun-noun sequences ‘*tins beans*’ or ‘*tins sliced peaches*’ in place of

an ‘of’ genitive also appears to be prompted by considerations of efficiency. This example shows that the goal of text messages, often perceived to emulate spoken communication, can in fact be to convey information. Due to the space constraints of the medium, but more likely time constraints for typing, the goal of information transmission is served well by noun-noun combinations.

However, while the place of this register on the dimension is indicative of the informational focus of text messages and the noun-noun sequences even show a variety of expressed meanings, this result should be interpreted with caution due to the extreme shortness of the texts and the subsequent inflation of scores (an issue discussed in Chapter 3). Thus, much lower scores of text messages on the dimension should not be interpreted as indicators of greater information density in this register than that observed in essays. It is also apparent that the nature of the noun phrases serving the goal of information presentation in text messages is quite different from those in essays: unlike essays, text messages focus on concrete, tangible objects rather than complex phenomena. This contrast may be attributed to the difference in setting and subject matter. Text messages represent casual communication between friends or family, who typically interact in a private setting; the assumed background knowledge is not specialist in nature, and the noun-noun sequences denote entities and objects found in such situational contexts. Thus, while noun-noun sequences can indeed be prominent in text messages, a register not typically associated with information presentation, the nature of the nominal sequences reflects the situational characteristics of the register.

Sample 6. Author 72, text message (text id: 20137): Dimension 1 score: – 2.31

Hi **trust** you are ok ,the shopping items **are** as follows, potatoes, bananas ,grapes, ice ceam, 2 cranberry juice, squirty cream, ,2 rice , Hollander sauce, 2 tins sliced peaches, 2xtins beans ,2 jars masala curry, 2 rice, fresh salmon, beef hamburgers, hot cross buns ,digestive biscuits, chocolate biscuits, custard cakes, 6 milk, 6 wine, sandwich bags, many thanks. Xxx

While information presentation may not be typically associated with text messaging, the informational focus of business memos directly relates to its goal of informing the addressee and reporting the results of the search, whose search items are specified by that addressee. The noun-noun sequences in Text Sample 7 serve to present the results of such a search and all involve proper nouns in geographic names or names of the researched facilities. The most common semantic relation between the nouns in the phrases is that of location (e.g., *Helsinki Strand*), but others include time (*evening activity*), and purpose or content (e.g., *souvenir shop*, *History Museum*).

The caveat regarding text length affecting the results applies to business memos as well, many of which are quite short. Nevertheless, noun-noun sequences play a central role in this register, as participants present information meant for a supervisor in a work context in a straightforward and efficient manner.

Text Sample 7. Author 61, business memo (text id 27135), Dimension 1 score: – 2.56

Hotel: Hilton Helsinki Strand - King Guest room with Sea View - 2-4th Sep 2021 = £193 per night
Breakfast Cafe: La Torrefazione, Hakaniemi - 8am-6pm - 3 min walk from hotel City's History Museum: The National Museum of Finland - 11am-6pm - 21 min walk/10 min drive from hotel National dish Restaurant: Kanstan Molja - 10 min drive from hotel Souvenir shop: Kankurin Tupa - 16 min walk/5 min drive from hotel Evening activity: Wallis' Skatta Karaoke Bar - 20 min walk/8 min drive from hotel

4.1.2 Social Group and Individual Variation on Dimension 1

The previous section presented the feature co-occurrence patterns associated with Dimension 1 in the texts of different registers. Register was shown to be a significant predictor of linguistic variation on the dimension, and it was illustrated how registers differ in their place on the dimension – the extent to which they gravitate towards one functional pole or the other. These differences were related to the communicative situation of the registers, namely the relationship between the participants, the amount of shared knowledge and the nature of this knowledge, the

communicative goal, or the affordances and restrictions of the production circumstances. In the same vein, this section plots the social groups of the study along the dimension and discusses the observed patterns.

Figures 4.2-4.3 plot the age and gender groups along Dimension 1 and Tables 4.4-4.5 below present the descriptive statistics (means and standard deviations) for each group. Neither age nor gender is a significant predictor of variation, not explaining any variance on the dimension (Age: $F(4,724) = 0.33, p > 0.003, R^2 = 0\%$; Gender: $F(1,720) = 0.05, p > 0.003, R^2 = 0\%$). As follows from the descriptive statistics in Tables 4.4-4.5, there are no differences between any of the groups ($p > 0.003$). This result suggests that the age groups of the study as well as the gender groups do not differ in their way of navigating the linguistic space between online oral elaboration and information density.

Table 4.4. Age groups on Dimension 1. Descriptive statistics

Age group	Dimension 1	
	<i>M</i>	<i>SD</i>
19	-0.11	1.14
20	-0.02	1.01
21	0	0.98
22	0.05	0.91
23	-0.06	1.02

Figure 4.2. Age groups on Dimension 1

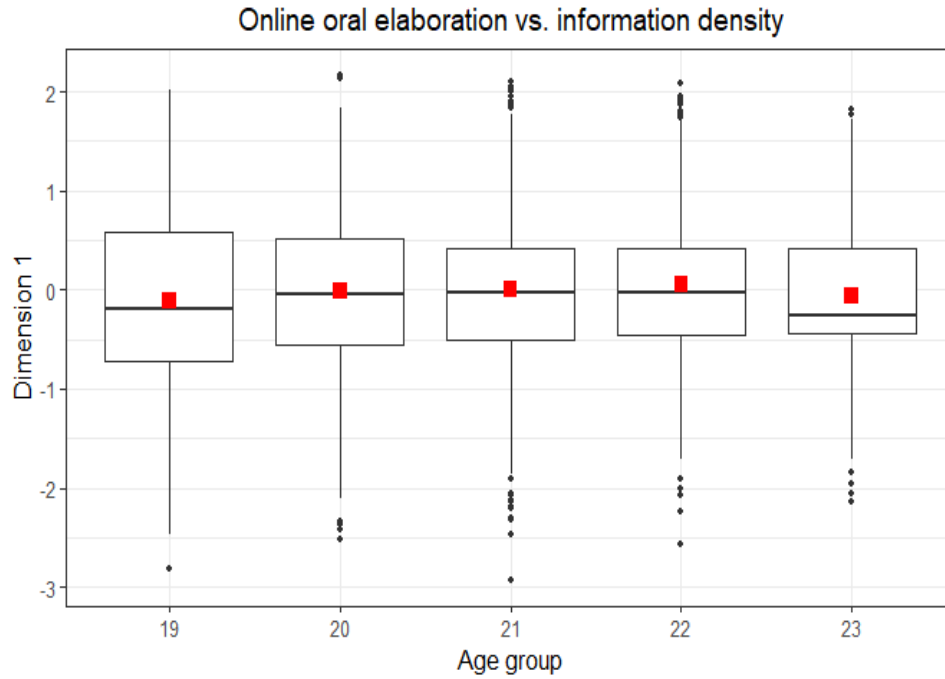
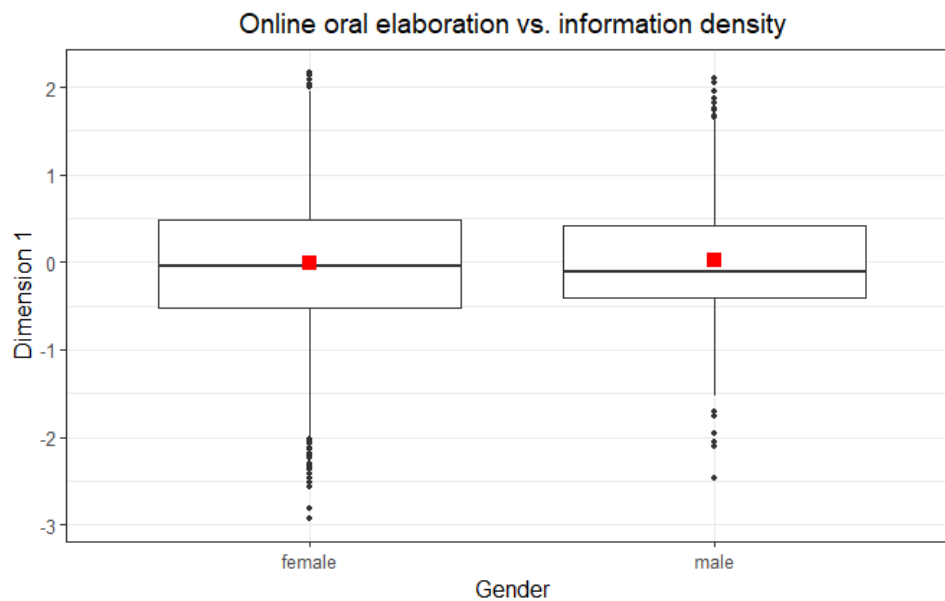


Table 4.5. Gender groups on Dimension 1. Descriptive statistics

Register	Dimension 1	
	<i>M</i>	<i>SD</i>
Males	0.02	0.97
Females	0	0.98

Figure 4.3. Gender groups on Dimension 1



The age groups included in the study range from 19 to 23-year-old undergraduate students, and while the difference in biological age is not sufficient to warrant differences in functional linguistic patterns (i.e., one year is not likely to result in age-grading), the one-year difference between the groups may represent the length of the students' immersion in the academic context. This length of immersion may have an effect on their performance in some academic registers, for which phrasal complexity has been shown to be characteristic, such as essays (e.g., Staples et al., 2016). However, an overall development of increased phrasal complexity across levels appears unlikely in view of the nature of the other registers in the corpus. Essays and emails are the only registers naturally produced in the academic context, in which case instruction or exposure could have an effect on where they fall on the cline (although emails, as was shown above, may be both oral and informational). In the case of the other registers, not academic in nature, there is no reason to anticipate a development in phrasal expression over time. Considering the varied nature of the registers in the corpus, it is thus not surprising that age, or university level, does not predict how the dimension space is navigated.

Likewise, gender is not a significant predictor of variation on the dimension (Figure 4.3; Table 4.5). This result suggests that the identified underlying function is equally characteristic for the speech of men and women and does not distinguish between the groups. Some of the features strongly associated with the dimension have been previously discussed as distinctive for gender groups. Coordination and subordination, for example, have been associated with gendered speech, the former seen as characteristic for women (Jespersen, 1922, cited in Coates, 2004); among the gender-distinctive features are also emphatics and hedges (Eckert & McConnell-Ginet, 2003; Coates, 2004; Brezina, Love & Aijmer, 2018), discourse markers, adverbial stance structures, and the amount of talk (Coates, 2004), all found prevalent in

women's speech. All these features form a strong association with Dimension 1, contributing to online oral elaboration. The results of this study, however, do not confirm the trends identified in investigations of these individual features. That is, while there may be a preference exhibited by a certain gender towards a particular feature, a more holistic approach undertaken by this study, in which several co-occurring features, each performing individual functions, jointly comprise an underlying general linguistic function, does not reveal the same trend. Rather, the underlying function elucidated by the dimension appears to manifest itself the same way in the speech of men and women. This finding, however, does not contradict conclusions of previous sociolinguistic research. The general functional pattern identified by the dimension is not equivalent to any of the individual functions of the features at its core. That is, the individual functions of emphatics, hedges, coordination, subordination, discourse markers, and adverbs all contribute to this general pattern, but the underlying function revealed through their co-occurrence is inherently different from each individual function of the features. As a result, while hedges may indeed be favored by women, leading to the conclusion that women show a tendency to tentative expression, tentativeness is not the function of Dimension 1 and is not equivalent to oral elaborate or informational expression.

Another reason for this outcome could be multifunctionality of linguistic features. Aries (1996) points out multifunctionality as a caveat for gender research relying on specific linguistic features, saying that while a feature may have been previously identified as distinctive, its functions may be different in other contexts. In this study, while a feature may be strongly associated with a dimension due to its particular role in the investigated domain, the reason for that strong association may be one particular function frequently performed by that feature in the domain; however, it may be its other functions that were found important in distinguishing

between men's and women's speech, and they may not be performed in the texts of this corpus and thus may not contribute to the dimension. Subordination, a feature encompassing numerous structural types and functional meanings, may be an example. An important consideration here is that features associated with gender differences were identified mainly in spoken registers (as discussed in Chapter 2), while the registers of this corpus represent very different domains. It was shown by previous research on gender (overviewed in Chapter 2) that literate or informational contexts did not reveal gender differences.

Overall, the lack of significance of gender observed here seems in line with other studies that have previously relied on a combination of features in gender research rather than individual features. The overview in Chapter 2 showed that those studies did not observe the gender differences found in studies based on individual features and instead consistently reported differential patterns heavily dependent on the social roles of the participants and context. These studies have reported a minor role of gender or a lack of differences. Similarly, this study shows that situational differences play a much more prominent role in navigating general communicative patterns than gender. Thus, it may be that while more specific functions of individual features may reflect gender differences (in some contexts), a general pattern like the one captured by the dimension is fundamental to communication and equally characteristic for men and women.

While social group does not contribute to the variation along the dimension, individual differences among the 112 people in the corpus account for 1% of variation ($F(111, 624) = 0.27$, $p > 0.003$) after the other variables have been accounted for. While this result is nonsignificant, it shows that additional variation on the dimension may be due to differences between individual authors.

4.2 DIMENSION 2. ABSTRACT VS. CONCRETE DISCOURSE

The poles of Dimension 2 represent a contrast between abstract and concrete discourse. This contrast is the most clear through the opposition between typical features of abstract discourse, such as nominalizations (0.56) and abstract nouns (0.34) on the one hand, and concrete nouns (–0.45) on the other. Alongside nominalizations and other abstract nouns, the positive pole features a range of other nominal features whose functions entail presenting complex abstract concepts. Such features include ‘of’ genitives (0.48), expressing a range of meanings and relations extending beyond possession (e.g., partative, defining, objective; see Biber et al., 1999, p. 303) and prepositional phrases, specifically nouns followed by prepositions other than ‘of’ (0.47). Abstract concepts conveyed through nouns are further elaborated through ‘to’ (0.41) and ‘that’ (0.35) noun complement clauses, seen as markers of “abstract nominalized stance towards a proposition” and “assessment of certainty of the proposition” (Biber et al., 1999, p.642). While noun complement clauses disclose the content of the head noun, relative clauses modify it, adding information about the head noun or serving as an “idea unit expansion” (Chafe, 1982, 1985, cited in Biber, 1988, p. 234). Non-finite ‘ed’ relative clauses (0.30) and pied-piping (0.28) serve this purpose.

While the abstract phenomena themselves are denoted through these nominal features and they are further developed through the noun complement and relative clauses, speaker evaluation of the proposition (Winter, 1982, cited in Biber, 1988, p. 231), objective characterization of the complex notions discussed or personal stance, is expressed through adjectival structures – adjective complement clauses (0.35) and predicative adjectives (0.41) both appearing on the positive pole.

Another feature commonly characterized as a marker of detached, impersonal, and abstract discourse – passive voice – is represented on the dimension as agentless (0.34) and

agentive (0.35) passives. The abstract nature of passive structures is attributed to the fact that the agent either does not occupy the central place in the discourse and is not connected to it as closely as the patient or deemed unimportant and is absent from the structure (Biber, 1988, p. 228). Finally, type-token ratio (0.51), the ratio of different words in a text to the total number of words, a measure of lexical diversity, appears to suggest that this kind of abstract discourse presenting complex concepts and ideas is highly specific, planned, and delivered with attention to detail.

It is important to note at this point that this dimension structure, namely the abundance of abstract, nominal, literate features, does not suggest that the abstract pole of Dimension 2 overlaps with the negative pole of Dimension 1. The identified dimensions are, of course, interrelated, and it is indeed likely that texts that prioritize information density over oral online elaboration (i.e., get negative scores on Dimension 1) are also abstract (i.e., positive on Dimension 2). As indicated by factor correlations, however, the dimensions capture unique variance (specifically, $r = -0.12$ between Dimensions 1 and 2), and are thus independent of each other.

The concrete end of the dimension is comprised by concrete nouns (-0.45) and indefinite pronouns (-0.40). While the opposition formed by concrete nouns with the abstract nominal features discussed above is intuitive, the role of indefinite pronouns, a feature usually contributing to the construction of vague, general discourse, is perhaps less obvious with relation to concreteness. Its place in concrete discourse will become apparent from the analysis of texts, particularly in some registers, such as image descriptions, presented below. Specifically, it will be illustrated that the feature does serve the function of referencing concrete objects, but the speaker expresses uncertainty about them: for example, these objects may not be in clear view

and thus are not clearly discernible to the speaker who is describing the visual. These seemingly vague, indefinite features therefore do serve the function of referring to concrete objects.

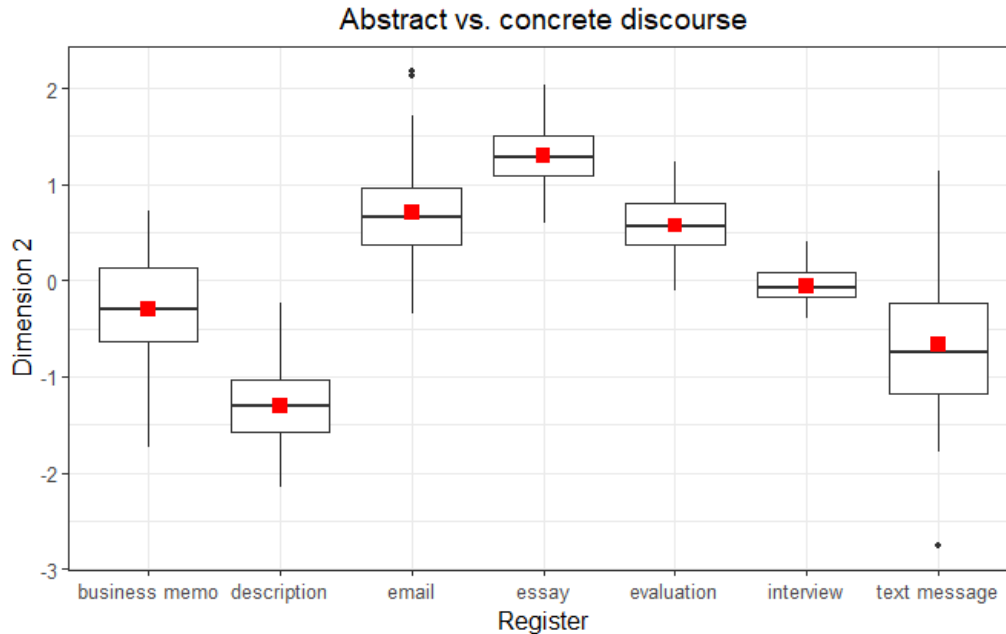
4.2.1 Registers along Dimension 2

Again, register is a significant predictor on the dimension, accounting for 79% of variance ($F(6, 729) = 428.1, p < 0.001$), with all register pairs significantly different from each other with the exception of emails and evaluations ($p > 0.003$, Cohen's $d < 0.5$). Figure 4.4 shows the registers on Dimension 2 and Table 4.6 presents the descriptive statistics (means and standard deviations) for each.

Table 4.6. Registers on Dimension 2. Descriptive statistics

Register	Dimension 2	
	<i>M</i>	<i>SD</i>
Business memos	-0.29	0.53
Description	-1.31	0.40
Emails	0.70	0.48
Essays	1.30	0.29
Evaluations	0.57	0.32
Interviews	-0.05	0.17
Text messages	-0.67	0.69

Figure 4.4. Registers on Dimension 2



Essays ($M = 1.30$; $SD = 0.29$) are the register with the highest positive score on the dimension. Text Sample 8, an essay written for a psychology class, illustrates the features associated with abstract discourse. The abstract concepts in the passage are conveyed through a variety of nominal features, such as numerous nominalizations (e.g., *definition*, *recognition*, *expressor*, *happiness*, etc.) and other abstract nouns (e.g., *menace*, *display*). Additionally, the passage elucidates the complex relationship between many of these abstract phenomena, denoted through these features, and other similarly complex notions. These relationships are conveyed through the prepositional phrases following many of these abstract nouns as postnominal modifiers. Thus, an extended noun phrase with multiple levels of embedding like *the **definition of correct recognition of facial displays of emotion*** contains three consecutive postnominal modifiers, each conveying the essence of the previous abstract concept through another. Even in this short excerpt, such multiple phrasal embedding is frequent and pervasive. The relationships conveyed this way are not limited to those expressed through genitive of-phrases, with other

prepositional phrases (nouns followed by prepositions other than ‘of’ – here, e.g., *uncertainty in its definition*) forming an equally strong association with the dimension. Such complex relations are also conveyed through the interplay of abstract nouns with other features important for the dimension: the predicative adjective *able* followed by an adjective complement clause in the following example – *the importance of being **able to** correctly decipher facial displays of emotions* – express an abstract idea the same way the nominalization *ability* would in its place.

Another feature of abstract discourse on the positive end of the dimension and prominent in the passage is passive voice. Agentless passive structures in the passage serve to mitigate the force of some statements by removing the agent, thus lending the author an opportunity to show less commitment to or less ownership of the proposition than the active counterpart would convey (*it can be said, can be interpreted*). Examples of agentive passives like *can only **be defined by the expressor*** or *has **been correctly interpreted by the intended receiver*** show that the agents in the by-phrase, expressed through nominalizations, are not known or concrete. Rather, *expressor* and *receiver* are undefined abstract referents in the described process of facial recognition. The passage thus illustrates how a number of positive Dimension 2 features co-construct abstract discourse in essays.

Text Sample 8. Author 86, essay (text id: 10976), Dimension 2 score: 2.04

The **definition of correct recognition of facial displays of emotion** greatly varies across a **multitude of circumstances and environments**. The heavily **disputed** topic holds **uncertainty in its definition of ‘correct recognition’**, however, it can **be said** that this can only **be defined by the expressor** and if their affective signal has **been** correctly **interpreted by the intended receiver** (Fridlund, 2017). This is due to the **nature of facial displays of emotion** and how one **display**, such as the smile, can **be interpreted in ways like happiness or menace**. This ultimately exemplifies the **importance of being able to** correctly decipher facial **displays of emotion** as it enables humans to convey **intentions and feelings**.

Still in the positive range of the dimension but generally less abstract are emails ($M = 0.70$; $SD = 0.48$) and evaluations ($M = 0.57$; $SD = 0.32$). Being much lower on the dimension

than the essay discussed above, Text Samples 9 and 10, an email and an excerpt from an evaluation, demonstrate a combination of features of abstract discourse and references to concrete objects. Text Sample 9 is an email written by a student to their supervisor, in which the student proposes a dissertation topic and requests feedback. The concrete references in this email relate to concrete experiences shared by the participants, such as references to recent communication (*week, weekend*) or denote concrete steps, elements, or products of the academic process (*questions, paragraph, works, aims, topic*) and proper names of particular prominent figures in the field (*Thomas Hobbes, John Locke, John Stewart Mill*), which constitute the shared professional knowledge between the participants. However, abstract notions still prevail in this email as a result of its topic and goal – namely, to propose a dissertation topic. In view of this, abstract phenomena are of central importance to this text, as they are used to state the topic of the proposed research (*liberalism, theories such as the hedonistic principle, the harm principle and utilitarianism, concept of ‘absolute freedom’ in Liberalism, origins of the Theory, Islam, Islamic perspective*). The passive structures in the excerpt accommodate the discourse need to omit the agents in the discussion of this abstract subject matter. It is the concepts central to these theories (e.g. *concept of freedom*) that receive prominence in this discussion rather than the agents who acted on them. In the clause *the concept of freedom is portrayed or even limited in the Quran*, for example, it is how the concept is presented in the religious text that is of interest to the writer.

Text Sample 9. Author 8, email (text id: 9465), Dimension 2 score: 1.29

Hi , I hope you had a great weekend! I just wrote the below paragraph as last week you suggested I write a paragraph on the aims and objectives of my **dissertation**. I would highly appreciate any **feedback**! In this **dissertation** I intend to explore the following question: ‘ . I will analyse the **concept of ‘absolute freedom’ in Liberalism**, looking at the **origins of the Theory** and where the **concepts** primitively arose from. I will scrutinize the works and **theories of philosophers** around the topic of liberalism over

the years such as Thomas Hobbes, John Locke and John Stewart Mill. Looking at **theories** such as the hedonistic **principle**, the **harm principle** and **utilitarianism** I intend to investigate the first **principles of Liberalism**. Comparatively I will analyse how the **concept of freedom is portrayed** or even **limited** in the Quran. I will then investigate the **theory of absolute freedom in liberalism** through an Islamic **perspective**. After researching the two **concepts of freedom** I will then be **able to** analyse whether they **are juxtaposed**. Is this **dimension of liberalism** applicable to **Islam** or can the two not go hand in hand? Of course in my **dissertation proposal** I will be referencing etc. Thank you,

The evaluation in Text Sample 10 abounds in abstract notions, although of a different nature. Through both abstract nominal features (e.g., *atmosphere*, *environment*, *scenery*, *greenery*) and predicative adjectives with adjective complements (*facilities are available*; *Although it is small, it is beautiful*; *being on campus was amazing*; *it was nice to enjoy time*), the author evaluates availability of resources, the general impression the place produces on visitors, and characteristics of the ambience. This evaluation is not possible, however, without references to visitors and students, who form these impressions, specific locations, or moments in time, achieved through the concrete nouns in the passage. This balance is prompted by the communicative goals of the register – in order to evaluate the place and draw conclusions regarding the resources and facilities, the authors need to provide illustrations and support their statements by descriptions of specific occasions, interactions, people, and objects. This balance of the two goals explains the intermediate position of the register on the dimension – in the positive abstract range, but considerably closer to the concrete range than essays and even emails.

Text Sample 10. Author 5, evaluation (id 122), Dimension 2 score: 1.24

Aston University has many excellent **facilities** that are **available** to all **staff** and most importantly all students. However, there are three that have really stood out to me whilst studying at Aston. Firstly, the campus offers students and passers-by a serene **atmosphere** and **environment**. Although it is **small**, it is **beautiful** as it has an amazing **scenery** and **greenery**. Being on campus in first year was **amazing** - and at summer, it was **nice to enjoy time with** friends on the grass, whilst almost **being attacked by** the geese - it is definitely their **territory**.

Surprisingly, interviews ($M = -0.05$; $SD = 0.17$), in which the participants discuss their eating habits and daily routines and which thus may be expected to gravitate towards concreteness, cluster close around zero, with a rather low standard deviation. This suggests that interviews tend to systematically represent a combination of concrete and abstract features rather than be exclusively comprised of references to concrete entities. Text Sample 11 illustrates this balance and reveals that, among other references (e.g., *food*, *time*), the most notable instances of abstract discourse in such texts are expressions of the participants' stance – in this text, evaluation of the content of the proposition in the form of predicative adjectives (*basic*, *nice*, *perfect*). Such expressions of stance serve as explanations or justifications of the choices and habits described in these interviews and in that contribute to their overall communicative goal.

Text Sample 11. Author 79, interview (text id: 173): Dimension 2 score: 0.25

Oh, my favourite restaurant. Oh **God**, this is going to sound really **basic** but I do, I feel like you can't go wrong with . Like whenever I'm **hungry**, I just go to and I pretty much get the same thing all the **time**. So, for starters, I get haloumi, erm, haloumi sticks and it's like with a sweet salsa dip, I love that, I could eat that all day. And I, I like to get one for myself cause I don't like sharing that, cause there's only five sticks and that's enough for me. Erm, and then for my mains, I used to love getting a sunset burger but now I've moved on to, I get four chicken thighs and I get the Peri-tamer sauce on it cause it's like a mix between hot with a tinge of sweet. It's kind of **nice**, it's like, I like mixing savoury and sweet **food** so that's **perfect** for me.

Further down the abstract/ concrete continuum are business memos ($M = -0.29$; $SD = 0.53$), occurring mostly in the negative range. Concreteness appears to be a natural feature of this register, whose goal is to present concrete recommendations for a specific business trip based on previously specified search criteria. However, while concreteness does prevail, business memos exhibit considerable variation across their texts and extend into the positive range (Figure 4.4), which suggests that the register is quite varied with respect to abstract and concrete features. Text Sample 12 is an example of a business memo prioritizing concrete suggestions of specific places and activities. Yet, some of the recommended activities entail references to abstract

concepts, such as art forms (*opera, photography*) or a style of art (*baroque*). Other abstract nouns denote concepts related to a visitor's experience of the suggested places (*closing time, entry per person*). Additionally, abstract discourse is manifested through ellipted passive structures, used to indicate locations (*based in*) or the nature of the deal being suggested (*with breakfast included*). Elements of concrete and abstract discourse thus jointly contribute to the communicative purpose of the register, providing specific information related to the itinerary as well as tourist experiences of art and culture.

Text Sample 12. Author 98, business memo (text id: 27175), Dimension 2 score: – 1.07

Hotel Indigo; **Based** in Bulevardi (26), Helsinki Finland. Includes 1 queen bed with breakfast included. Price is 123.13 euros per night. Brooklyn cafe; 9AM start **based** in Fredrikinkatu 19,00120. The finnish Mes of photography - museum based in Helsinki kapelitehdas. 12 euro entry per person. Restaurant; Ravintola, The Glass. **Based** in Porontakymtis Kamp Galleria, K1 Shopping Mall; Kamp, Galleria. 8PM closing **time**. **Based** in Pohjoisesplandia (33). Evening activity; Opera concert. FIBO- Finish barouque orchestra. 6PM closing **time**. **Based** in The German Church, Helsinki.

Text messages ($M = -0.67$; $SD = 0.69$) exhibit substantial variation on the dimension, although occur primarily in the concrete range. Text Sample 13 is an example of a text message describing an event witnessed by the author. The concrete nouns refer to the participants of this event (*police, ambulance, first year (student)*) and objects relevant to the story. The main goal of a message like this is to convey information with attention to detail as it is this detail that is of interest to the addressee and the focus of the writer's observations. This goal and the topic – an occurrence in the writer's daily life – are conducive to concrete expression rather than abstraction.

Text Sample 13. Author 63, text message (text id 19066), Dimension 2 score: – 1.34

Yeah I'm **fine** I had my airpods in on **noise cancelling** so I didn't know what was going on. But the police and the ambulance came and took her away. They were a first year living downstairs so don't know them xx

Image descriptions ($M = -1.31$; $SD = 0.40$) are entirely in the negative range of the dimension. Text Sample 14 is a typical example, in which the description relies heavily on concrete references to people and objects in the visual. As noted earlier, indefinite pronouns serve this function alongside concrete nouns in this register when the described objects are unclear to the author or hidden from view (*looking down at something*). Alternatively, indefinite pronouns serve to express the author's assumption or a guess regarding something that is not in the visual but relates to the observed scene (*seem to be laughing at something*). Some of such assumptions are made about the emotions experienced by the people in the visuals (*be very **happy***) or characteristics, such as race and ethnicity (***Asian***), expressed through predicative adjectives – a positive feature. Such references, however, are rare overall (especially with regard to emotions and feelings), as there is usually no evidence supporting such statements. The texts are therefore dominated by concrete features, the primary linguistic means necessitated by the task.

Text Sample 14. Author 87, image description (text id: 3586), Dimension 2 score: – 2.15

Okay. So, for image number one, there are two girls, erm, they look to be students and they seem to be laughing at something, looks like they're laughing at a, a laptop, maybe a phone. Erm, one of the, erm, girls is a Black girl with black hair and the other one is a white girl with, erm, brown hair. And yeah, they seem to be very **happy**. They're, it looks like they're laughing at something and looking down at something. Now, on number two, this has two boys, erm, si-sitting together. It looks like they're in, maybe, a café or a student room and they're, they're looking, that one of, there's one of the boys is **Asian** and the other one is a white boy. And the Asian boy's looking at the screen, the M-MacBook, erm, screen and the white boy is looking at the Asian boy.

4.2.2 Social Group and Individual Variation on Dimension 2

Age and gender are nonsignificant on Dimension 2, age not explaining any variance ($F(4, 724) = 0.25$, $p > 0.003$; $R^2 = 0\%$) and gender ($F(1, 720) = 0.42$, $p > 0.003$) explaining 0.05% ($R^2 = 0.05\%$). Figures 4.5-4.6 and Tables 4.7-4.8 show that all the age and gender groups appear to

‘behave’ very similarly on the dimension, using its full range, again extending from the abstract to the concrete pole and not showing a tendency towards concrete or abstract discourse.

Table 4.7. Age groups on Dimension 2. Descriptive statistics

Age group	Dimension 2	
	<i>M</i>	<i>SD</i>
19	-0.02	1.02
20	-0.06	0.98
21	0.03	0.95
22	0.02	0.92
23	0	0.95

Figure 4.5. Age groups on Dimension 2

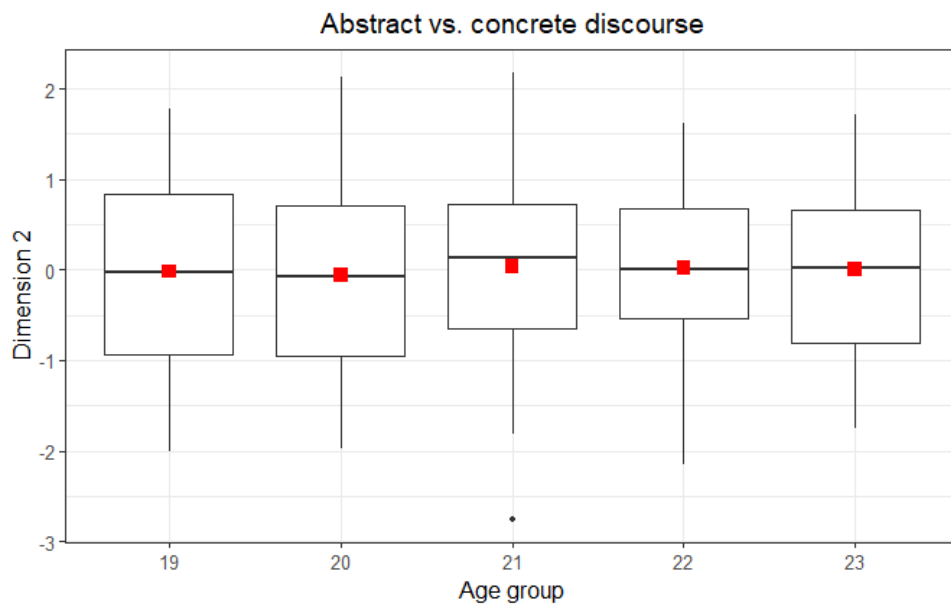
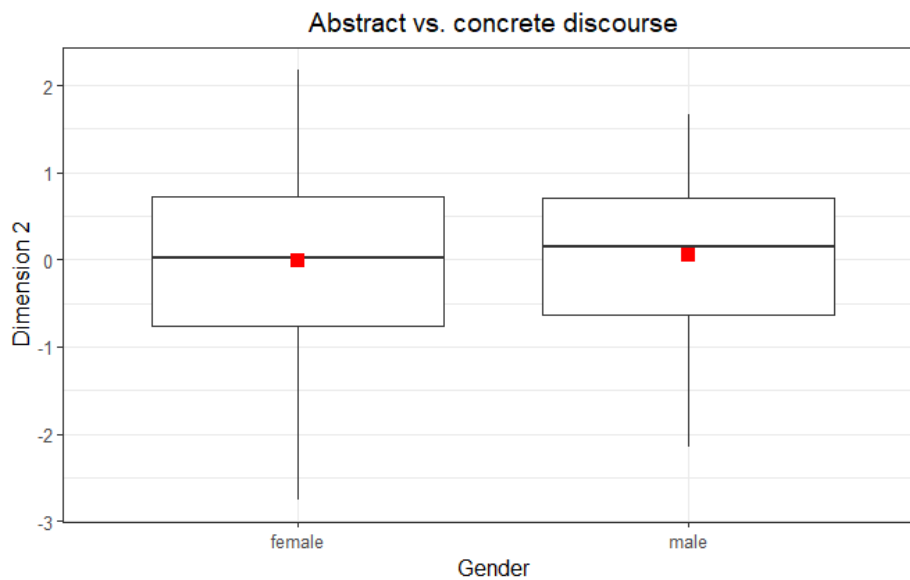


Table 4.8. Gender groups on Dimension 2. Descriptive statistics

Register	Dimension 2	
	<i>M</i>	<i>SD</i>
Males	0.05	0.9
Females	-0.01	0.96

Figure 4.6. Gender groups on Dimension 2



Again, certain features on the dimension, could be viewed as structures associated with different social dynamics and power. Impersonal structures, for example, including features such as passive voice (Brown, 1998, cited in Coates, 2004, p. 107), have been hypothesized to be associated with women’s speech (Lakoff, 1975) and later with ‘powerless’ language (O’Barr & Atkins, 1980) – typically a reflection of social class rather than gender, although frequently occurring in female speech. On the other hand, Lamb (1991) investigates passive voice and complex nominal features, such as nominalizations, as language indicative of vague expression and accountability avoidance in gender discourse and finds them to be more typical of male writers.

However, while passive voice as well as nominalizations and other nominal features are important contributors to the abstract pole of Dimension 2, it is patterns of co-occurrence with the other features on the dimension that comprise abstract discourse in this study. As discussed in Section 4.1.2, this difference between results observed by studies of individual features and the present study relying on a macroscopic analytical method again suggests that while the functions

of individual features may have been more or less prominent in the speech of a particular social group and contributed to a certain effect (e.g., vague expression, lack of confidence, etc.), these specific functions are not equivalent to the functional pattern captured by the dimension. Specifically, in the case of the abstract discourse described by this study, major contributors to the dimension alongside passive voice and nominalizations are other complex nominal structures such as different kinds of prepositional phrases, further enhanced by various types of noun complement clauses and predicative adjectives. This general pattern is not restricted to the individual function of passive constructions or individual functions of nominalizations and the kind of ‘vagueness’ they contribute. Thus, the lack of significance of gender on the dimension is again informative: unlike a specific feature contributing to this general pattern, which has been shown to reflect gender differences, this general functional pattern is used to an equal extent and in similar ways by men and women. Another explanation for a lack of significance of gender appears to be the nature of the domain. In many investigations, including the ones cited here, differential language use is examined in high-stakes situations, such as performance in the role of a courtroom witness (O’Barr & Atkins, 1980), and/or on sensitive topics, such as journal publications on abuse (Lamb, 1991). The results observed in those domains, which may trigger a certain linguistic behavior, are likely to differ from the language used by university students in natural university settings or research settings on neutral general topics.

Considering the pattern revealed by the dimension, i.e., the continuum between abstract and concrete expression, it is also intuitive that age groups do not differ in their use of its linguistic resources, but use the full scope of the dimension; the distinction between abstract and concrete language has not been found or anticipated by previous research on age.

Finally, individual characteristics ($F(111, 624) = 0.23, p > 0.003; R^2 = 0.3\%$) have minimal effect on the dimension, adding 0.3% of variance to that explained by register ($R^2 = 79\%$) and gender ($R^2 = 0.05\%$).

4.3 DIMENSION 3. OTHERS-ORIENTED DESCRIPTIVE VS. SELF-ORIENTED OR INTERACTIVE DISCOURSE

The features that offer the most immediate insight into the nature of Dimension 3 are 3rd person pronouns (0.52) on the positive end and 1st (-0.61) and 2nd (-0.37) person pronouns on the negative pole. It is this opposition between 3rd person references on the one hand and self-references (as follows from the large negative loading of 1st person pronouns) on the other that defines Dimension 3, whose positive pole represents discourse centered around exophoric (mainly animate) referents and whose negative pole constitutes largely self-oriented speech. However, while the role of 2nd person pronouns is much less substantial as can be seen from a much lower factor loading, the interpretation of the dimension's negative end would be incomplete without accounting for the interactivity added by 2nd person references if they are present. 2nd person pronouns are interspersed with self-references in some texts, while they are completely absent in others, where self-references prevail. The part of the label proposed for the negative pole (self-oriented or interactive discourse) reflects this variability.

Other features on the positive end add substance to its 3rd person orientation. It becomes apparent from the high weight of the progressive aspect (0.53) that it is the current actions of 3rd person referents that are the focus of this discourse. Additional characterizations of these referents are offered through 'ing' relative clauses (0.34) as well as 'wh' relative clauses on subject position (0.28), and their physical or mental attributes are conveyed through attributive adjectives (0.38). Existential 'there' (0.36) serves to construct statements regarding these

referents and contributes to descriptions of settings, events, or objects in the discourse surrounding them. Finally, an interesting feature of others-oriented discourse in the corpus is its tentative nature, the uncertainty speakers express when discussing other people or entities, hedging their statements through the verbs ‘seem’ and ‘appear’ (0.52). The feature is strongly associated with the dimension and is consistently observed in the texts.

On the other end of this cline, 1st and 2nd person pronouns are complemented by the verb ‘have’ (−0.38), predictive modal verbs (−0.31), and time adverbials (−0.31). The functions of these features in self-oriented or interactive texts seem to vary, as is illustrated further. In some texts, these features appear in accounts of the speakers’ habitual states and actions, their routine and lifestyle. In others, they serve to refer to expected, imagined, or hypothesized future prospects and possibilities in the speaker’s life. This variety is primarily due to the diverse functions of predictive modals, not distinguished by the tagger and thus merged under that superordinate label. As a result, meanings such as denoting a future intention, a hypothetical possibility in the past or present, a habitual past action, and others are expressed through the same feature. Regardless of the exact meaning of the modals, however, this discourse tends to revolve around the speaker or the 2nd person referent, the feature thus contributing to self-focus or interactivity.

4.3.1 Registers along Dimension 3

Register explains 77% of variance on Dimension 3 ($F(6, 729) = 372.4, p < 0.001$), and all register pairs are significantly different from each other. Figure 4.7 plots the registers along the dimension, and Table 4.9 shows the descriptive statistics (means and standard deviations).

Table 4.9. Registers on Dimension 3. Descriptive statistics

Register	Dimension 3	
	<i>M</i>	<i>SD</i>
Business memos	-0.19	0.44
Description	1.57	0.41
Emails	-0.43	0.45
Essays	0.68	0.25
Evaluations	0.11	0.25
Interviews	-0.67	0.18
Text messages	-1.02	0.91

Figure 4.7. Registers on Dimension 3

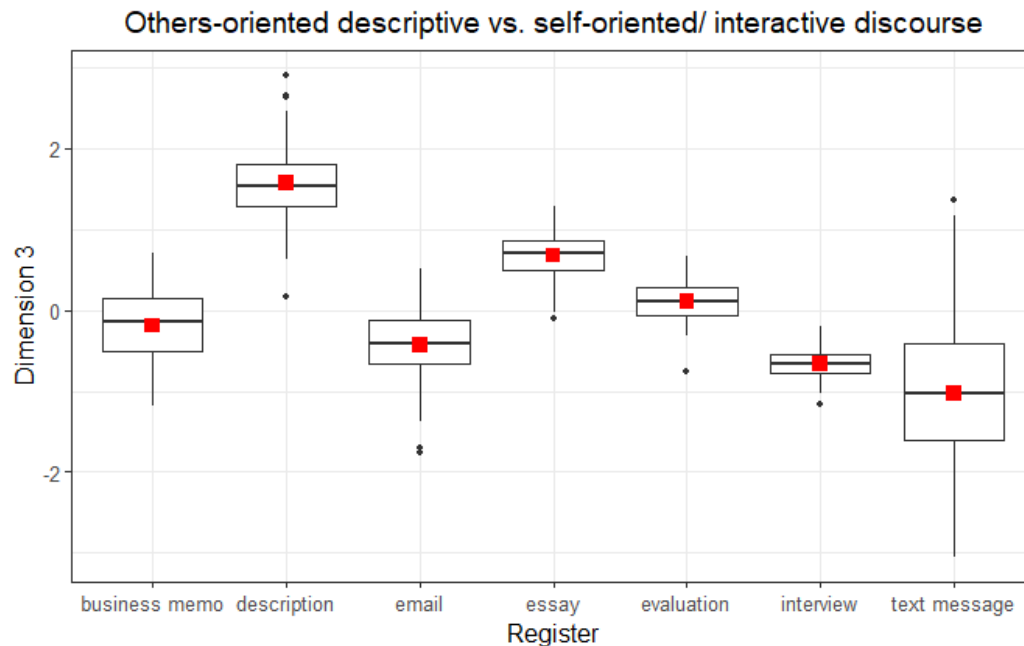


Image descriptions ($M = 1.57$; $SD = 0.41$) have the highest positive scores on the dimension and are found entirely in the positive range. This result is intuitive in view of the nature of the task – as the participants describe the people in the image and focus on their actions, there is nothing in their speech related to their own experience. Their speech is monologic and thus non-interactive, which also tends to result in a lack of 2nd person references. Text Sample 15 demonstrates this focus on others, where 3rd person pronouns refer to the characters in the image, attributive adjectives provide traits of their physical appearance,

progressive verbs convey their actions at the moment captured by the visual, and ‘wh’ relative clauses on subject position add information about the described individuals at the center of this discourse (*And another lady, who’s sitting down...*). Several instances of the verb ‘seem’ throughout the text serve the function of hedging as the author adds detail to the description inferable from the visual (*she seems to be a lot happier; they seem to be in some kind of joke*), but not factual or based on evidence. The participants were presented with the same three images for the task of description, all of which depicted scenes from university life, with students as their focus. Others-oriented discourse is thus central to the register.

Text Sample 15. Author 70, image description (text id: 3560), Dimension 3 score: 2.90

So **there is** one lady in image one that’s **standing** over another one. **She’s wearing** a **black and white striped** shirt with **black** jeans and a belt. **Her** hair’s almost like an **ombré brownish** blonde. **She seems to be smiling**. And another lady, **who is sitting down, she’s wearing** a denim jacket with a **black** top. She’s got braids and **she seems to be a lot more happier** than the other one. **They seem to be** in some kind of joke or **laughing** about a situation. And the area that **they** are in looks like **they’re** at work. In image two, it looks like two students, possibly **revising, they’re sitting** in a chair with **green** polka dots on there. It’s like a **big** sofa chair. One has a MacBook that **he seems to be showing** the other one. **He’s wearing** a **red and black chequered** shirt with a **white** top underneath. **He has black** hair, **he seems to be smiling** next to the other one **who is wearing** an also like a **black blue** shirt with really **huge blue** roses on there, with a c- a text on the shirt. **His** ears are pierced, **he’s** got a **black** earring. **He seems to also be enjoying** the session, what looks like a revision session, that **they’re having**; all of this at a desk and the **green** couch.

Essays ($M = 0.68$; $SD = 0.25$) are much lower on the cline from others-oriented to self-oriented discourse, but are also found predominantly in its positive range, with exceptions of two texts. Expectedly, the academic subject matter of essays and their goal of researching, synthesizing, and discussing questions of scientific nature does not lend itself to interactivity or self-focus. However, others-orientedness may appear just as unlikely in academic essays, which do not typically aim to describe others or narrate events they participate in. Analysis of high-scoring texts on the dimension revealed that others-oriented discourse in essays tends to serve the

purpose of contributing to the construction of a larger argument or making a larger statement. That is, in contrast to image descriptions, where others-oriented discourse is the essence of the register and is in itself the ultimate product of the task, in essays it is a constituent within a more complex textual structure. Text Sample 16 illustrates the functions of others-oriented discourse integrated into the texture of an essay on the topic of gender fluidity and its perceptions. The text first offers general discussions of particular gender groups, which are the subject of analysis. Among the positive features on the dimension, attributive adjectives are the main linguistic tool comprising this discussion and describing the 3rd person referents, i.e., the gender groups, as well as the social thought and theories surrounding the topic (*non-binary people, non-binary individuals, higher power, higher respect, human lives, previous positive symbolism, general consensus, progressive society*). The essay then provides a report of an interview conducted by the author with a research participant and a detailed qualitative analysis of the statements made by this participant. 3rd person references in this part of the text become concrete (*she, her true opinions, Anna was herself*). The report also centers around the respondent's actions during the interview, conveyed through progressive structures (ellipted progressive *when explaining, when repeating*). In both the general discussion of the phenomenon of gender and the report of the interview, the author's conclusions are rendered tentative by the use of 'seem/ appear'; animate entities as well as inanimate objects which are the focus of attention are enhanced through 'wh' relative clauses on subject position (*people who identify...; the word "definitely" which has strong intentions; her true opinion which may have then been altered*). The content of the interview is thus a major portion of the essay and serves the goal of illustrating the construction of social identity through the lens of an individual experience. Narrating this experience or retelling the content of the interview is thus not among the essay's ultimate goals. Rather, others-

oriented discourse is prominent in essays like this due to their major contribution into the overall argument of the piece.

Text Sample 16. Author 71, essay (text id: 10943), Dimension 3 score: 1.21

It **seems** to suggest that **non-binary** people have had an association with a **higher** power **which** had led **them** to think that **they** are more than male or female. As gender governs **human** lives, perhaps **non-binary** individuals are likened to being more than human. This comment is significant as it regards **non-binary** people in a **higher** respect than people **who** identify with just one gender. [...]

Despite the **previous positive** symbolism, Anna goes on to express the need for another category for **non-binary** people. **She** initially used the word “definitely”, **which** has **strong** intentions, when **explaining** that this should be the case. This was retracted when **repeating** the end of the sentence suggesting that this was misspoke. This could indicate that this was **her true** opinion **which** may have then been altered in order to **appear** more socially acceptable and to fit in with the **general** consensus of a **progressive** society. This signifies that society as a whole may not be as complicit in ‘othering’ behaviour as Anna was **herself**.

Evaluations ($M = 0.11$; $SD = 0.25$) incorporate features of both others-oriented descriptive and self-oriented discourse. Texts in this register aim to describe the facilities of the Aston campus that the students like or dislike, express opinions, and provide reasoning. This goal is impossible to fulfil without references to these facilities and students and staff who use them or manage them. At the same time, however, these descriptions are subjective and are based on the authors’ past or present first-hand experiences of the place, which gives rise to expression of stance and personal anecdotes used as evidence. Text Sample 17 illustrates this balance. Frequent 3rd person pronouns and ‘wh’ relative clauses on subject position refer to students and their needs served by the Aston facilities (*students **who** want to converse together; students **who** just want to do **their own** thing in silence*). The positive pole of the dimension in this register is characterized by inanimate as well as animate references and is thus viewed as both others-oriented and descriptive. In Text Sample 17, some ‘wh’ relative clauses on subject position add information about the described facilities (*aid to students, **which** includes asking about **general** health/concerns*), and the attributive adjectives all serve to describe Aston’s

resources (*sufficient library; different floors for different purposes*, etc.) The most salient among the negative features in the passage is the 1st person pronoun ‘I’, which allows the author to share (attitudinal) stance toward the described facilities (*I like; I think it is even better; I believe it is also very efficient*). Others-oriented descriptive discourse and self-orientation thus converge in the register of evaluations as authors describe the services offered by the university to its students and, as the recipients of these services, also evaluate them.

Text Sample 17. Author 55, evaluation (text id: 133), Dimension 3 score: 0.16

One thing I like about the Aston University Campus (AUC) is the library. Every university needs a **sufficient** library that makes students feel like **they** can survive all-nighters and complete **their** coursework without fail. The Aston library provides this comfort and I think it is even better as it has **different** floors for **different** purposes. The **different** floors allow students **who** want to converse together to enjoy **their** time at the library, it also supports the students **who** just want to do **their own** thing in silence and focus. Having options makes going to the library an **easier** task. A second thing I like about the AUC is the student union and its facilities. Again, the student union is a university necessity and provides a lot of aid to students, **which** includes asking about **general** health/concerns or even going for more **specific** issues. I believe it is also very efficient to have so close to the university buildings and **there** are several facilities to use since they created the **new** building - **which** also provides a space to relax and commune with other students.

Business memos ($M = -0.19$; $SD = 0.44$), still lower on Dimension 3, exhibit a more pronounced tendency toward self-focus and interactivity and hardly rely on references to entities other than the addressor or the addressee. Interactivity as well as self-focus may seem unusual in a text written to a supervisor in a work context with the goal is to inform. However, Text Sample 18 shows that since some business memos assume a working relationship between the addressor and the addressee, they also adopt a more personal and direct tone. 1st person in the passage is used to refer to the work done by the author or future services (*I have planned; I will make sure*), while 2nd person pronouns occurring with predictive modals address the reader directly as the itinerary of the trip is presented (*you will drive into; You will be staying*). Time adverbials, another negative feature, mark the specific point in the trip itinerary. Several positive features,

attributive adjectives, 'wh' and 'ing' relative clauses, add descriptive detail to the passage, but the emphasis on interactivity, reflecting the collaborative relationship between the addressor and the addressee, conveyed through 1st and 2nd person pronouns is considerably more prominent.

Text Sample 18. Author 14, business memo (text id: 27083), Dimension 3 score: – 0.20

Hi Sarah, I have planned your trip to Helsinki in September and have gathered as much information as possible. Please see below. if you need more or would like anything else incorporated please let me know. FLIGHTS: Unfortunately, the **best** place to fly to Helsinki was from Manchester, however, I will make sure that you have a taxi **waiting** for you from Birmingham to take you there and back. it will be an **early** flight: 05:55am on 2nd September 2021 from Manchester airport with Luftstansa with one stop in Amsterdam and you will drive into Helsinki airport at 13:10pm Your flight back is an evening one from 18:15pm to 21:35pm -to give you most of that day to enjoy Helsinki and what is has to offer - on the 4th September 2021 Total cost: £138.29 - optional to add carriage luggage HOTEL You will be staying in a **lovely** waterfront Hotel called Clarion Hotel Helsinki **which** is only a mile from the city centre.

The prevalence of self-focus and interactivity in emails ($M = -0.43$; $SD = 0.45$) can be attributed to communicative goals of the texts. The emails in the corpus are written by students to their professors, academic advisors, or other university officials to ask questions, request help, assistance, or clarification, share an idea and request feedback, ask for guidance or advice, or explain personal circumstances and request an academic accommodation. All of these goals warrant focus on the writer and the addressor. The email in Text Sample 19, for example, exemplifies self-orientedness and interactivity as the author updates the addressee on the progress they have made towards a task (*I wanted to update you; I have put my study up, I've transcribed two*, etc.), informs the addressee about the circumstances that will affect this progress in the future (*my exams open...it'll be hard for me...*), asks questions and involves the addressee by expressing a wish to discuss a course of action together (*I don't know whether that means we should keep going; I would like to go through **them** with you*). 1st and 2nd person pronouns, predictive modals, and time adverbials, the latter contributing concreteness to the

progress report (*I've done 4 of my friend interviews now*), are the negative features that construct this discourse. The text does contain positive features, most notably reference to prospective study participants and conducted interviews, but these references do not detract from the heavy interactive focus of the email.

Text sample 19. Author 55, email (text id: 7938), Dimension 3 score: -0.87

Hi , Hope you had lovely holidays! I just wanted to update you, I've done 4 of my friend interviews now, aside from one 50 minute one, the rest are between 1-1.5hrs, so I don't know whether that means we should keep going or try to do the analysis first. But that being said, I've put my study up on anyway but no one has signed up! If they don't it's fine I can still send my advert to friends I guess but I just wanted to ask what to do from here (aside from start analysing). I would like to go through **them** with you (I've transcribed two) but my exams open this and I'm kind of swamped with revision so it'll be hard for me but hopefully when it's over we can do that, but I can still conduct interviews if I need to. Hope to hear from you soon, **Final Year**

Interviews ($M = -0.67$; $SD = 0.18$) occur entirely in the negative range of the dimension and exhibit minimal internal variation as shown by the lowest standard deviation among the registers (Table 4.9; Figure 4.7). This definite tendency towards self-oriented discourse is to be expected in view of the personal topic of the interviews as well as their communicative goal. The topic and the purpose of the interviews being held constant (with the same questions asked) may explain the relative lack of variation across texts. Text Sample 20 demonstrates how 1st person pronouns, predictive modals denoting habitual actions, and the verb 'have' construct discourse centered entirely around the speaker.

Text Sample 20. Author 31, interview (text id: 93), Dimension 3 score: -1.17

Okay. Erm, so in a **typical** day, erm, I don't normally eat breakfast because I don't wake up, erm, early enough to eat breakfast. So, I would go straight to lunch. So, for lunch, I'd always make something quite quick, erm, because I just don't have the time to **be making** a long lunch, erm, because I, obviously, I want to do the rest of my day, I don't want to be in the kitchen for ages. So, I'll just put something in the oven, normally. Or I'd have like a bowl of oats or, erm, oatmeal or something like that. Erm, if I put something in the oven, normally it's like a pizza or like, erm, anything like waffles or like burger, erm, just something quick. Erm, and I'll probably have like salad with it as well or and like a sauce or something like that. Erm, and after that, erm, I do snack a lot. Erm, I

don't like sweets, I don't like chocolates, I don't like anything like that, I don't eat anything like that.

The last register on the negative end, although showing considerable variation along the dimension and extending into its positive range, is text messages ($M = -1.02$; $SD = 0.91$). As was noted in the discussion of text messages previously, low negative dimension scores in the case of this register should not suggest that its self- or interactive focus exceeds that of business memos, emails, and interviews, as the presence of negative features even with rather low rates of occurrence in a short text inflates the score. Nevertheless, self-focus and interactivity are clearly characteristic for the register comprised mostly of personal exchanges between friends and family. Text Sample 21 is an example of such interactive discourse in text messages.

Text Sample 21. Author 107, text message (text id: 11723), Dimension 3 score: -3.05

i'll let you know once i have my schedule for **next** week cos i've got an assignment due this xx

4.3.2 Social Group and Individual Variation on Dimension 3

Age ($F(4,724) = 0.06$, $p > 0.003$, $R^2 = 0\%$) and gender ($F(1,720) = 0.003$, $p > 0.003$, $R^2 = 0\%$) do not predict variation on Dimension 3, with groups showing the same pattern of a wide spread of scores along the dimension as seen on Dimensions 1 and 2. Figures 4.8-4.9 and Tables 4.10-4.11 show that none of the groups show a preference for others-oriented descriptive or self-oriented interactive discourse, but rather all of them employ the full linguistic spectrum from texts that describe and discuss the actions of others to those that reveal self or present an interaction.

Table 4.10. Age groups on Dimension 3. Descriptive statistics

Age group	Dimension 3	
	<i>M</i>	<i>SD</i>
19	0	1.12
20	-0.01	0.98
21	-0.01	0.93
22	0.01	1.01
23	0.05	0.91

Figure 4.8. Age groups on Dimension 3

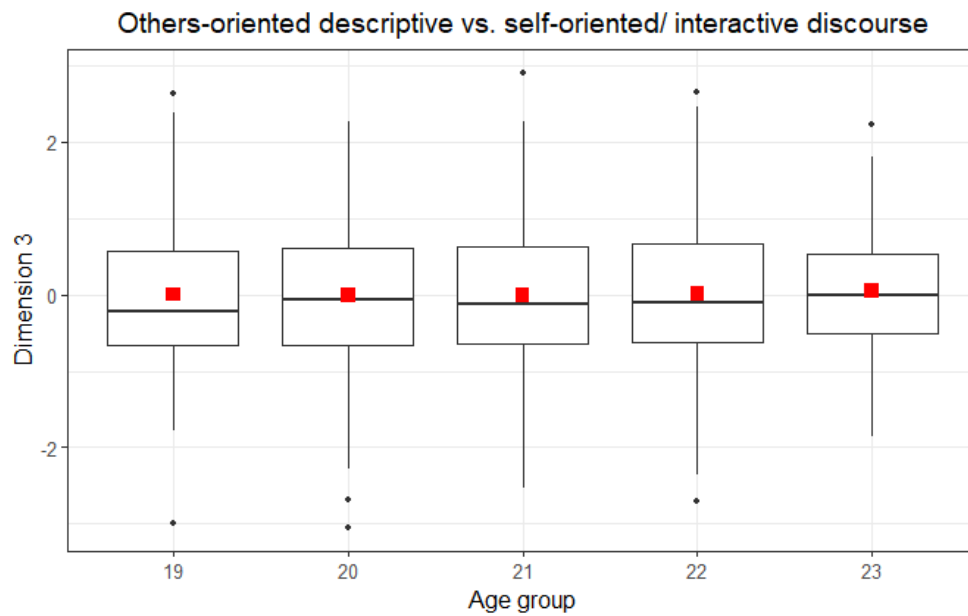


Table 4.11. Gender groups on Dimension 3. Descriptive statistics

Register	Dimension 3	
	<i>M</i>	<i>SD</i>
Males	-0.01	0.98
Females	0	0.97

Figure 4.9. Gender groups on Dimension 3



1st, 2nd, and 3rd personal pronouns are among the highest-loading features on the dimension poles and among the features at the center of many discussions around gender. The divide between 1st and 2nd personal pronouns on the one hand and 3rd person pronouns on the other has been discussed in light of the general need for gender attribution or ‘genderizing’ referents (Eckert & McConnell-Ginet, 2003). Among other examples in this discussion, Eckert and McConnell-Ginet cite an author who intentionally did not reveal the gender of their main character, never referring to them in the 3rd person and systematically resorting to only gender-neutral options (1st person narrative and 2nd person addresses by other characters). This deliberate choice is counterposed with the tendency of gender attribution fundamental to the English language, as the 3rd person singular pronouns are inherently gendered. While Dimension 3 does capture this divide between gender-neutral and gendered pronouns on the opposing poles (although it is important to note that the 3rd person pronouns include the plural form, which is gender-neutral), the functional pattern at its core is not limited to this specific function of the pronominal features. As was illustrated in the above discussion of the dimension structure and

the texts exemplifying others-oriented language, description and the focus on others' actions are important component of this discourse, which often serves larger purposes fundamental to the texts, such as proving a point, constructing an argument, providing a concrete example of an abstract phenomenon, or integrating others' views in a discussion of a controversial topic.

This functional pattern thus also differs from another common phenomenon investigated in connection with gender which involves “non-present others” – namely, gossip as a building block of social cohesion, seen as more typical of, although not limited to, women (Coates, 2004, p. 104). While the features comprising the positive pole do focus on actions and descriptive detail, there are no (or there may be minimal) 3rd person references in the corpus that perform this function. The general function captured by the dimension therefore again extends far beyond the individual features in its structure. Similarly, the negative pole, which marks self-focus or interactivity, represents a more complex functional configuration (e.g., enhanced by adverbials and modal verbs) than that of 1st and 2nd pronouns. Variation along this cline and the use of the full spectrum of its linguistic resources again appears to be equally characteristic for all the social groups of the study.

Individual differences ($F(111, 624) = 0.25, p > 0.003, R^2 = 0.5\%$) do not account for much variance on the dimension, contributing an additional 0.5% to the 77% explained by register.

4.4 DIMENSION 4. EVIDENCE-BASED STANCE

Dimension 4 is composed of a wide variety of verbal features of stance on its positive end and no negative features. These include ‘that’ verb complement clauses controlled by stance verbs (0.76), public verbs (0.56), ‘that’ deletion controlled by public verbs (0.48), ‘to’ verb complement clauses controlled by stance verbs (0.44), private verbs (0.38), verb complement clauses controlled by public verbs (0.32), ‘that’ deletion controlled by private verbs (0.28), and

verb complement clauses controlled by private verbs (0.26). The semantic classes of verbs represented on the dimension, namely private and public verbs, the former expressing thoughts, feelings, emotions, attitudes, judgments, conclusions, and other mental processes and intellectual states (Biber, 1988, p. 242) and the latter serving to announce the speaker's position publicly, both contribute to the expression of stance. Complement clauses controlled by these verbs, such as 'that' verb complements, include two constituents – the verb expressing the speaker's stance and the proposition towards which stance is directed (Biber et al., 1999, pp. 969-972). These clauses thus serve as a means of expressing evaluation of the proposition contained in the clause (Winter, 1982, cited in Biber, 1988, p. 231). The same function is served by the complement clauses controlled by these verbs with an omitted complementizer (i.e., 'that' deletion). Biber (2006) distinguishes between stance structures directly attributed to the speaker (or the holder of the expressed belief, opinion, etc.) and structures in which stance does not have an explicit attribution and is therefore impersonal. Various types of complement clauses, such as the ones associated with this dimension, represent the first category as they are always preceded by the subject, whose stance is expressed. Modal verbs of possibility (0.28), on the other hand, belong to the second category as the speaker may not be explicitly stated. The feature is less strongly associated with the dimension, but does contribute to its overall functional pattern and can be viewed as speaker evaluation of the proposition: according to Chafe (1985, cited in Biber, 1988, p. 241), for example, possibility modals are evidential structures that "mark reliability."

A feature not immediately associated with stance but loading on the dimension is the past tense (0.37). Its function becomes clear when patterns of its cooccurrence with the other features on the dimension are examined in context, where apart from expressing a past intellectual state or speech act, the past tense is used to recount events that often serve as evidence supporting the

propositions of the stance clauses. The underlying function served by all the features of stance in combination with past tense has therefore been interpreted as evidence-based stance, which gives the dimension its label.

4.4.1 Registers along Dimension 4

Register accounts for 38% of variance on the dimension ($F(6,729) = 69.86, p < 0.001$), all the register pairs significantly different from each other with the exception of emails and descriptions ($p > 0.003$, Cohen's $d < 0.5$). Figure 4.10 shows the registers on Dimension 4, and Table 4.12 contains the descriptive statistics for each.

Table 4.12. Registers on Dimension 4. Descriptive statistics

Register	Dimension 4	
	<i>M</i>	<i>SD</i>
Business memos	-0.85	0.45
Descriptions	0.28	0.58
Emails	0.39	0.65
Essays	-0.02	0.43
Evaluations	-0.28	0.37
Interviews	-0.59	0.26
Text messages	0.95	1.63

Figure 4.10. Registers on Dimension 4

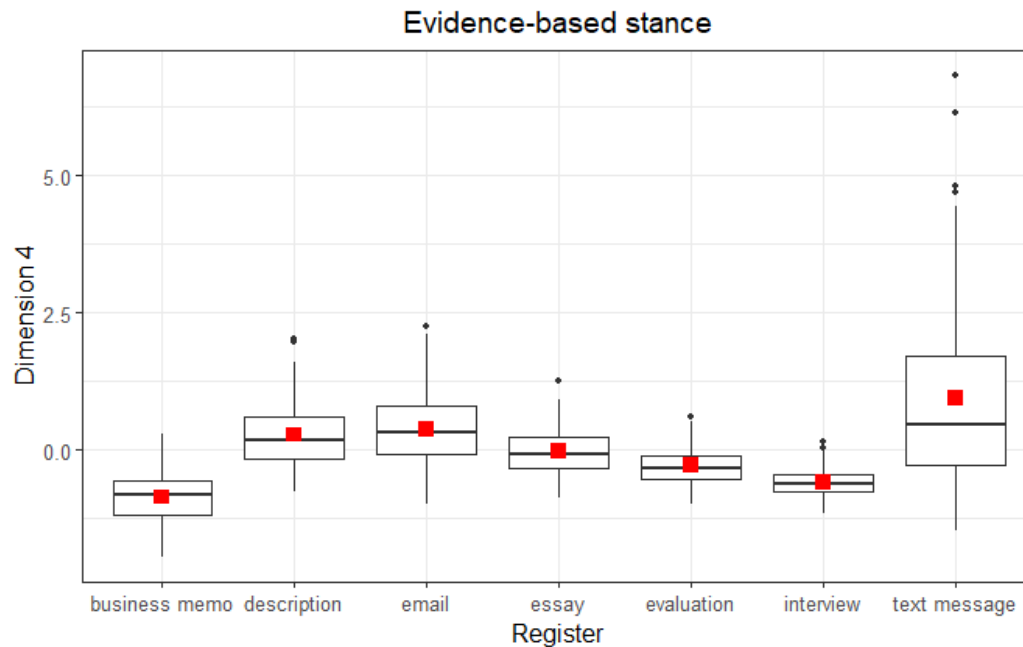


Figure 4.10 shows that text messages ($M = 1.63$; $SD = 0.95$) extend the most into the positive range of the dimension. The tendency to express stance may be attributed to the nature of the exchanges typically occurring through text messages and the relationship between the interactants, who may share their thoughts and observations about the events of their lives or report others' speech, both goals achieved through the structures comprising the dimension. Text Sample 22 is an example of such an exchange, in which the student tells a friend or a relative about the interaction with a professor and its outcome. In this interaction, both the student and the professor express their stance regarding the student's progress in the course. This short text makes use of two stance structures – clauses controlled by a public verb with and without a complementizer, integrating the speech of the student and the professor. Since the text describes a past interaction, both reporting verbs are in the past tense. Three positive features associated with the dimension thus jointly contribute to the task of reporting this exchange, in which the expression of stance is an important goal for both participants.

Text Sample 22. Author 49, text message (text id: 17279), Dimension 4 score: 6.83

but i **said (that del)** i was aiming for top marks and he **said that** the improvements would go a long way to being top marks

Emails ($M = 0.39$; $SD = 0.65$) are another register, in which the expression of stance is quite prominent, as students share their thought process about the assignments, topics, and other class-related matters with their professors or advisors. However, the general goals of emails in most cases extend beyond simply expressing thoughts or reporting others' position on a topic. Rather, stance tends to contribute to the fulfillment of a larger purpose, and emails with many quite different overarching purposes incorporate expression of stance. Text Sample 23, for example, aims to resolve an issue with an electronic submission and ensure that a later submission will be possible. To this end, the author describes their experience with the system, where past tense verbs denote the actions meant to serve as evidence of the author's efforts, and accompanies this description with their thought process (*I **thought that** I would be able to; I really **hope that** I am somehow able to*). Private verbs are also used to invoke the addressee's response (*if you **wish**; please let me **know***) and their anticipated stance on the matter. The possibility modals in the passage in combination with stance verbs (*it **said (that del)** I could complete it; let me **know** how I **can** resolve this issue*) refer to a possibility of a resolution and in that appear to contribute to the author's stance as well.

Text Sample 23. Author 112, email (text id: 4755), Dimension 4 score: 2.25

Dear ,

I recently **had** to complete a quiz for " " () (second year). However, I **did** have an issue because when I **clicked** to complete it on it **was** no longer available. I **thought that** I would be able to complete it because the last time that I **clicked** on the test it **said (that del)** I **could** complete it by Second week " " but it is gone. I really **hope that** I am somehow able to work something out as I need 40% on each test in order to pass this module. I also have a Screen shot of where it **says** " " therefore I can present it to you if you **wish**. Please let me **know** how I **can** resolve this issue. Thank you. Kind regards, () second year student.

Expression of stance in image descriptions ($M = 0.28$; $SD = 0.58$) tends to signal the speaker's uncertainty as they make assumptions or conclusions regarding the content of the visuals, with private verbs being particularly common in texts. Supporting these statements through past references is not a feature of this register as all the statements are based on the described visual, and the task does not call for references to past experiences. Text Sample 24 exemplifies such expressions of uncertainty through private verbs followed by a complement clause with 'that' deletion, a 'that' complement clause with the complementizer present, and a 'to' complement clause, all three verbs used to hedge the propositions. While this hedging reflects the communicative situation of the task performed by the speakers, who cannot be certain about the events and people depicted in the visuals, features associated with the dimension appear to also reflect different situational characteristics – those of the fictional reality of the visual. In these cases, the features of the dimension refer to the speech and mental states of the characters shown by the visual (*pointed out, speaking, talking, reciprocating*). Thus, as the speaker placed in one situational context produces a text describing another, the features associated with the dimension appear to be a response to both the actual situation of the task and the reality being described.

Text Sample 24. Author 88, image description (text id: 3587), Dimension 4 score: 2.03

And it **appears that** one student has **shown** another student something that they may have missed beforehand. And this picture des-, **shows** the amazement of the student who missed the, erm, who missed what the other student has just **pointed out** to them. And erm, now both students are laughing at the mistake. Er, for picture two, it **appears to** be two students, again, working together. Erm, one student **seems to** have been **talking** and the other one **seems to** be a bit befuddled by what the initial student has just made. Erm. They **seem to** be in a, a group based context, erm. Yeah, I **think (that del)** that's it for that one. Er, for picture three, erm, it's a picture of two students who **seem to** be both engaged in listening to someone who is **speaking** right now. Erm, it **appears that** the person who is **speaking** has just made a joke and both students are **reciprocating** by laughing at said joke.

As texts produced with the goal of researching a topic, presenting scientific facts, and objectively discussing existing perspectives, essays ($M = -0.02$; $SD = 0.43$) gravitate towards the negative pole, with fewer features of stance. However, the subject matter of essays often entails discussions of views expressed by prominent scholars, historical figures, or political leaders. Thus, when Dimension 4 features are frequent, they often contribute to their expression of stance and the interpretation of their views by the readers and the general public. Text Sample 25 is an example, in which several private verbs function to express impersonal stance of unspecified entities when the researched phenomenon is discussed (*when **examining** sociological roots and religion; **understanding** how the field has been defined; **can be seen to** have more of a positivist outlook; this **infers that** it is in fact faith, which allows societies to develop*). Additionally, private (***viewed, considered, regarded***) and public verbs (***states, claims, claiming***) convey expert views and hypotheses. Past tense in the essay occurs as these expert claims are cited (***viewed, considered, regarded***), thus serving to integrate evidence to the author's own statements (*the perceptions of classical theorists are crucial in **understanding** how the field has been defined for many years*). The register of essays thus tends to rely more on the expert stance of the cited sources rather than the personal stance of the author.

Text Sample 25. Author 8, essay (text id: 10956), Dimension 4 score: -0.08

When **examining** sociological roots and religion, the perceptions of classical theorists are crucial in **understanding** how the field has been defined for many years. One of the earliest classical thinkers in this region, Emile Durkheim, **viewed** religion as a 'social force' (Press and Swatos, 1998). He **saw** religion as an almost social glue which **reinforced** society. [...] In his book 'The Division Of Labour In Society', Durkheim **states that** religion is simply 'something social' (Durkheim, 2013). This aligns with James Beckford outlook, who **claims that** religion is constructed (Beckford, 2003). These two individuals simply **considered** religion **to** be a product of society. Durkheim **can be seen to** have more of a positivist outlook on religion, **claiming (that del)** it **was** an expression of 'social cohesion' (Durkheim, 2013). [...] This **infers that** it is in fact faith, which allows societies to develop, prosper and thrive. However, unlike Durkheim,

Beckford **regarded** religion as a tool which is used by several agents, for instance, in areas such as education, law and health (Beckford, 2003).

Expression of stance may appear to be crucial to the register of evaluations ($M = -0.28$; $SD = 0.37$) as it directly reflects the main communicative goal of the register. Indeed, as students describe the facilities of Aston, they often express their attitudes and impressions, as can be seen in Text Sample 26, where private verbs are used to express the author's positive impressions (*like, enjoy*) and possibility modals (*can then go onto playing; can come and relax; can feel at ease*) describe the opportunities enjoyed by the students. The authors sometimes report on recent changes, innovations, and improvements, thus offering reasons for their likes or dislikes and using the past tense to integrate evidence (*Aston recently **changed** the Students Union*). However, while stance is definitely present in the register, Text Sample 26 also shows that the clausal structures, strongly associated with the dimension, are not common. The goals of the register do not call for reported speech, and listing the strong and weak points of the campus facilities does not seem to require integration of complex idea units. As a result, it is mainly attitudinal stance expressed through private verbs, possibility modals, and occasional references to past events supporting statements with evidence that account for the relatively low place of the register on the dimension.

Text Sample 26. Author 53, evaluation (text id: 129), Dimension 4 score: -0.52

One thing I **like** about Aston Campus and its facilities is the sports that Aston University provides. There is a swimming pool, gym, badminton and many more sports. This is very good as it **can** allow all sports enthusiasts to **enjoy** their sports and to also keep active. From these many sports Aston offers you **can** then go onto playing nationally or professionally at a sport you are good with. With sports like badminton and tennis you have the chance to hire rackets which is very beneficial if you do not have the equipment but would love to play. The second thing I **like** about Aston Campus and its facilities is the Student Union. The Student Union is a safe place where students **can** come and relax, **enjoy** some peace and quiet and also catch up with their friends. Aston recently **changed** the Students Union into a new building where there is now a bar, restaurant, pool tables, games consoles and giant bean bags where students **can feel** at ease.

Similar to evaluations, in the case of interviews ($M = -0.59$; $SD = 0.26$), it may appear that stance is to be expected in the register, as speakers answer questions about an aspect of their personal life. It becomes clear, however, that the stance expressions prominent in interviews differ markedly from the stance comprising the dimension. Text Sample 27 lacks the clausal features of stance associated with the dimension and only features occasional modal verbs to express possibility of a certain choice and one private verb (*you know*) in a structure aiming to involve the listener rather than express stance. However, the passage contains several adverbs of stance, expressing certainty, such as the frequent use of *obviously*, or intensifying the propositions, such as *really* used throughout the turn. These adverbial features of stance, not associated with the dimension, indicate that a different kind of stance expressions is prevalent in the register of interviews. Similar to the register of evaluations, the scarcity of clausal structures may reflect the goals of the register, where presenting and integrating complex ideas, whether the authors' own or others', is less common. The nature of the shared knowledge with the interviewer, general knowledge about a relatable topic, on the other hand, may make qualifying statements with adverbs a natural choice (*obviously* implies that something is obvious to both). Supporting such statements with past evidence is typically unnecessary, as the interview questions concern the present rather than past habits and do not require proof, which results in a scarcity of past tense in the register.

Text Sample 27. Author 45, interview (text id: 115), Dimension 4 score: -0.70

Okay. Erm, so normally, every single day I have around two full meals. So, that would be breakfast and dinner. So, for breakfast, I normally have either baked oats, erm, and that will be with different toppings so, for example, be like golden syrup, white chocolate. And then I **can** have like a cup of tea or a hot chocolate with it. Other times, I **might** be having scrambled eggs, toast and, erm, tea. I don't really do too much breakfast because I normally just kind of make it to the bus because I'm rushing for work. Erm, and then, o- obviously, I'm working most of the time during the day so I don't really have lunch often. So, like I come back home, that's when, you **know**, I start preparing my dinner. Erm,

obviously, kind of snacking throughout, erm, normally just salt and vinegar crisps, nothing too exciting. But yeah, I start preparing my dinner.

Finally, stance is rare in business memos ($M = -0.85$; $SD = 0.45$) – the feature of the register that may be attributed to the unequal relative status of the addressor and the addressee and the informational focus. As business memos are in essence a task completed by an employee for their employer (although in a contrived context) and contain recommendations for the employer's trip, expression of the employee's personal stance regarding the trip appears out of place. Text Sample 28 is an extreme example of a business memo that presents the researched suggestions in a list format, with no interaction with the addressee or reference to the author's work. It is not surprising that such responses to the task do not contain expressions of stance, especially the clausal idea integration structures associated with the dimension.

Text Sample 28. Author 112, business memo (text id 27080), Dimension 4 score: – 1.95

Hotel- 'Arctic DeLuxe.' Early bird rate- 378 euros for 2 nights. Cafe- Address- Toolonlahdenkatu 2, 00100, Helsinki Helsinki City Museum-Address- Aleksanterinkatu 16, 00170 (opens 11am). Restaurant Konstan Moija- Address- Hietalahdenkatu 14, 00180. Souvenir shop- Annensoppi- Address- Fredrikinkatu 68, 00100. Live music- Storyville- Address- Museokatu 8, 00100.

4.4.2 Social Group and Individual Variation on Dimension 4

Age ($F(4, 724) = 0.45, p > 0.003, R^2 = 0.1\%$), gender ($F(1, 720) = 0.13, p > 0.003, R^2 = 0\%$), or individual differences ($F(111, 624) = 0.66, p > 0.003, R^2 = 0.5\%$) are again nonsignificant predictors of variation and add minimally to the amount of variance explained by register ($R^2 = 38\%$). Figures 4.11-4.12 and Tables 4.13-4.14 show that the social groups of the study again display the same tendency observed on the other dimensions: there is no marked preference for one pole or the other, but rather all social groups occupy the entire linguistic space of the dimension, with texts rich in expression of evidence-based stance as well as texts lacking in it.

These results suggest that clausal speaker-attributed evidence-based stance is not a linguistic phenomenon developed over the course of students' university careers or a phenomenon that distinguishes between genders.

Table 4.13. Age groups on Dimension 4. Descriptive statistics

Age group	Dimension 4	
	<i>M</i>	<i>SD</i>
19	0.01	1.11
20	-0.08	0.81
21	0.04	1.04
22	0.01	0.93
23	0.04	1.22

Figure 4.11. Age groups on Dimension 4

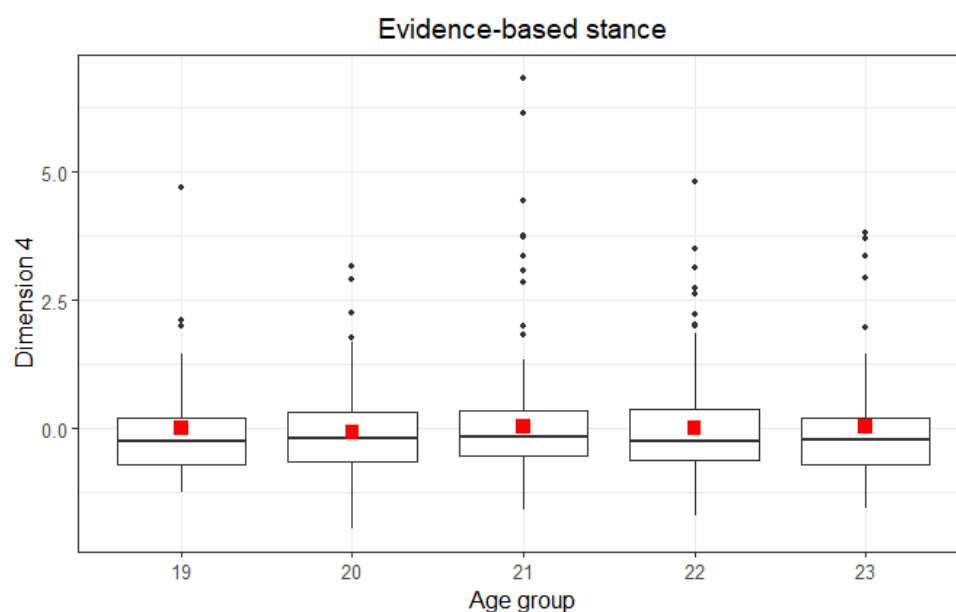
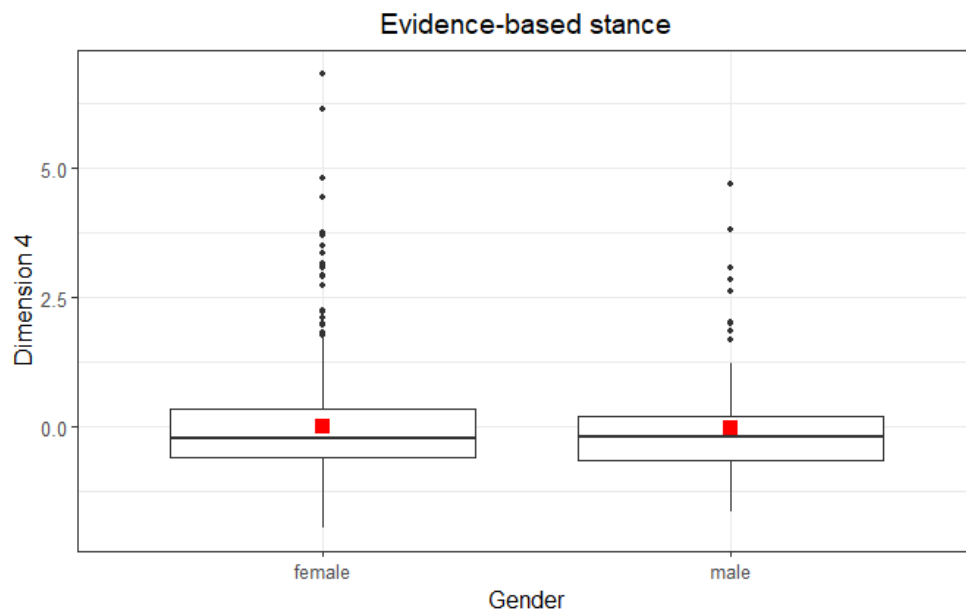


Table 4.14. Gender groups on Dimension 4. Descriptive statistics

Register	Dimension 4	
	<i>M</i>	<i>SD</i>
Males	-0.02	0.98
Females	0.01	0.97

Figure 4.12. Gender groups on Dimension 4



In the case of gender, stance has been the subject of research, as it tends to be seen as an expression of confidence and thus be associated with stereotypical perceptions of social gender positioning (Precht, 2008). However, the most common expressions of stance tied to gender rhetoric seem to be adverbial structures, such as hedges or emphatics, expletives and other markers of affect, attitudinal adjectives, negation, or personal pronouns (e.g., Eckert & McConnell-Ginet, 2003; Coates, 2004; Precht, 2008; Brezina, Love & Aijmer, 2018; Durmus & Cardie, 2018; Li & Li, 2022), and features not immediately associated with stance, such as sentence length (e.g., Mulac & Lundell, 1994). The features comprising Dimension 4 do not appear to have been examined with relation to gender. Some relevant studies include Grob, Meyers and Schuh (1997), who investigate mental verbs used by men and women, although they do not observe a difference, and Precht (2008), who investigates mental verbs and mental verbs followed by ‘that’ complements and observes significant differences between gender groups in the verbs used. As discussed in sections 4.1.2, 4.2.2, and 4.3.2, however, the functional pattern identified by the dimension represents a complex interaction of the features at its base and is not

equivalent to the function of each. Thus, the result observed by this study, namely the lack of difference between the gender groups on the dimension, does not discount the findings of studies that focused on individual features. Rather, the present study identifies a general functional pattern in discourse that, unlike more specific functions of individual features, appears to be fundamental to both men's and women's communication.

4.5 CONCLUSION

This chapter has described the four general underlying patterns of functional linguistic variation in the corpus. It has described these patterns in terms of their linguistic composition, namely the functional contribution of each of the features to the overarching function elucidated by the dimensions. Feature co-occurrence and their contribution to this general function was then illustrated in context in each register on each dimension. It was noted at the beginning of this chapter that while these texts illustrate the ways linguistic features co-occur in texts of different registers, the varying extent to which they co-occur, and register differences on the dimensions, they do not represent any register in its entirety. Many of the registers exhibit substantial variation within their scope, extending along the dimensions and incorporating the features of both poles. This internal variation and its sources are examined closely in the following chapters. Specifically, social group (Chapter 5) and communicative differences among texts (Chapter 6) will be investigated within registers as the potential sources of the observed internal variation.

The present chapter has also observed extensive variation within age and gender groups. A consistent pattern has emerged among the social groups, uniform in their performance on all four dimensions. None of the groups show a distinct preference for either of the dimension poles; rather, all groups make use of the full linguistic range of the dimensions, which may indicate that all social groups observe the stark register differences on all dimensions. The following chapter further examines each age and gender group's cross-register trends.

CHAPTER 5. SOCIAL GROUP AND REGISTER

This chapter explores the relationship between social group and register with regard to the identified underlying functional patterns of linguistic variation in the corpus. Specifically, the study targets social group variation across registers (RQ 3; Section 5.1) and within registers (RQ 4; Section 5.2).

5.1 AGE AND GENDER GROUPS ACROSS REGISTERS

RQ 3: How do the social groups of the study vary *across* registers?

- Within each age group, to what extent does register predict linguistic variation?
- Within each gender group, to what extent does register predict linguistic variation?

It was established in the previous chapter that age and gender do not predict linguistic variation on the identified dimensions, accounting for minimal amount of variance (0%-0.1%), while the contribution of register is consistently major on all dimensions (76%-79% of variation on 3 dimensions and 38% on 1 dimension). The major effect of register and the lack of effect of age or gender suggest that all social groups in the corpus systematically respond to the demands of the situation of use. This is confirmed through ANOVA tests run within each age and gender group, which show that register is always a significant predictor of variation within each social group on all dimensions (an exception is Dimension 4 in the 19-year-old group, $p = 0.015$). Appendix 3 contains the descriptive statistics and the amount of variance explained by register in each social group on each dimension. All groups also demonstrate the same pattern of register differences as the overall pattern of register variation on the dimensions discussed in Chapter 4. (The Figures illustrating these patterns within each social group on each dimension can also be

found in Appendix 3). This discussion is therefore not repeated here. This analysis is illustrated briefly on Dimension 1 in the next section.

5.1.1 Age Groups across Registers on Dimension 1

Tables 5.1 and 5.2 present the descriptive statistics and the amount of variance explained by register in each age group. The Tables show that in each case register is a significant predictor of variation. Figures 5.1–5.5 illustrate register differences within each age group.

Table 5.1. Registers within age groups on Dimension 1. Descriptive statistics

Register	19 y. o.		20 y. o.		21 y. o.		22 y. o.		23 y. o.	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-1.37	1.15	-1.16	0.83	-0.94	0.82	-0.85	0.63	-1.11	0.74
Description	0.68	0.44	0.42	0.48	0.39	0.53	0.47	0.49	0.4	0.67
Emails	-0.37	0.51	0.07	0.3	0.04	0.36	0.01	0.38	-0.07	0.43
Essays	-0.24	0.32	-0.25	0.29	-0.24	0.24	-0.25	0.31	-0.19	0.15
Evaluations	0.18	0.01	0.28	0.22	0.11	0.25	0.13	0.26	0.12	0.07
Interviews	1.61	0.31	1.62	0.21	1.65	0.22	1.62	0.21	1.57	0.19
Text messages	-1.09	0.43	-0.93	0.61	-0.97	0.76	-0.73	0.6	-1.01	0.7

Table 5.2. Variance explained by register within age groups on Dimension 1

Age group	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
19 y. o.	6	17.98	<0.001	77%
20 y. o.	6	100.3	<0.001	77%
21 y. o.	6	109.4	<0.001	73%
22 y. o.	6	109.5	<0.001	76%
23 y. o.	6	23.01	<0.001	76%

Figure 5.1. Register differences within the 19-year-old group. Dimension 1

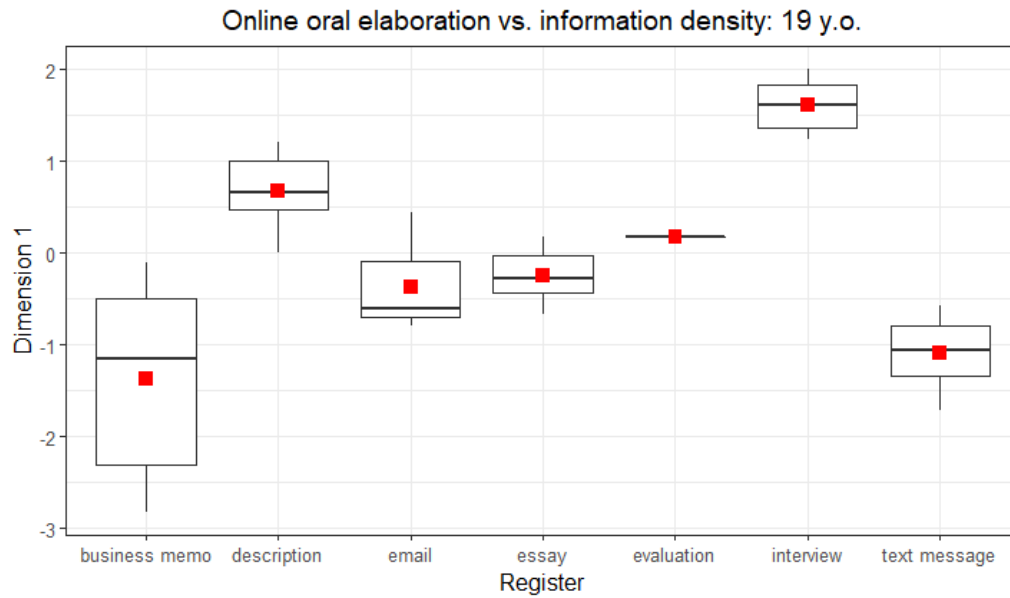


Figure 5.2. Register differences within the 20-year-old group. Dimension 1

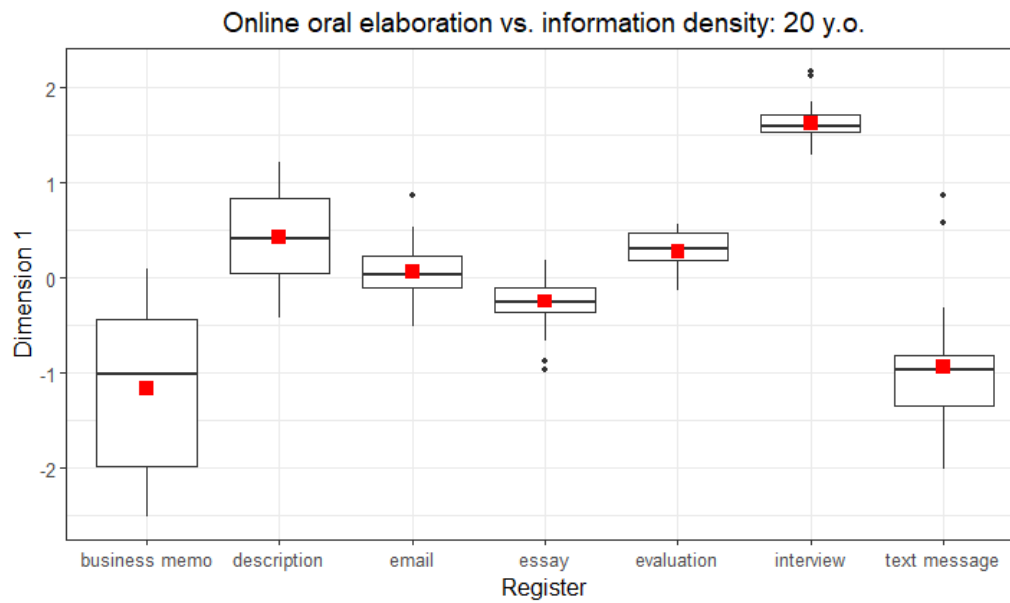


Figure 5.3. Register differences within the 21-year-old group. Dimension 1

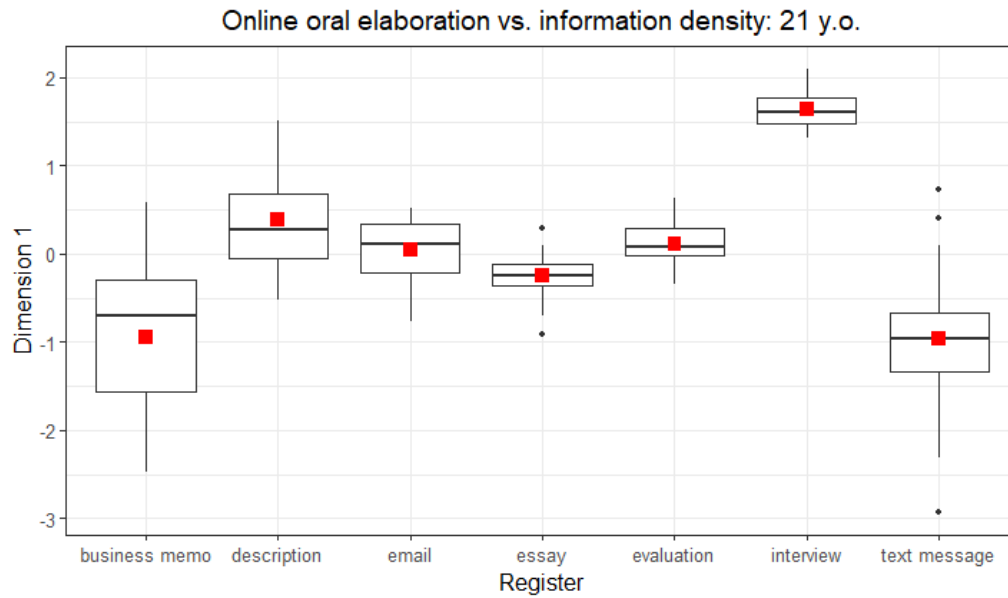


Figure 5.4. Register differences within the 22-year-old group. Dimension 1

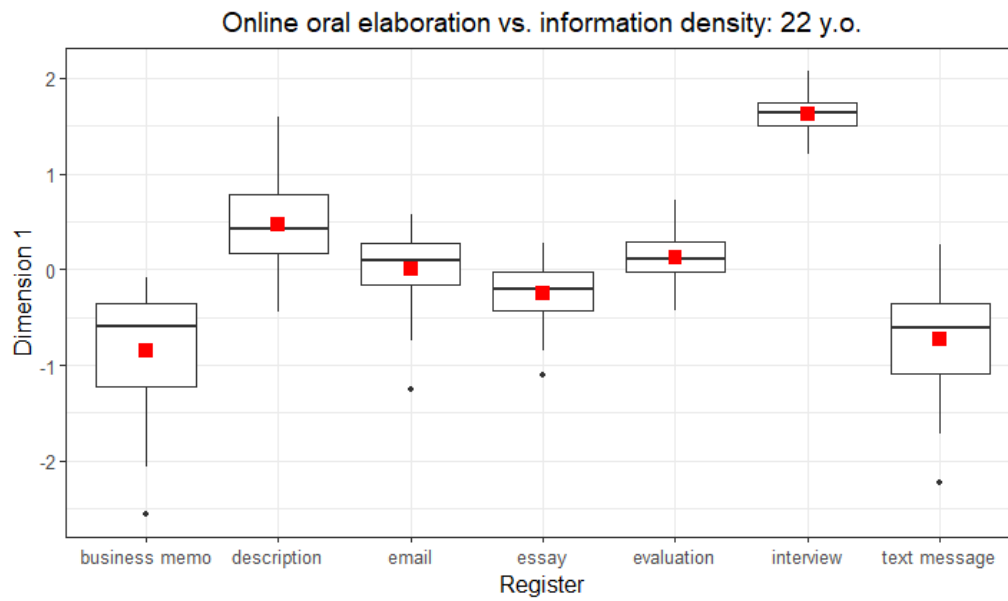


Figure 5.5. Register differences within the 23-year-old group. Dimension 1

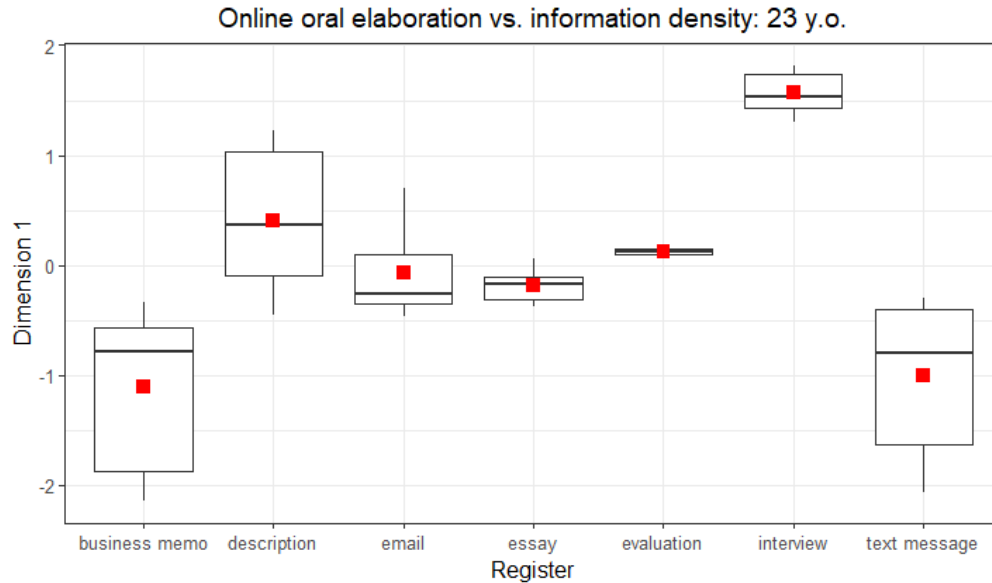


Table 5.3 summarizes the results of posthoc analyses carried out within each age group on Dimension 1 and shows register differences within each group. Due to small and unequal sample sizes of the groups, in pairwise register comparisons the study relies on Cohen’s d measures of effect size, considering an effect of $d > 0.5$ a significant register difference. Table 5.3 also allows a comparison of register patterns among age groups – i.e., it shows whether the same register pairs are found significantly different by all age groups (or if certain register pairs are nonsignificant for all groups) or whether some age groups differ in how they view register distinctions.

Table 5.3. Age groups across registers on Dimension 1: Pairwise comparisons; Cohen’s d

Register pair	Age				
	19	20	21	22	23
description-business memo	2.3	2.3	1.9	2.3	2.1
email-business memo	1.1	2.0	1.5	1.6	1.7
essay-business memo	1.3	1.4	1.1	1.2	1.7
evaluation-business memo	1.9	2.4	1.7	2.0	2.3
interview-business memo	3.5	4.6	4.3	5.2	5.0
text message-business memo	0.3	0.3	-0.03	0.2	0.1
email-description	-2.2	-0.9	-0.7	-1.0	-0.8

essay-description	-2.4	-1.7	-1.5	-1.7	-1.2
evaluation-description	-1.6	-0.4	-0.6	-0.8	-0.6
interview-description	2.4	3.2	3.1	3.0	2.4
text message-description	-4	-2.4	-2.0	-2.2	-2.0
essay-email	0.3	-1.1	-0.9	-0.7	-0.4
evaluation-email	1.5	0.8	0.2	0.3	0.6
interview-email	4.7	9.0	5.3	5.2	5.0
text message-email	-1.5	-2.1	-1.6	-1.5	-1.6
evaluation-essay	1.8	2.1	1.4	1.3	2.6
interview-essay	5.8	7.4	8.2	7.0	10.3
text message-essay	-2.2	-1.4	-1.3	-1.0	-1.6
interview-evaluation	6.5	6.2	6.5	6.3	10.1
text message-evaluation	-4.2	-2.6	-2.0	-1.8	-2.3
text message-interview	-7.2	-5.6	-4.7	-5.2	-5.0

It becomes apparent from Table 5.3 (and Figures 5.1-5.5) that age groups almost always recognize the same register distinctions. That is, significant and nonsignificant register patterns are mostly the same in all age groups, with only a few exceptions (bolded in Table 5.3). These exceptions include the effect of the difference between evaluations and image descriptions in the case of the 20-year-old group being slightly lower ($d = -0.4$), while in all other groups the effect is medium or large; small effect of the difference between essays and emails in the case of 19-year-olds ($d = -0.3$) and 23-year-olds ($d = -0.4$), while in other groups the effect is large; and small effect of the difference between evaluations and emails in the case of 21- and 22-year-olds ($d = -0.2$ and $d = -0.3$, respectively), while the effect is large or medium in the other groups (Table 5.3).

While these patterns show that some groups may differ in the extent to which the registers are distinct from one another, these results are hard to illustrate with text analysis as these differences are barely apparent in texts. Additionally, texts of a certain register produced by a particular age group vary to a considerable extent. For example, a particularly oral and elaborated evaluation written by a 20-year-old participant will affect the register mean score in

that age group. However, selecting that text to illustrate that this group’s evaluations are more distinct from their image descriptions compared to other groups is problematic, as the presence of such a text does not mean that all evaluations written by 20-year-old authors reflect that trend.

Overall, the results observed by the study indicate that distinct interpretations of register differences by age groups are rare. Register is a predictor of variation in all age groups, and as a rule, age groups distinguish or do not distinguish between the same register pairs. Groups that may differ in the extent to which they differentiate between registers include all and any age groups – from 19 to 23-year-olds, thus not suggesting that younger and older groups reveal different patterns in their approach to register variation.

5.1.2 Gender Groups across Registers on Dimension 1

Register was found a significant predictor within both gender groups (Tables 5.4-5.5), and Figures 5.6–5.7 show the register patterns within each group.

Table 5.4. Gender groups across registers. Dimension 1

Register	Men		Women	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.91	0.77	-1.04	0.79
Description	0.45	0.58	0.43	0.49
Emails	-0.06	0.41	0.02	0.37
Essays	-0.17	0.21	-0.26	0.28
Evaluations	0.12	0.21	0.15	0.25
Interviews	1.66	0.22	1.62	0.22
Text messages	-0.92	0.63	-0.89	0.66

Table 5.5. Variance explained by register within gender groups. Dimension 1

Dimension	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Men	6	63.92	<0.001	75%
Women	6	291.5	<0.001	75%

Figure 5.6. Register differences within the male group

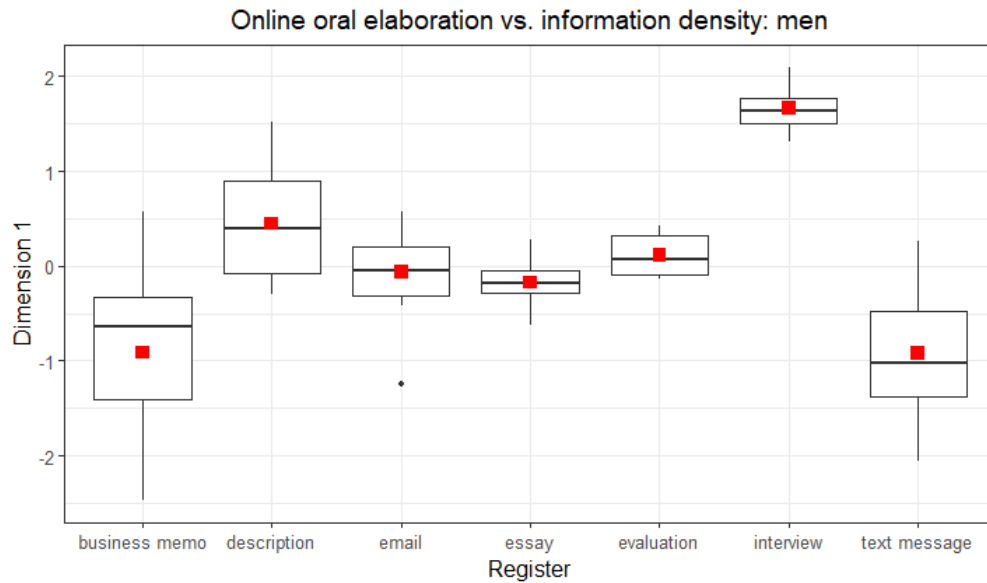
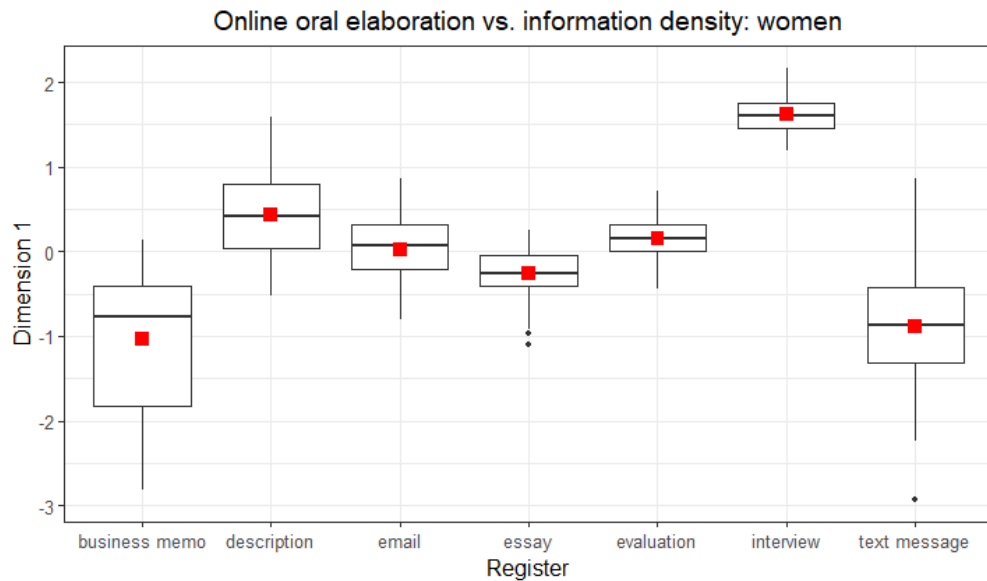


Figure 5.7. Register differences within the female group



Posthoc analyses showed that the patterns of register differences within the groups are mostly the same. Table 5.6 below summarizes these significant register differences for men and women. Table 5.6 shows that overall men and women follow the same register patterns, and only the pair ‘essays-emails’ shows noteworthy differences. Specifically, essays and emails are less distinct in the case of men ($d = -0.3$) than in the female group ($d = -0.8$). Women’s emails are

overall slightly more oral and elaborated ($M = 0.02$; $SD = 0.37$) than men's ($M = -0.06$; $SD = 0.41$), while women's essays ($M = -0.26$; $SD = 0.28$) are slightly more informational than men's ($M = -0.17$; $SD = 0.21$), which accounts for the larger effect in the female group (Table 5.4).

Table 5.6. Gender groups across registers on Dimension 1: Pairwise comparisons; Cohen's d

Register pair	Males	Females
description-business memo	2.0	2.2
email-business memo	1.4	1.7
essay-business memo	1.3	1.3
evaluation-business memo	1.8	2.0
interview-business memo	4.5	4.6
text message-business memo	-0.01	0.2
email-description	-1.0	-0.9
essay-description	-1.4	-1.7
evaluation-description	-0.7	-0.7
interview-description	2.7	3.1
text message-description	-2.2	-2.3
essay-email	-0.3	-0.8
evaluation-email	0.5	0.4
interview-email	5.2	5.2
text message-email	-1.6	-1.7
evaluation-essay	1.4	1.5
interview-essay	8.5	7.4
text message-essay	-1.6	-1.2
interview-evaluation	7.1	6.2
text message-evaluation	-2.2	-2.1
text message-interview	-5.4	-5.1

Such exceptions, however, are rare as can be seen from Table 5.6, and again these differences in the extent to which a gender group distinguishes between the registers are not apparent through text analysis. Additionally, as noted above, variation within a gender group with regard to the emails and essays they produce precludes generalizations to the entire group. Thus, generally, genders show the same cross-register trends, recognizing the same register differences or not distinguishing between the same registers (e.g., text messages and business memos).

5.2 AGE AND GENDER GROUPS WITHIN REGISTERS

RQ 4: How do the social groups of the study vary *within* registers?

- Within each register, to what extent does age predict linguistic variation?
- Within each register, to what extent does gender predict linguistic variation?

This section investigates social groups within registers. By doing so, it examines the possibility that register-internal variation observed on the dimensions (Sections 4.1.1, 4.2.1, 4.3.1, 4.4.1) may be due to social group differences within registers. To this end, age and gender are considered as predictors of linguistic variation within each register. Appendix 4 contains the descriptive statistics for each age and gender group within each register, the amount of variance explained by age and gender on each dimension, and the Figures illustrating these results.

Neither age nor gender is found to be a significant predictor of linguistic variation within registers on any of the dimensions. That is, age or gender groups do not differ from each other, and none of the social groups shows a preference towards a particular dimension pole on any of the dimensions. However, another important pattern emerging from these results is that of extensive variation within each age and gender group, as shown by quite high standard deviations, often exceeding the means. That is, while age and gender groups do not differ from each other to a significant extent, they also consistently show variation within their scope. Thus, all groups employ a range of the linguistic resources and functional possibilities associated with the dimensions in all registers. For example, the same age or gender group can produce oral elaborated or informational texts rather than showing a preference for one extreme or the other.

This result, consistently observed within each register, suggests that there are other factors at play that account for this extensive variation within the groups (the subject of Chapter 6). As texts by all age groups and both gender groups from the opposing dimension poles are

examined, it is observed that specific communicative factors, such as the communicative goals and topics that vary across texts, may play a part in the internal variability that all groups exhibit within registers. For example, authors of all age groups and both genders produced business memos ranging from highly elaborated texts to texts with no features of elaboration and a high density of nominal features instead. These distinct approaches to the register, resulting in strikingly different linguistic representation, may be attributed to factors such as the authors' understanding of and familiarity with the situational characteristics of the register, their individual ability to envision a relationship with a superior in a corporate context, their prior exposure to similar contexts and similar texts produced by speakers and writers in such professional relationships, among others. The results indicate, however, that these varying interpretations of the register do not correspond to authors' age or gender, but rather represent a communicative choice of how to accomplish the goals of the register. Such different communicative choices were made by authors of the same age groups and the same genders.

Likewise, social group did not predict linguistic variation in any of the registers on Dimension 4 (the other dimension chosen for all analyses, as explained in Chapter 3), and clausal features of stance are used to varying extents by all the groups. The extent to which these features are employed in texts within each of the groups again appears to depend on additional communicative factors. For example, text messages whose focus is on the author's interaction with a third party, its content, and outcome are more likely to resort to stance features than those that describe the author's plan for the day or a car accident. As these distinctions are systematically observed within all age and gender groups, it is again likely that it is these factors that account for register-internal variation on the dimension.

The next section presents analysis of texts within registers produced by each age group that represent opposite extremes of the dimensions, demonstrating that each age group produces a range of texts rather than gravitates towards a particular end of the dimension. These analyses are demonstrated within business memos, the register with the highest standard deviation on Dimension 1 ($M = -1.01$; $SD = 0.79$), and within text messages – the most varied register on Dimension 4 ($M = 0.95$; $SD = 1.63$).

5.2.1 Text Analysis: Age within Business Memos on Dimension 1

The register of business memos has been consistently shown to vary considerably along all dimensions (Chapter 4), and as noted above, on Dimensions 1, business memos are the most varied register. This section examines the extent to which this linguistic variation within the register is due to variation by age. Table 5.7 presents the descriptive statistics for each age group (means and standard deviations), Table 5.8 summarizes the results of the one-way ANOVA measuring the effect of age, and Figure 5.8 illustrates these results. The results indicate that overall age is not a significant predictor of variation within the register, accounting for only 3% of variance.

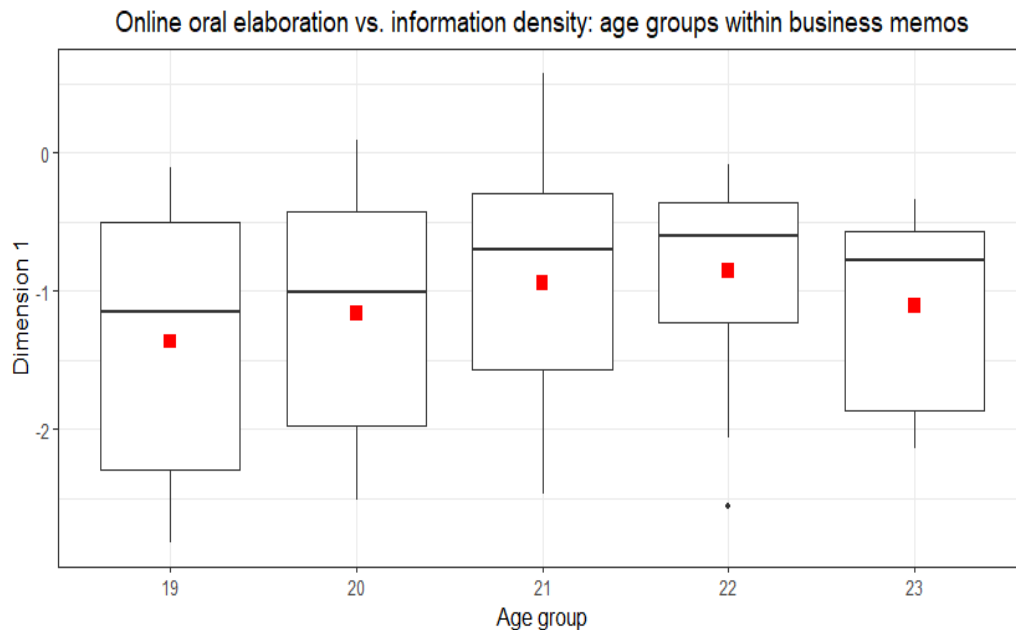
Table 5.7. Age groups within business memos. Descriptive statistics

Age	Dimension 1	
	<i>M</i>	<i>SD</i>
19 ($n = 6$)	-1.37	1.15
20 ($n = 28$)	-1.16	0.83
21 ($n = 37$)	-0.94	0.82
22 ($n = 32$)	-0.85	0.63
23 ($n = 8$)	-1.11	0.74

Table 5.8. Variance explained by age within business memos

Dimension	<i>df</i>	<i>F</i>	<i>P</i>	<i>R</i> ²
Dimension 1	4	0.99	>0.001	3%

Figure 5.8. Age groups within business memos: Dimension 1



Text Samples 29-33 below are examples of business memos produced by each age group.

Text Sample 29. Author 28, business memo; 19-year-old group (text id: 27098); Dimension 1 score: -0.10

As our CEO is avid of water, this hotel is **only** a 10 minute walk from the 'Market Place' where they will find long streams of water with accessible footpaths to stroll on. To add to that, the 'Pizzeria Van Tribunal'- **only** 5 minutes away from the hotel by car would be a great restaurant to dine at, **specially because** the other options were **predominantly** seafood based **and** our CEO is allergic to such. [...] **Furthermore**, on the 3rd of September- the middle day of their trip, there is a concert for **only** 25 euros with a popular metal band 'Korpiklaani' in central Helsinki. 'Taiga Colors' is a **highly** rated gift and specialty store with various souvenirs and post cards for our CEO to shop at and gift their children **with (stranded)**.

Text Sample 30. Author 13, business memo; 20-year-old group (text id: 27082); Dimension 1 score: -0.04

I **have found** a 4-star hotel situated 0.4 miles away from the city centre. For two nights it will cost £216. It is **also** situated close to the Seurasaraselka waters so, if we **book** a room with a view you will **always** get to see the water. This is £16 over budget **but it** was the best option in regards to tourist reviews and location. Caferino Obas is a cafe, which is **only** a ten-minute walk away from the hotel **and** they **offer** a great breakfast menu.

Text Sample 31. Author 86, business memo; 21-year-old group (text id: 27162); Dimension 1 score: -0.08

There **is** a hotel called Clarion Hotel Helsinki. **It is** near the water **and** i believe **has** business type rooms. [...] **It is roughly** £120 a night. I **have checked** amenities in the hotel **and** there **is** a rooftop bar and pool as well as a restaurant. In terms of a cafe, there **is** a cafe in the hotel if the CEO **does** not feel like moving **very** far that day. **However**, if they **want** to have a little walk around the city, there **is** a cafe called 'little big cafe' **and it is** an 8 minute walk from the hotel or a 5 minute drive, **so very** close. For the museum, there **is** a whole museum dedicated to the city of Helsinki called Helsinki City Museum.

Text Sample 32. Author 14, business memo; 22-year-old group (text id: 27083); Dimension 1 score: -0.08

Your flight **back is** an evening one from 18:15pm to 21:35pm -to give you most of that day to enjoy Helsinki and what **it has** to offer - on the 4th September 2021 Total cost: £138.29 - optional to add carriage luggage HOTEL You will be staying in a lovely waterfront Hotel called Clarion Hotel Helsinki which **is** only a mile from the city centre. **It is** a bed and breakfast **and** the total **comes to only** £268 for the two nights. **It has** fantastic reviews and views.

Text Sample 33. Author 24, business memo; 23-year-old group (text id: 27094); Dimension 1 score: -0.33

It is a 30 minute walk **away or alternatively** you can get a taxi **there** in 10 minutes. [...] **it is** in the vicinity of your hotel. - **Also** in the vicinity **is** a souvenirs shop called..... Souvenirs shop to buy a little memento for your children. **It is** a short 3 minutes walk **away**. -In the evening, if you **are feeling** the need for a down to earth, intimate experience, i **suggest** Koko Jazz Club. If you **want** to see some amazing views of the city, I **recommend** Ateljee Rooftop Bar, where you can take **it** all in and mingle with the locals if you would like **that**.

It is apparent from these excerpts that the businesses memos, produced by the five age groups, are not distinct in how they incorporate the linguistic features of oral elaboration and information density. All the texts represent a balance of informational features and features of oral elaboration. All the informational features in the excerpts denote locations (city centre) and proper names (Ateljee Rooftop Bar, waterfront Hotel called Clarion Hotel Helsinki), measures of distance (8 minute walk), or a type of place through noun-noun sequences of purpose (Koko Jazz Club, souvenirs shop). These informational features are therefore indispensable in view of the communicative goals of business memos, which are centered around these entities and aim to inform about them. On the other hand, features of oral elaboration include the present tense to express recommendations and suggestions (*I suggest, I recommend*) or provide information

about the proposed options (*It is a bed and breakfast and the total comes to...*), causative adverbial clauses (*especially because the other options were predominantly seafood based and our CEO is allergic to such*) providing the rationale for the choice, adverbs expressing stance and qualifying propositions (*were predominantly seafood based, you will always get to see the water*) as well as organizing the discourse (*or alternatively you can...*). Clausal coordination (*and our CEO is allergic to such; and it is an 8 minute walk from the hotel*) serves to enhance the statements with additional detail.

The five texts above, each written by a different age group, all incorporate features of elaboration to an equal extent. These examples show that all authors tend to elaborate on the proposed itinerary, share their thought process, rationale in making these choices, and the logical relations between ideas. While students may have had various amounts of exposure to business communication or other registers that may have prepared them for similar tasks, it does not follow from the results here that the younger age groups have revealed a lack of such exposure in comparison to older groups.

On the other hand, there are business memos which do not rely on features of elaboration, as shown in Text Samples 34-38 below. Again, age does not appear to be a factor determining this choice. This is evidenced by the fact that the same lack of differences between the age groups which is observed in the texts above, all employing a balance of informational and oral features, is found in texts which are much less elaborated. These texts, again produced by the different age groups of the study, lack not only in oral elaboration, but often fail to provide all the information required by the business memo. The reasons for this choice may range from this decision serving as an avoidance strategy and a lack of competence in the task to a lack of motivation or interest, as well as external circumstances, such as a lack of time. What is

apparent, however, is that this communicative choice is made by all the age groups of the study. This trend thus again precludes the conclusion that age groups demonstrate varying levels of competence in the register due to the advantage of immersion into the business context or classroom instruction older participants may have had over the course of their university careers. Rather, the lack of elaboration and an overall minimalist approach to the task shown in the texts below appears to be a choice made by all age groups for reasons other than demographic.

**Text Sample 34. Author 69, business memo; 19-year-old group (text id: 27144);
Dimension 1 score: -1.77**

Hotel Recommendations: Clarion Helsinki Hotel - £118 per night and 2 nights £235 - Near the water -4 star hotel, location 7.9/10, rooms 7.9/10, service 8/10, cleanliness 8.8/10, value for money 7.9/10 public cafes - breakfast place opens at 10am 5/5 rating The national Museum of Finland - cities history is covered.
Kankurin Tupa - reviewed as best souvenir shop in Helsinki; 4.3/5 rating Restaurant Savotta 4.4 rating - **serve** charred white fish and cucumber lizard -cottage cheese salad and wild herb pesco mini Karelian pies Helsinki music centre - rated **highly** on trip advisor

**Text Sample 35. Author 30, business memo; 20-year-old group (text id: 27101);
Dimension 1 score: -1.66**

- Hotel: Radison Blu Seaside Hotel, Helsinki 2nd-4th September 2021= £181 0.9m from the city centre and situated on the boarder of the sea. - Cafe (for breakfast): Fazer Café Kammpi (open 8am-8pm) in Kammpi centre. - Museum: Finnish Museum of Natural History. - Restaurant: booking at Lappi, Helsinki- serving authentic dishes (such as Paronkaristys- sautéed reindeer and many other non-sea food authentic dishes). - Souvenir: 'Port of Souvenir' - Helsinki. - Live music: Keltanien Kissa

**Text Sample 36. Author 93, business memo; 21-year-old group (text id: 27170);
Dimension 1 score: -1.86**

Itinerary (2- day conference to Helsinki) Hotel: Clarion Hotel (£166 a night with a view of the sea) Breakfast Café: Johan Nystrom- Kanavaranta Museum: Helsinki City Museum (Spiritual, material and architectural heritage) Restaurant: Lappi Restaurant (Menu **consists** of meat such as Grilled Fillet of Reindeer) Souvenir shop: Kankurin Tupa (not far from hotel) Things to do in the evening: Wallis Skuha Karaoke Bar

**Text Sample 37. Author 22, business memo; 22-year-old group (text id: 27092);
Dimension 1 score: -1.91**

Hotel - U14 Autograph Collection - 4 star hotel £147 per night - check in 15:00 - check out 12:00 - 2 blocks away from the south harbour Cafe for breakfast - Fazer cafe -

Kluuikatu 3 National Museum of Finland - opening times - Tue - Sunday - 11:00 - 18:00
- Mannerheimintie 34, Itelsinki 00100, Finlnad National dish - Sauteed Reindeer
'Poronkarstys' Restaurant Saaga - Bulevardi 36, 00120 Helsinki Gift shop - Lioporo -
Kauppatori - 00130 Helsinki, Finland Savoy Theatre - Kasarmikatu 46, 00130 Helsinki,
Finland

**Text Sample 38. Author 29, business memo; 23-year-old group (text id: 27099);
Dimension 1 score: -1.95**

Helsinki trip- Sep 2-4 2021 Hotel option 1- Raddison blu plaza hotel- £119- 4.4 stars
Close to museum of contemporary art **Offers** upscale resteraunt and bar 30 mins from
airport Close to Taiseneini park Hotel option 2- Hotel U14, Autograph collection- £146-
4.9 stars Harborside district, boat cruises- **only** 2 blocks **away** Close to Esplenedi Park-
Espi stage- concerts held during summer; Sep 3rd 'Torpikleni' holding a concert- Finnish
rock Hotel option 3- Hilton helsinki

Overall, it becomes apparent from these texts (as well as seen from Table 5.7 and Figure 5.8) that each age group exhibits substantial variation within its scope, which encompasses texts with minimal elaboration as well as those that provide detailed and thorough responses to the task. Thus, it is not possible to assert that this different interpretation of what constitutes a business memo stems from differences in age. In fact, the amount of variation observed within all groups suggests that the role of individual differences or other factors may be more prominent than that of the age group. Internal variation within business memos and the possible basis for these differences within the register will be investigated in the next chapter.

5.2.2 Text Analysis: Gender within Business Memos on Dimension 1

It was further tested whether gender accounts for the linguistic variation within business memos. Tables 5.9-5.10 below summarize the results and Figure 5.9 shows the gender groups along Dimension 1. These results indicate that gender does not predict variation within the register, and the Figure shows that the groups do not differ. Both gender groups, however, consistently display variation within their scope, suggesting that both men and women tend to exhibit variation within the register due to individual preferences or other factors.

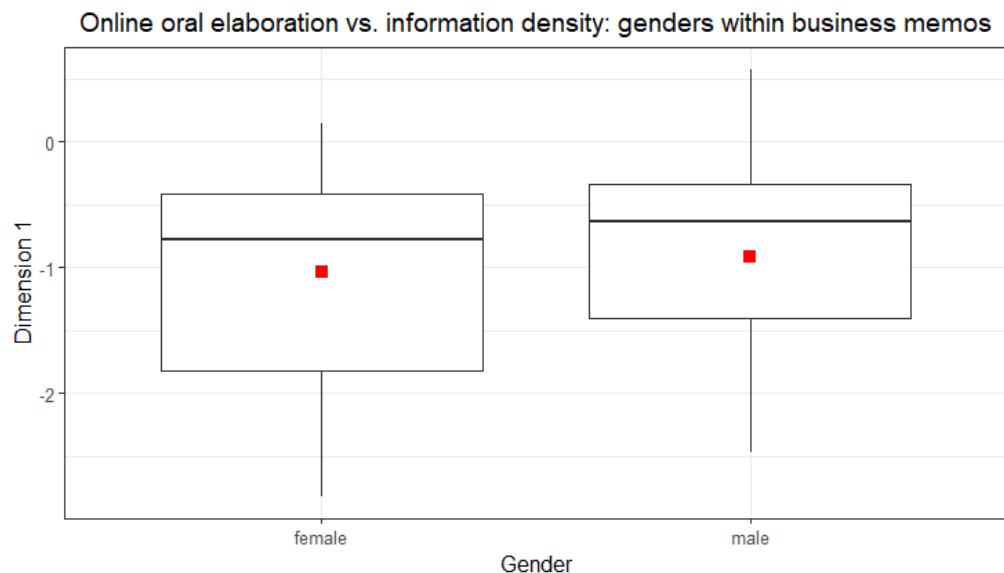
Table 5.9. Gender groups within business memos

Gender	Dimension 1	
	<i>M</i>	<i>SD</i>
Men (<i>n</i> = 20)	-0.91	0.77
Women (<i>n</i> = 90)	-1.04	0.79

Table 5.10. Variance explained by gender in business memos

Dimension	<i>df</i>	<i>F</i>	<i>p</i>	<i>d</i>
Dimension 1	1	0.39	>0.001	0.2

Figure 5.9. Gender groups within business memos: Dimension 1



Text Samples 39-42 exemplify a tendency towards oral elaboration (Text Samples 39-40) as well as information density (Text Samples 41-42) by male and female groups. Text Samples 49-50 illustrate such features of oral elaboration as present tense, contractions, adverbs, clausal coordination, and demonstrative pronouns and the pronoun ‘it’ used by both male and female groups, which equally contribute to the immediacy of the texts, situate them in time and space, allow for a detailed and thorough presentation with reasons for the made choices and an expression of authors’ stance towards the propositions (e.g., through emphatics – *it is **only** 1.8 miles from...*).

Text Sample 39. Author 100, business memos, men (text id: 27067), Dimension 1 score: 0.57

Dear Mrs Botago, I've done some research into some of the top places in Helsinki you might want to try whilst you're there, and I'll copy in the bookings team so they have the addresses if you want them. The hotel I'd recommend is the Hilton Helsinki Strand, it was that or the Radisson Blu but the Hilton had better rooms with a view (can't beat a window-facing bed in the morning!). - 217 Euros per night - Has all your basic hotel stuff, plus Hilton is great anyway - John Steinbergin ranta 4, Helsinki, 00530, Finland.

Text Sample 40. Author 106, business memos, women (text id: 27073), Dimension 1 score: 0.14

Regarding the information I have gathered for the CEO's trip to Helsinki, I have managed to find a multitude of things ranging from hotels to souvenir shops for her children. To start of with, the hotel I thought would be the best fit is the Radisson Blu Seaside Hotel, due to the fact it overlooks the scenic Gulf of Finland. The hotel further has access to gyms, saunas and serves traditional and fresh food, only for £74 pn. It is only 1.8 miles from the centre.

In contrast, Text Samples 41-42 by male and female authors illustrate a different tendency and thus exemplify variation within each group. These texts do not offer the level of detail shown in the passages above and merely state the selected options, expressed mostly through nominal sequences.

Text Sample 41. Author 9, business memos, men (text id: 27165), Dimension 1 score: -2.47

Helsinki Conference, 2-4th September 1. Radisson Blu Seaside Hotel - £180 per night 2. Cafferino Oba - cafe 3. Sinebrychoff Art Museum - for culture and history 4. Kosmos Restaurant 5. Trimeeri Oy - gift shop 6. Storyville - Jazz club for evening entertainment

Text Sample 42. Author 61, business memos, women (text id: 27135), Dimension 1 score: -2.56

Hotel: Hilton Helsinki Strand - King Guest room with Sea View - 2-4th Sep 2021 = £193 per night Breakfast Cafe: La Torrefazione, Hakaniemi - 8am-6pm - 3 min walk from hotel City's History Museum: The National Museum of Finland - 11am-6pm - 21 min walk/10 min drive from hotel

It follows from these results that both gender groups are quite varied with regard to their preference for features of oral elaboration or information density, and both employ a wide range of Dimension 1 linguistic resources. Thus, it is factors other than social group affiliation that are likely to account for the variation within business memos.

5.2.3 Text Analysis: Age within Text Messages on Dimension 4

Text messages are consistently among the most varied registers on all dimensions. As it was mentioned earlier, however, the shortness of text messages leads to dimension score inflation, especially on the negative pole, where many text messages simply do not contain either positive or negative features of the dimension, or contain some positive features. As explained in Chapter 3, the number of positive features on the dimension is so large that the occurrence of several of these features in short texts, where the other features did not have a possibility to occur, does not result in these texts' higher position on the dimension. While overall the positions of text messages on the dimensions are interpretable and reflect the dimension functions, the results need to be interpreted with caution, especially as internal variation in the register is analyzed.

The present section examines the effect of age in this internal variation in the register on Dimension 4. Tables 5.11-5.12 present the descriptive statistics (means and standard deviations) for each age group and the amount of variance explained by age, and Figure 5.10 illustrates the age groups' use of evidence-based stance features. The Tables show that age was not found to be a significant predictor of variation within the register, explaining 5% of variance. Instead, as in the case of business memos, all age groups show substantial variation on all the dimensions, using a wide linguistic range rather than showing a preference for either pole. This variability is again demonstrated by the text messages showing opposite tendencies written by authors of the same age groups.

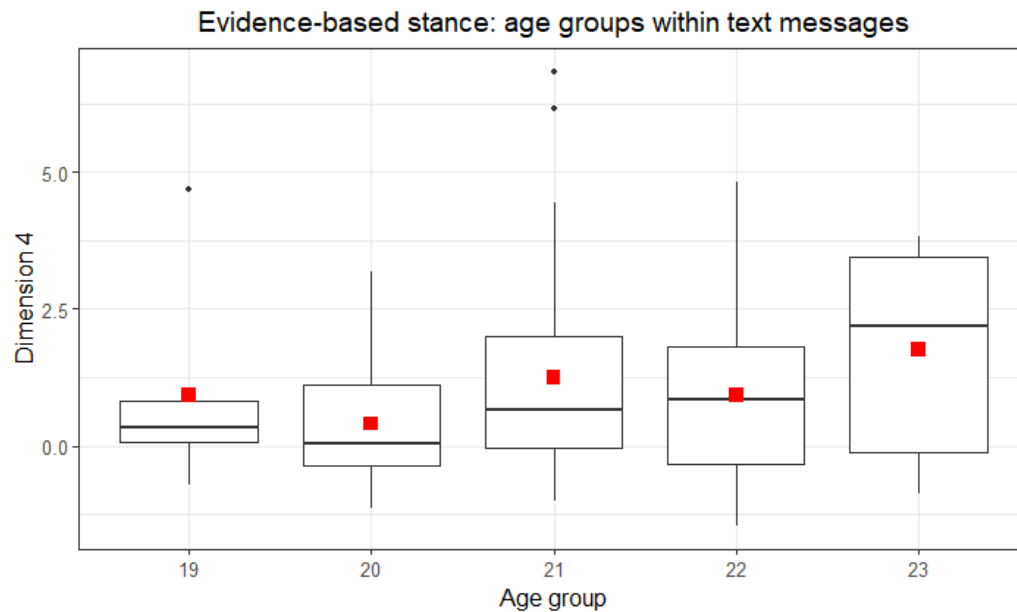
Table 5.11. Age groups within text messages. Descriptive statistics

Age	Dimension 4	
	<i>M</i>	<i>SD</i>
19 (<i>n</i> = 6)	0.94	1.92
20 (<i>n</i> = 28)	0.41	1.08
21 (<i>n</i> = 37)	1.25	1.91
22 (<i>n</i> = 32)	0.93	1.50
23 (<i>n</i> = 8)	1.75	1.97

Table 5.12. Variance explained by age within text messages

Dimension	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Dimension 4	4	1.59	>0.001	5%

Figure 5.10. Age groups within text messages: Dimension 4



Text Samples 43-47 are text messages written by the different age groups of the study. Clausal stance features, possibility modals, and the past tense, the latter typically supplying evidence for the authors' stance, are quite prominent in all these texts, where these features denote mental processes and report the speech of participants of the described events. These texts focus on instances of communication between the author and other entities, and it is the content of these communicative events and their outcome that is at the center of the text message.

Clausal features of speaker-attributed stance are therefore crucial to the communicative goal of such texts.

Text Sample 43. Author 9, text message, 19-year-old group (text id: 22194),

Dimension 4 score: 4.69

don't **forget** to send me your timetable so I **can** book a ticket to come grab the stuff

Text Sample 44. Author 112, text message, 20-year-old group (text id: 12843),

Dimension 4 score: 3.17

Say to them (**that del**) ur happy to help but they **need** to **ask** not **demand**

Text Sample 45. Author 49, text message, 21-year-old group (text id: 17279),

Dimension 4 score: 6.83

but i **said** (**that del**) i was aiming for top marks and he **said that** the improvements would go a long way to being top marks

Text Sample 46. Author 12, text message, 22-year-old group (text id: 12870),

Dimension 4 score: 4.82

They **wanted** it for their kid and **complained** (**that del**) its finished I **said** (**that del**) people who are adults **can** play as well and its their money

Text Sample 47. Author 29, text message, 23-year-old group (text id: 14943),

Dimension 4 score: 3.83

I may have to **say** until he **loved** our idea so much he **said** he will email us today if he **can** organise meetings with some managers to teach us and to make connections

In contrast, features of stance associated with the dimension are absent in Text Samples 48-51 below, again produced by different age groups. While some feature of stance not associated with the dimension do occur in these text messages (adjective complement clauses in *nervous to send off* or *I'm sorry to keep changing this event*), their general scarcity seems to be determined by the subject matter and the goals of these texts. These text messages do not relay an interaction the author was involved in or a thought process. Rather, the central goal of informing the addressee of a specific fact, such as the location where the author's relative lives, the task components of 'coursework', the plan for the day, and cancellation of an event, respectively, does not warrant frequent stance expression. This variability within text messages is

again observed in the case of each age group, as authors of all ages are guided by what appears to be communicative considerations within the register.

**Text Sample 48. Author 69, text message, 19-year-old group (text id: 19896),
Dimension 4 score: -0.71**

Bc then my cousin lives there too and I won't be like alone alone gym

**Text Sample 49. Author 89, text message, 20-year-old group (text id: 22377),
Dimension 4 score: -1.13**

They were good! Nervous to send one off. Its called Coursework over here. Where you have the exam question/title and you have to write 1200/1500 words.

**Text Sample 50. Author 43, text message, 21-year-old group (text id: 16680),
Dimension 4 score: -1.00**

yeah all good on my way to the library for a couple hours then linguistic interview for an hour then lunch then librsry again

**Text Sample 51. Author 32, text message, 22-year-old group (text id: 15449),
Dimension 4 score: -0.81**

Hello everyone, I'm so sorry to keep changing this event but due to lockdown in starting on , the whole celebration is gonna have to be cancelled, thank you all for your would have be attendance, God bless you all

5.2.4 Text Analysis: Gender within Text Messages on Dimension 4

This section examines gender as a predictor of linguistic variation within text messages. Tables 5.13 and 5.14 below present the descriptive statistics (means and standard deviations) for each gender group and the amount of variance explained by gender. Gender is not found to be a significant predictor of variation within the register, and both male and female groups use a range of its linguistic resources rather than showing a preference for either pole.

Table 5.13. Gender groups within text messages. Descriptive statistics

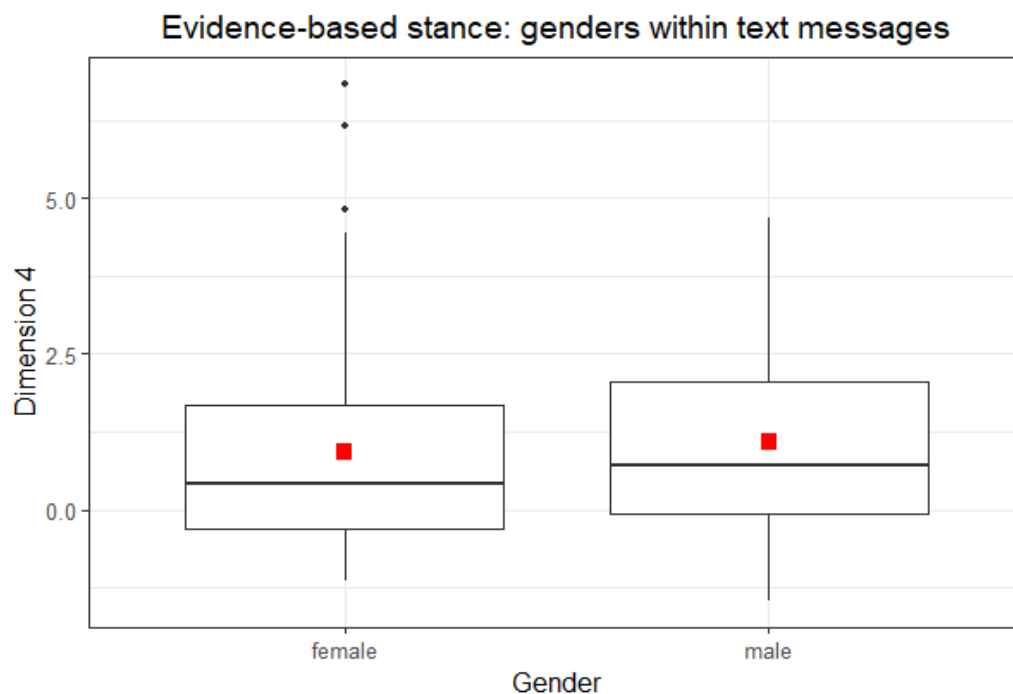
Gender	Dimension 4	
	<i>M</i>	<i>SD</i>
Men (<i>n</i> = 20)	1.09	1.62
Women (<i>n</i> = 90)	0.93	1.66

Table 5.14. Variance explained by gender in text messages

Dimension	<i>df</i>	<i>F</i>	<i>p</i>	<i>d</i>
Dimension 4	1	0.15	>0.001	0.1

Figure 5.11 plots the gender groups on Dimension 4 and shows that the groups do not differ with regard to stance features and reveal the same pattern of extensive internal variation. Text Samples 52-53 contain text messages which employ salient Dimension 4 features, such as private and public verbs followed by complement clauses to refer to mental and communicative processes.

Figure 5.11. Gender groups within text messages: Dimension 4



Text Sample 52. Author 9, text message, men (text id: 22194), Dimension 4 score: 4.69

don't **forget** to send me your timetable so I **can** book a ticket to come grab the stuff

Text Sample 53. Author 112, text message, women (text id: 12843), Dimension 4 score: 3.17

Say to them (**that del**) ur happy to help but they need to **ask** not **demand**

In contrast, mental processes or communicative events are not the focus of the texts in Text Samples 54-55, which convey concrete information rather than express stance towards it. Dimension 4 features are thus not required for this goal and this content of the message. It is apparent that this lack of need for stance features is recognized by male and female writers, and this communicative choice is made by authors of both genders. The systematic variation on the dimension shown by both gender groups and the lack of significant difference between them again suggest that it may be communicative factors, such as the authors' goals and topics, consistently shown to distinguish between texts by the authors of the same genders, that may account for the observed variation within the register.

Text Sample 54. Author 88, text message, men (text id: 22088), Dimension 4 score: -0.59

I **got** in half an hour ago too. Massive delays coz of the trains in

Text Sample 55. Author 43, text message, women (text id: 16680), Dimension 4 score: -1.00

yeah all good on my way to the library for a couple hours then linguistic interview for an hour then lunch then library again

5.3 CONCLUSION

This chapter explored the relationship between social group (age and gender) and register. The study investigated cross-register patterns within each social group and overall found these patterns to be the same across groups. That is, all age and gender groups of the study systematically distinguish between the registers and rarely differ in the ways they approach register distinctions.

This chapter then demonstrated that age and gender do not contribute to linguistic variation within registers, and each social group further varies substantially within its scope and produces a variety of texts within a certain register. That is, all social groups produce texts that are more oral and elaborated or texts that are more informational (Dimension 1); similarly, all social groups produce texts with varying amounts of stance features (Dimension 4). The analysis presented in this chapter suggested that this systematic variation observed within each age and gender group may be attributed to specific communicative distinctions recognized by authors of all ages and both genders. This possibility is further explored in Chapter 6 on communicative variation within registers.

CHAPTER 6. COMMUNICATIVE VARIATION ACROSS TEXTS

The previous chapter has investigated the relationship between register and social group. Specifically, social variation by age and gender was examined as possible sources of linguistic variation within registers. Chapter 5 showed that age and gender are not significant predictors of this register-internal variation on any of the functional dimensions of variation. That is, age and gender groups do not differ from each other within registers, and none of the groups shows a preference for either of the dimension poles. Rather, all groups exhibit extensive variation and use a range of the linguistic resources of each dimension within all the analyzed registers. Analysis of texts demonstrating this variability *within the same* age and gender group (i.e., oral elaborated texts vs. more informational ones and texts expressing stance to various degrees) revealed that there are likely to be other factors that play a part in functional language use and result in linguistic variation across texts within a register. This chapter examines this possibility further as it explores communicative variation within registers and addresses the following research question:

RQ 5: To what extent do communicative differences across texts of a register explain linguistic variation within registers?

Sections 6.1 and 6.2 discuss variation across texts within registers on Dimensions 1 and 4, respectively.

6.1 COMMUNICATIVE VARIATION ACROSS TEXTS: ORAL ELABORATION VS. INFORMATION DENSITY

To investigate variation within registers, the study identifies the position of each text relative to the register mean. Specifically, the study examines the distance between each text and the

register mean in each register (procedure described in Chapter 3, Section 3.3.4). As discussed in Chapter 4, Dimension 1 is a cline from oral elaboration on its positive end to information density on the negative end. The focus of this analysis is then on three groups of texts: 1) texts that represent the central tendency of the register; 2) texts that deviate from the central tendency and are positioned *above* the register mean on the dimension, i.e., show a preference for oral elaboration – the functional pattern associated with its upper end; 3) texts that deviate from the central tendency and are positioned *below* the register mean on the dimension, i.e., show a preference for information density – the functional pattern associated with its lower end (Appendix 5 contains the individual text scores of each text in each register, the respective register means, and the measures of distance between them).

Table 6.1 below summarizes these results for each of the six registers included in the analysis – essays, emails, interviews, image descriptions, evaluations, and business memos – and shows the number of texts within each distance group (i.e., how far the text is from the mean) and the position of these texts relative to the register mean (above or below).

Table 6.1. Number of texts per distance and direction group within each register. Dimension 1: Oral Elaboration vs. Information Density

Register	Direction of deviation	Distance groups				
		<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Essays ($M = -0.24$; $SD = 0.27$)	Above M	11	30	12		
	Below M	22	25	6	5	
Emails ($M = 0$; $SD = 0.38$)	Above M	10	23	26	2	
	Below M	13	13	17	6	1
	Above M	19	23	12		

Register	Direction of deviation	Distance groups				
		<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Interviews ($M = 1.62$; $SD = 0.21$)	Below M	21	30	7		
Descriptions ($M = 0.44$; $SD = 0.51$)	Above M	7	14	19	10	5
	Below M	11	11	20	15	1
Evaluations ($M = 0.15$; $SD = 0.24$)	Above M	11	14	7		
	Below M	11	16	5		
Business memos ($M = -1.01$; $SD = 0.79$)	Above M	2	10	23	22	9
	Below M	3	6	8	7	22

It can be seen from Table 6.1 that registers differ in the amount of internal linguistic variability within them, with some registers allowing considerably more freedom and others being noticeably more restrictive. For example, Table 6.1 shows that business memos are the register with the largest number of peripheral texts (in both peripheral distance groups – 0.6-0.9 and >0.9). It follows that despite having a specific task to follow – to research possible options that meet the specified criteria (e.g., the allocated budget, the addressee’s personal preferences with regard to locations, dietary needs, dates of the trip) – the texts still reveal an unusual amount of linguistic variation in how the task was approached and employ a wide variety of linguistic means on the cline from more oral elaborated to more informational discourse. Furthermore, the position of these peripheral texts above or below the mean in business memos suggests that deviations from the central tendency are equally likely in both directions, and business memos range from markedly informational (since the mean in business memos is in the negative range, these peripheral texts are extreme examples of information density) to quite elaborated (a

deviation that runs counter to the overall informational tendency of the register). These results show that despite the clear informational focus in this register overall ($M = -1.01$; $SD = 0.79$), there is a wide scope for variation in business memos with regard to Dimension 1 features.

Business memos are followed by image descriptions, emails, and essays – the registers in which texts also appear to vary quite widely. These registers all have several texts in the ‘Peripheral’ groups, although their number is not as high as it is in business memos. Table 6.1 shows that variation in image descriptions is again in both directions – an equal number of texts in the ‘Peripheral’ distance groups is positioned above and below the mean, indicating that image descriptions, while overall being an oral and elaborated register ($M = 0.44$; $SD = 0.51$), allow both possibilities, namely quite elaborated texts with features of oral discourse as well as texts with features of information density. The peripheral texts in essays ($M = -0.24$; $SD = 0.27$) and emails ($M = 0$; $SD = 0.38$), on the other hand, appear to gravitate towards information density (5 peripheral essays positioned below the mean and none above the mean; 7 peripheral emails positioned below the mean and 2 above the mean). While the texts in the ‘Moderately Removed’ distance groups (0.3-0.6) of these registers do show a tendency towards oral elaboration, it appears that the pattern of extreme deviation in these registers (shown by the two ‘Peripheral’ distance groups) is associated with information density.

Evaluations ($M = 0.15$; $SD = 0.24$) and interviews ($M = 1.62$; $SD = 0.21$) do not have any texts in the ‘Peripheral’ distance groups and are thus considerably less varied. Table 6.1 shows that most of the texts in these registers are quite close to the mean – found in ‘Central’ (<0.1) and the ‘Minimally Removed’ (0.1-0.3) distance groups. Thus, unlike the other registers, evaluations and interviews do not seem to permit as much communicative freedom, and texts are relatively uniform in their linguistic make-up.

The pattern emerging from this classification, namely some registers allowing extensive variation and others being considerably less varied, is not surprising if the registers are analyzed with regard to the situational parameters that define their texts. Situational analysis of the texts within each register is thus the next step taken by the study. In accordance with the notion of functional correspondence between situation and language (Biber & Egbert, 2023), this situational analysis of texts is expected to inform the observed linguistic variation across texts. That is, if texts display linguistic variation, there is reason to believe that this linguistic variation is explained through the value of a particular situational parameter. The situational parameters that may play a part in linguistic variation in each register are the subject of the next section.

6.1.1 Application of a Framework for Situational Analysis

Proposing the situational framework for register analysis, Biber (1994) discusses the different levels of generality or specificity of registers. This distinction hinges on the number of situational parameters that are possible to define for different registers. Registers are not “equally well-defined” in their situational and linguistic characteristics (Biber, 1994, p. 36); rather, more general registers are likely to be undefined with regard to a larger number of situational parameters, while the more specific registers usually are more well-defined situationally, with the values for a larger number of parameters specified. For example, applying the framework to texts of expository prose, it is possible to specify that time and space are not shared between the participants, that the mode is written, production is offline, allowing opportunities for editing and revision, and that the main purpose of such texts is typically to inform the addressee. Such texts, therefore, show substantial internal variation and encompass a wide variety of more specific registers from academic prose to business reports or newspaper articles. However, for a more specific register, such as written academic prose, in addition to the parameters above, it is also

possible to specify the domain (always academic) and the level of discussion (always dealing with specialized topics). If an even more specific register is analyzed, such as a psychology article, the additional situational parameters that can be specified are the participants and subject matter – that is, readers and writers versed in psychology. This application of the framework thus suggests that registers on a higher level of generality, such as expository or even academic prose, are likely to encompass a wider range of variation than registers on a much lower level of generality, such as academic articles and, even more specifically, academic articles in psychology. This comprehensive framework for situational analysis thus enables us to compare registers in terms of the degree of generality, which in turn explains the amount of internal variation they permit.

The study draws on the initial version of the framework (Biber, 1994) and its subsequent renditions (Biber & Conrad, 2019) presented in Chapter 2 (Table 2.1) and analyzes the six registers of interest in terms of the following situational parameters: the addressee (singular, plural, or unenumerated), the relationship between the participants (personal, professional, academic), and their relative status (lower => higher or equal), the degree of interactiveness (low, medium, or high), the degree of shared knowledge between the participants (low, medium, or high) and its nature (specialist or general), the domain (school, work, or personal), the setting (public or private) and whether the time and space are shared by the participants, the mode (written or spoken), the specific medium (permanent, such as printed or recorded and transcribed, or transient, such as face-to-face real-time communication), production and comprehension circumstances (immediate online communication or offline with opportunities for editing and revision), general and specific purpose (open parameters, i.e., a unique value identified for each text rather than a choice made from limited possibilities), and topic (also an

open parameter). Table 6.2 below summarizes the application of this framework to the registers of the corpus.

It is important to note that the values of the situational factors in Table 6.2 have been identified for the specific registers in the corpus rather than for these registers generally. This is an especially important distinction in the case of some registers. For example, the essays or emails contained in the corpus are naturally occurring registers (the participants of the Aston ‘100 Idiolects’ project submitted their actual essays and emails produced in real-world contexts). Thus, the situational characteristics of these registers in this corpus may more closely reflect the situational characteristics of other academic essays or other (mostly school) emails. In contrast, the texts in the register of business memos are a response to a scenario and thus do not represent the naturally-occurring register. It becomes clear from Table 6.2 that the situational parameters that are possible to specify for these texts differ from the ones that would define naturally-occurring business memos. Specifically, the business memos in the corpus are written on a known topic (business trip to Helsinki), which is further restricted by highly specific search criteria included in the task (e.g., the kind of hotel preferred, the dietary restrictions, etc.), and are addressed to a particular audience (the imagined CEO). This information is not available if business memos in the real-world context are analyzed with respect to these parameters. While all business memos are written within a work domain and in a professional (corporate) context, topics as well as who exactly the texts are meant for may vary widely.

Similarly, the registers of evaluations, interviews, and image descriptions are elicited rather than naturally occurring, which is reflected in the values for some situational parameters. The interviews in the corpus, for example, were quite restricted in that the topic was held constant across interviews, and it is apparent from the responses that even the questions did not

vary from one participant to another. This, of course, makes the interviews in this corpus much more specific than the register of interviews could be otherwise, in which case topic and even the communicative purpose could be variable. In the case of evaluations and image descriptions, which arguably do not exist in the natural context (unless they occur within another natural register – e.g., a task in a foreign language class), the situational parameters were specified based on the task presented to the participants.

Table 6.2. Application of the framework for situational register analysis (Biber 1994; Biber & Conrad, 2019)

Register	Addressee	Relationship	Relative status/power	Interactiveness	Shared knowledge	Domain	Setting; Shared time/ place	Mode	Specific medium	Production & comprehension	General Purpose	Specific purposes	Topic
Essays	Sg (course instructor)	Academic	Lower >higher	Low	High (specialist)	School	(Semi)public /no	Written	Permanent electronic	Revised & edited/ careful reading	–	–	–
Emails	Sg	1.– (undefined for non-academic emails) 2.academic	1.– (undefined for non-academic emails) 2.Lower > higher	High	1.— (undefined for non-academic emails) 2. med/high	1. – (undefined for non-academic emails) 2. school	Private/ no	Written	Permanent electronic	Revised & edited/ careful reading	–	–	–
Interviews	Sg	– (prof/academic or 1 st time encounter)	– (lower >higher or equal)	High	Low (general)	Research project	Public/ yes	Spoken	Transient face-to-face	Online	Answer questions/ inform	Share personal experience, give examples	Food, eating habits
Descriptions	Sg/pl	Participant-researcher	– (lower >higher or equal)	Low	Tied to the visual	Research project	Public/ no	Spoken	Permanent recorded	Online production/ offline comprehension	Describe	Other more specific purposes are unlikely	Content of the visual
Evaluations	Sg/pl	Participant-researcher	– (lower >higher or equal)	Low	– (general)	School, research project	(Semi)public /no	Written	Permanent electronic	Revised & edited/ careful reading	Describe Evaluate	–	Aston campus facilities
Business memos	Sg	Professional	Lower >higher	Low	High (general)	Work	Private/ no	Written	Permanent electronic	Revised & edited/ careful reading	Inform	–	CEO's trip to Helsinki

In the context of this study, whose goal at this stage is to identify the basis for the observed linguistic variation across texts within each register, this framework serves as a starting point for analysis. The study first applies this framework to the six registers and specifies the value of each parameter for the texts in these registers if this is possible. That is, the values of some situational parameters are expected to be the same for all texts of a register – e.g., all essays are written texts. Mode (written or spoken) is therefore held constant for all texts within essays. In other cases, however, it is not possible to say that all texts within a register are the same with regard to a particular parameter. For example, essays vary in topic and purpose. These parameters are then variable rather than constant for that register. These variable slots in Table 6.2, i.e., the situational parameters that may vary across texts of a register, provide a natural starting point in the search for possible reasons for linguistic variation. If we assume that situational variation corresponds to linguistic variation – the key tenet of register research – it is likely that identification of variable situational parameters may explain the observed linguistic variation across texts. That is, if a certain situational parameter varies across texts of a register, it would follow that variation in that situational parameter may correspond to the linguistic variation across texts of that register.

For example, in the case of essays and emails, the variable parameters are purpose and topic. This suggests that the observed linguistic variation within these registers may be due to textual variation on the basis of these parameters as essays and emails fulfil different communicative goals and address a variety of topics. In the case of evaluations, while the general purpose was possible to identify for all texts of the register – to describe and express stance – specific purposes within evaluations may still vary.

The framework for situational analysis applied here (Biber, 1994; Biber & Conrad, 2019) has been used extensively for register analysis of a variety of domains and has proved especially useful in explaining between-register differences (e.g., overview of studies in Appendix 1 to Biber & Conrad, 2019). However, despite its comprehensive nature the framework does not present an exhaustive set of situational parameters that determine language use, and it is likely that additional factors, not accounted for by this existing framework, may be at play when variation within registers is examined. For example, Gray (2015) adapted the framework as she identified additional communicative factors accounting for variation in journal article types.

These considerations are key in the context of this study which examines variation across texts within registers. That is, the existing framework with previously established situational parameters may not be sufficient to explain variation within registers, and it may be additional situational factors that account for this register-internal linguistic variation. In the case of essays and emails, for example, while it is natural to assume that purpose and topic contribute to variation within these registers as these are the parameters that vary across texts (Table 6.2), it is still possible that other factors may be at play as well. An example of a register where the likelihood of such additional factors contributing to linguistic variation across texts is high is image descriptions. Table 6.2 shows that it is possible to identify the values of all the situational parameters of the existing framework for all texts of this register. This large number of specified parameters implies that this register is highly specific (Biber, 1994) and, consequently, should not exhibit as much internal variation as registers with a large number of unspecified parameters. However, this is not the case for the image descriptions in this study – a highly varied register, with peripheral texts both above and below the mean (Table 6.1). The existing framework thus

does not provide a basis for such extensive linguistic variation and points to the fact that the reason for the observed linguistic variation across texts lies elsewhere.

The following sub-sections present the results of text analysis, whose goal is to identify the basis for such variation across texts of a register. It will be shown that these situational reasons for linguistic variation across texts vary among registers. This analysis will focus on four of the six registers: essays, emails, image descriptions, and business memos, as these registers exhibited wide internal variation across their texts. Interviews and evaluations, on the other hand, did not reveal extensive variation across texts (as shown by a lack of texts in the ‘Peripheral’ distance groups – Table 6.1) and are therefore not discussed further.

6.1.2 Variation within Essays

Table 6.2 shows that for academic essays the addressee is single if the addressee is the instructor teaching the class but may be unenumerated if the audience includes any other unknown addressees (e.g., classmates or students of future classes using the essay as a sample, researchers working with the corpus, etc.). Since the most immediate and the most likely audience is the instructor, the values of the other parameters are specified for that option. The relationship is thus labeled ‘academic’, and the general domain is labeled ‘school’. The status of the author of the essay is thus lower relative to the status of the addressee. The level of interactiveness is low as the addressee is not in any way involved in the construction of the text. The knowledge shared by the participants is specialist as all the essays are written on a topic covered in a college class in (cognitive or behavioral) psychology, politics and sociology, business and marketing, or law. The degree of this shared specialist knowledge is high. The setting of this communication is public, but at the same time is limited to a specific class and as a rule is not shared with other participants. The time and space are not shared by the interactants; the mode is written, and the

specific medium is permanent electronic. The production circumstances thus allow ample opportunity for planning, revision, and editing. All the discussed situational parameters are held constant for all texts of the register. That is, it is safe to assume that the values for the relationship between the author and the addressee, the level of interactivity, degree and nature of the shared knowledge, the setting, the mode, the specific medium, and the production circumstances are the same for all the academic essays in the corpus.

This is not the case, however, for the purposes of these texts and their topics. It is not possible to say that all essays set the exact same communicative purpose (e.g., that of informing the audience). Rather, it becomes apparent from text analysis that a variety of purposes are set in essays. The same is true with regard to the essay topics. It is not possible to specify the topic of all academic essays in the corpus as texts are written on a range of topics. Purpose and topic are therefore the variable situational parameters in the register. These variable situational parameters become the focus of this study in the following analyses as the factors that could explain the linguistic variation within essays.

These analyses consisted in recording the purpose and topic for each essay in the three groups of texts of interest to the study: texts that represent the central tendency of the register (the 'Central' distance group), texts that are considerably higher on the dimension than the register mean (texts *above* the mean, which in the case of essays are texts in the 'Moderately Removed' group), and texts that are considerably lower on the dimension than the register mean (texts *below* the mean in the 'Peripheral 1' distance group and 'Moderately Removed' group, the latter inspected for additional evidence). Texts that are above the register mean on Dimension 1 show a trend towards increased oral elaboration, while texts that are below the register mean

gravitate towards information density. Table 6.3 below contains the number of essays in these groups.

Table 6.3. Number of essays per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density.

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Essays ($M = -0.24$; $SD = 0.27$)	Above M	12	30	12		
	Below M	21	25	6	5	

As purpose and topic were recorded for texts in the three groups of interest and the assigned purpose and topic labels were inspected, an interesting pattern emerged in the way essay topics clustered on the dimension. These topics are summarized in Table 6.4 and are organized by distance group and direction of deviation. While the essays in the corpus are not annotated for any additional characteristics, such as discipline, a discipline label was assigned in the process of this analysis (also shown in the Table). The essays in Table 6.4 are thus grouped according to discipline the essays appeared to fit the most closely, with the concrete topics listed under each discipline. The identified disciplines include cognitive and behavioral psychology, sociology, politics and sociology, law, and business and marketing.

Table 6.4. Essay topics within each distance and direction group. Dimension 1: Oral Elaborations vs. Information Density

Distance group	Above M	Below M
Central tendency (<0.1)	1. Sociology – 3 essays <ul style="list-style-type: none"> • Mass protests – theories, causes, evidence • What is religion (in sociology and international relations) 	1. Sociology – 5 essays <ul style="list-style-type: none"> • Job protection, pay, working conditions of house girls (in Nigeria and generally) • Sport, hyper-consumerist culture & globalization

Distance group	Above <i>M</i>	Below <i>M</i>
	<ul style="list-style-type: none"> • Employment relations; gender equality in the work place 	<ul style="list-style-type: none"> • Covid-19 & its repercussions • Gender stereotyping & roles; social & cultural differences • Mental & physical health during Covid-19
	<p>2.Politics/ sociology – 2 essays</p> <ul style="list-style-type: none"> • Emergence of the modern state, Machiavelli’s views applied to modern politics X 2 	<p>2.Politics/ sociology – 3 essays</p> <ul style="list-style-type: none"> • Liberalism, its types, and key thinkers • Gender in global politics, feminist theory, international relations; behaviorist & social constructivist perspectives • Costs and benefits of immigration
	<p>3. (Cognitive/ behavioral) psychology – 1 essay</p> <ul style="list-style-type: none"> • Case study of a specific patient’s condition (asthma, sleep deprivation, stress, anxiety, depression, isolation) 	<p>3.(Cognitive/ behavioral) psychology – 7 essays</p> <ul style="list-style-type: none"> • Autism spectrum disorder; efficiency of cognitive interviews with AI patients • Characteristic features of and key theories explaining the Autism spectrum disorder • Post-traumatic stress disorder, Obsessive-compulsive disorder, dependence as a reason for substance abuse (biological & cognitive effects) • Maternal depression, risks for child development, parenting styles, types of reasoning • Biological & Psychoanalytical approaches to behavior; determinism (x 3)
	<p>4.Law – 4 essays</p> <ul style="list-style-type: none"> • Good faith obligations in commercial law (x 2) • Claim against a former employer (employee’s circumstances, reasons for claim, company policies) 	<p>4.Law – 3 essays</p> <ul style="list-style-type: none"> • UK international sales laws • Good faith obligations in commercial law • Claim against a former employer (employee’s circumstances, reasons for claim, company policies)

Distance group	Above <i>M</i>	Below <i>M</i>
Moderately removed (0.3-0.6)	<ul style="list-style-type: none"> • Analysis of specific court case & court ruling: appeal of murder case <p>5.Business/ Marketing – 2 essays</p> <ul style="list-style-type: none"> • Gaming industry: eSports entering new markers • Company performance assessment (customer service at Kenya Airways) <p>1.Sociology – 10 essays</p> <ul style="list-style-type: none"> • Male & female language use (hedges & intensifiers, positive/negative adj) in cooking competitions; situation formality; politeness theory; dominance theory • Gender & sexual power; gender identity • Race as socially constructed reality • Leadership, human behavior, interpersonal processes, personal approach to team management • Stress, depression & anxiety among students during Covid-19 • Virtue & vice as fundamental elements of human nature; ethical & moral duties; religion & law within morality systems; ideals of happiness • Social media & its effects on self-perception & personal relationships 	<p>5.Business/ Marketing – 3 essays</p> <ul style="list-style-type: none"> • Marketing strategies of 2 brands (evaluation & proposal to strengthen the strategies) • Corporate social responsibility (CSR) and political influence; CSR as a means of constituency building and facilitating access to policy elites • Middle management of a company & its roles <p>1.(Cognitive/ behavioral) psychology/ neurobiology – 6 essays</p> <ul style="list-style-type: none"> • Developmental dyslexia, core deficit theories, temporal processing hypothesis, magnocellular theory • Core assumptions and key features of psychoanalytical and biological perspectives in psychology • Alpha Inhibition Hypothesis; disengagement of task-irrelevant brain areas • Adolescent self-inflicted harmful behaviors. Stroop effect • Word reading automaticity (tested through (in)congruence of presentation & meaning) • Adolescent psychopathologies

Distance group	Above <i>M</i>	Below <i>M</i>
	<ul style="list-style-type: none"> • Breach of ethics: FIFA’s response to corruption scandal & bribery allegations • Social media & societal norms (self-esteem, self-worth, and gender norms) • Experiences during Covid-19 – qualitative analysis of interview data <p>2. Marketing/ Business – 1 essay</p> <ul style="list-style-type: none"> • Workplace conflict & employee motivation <p>3. Law – 1 essay</p> <ul style="list-style-type: none"> • English contractual law in relation to good faith obligations 	
Peripheral 1 (0.6 – 0.9)		<p>1.(Cognitive/ behavioral) Psychology – 5 essays</p> <ul style="list-style-type: none"> • Neuroticism & unhealthy eating • Dual-systems modeling • Visual attention task, anticipatory attention, attentional selection and suppression • Alpha Inhibition Hypothesis; disengagement of task-irrelevant brain areas • Alpha Inhibition Hypothesis; visual attention task

Direction of deviation was not accounted for in the central group. The distance between individual text scores and the register mean is minimal in this group (< 0.1); inspecting the direction of this deviation (above or below the mean) is therefore not informative, and there is no demonstrable difference between texts above or below the mean. Rather, this group representing the central tendency of the register is viewed holistically as texts conforming to the general situational characteristics of the register and not showing any linguistic or situational

distinctions. Table 6.4 shows that texts of the ‘Central’ distance group come from a variety of topics from a variety of disciplines, and the same disciplines occur on either side of the mean. Furthermore, within these disciplines the same topics are often found above or below the register mean. There are thus no identifiable patterns with regard to topic-purpose distinctions in either direction in the texts of the ‘Central’ group.

When deviation from central tendency in the register is inspected, however, the emerging pattern reveals a quite stark contrast between the disciplines of sociology and psychology. It can be seen in Table 6.4 that 10 of the 12 essays in the ‘Moderately Removed’ distance group *above* the register mean are essays on topics in sociology, while all the essays in the ‘Peripheral’ group ($N = 5$) as well as all the essays in the ‘Moderately Removed’ group ($N = 6$) *below* the register mean are topics in cognitive or behavioral psychology. These results suggest that essays in sociology consistently require considerably more features of oral elaboration, while essays in psychology clearly prioritize informational features and are markedly less oral and elaborated.

It is useful at this point to turn to previous research that has offered insights into disciplinary distinctions and the linguistic differences they entail. Discipline has been investigated extensively and found to be a predictor of linguistic variation within university textbooks (e.g., Conrad, 1996; Biber, Conrad, Reppen, Byrd & Helt, 2002; Egbert, 2015; Egbert & Gracheva, 2023), journal articles (Gray, 2015), and university student essays (Gardner, Nesi & Biber, 2019). All of these studies have used dimensions of variation either identified by Biber (1988) or by their own analyses as linguistic variables.

The distinction between disciplines is often made on the basis of their characterization as ‘soft’ or ‘hard’ sciences (e.g., Hyland, 2002a; Swales, Ahmad, Chang, Chavez, Dressen & Seymour, 1998, cited in Gray, 2015, p.12). This distinction does reveal important linguistic

differences between disciplines: specifically, with regard to Dimension 1 features, Gardner et al. (2019) note that hard sciences have been shown to call for information density, while soft sciences are more involved. Hardy and Roemer (2013, cited in Gardner et al., 2019), for example, observe such contrasts in student writing (Michigan Corpus of Upper-level Student Papers) in philosophy and physics.

However, on this basis, psychology and sociology are both grouped into a single category of ‘soft’ sciences and do not represent these opposite extremes. Rather, the subject matter of these disciplines suggests a common ground between them. In the BAWE corpus, used in Gardner et al.’s (2019) study, however, the disciplinary distinction is much more fine-grained. Psychology is classified in a group of Life Sciences along with agriculture, biology, food science, and health, while sociology is categorized as a Social Science together with business, economics, law, and politics. While to an extent the disciplines of psychology and sociology share a research agenda, both focusing on individual characteristics and their interaction with those of others, Gardner et al. identify linguistic differences in these two relatively close fields of study. In fact, the results observed by this study mirror those of Gardner et al. (2019) quite closely.

Dimension 1 in Gardner et al. (2019), interpreted as ‘Compressed Procedural Information vs. Stance towards the Work of Others’, resembles Dimension 1 of this study and the typically identified oral/literate divide. Specifically, the ‘Compressed Procedural Information’ end of that dimension is comprised of nominal features of information density, pre-modifying nouns (also conveying informational focus in this study) being the feature most strongly associated with the dimension, while ‘Stance towards the Work of Others’ contains verbs and adverbial features, which contribute to oral elaboration on Dimension 1 of this study. Gardner et al. (2019) find that

Life Sciences, which include psychology, occur quite high on the informational end, requiring substantial information density, while Social Sciences, which include sociology, gravitate toward the other pole. The relative position of these disciplines on the dimension thus reflects the same trend observed in the present study – namely, psychology incorporating more informational features and sociology calling for increased oral elaboration. As the dimension is discussed, however, the contrast between these particular disciplines is not examined by Gardner et al.’s study.

To explore further what accounts for these disciplinary differences, the present study focuses on the specific topics within each discipline listed in Table 6.4. It was mentioned earlier that topic was found to be the most informative situational parameter as texts within the three distance groups of interest were analyzed. It should be acknowledged that while topic is the main focus, it becomes clear from this analysis that topics do not exist in isolation from communicative purpose, and while a list of identified purposes for each essay did not reveal a pattern as clearly interpretable as was shown by topic, a close interrelationship between topic and purpose will be discussed later in this chapter and further in Chapter 8.

The topical distinction that emerges from Table 6.4 appears to be that between more and less personal topics (or cohesive groups of topics that may be thought of as ‘topic types’). Topics more closely related to human experience, human qualities, life skills, and behavior guidelines are predominantly found within sociology, while psychology features topics at the center of which is the scientific enterprise. One major contributing factor to this disciplinary distinction (at least clearly observed in the corpus of this study) appears to be the likelihood of empirical research conducted on the subject matter studied in the discipline and the use of scientific approaches to explaining the phenomena observed through scientific testing (e.g., psychoanalytical and

biological approaches to a study of behavior). Such content is prominent in cognitive or behavioral psychology but rare in sociology essays, where topics involve theorizing and abstract thought (e.g., key thinkers in Liberalism and their vision of social arrangement). As a result, empirical research is rare on such matters, which do not seem to lend themselves to concrete measurement, objective judgment, or empirical testing.

Psychology essays only feature topics rooted in scientific matters – for example, an implementation of the visual attention task, anticipatory attention and attentional selection and suppression, Alpha Inhibition Hypothesis, or empirical testing of the relationship between neuroticism and eating habits. Such topics do investigate the human condition and in that sense provide insights into human nature. However, while the discussed hypotheses and the end results of the scientific tests shed light on human condition, these topics do not involve personal subject matter immediately relatable to an audience with no or little specialized knowledge of cognitive processes. Reaching such a reader by relating the concepts to daily experiences is not the goal of these texts and not something the author could do. Instead, these topics involve focused discussions of cognitive phenomena such as brain regions and brain activity, the steps taken in an experiment targeting a particular construct (e.g., measured through reaction times in response to a stimulus), and an interpretation of findings. Alternatively, such essays may discuss the causes and key characteristics of cognitive phenomena postulated by approaches to a study of psychology. These essays make use of terminology denoting complex cognitive processes, scientific theories, and research procedures. It is this scientific, technical focus that in the results of this study is associated with increased information density.

In stark contrast are the topics in sociology, all of which elucidate daily human experience or matters that constitute the background knowledge of a general audience and lend

themselves to a discussion rather than a scientific investigation. Gender and race dynamics, leadership qualities formed through behavior, increase in stress and anxiety in response to the pandemic, questions of ethics and morality as key foundations of human social experience and instances of breach of ethics, the effects of social media on self-perception and self-esteem, and the societal norms perpetuated by social media are all topics at the center of numerous discussions in today's society, readily accessible to a general unspecialized audience as well as the author writing about them, and overall more personal than the scientific matter at the core of the psychology essays in the corpus. This personal nature of the topics lends itself to increased oral elaboration and does not necessitate dense information presentation to the same extent.

A study that lends additional support to this interpretation is Egbert and Gracheva (2023), which observes a rather strong effect of discipline ($R^2 = 0.28$) within the register of introductory university textbooks on Biber's (1988) Dimension 1 ('Involved vs. Informational Production'), whose structure mirrors Dimension 1 of the present study. The disciplines of that study, geology and psychology, represent a distinct contrast in terms of their typical focus and research agenda. The exact position of the disciplines, geology and psychology, on the dimension also mirrors the conclusions of the present analysis regarding the nature of the topics. Psychology ($M = -8.22$; $SD = 6.94$) is positioned much higher on the dimension, showing a relative preference for oral involved production, while geology ($M = -13.99$; $SD = 5.82$) is noticeably more informational. The higher position of psychology on the dimension should not, however, appear contradictory to the results presented here, as the psychology essays in the present corpus with their strong scientific focus are in stark contrast with the passages from introductory psychology textbooks. While the psychology essays in this corpus are impersonal, technical, and focused solely on the research procedures they describe, hypotheses, and key concepts in the discipline, introductory

psychology textbooks explain every topic (personality, perception, learning, motivation, and stress) through references to daily experiences familiar to the reader, relatable examples, and personal addresses that invite the reader to experiment and find out the described effect for themselves. Expectedly, the personal nature of these texts results in higher Dimension 1 scores. Texts in geology, on the other hand, do not integrate such references due to the overall less personal subject matter and the lack of connection with the reader's or writer's everyday experiences. The topics of these texts are far removed from human experience and include plate tectonics, rivers, glaciers, metamorphic rocks, and radiometric dating. Involvement is not characteristic for these texts, which are much more informationally dense. Overall, these results not only support the fact that the effect of discipline is apparent within a register (academic essays as well as introductory textbooks), but also lend additional evidence to the basis of the observed disciplinary distinction, namely the personal nature of their subject matter.

So far, these topical differences between personal and less personal, technical topics have been discussed with relation to particular disciplines consistently calling for increased oral elaboration or placing emphasis on information presentation. The study finds, however, that the effect of topic extends beyond these particular disciplinary distinctions as essays from other disciplines are analyzed – Marketing and Business and Law – both found above the mean in the 'Moderately Removed' group. It will be shown in the analyses below that in these disciplines too the authors draw on their personal judgments, interpretations, and views, which calls for features of oral elaboration. The essays from these disciplines thus overall support the fact that it is the nature of the topics that appears to determine the tendency to oral elaboration and information density in the register.

Thus, the study finds that the disciplinary distinctions explain the emerging pattern only in part. On the one hand, the underlying reason for the linguistic differences appears to be the possibility for empirical research or different degrees of empirical focus in a particular discipline. On the other hand, however, a related question is what accounts for the possibility of empirical research in a discipline. While this question is not addressed here directly (i.e., there is no evidence presented for a lack of empirical research in sociology, marketing and business, or law), it is suggested that it is differences in the specific topics these disciplines commonly address that result in the observed trends. That is, topics of a more personal nature that draw on life experience and general background knowledge are more likely to lend themselves to oral elaboration than topics far removed from such experiences. This distinction should not be bound to a particular discipline, but rather could be found within any disciplines (as in the above examples of philosophy and physics from Hardy & Roemer, 2013, or psychology and geology in Egbert & Gracheva, 2023). Furthermore, it will be illustrated later in this chapter as variation within emails is examined that the pattern of topic determining the balance between oral elaborated and informational features extends beyond academic discourse and therefore beyond disciplinary distinctions.

It was stated earlier, however, that topics are not isolated from communicative goals. In fact, the analysis showed that it is more likely that a particular topic is presented through an associated set of communicative goals that can effectively convey it, and particular goals tend to be set with regard to specific topics. For example, the text analysis in the next section will show that the goal to discuss, analyze, and evaluate appears to be associated with the topics which tend to draw on human experience and thus lend themselves more readily to a discussion, analysis, and evaluation. Conversely, highly specialized topics that are technical in nature and thus

removed from daily experience do not seem to co-occur with the goals to discuss, analyze, evaluate as often, or at least, the type of analysis in psychology essays is not subjective, but rather resembles a research synthesis. The goals accomplished in those essays were typically to report results or present already established findings. These patterns thus suggest that certain topic types (personal or technical) entail certain communicative purposes, and this association between the topic type and communicative purpose determines the extent of oral elaboration and information density in academic essays.

The next section presents detailed linguistic analysis of essays found in the central group and groups deviating from the central tendency in both directions, illustrating the patterns discussed above. –

Text Analysis: Central and Peripheral Essays above and below the Register Mean

Central Tendency

Text Sample 56 is an example of an essay at the intersection of Politics and Sociology on the topic of the origins, schools of thoughts, and the modern trends in liberalism. The passage makes use of several nominal sequences, which refer to the key concepts in politics and social studies and are central for the topic (*market liberalism; working class; rationalist assumptions*). On the other hand, features of elaboration, mainly the present tense and demonstrative pronouns, are used to state the beliefs of the prominent thinkers and then refer to these beliefs in the following discussion. This essay overall conforms to the expectations of the register – it informs the reader by presenting and problematizing a complex topic, discussing the relevant theories, analyzing the relationships between them, and drawing conclusions. That goal is achieved through a balance of features of oral elaboration and information density, with features of oral elaboration used to

describe the state of the art in the field and connect ideas and informational features integrating the necessary complex constructs central to the topic.

Text Sample 56. Author 9, sociology essay – central (text id: 10983), Dimension 1 score: –0.28

Karl Marx, a well-known socialist, was known to criticise liberalism – we can see Marx’s view on the state in his book *The German Ideology* (1846) as he **claims** the state **is** a creature of the bourgeois economic interest. When looking at contradictions such as free market liberalism, industrialization and urbanization **it is** easy to understand why he believed **this**, Marx believed that liberalism was a cover for the economic exploitation of the working class. “The executive of the modern state **is** nothing but a committee for managing the common affairs on the whole bourgeoisie.” (Marx:1846) Due to responses like **these** from Marx and other socialists, a new strand of liberal thought was introduced known as social liberalism, new liberalism, modern liberalism and sometimes **even** welfare liberalism – Thomas Hill Green was the reason for **this**. Thomas Hill Green (1836-82) was a British philosopher who rejected utilitarianism and rationalist assumptions which were that humans **are** driven by self-interest emphasising instead the complex circumstances that **are** involved in the evolution of people’s moral character.

Peripheral Texts *above* the Register Mean

Text Samples 57-62, excerpts from essays occurring above the register mean (‘Moderately Removed’ distance group), exemplify the trend of personal topics necessitating oral elaboration.

Text Sample 57 is an excerpt from an essay discussing virtue and vice as fundamental elements of human nature, ethical and moral duties, the roles of religion and law within morality systems, and ideals of happiness. This essay presents the existing views on these matters (*Cunningham suggests*), analyzes the meaning behind these notions (*Though it seems easier to take a route which overlooks virtue, it is no way to live in the long-term, and to engage in corruptive behaviours compromises this*), and draws on concrete real-world examples (*ethical practices vary globally, and the desire to do good remains the same. For example, in the United States, shaking hands shows a sign of friendship, but within Asian societies, individuals bow to communicate their friendship and respect*). Features of oral elaboration (bolded) include present

tense, through which these general statements are made, conjuncts that explicate the relationships between ideas, clausal coordination used to develop ideas further and make logical links, and adverbs which qualify the propositions conveying critical information about the views discussed (*something that is reflected as being **inherently morally** evil*). Elaboration is thus not only possible in an essay of this kind, but aids the author in conveying the ideas of the key thinkers, providing examples, and relating abstract notions to specific experiences.

Sociology. ‘Moderately Removed’ Group *above* the Register Mean: 0.3-0.6

Text Sample 57. Author 51, sociology essay – above *M* (text id: 10907), Dimension 1 score: 0.09

However, it can be stated that ethical practices **vary globally**, and the desire to do good **remains** the same. **For example**, in the United States, shaking hands **shows** a sign of friendship, **but** within Asian societies, individuals **bow** to communicate their friendship and respect for one another. [...] To not ‘encroach on anyone’s behaviour’, one must identify which actions **constitutes** virtuousness. Cunningham **suggests** that although Virtue ethics **does** not provide a concrete structure for decision making, there **are** some scenarios which **provide** enough clarity to **suitably** reflect such ideologies. In such scenarios, one should acknowledge the virtues present and respect them whilst **also** acknowledging vices, such as corruption. **Holistically, even** if one course of action **seems** advantageous, to reflect on this personal narrative **is** most appropriate. Though **it seems** easier to take a route which **overlooks** virtue, **it is** no way to live in the long-term, and to engage in corruptive behaviours **compromises this**. To obtain a sense of inherent morality **is** a lifelong haul **and** virtue cannot be displayed on a selective basis. Virtues and vices **are quite** the opposite. Virtues **are** entrenched in honesty and truthfulness. **Therefore**, individuals within administrations that **engage** in dishonesty by taking bribes **for example, is** serious corruption; something that **is** reflected as being **inherently morally** evil.

The essay in Text Sample 58 discusses personality traits necessary for successful leadership and the relative contributions of nature and nurture into the development of such traits. Analyzing a previously discussed theory of personality development leads the author of this essay to use demonstrative pronouns for discourse organization, referring to various aspects of the theory and relating them to leadership. Present tense is again used to state the general truths (*it is often the case that some individuals **retain**...*), and adverbs again serve to qualify the

propositions and organize ideas (*which ultimately forms our personality; comparatively to Erikson's theorised age of 2-6*) as well as convey detail central for the discussed notions, i.e., how leadership skills are formed and what comprises them (*develop certain traits consciously; aptitude to think logically; Some are instinctively creative*). While the text does use features of information density to refer to important concepts – *personality traits, personality development, process knowledge, leadership role* – the essay's central goal is to analyze them and explicate the relationship between these concepts, which is achieved through elaboration.

Text Sample 58. Author 33, sociology essay – above M (text id: 10877), Dimension 1 score: 0.28

As a human being, we **are** born into different environments which **form** our psyche. Erikson (1950) theorised that the “Personality Development: Age 2–6”. **Therefore, this is** as good as being **innately** born with personality traits, **as** by this age we **have** no control **generally** of our environment, which **ultimately forms** our personality. **This is** relevant to leadership, **as** some traits can either be effective or a hinderance to being a successful leader. Despite **this however**, many people's developmental teenage / young adult years can serve them numerous learning curves, which **ultimately allow** them to develop certain traits **consciously, comparatively** to Erikson's theorised age of 2-6, where **it** will be **sub-consciously**. [...] Although, this knowledge can be picked up through experience with different types of people, **it is often** the case that some individuals **retain** certain behavioural process knowledge in a superior fashion to others. **Thus**, despite **this** being **slightly** less definitive in a person's understanding of what good Conceptual Skills **are, generally** we will find that **those** exposed to more types of people, will possess the best awareness of this. **Once again** suggesting that a leader **is** made by learning how others **behave**. Interpersonal skills **are** more **ambiguously** tied to an individual's 'made' skills. **As this is** a person's general analytical ability and aptitude to think **logically**. Some **are instinctively** creative in their problem-solving ability, suggesting that if a leadership role **requires** constant problem solving, a person may need to be 'born' with this trait.

The topic of the essay in Text Sample 59 is the effect of social media on self-perception and personal relationships, and the essay presents a discourse analysis of interview data with a research participant. Personal views, both the interviewee's and the author's, are therefore the primary focus throughout the text. Present tense is again the feature that contributes the most to expressing them, and clausal elaboration is commonly used to enhance and explicate ideas (*and*

this can affect an individual due to...). Conjuncts too serve the analytical purpose of such texts and introduce judgments and conclusions (*make them feel low, therefore affecting how they behave*). Since the text reports interpersonal content and the topic involves the author's and the interviewee's first-hand experiences, adverbs as well as some instances of present tense verbs acquire an emphatic tone (*how much social media is used and how it does cause problems; social media is constantly being used; particularly through cyber-bullying*). This expressivity appears key to the task addressed by the essay – identifying persistent discursive themes in the interviewee's views on a topic. The personal nature of these views resulted in the text's clear tendency towards involved oral expression.

Text Sample 59. Author 60, sociology essay – above M (text id: 10926), Dimension 1 score: 0.07

Eliza **understands** how **much** social media **is** used and how **it does** cause problems despite her lack of use of social media “I **rarely use** social media”. Eliza **portrays** the discursive object as social media and the discourse of social media being negative. Eliza **knows** that social media **is constantly** being used and that people on social media may behave **inappropriately** leading to serious consequences. She **understands** the effects **it** can have on a person and the reason why people may behave this way. Eliza **realises** that on social media people can act in a way that **is** not the same as how they **behave** in real life, **but** their actions can **ultimately** hurt the people they **love**. She **explores** that people **use** social media in the wrong way knowing that what they **are** doing **is** wrong however some people **also get** treated bad on social media for no reason **and this** can affect an individual due to increase of negativity. Eliza **points out** that not only **does** negativity affect communication between people but **also interferes** with how an individual **feels** about themselves, **particularly** through cyber-bullying which can destroy how someone might feel about themselves and make them feel low, **therefore** affecting how they behave.

The topic of the essay in Text Sample 60 again draws on the audience's experience with two phenomena the general public and the author are closely familiar with – social media and gender roles – as it discusses the construction of gender identity through social media-promoted stereotypes and the role of technology in the social shift in gender representation. Features of oral elaboration – present tense, clausal coordination, and adverbs – are used to construct the key

arguments (*In the west, society is largely led by choice, the choice of what to wear, what to eat, how to socialise, work, and live, but does this level of autonomy extend to gendered identity and sexual power*) and in many instances express stance (*society is largely led by choice; topics that are arguably very subjective*) towards societal norms. Conjuncts contribute to reasoning and discourse organization. Features of information packaging, on the other hand, while present (*media outlets/ platforms, gender identity*), are sparse, as it is the analysis of the role of media in modern gender portrayal that is at the center of the essay – a goal mostly served by oral elaboration of these notions rather than frequent references to them through nominal features.

Text Sample 60. Author 26, sociology essay – above M (text id: 10861), Dimension 1 score: 0.25

Not **just this, but** media outlets **also shine** a light on the current social climate of gender and sexual power from the likes of popularized magazines such as Cosmopolitan and social media platforms such as Instagram. Whilst both **have** their flaws, they **offer** a wide-ranging account of what these concepts **mean** to other people. **It is** safe to say that women and girls in the west **are** no longer bound to the ultra “feminine”, one-dimensional gender identity **once** placed upon them. Men **now also have** an array of diverse masculinities different to the conventional hegemonic image that was **once** the norm in western society. **However**, to what extent **are** these modern labels and power dynamics being adopted. In the west, society **is largely** led by choice, the choice of what to wear, what to eat, how to socialise, work, and live, **but does** this level of autonomy **extend** to gendered identity and sexual power. I will explore the extent to which gender and sexual power **is** changing in a postmodern society. I will investigate **this** [...] Gender and sexual power **are** topics that **are arguably very** subjective, definitions of both **often vary** depending on social groups or individuals.

Text Sample 61 is an excerpt from another essay conducting a discourse analysis of interview data – this time on students’ experiences during the pandemic. Again, both the author of the essay and the interviewee, whose position on the topic is the subject of the analysis, bring their first-hand experiences to the interview, which are then interpreted in the text. This analysis abounds in quite emphatic stance features, some of which are found in the excerpt (*Mia cannot help but call the circumstance unfortunate as it seems like it is the only key word that can truly*

capture what the feeling is like; WAY WAY better; announces how much happier Mia was before). The passage also features frequent instances of causative adverbial subordination (*as it seems like; and as this is an unchanging circumstance..., This is in contrast to how life was pre-pandemic as she describes it as being...*) as the author offers their interpretation of the interviewee's statements throughout the essay. This concomitant interpretation of the content, whether it is of interview data or ideas postulated by the key thinkers in the field, appears to be a prominent characteristic of essays on topics in sociology. The personal, relatable nature of such topics, which lend themselves to abstract theorizing rather than concrete presentation of specific findings or facts, appears conducive to such subjective interpretation. The features comprising oral elaboration are the natural linguistic means that contribute to that goal.

Text Sample 61. Author 73, sociology essay – above M (text id: 10947), Dimension 1 score: 0.17

Mia is unable to shake the ill feelings off **and** due to this, **it is** likely that Mia cannot help but call the circumstance unfortunate **as it seems** like **it is** the only key word that can **truly** capture what the feeling **is** like. Social interactions due to government lockdowns **has** meant that Mia **has** been unable to **socially** interact with others as she would have done **previously and as this is** an unchanging circumstance, Mia **has** likely come to terms with **it** and accepted **it** through the “unfortunate” being paired with “...**it is** what **it is**”. **It suggests** her experience **serves** some kind of ambiguity and **later elaborates** the effects of this ambiguity by saying the situation “**is** out of my control”, rendering her useless in a sense. Mia’s emphasis on the word “much” **denotes** a significant emotional toll on her, regardless of anything else that may be happening in her life. **This is** in contrast to how life was pre-pandemic **as she describes it** as being “**WAY WAY** better”, the expressed emphasis on “**WAY**” **announces how much** happier Mia was before **this** happened, showing the drastic shift in life experiences for Mia.

Marketing & Business. ‘Moderately Removed’ Group above the Register Mean: 0.3-0.6

Text Sample 62 is an evaluation of a specific company and its new leadership in terms of their efficiency in meeting employee needs and maintaining high levels of motivation. While many topics in Marketing and Business may require substantial amounts of specialist knowledge and

may be impossible to directly relate to personal experiences of a broad audience, the topic of this essay – employee motivation and needs, work conditions, and salary as factors that correlate with motivation – is likely to be relatable to many. It is not surprising then that the goal associated with this topic is that of analysis and evaluation – tasks that are fulfilled through oral elaboration. Additionally, part of the task addressed in this essay is giving recommendations and proposing ways for the new leadership of the company to improve employee satisfaction based on the evaluation of the company’s work. It is apparent that this essay does not make extensive use of complex terminology, does not report on any empirical investigations, and most of the nominal sequences (*employee motivation, employee engagement, job security, work conditions, company policies, employee needs*) denote notions expressing concepts the author is able to discuss relying on their own analysis of the company’s activity and interpretation of statements made by authorities in the field. This analysis calls for features of oral elaboration – present tense (used especially frequently due to the expression of recommendations), clausal coordination, causative subordinate clauses, demonstrative pronouns and the pronoun ‘it’, conjuncts, and other adverbs. While several of these features play an important part in the passage (causative adverbial clauses are used several times, for example, to justify statements), adverbs seem especially prominent, serving not only to organize ideas and provide the necessary modification (e.g., *work remotely*), but notably, to express stance and make the essay quite emphatic overall (*This is rather a very important component; Abi should get to know each of her employees individually as best as she can in order to truly understand*). The personal and rather emphatic nature of the essay may also lead the author to use features of inexplicit oral discourse, such as demonstrative pronouns (instead of explicitly naming the phenomena just discussed – *This is rather a very important component; This is because...; those who work remotely*), the pronoun ‘it’ (*the nature of the*

work *itself*; but if they *are* not present or accessible to the worker *it* could cause them to become unmotivated), and even contractions. Thus, similar to the sociology essays discussed earlier, this essay in Marketing and Business abounds in features of oral elaboration, all of which serve the goal of analysis and evaluation of a topic, whose understanding is to a substantial degree informed by personal experience.

Text Sample 62. Author 42, marketing and business essay – above *M* (text id: 10888), Dimension 1 score: 0.09

Firstly, I'd like to begin talking about employee motivation and **ultimately** employee engagement. **This is rather** a very important component for this business **because** there are two sets of employees, **those** who **work remotely**, home based, **and** we **also** have those employees that **work** on location at the business. [...] I **recommend** for the two sets of staff, **as** they **have** different factors and environments, which the two components of the theory can target. [...] Hygiene factors will not encourage staff to work **harder** or work **excessively** but if they **are** not present or accessible to the worker **it** could cause them to become unmotivated and not work **at all**. These factors **include** job security, work conditions, company policies or salary. Whereas 'Motivators', as put by psychologists Woods and West **are** "factors that **tap** into internal or intrinsic needs", **these include** recognition, promotion, achievement, responsibility at work and the nature of the work **itself**. One advantage of this model **is** that **it allows** managers to **better** direct their motivation and understand employee needs. Abi should get to know each of her employees **individually as best as** she can in order to **truly** understand the different factors that will motivate them to work and engage with business **blissfully** and **successfully**. **This is because** one motivational factor could work for one employee but wouldn't motivate another. **For example**, for an employee that **works** from home, working conditions wouldn't be as **much** of a motivating factor them as **it** would be for an employee that **works** on location **because** they **are** working from the comfort of their own home.

Law. 'Moderately Removed' Group *above* the Register Mean: 0.3-0.6

Another example of texts showing a tendency to oral elaboration comes from Law (Text Sample 63), which is an essay on the topic of good faith obligations in commercial law in Great Britain. In the essay, the author explains the principle of good faith and argues that the existing laws do not adequately implement it. The subject matter of the essay thus may appear heavily specialized and quite removed from personal experience. A closer look at the features of oral elaboration in

this essay shows, however, that the features contribute to the goal of argumentation (*It will be disputed in this assignment; This is the consistent position that the assignment will take*) and, similar to the essays in sociology and marketing and business, are used in the author's own analysis of the efficiency of the laws and their application to express stance. This subjective analysis is constructed through oral elaboration in the form of present tense and frequent use of adverbs in statements like *However, most important is the lack of statute and legal definition, making it merely impossible to take the principle seriously. Unfortunately, pockets of good faith-like principles are dissatisfactory in commercial law; Nevertheless, the law will penalise (in the form of remedies) parties who actively deceive the other party, thus, it may be questioned, what is the need for good faith and what purpose does it serve? or However, it has been explicitly stated numerous times*, in which features of elaboration directly contribute to the author's argument as they question the efficiency of the law or expose its flaws. This shows that the nature of the topic, which may at first seem specialized, lends itself to subjectivity and individual interpretations. Thus, again, it appears that matters such as human judgment, decisions made in specific court cases, and people's actions (*parties who actively deceive*) lend themselves to elaboration more readily than strictly scientific or technical topics.

Text Sample 63. Author 30, law essay – above M (text id: 10871), Dimension 1 score: 0.10

It will be disputed in this assignment that good faith, though **encourages** “honesty and integrity”, **does** not permit certainty and predictability (which **is** perceived as being more important than a just outcome). **This is** the consistent position that the assignment will take. [...] **It is** evident that there **is** a need for compulsory good faith in order prevent the risk of “falling behind its competitors” (international jurisdictions) and remain up-to-date in terms commercial law. **However**, most important **is** the lack of statute and legal definition, making **it merely** impossible to take the principle **seriously**. **Unfortunately**, pockets of good faith-like principles **are** dissatisfactory in commercial law and may be the catalyst for hostility. **Namely**, “English law **generally** steered clear of overreaching principles”- **for instance**, defining and measuring good faith and deciphering what

amounts to good faith, **this forms** the basis of speculation. **Nevertheless**, the law will penalise (in the form of remedies) parties who **actively deceive** the other party, **thus, it** may be questioned, what **is** the need for good faith and what purpose **does it serve?** (other than incurring further expense to the court system). **It is** for that reason; hostility towards good faith can be justified, most **importantly**, above the need for defeating other jurisdictions. **Moreover**, the basis of commercial law **is** certainty, whilst “the predictability of the legal outcome of a case **is** more important than absolute justice”. Good faith would ensure the contrary, justice at the expense of certainty. **However, it has been explicitly** stated numerous times that justice (**potentially** in the form of good faith) will not (**presently** and in the past) supersede the importance of a certain and predictable court system.

Peripheral Texts *below* the Register Mean

To illustrate the contrasting pattern of topics heavily focused on the scientific endeavor, several passages from essays in psychology are presented below. Text Sample 64 is an excerpt from an essay investigating the relationship between neuroticism and eating habits. Oral elaboration is unlikely on such a topic as, in view of the goal of such scientific reports, the author is not expected to and may not be able to discuss this topic without the support of experimental data, research, and concrete results. It is these external objects that comprise the essence of the essay rather than the author’s internal thought process. References to the research process, data, and results are all complex nominal sequences: *Neuroticism score* and *conscientiousness, openness, agreeableness score*, for example, refer to methods the author uses to operationalize the constructs; similarly, *eating habits, eating behavior, food consumption, and food categories*, repeated several times, are the phenomena targeted by the study or notions directly related to them, and *study results* refer to concrete measurements. It is thus concrete, objective findings that are the focus of such reports, and there is no room for the kind of speculation possible in essays whose topics are related to the author’s and audience’s personal experience.

Cognitive psychology. ‘Peripheral 1’ Group *below* the Register Mean: 0.6-0.9

Text Sample 64. Author 27, psychology essay – below *M* (text id: 10862), Dimension 1 score: –0.85

Researcher **predicts** neuroticism score will have a positive relationship with unhealthy eating habit; there **is** a negative correlation between conscientiousness, openness, agreeableness score and unhealthy eating habits. Finding from correlation and regression **shows** that there **is** a significant negative relationship of conscientiousness and unhealthy eating habit. There **are** no other evidence found between other personality and unhealthy eating behaviour. To look at the relationship between personality and eating behaviour, Keller and Siegrist (2015) study result **show** that neuroticism had a significant positive indirect effect on sweet and savoury food consumption, whereas conscientiousness had a significant negative indirect effect on consumption of food categories that **are** indicators of an unbalanced diet. **This shows** that there might be link between personality and food consumption and eating habits.

Text Sample 65 illustrates a different psychology topic from an essay with a similar goal – report on a research process and present research findings. The text is again heavily focused on the hypothesis tested (*alpha inhibition hypothesis*), the experimental task and tests applied (*attention task; Wilcoxon test*), the measured constructs (*mean alpha amplitudes; alpha range activity; response preparation cues*), and terms of physiology (*relevant brain areas*) – complex scientific or research-related notions conveyed through nominal sequences. While some features of oral elaboration are present, the function of some of these features, such as adverbs, changes as they are used to refer to qualify descriptions and processes related to research subjects and stimuli (*processed in the brain contralaterally; typically developing adolescents; will be actively processed in the left-hand side of the brain*) rather than to express emphatic author stance as in sociology essays. The author’s persona is not revealed in psychology essays, solely focused on the experimental procedure at hand.

Text Sample 65. Author 79, psychology essay – below *M* (text id: 10957), Dimension 1 score: –0.87

The alpha inhibition hypothesis was tested in the experiment **and it** was predicted that by recording from P3, there would be higher alpha amplitudes in cue left condition rather than right. A cued visual attention task was used where a participant was presented with cues and expected to respond according to the direction of the stimulus. After rectifying the data, the results found mean alpha amplitudes were higher in the left condition **and** the Wilcoxon test found the difference between left and right alpha range activity to be significant. The alpha inhibition hypothesis (AIH) **proposes** high alpha **is** indicative of

inhibition and disengagement of task irrelevant brain areas (Klimesch et al., 2007). **Therefore**, alpha is reduced in task relevant brain areas, and can be modulated by attention. Stimuli **are** processed in the brain **contralaterally**: stimuli presented (cued) on the right side will be **actively** processed in the left-hand side of the brain, **so** more alpha range activity on the right and vice versa. [...] participants showed alpha inhibition following response preparation cues **and** that **this** was highest for **typically** developing adolescents and lowest for ADHD inattentive participants. In (Sauseng et al., 2005) participants completed a cued visual attention task. When looking at valid trials alpha activity decreased at parietal sites contralateral to the hemifield stimuli were presented to.

Text Sample 66 is an excerpt from an essay overviewing the existing models which offer explanations for common heightened risk-taking behavior in adolescents. The essay does not report on concrete findings from a conducted experiment; however, frequent informational features are used to refer to the theoretical models in psychology (*dual-systems model*; *maturational imbalance model*), the cognitive processes explained by the models that underlie the discussed behavior (*regulatory inhibitory control process of the cognitive control system*; *development imbalance*; *decreased impulse control*), terms of physiology (*brain regions*), and the research process and its components (*future research*; *research findings*). Again, a topic so strictly tied to concrete scientific models and neurological processes, investigated and established through empirical research, is not conducive to subjective theorizing. Features of oral elaboration are minimal, with present tense in the general statements about the model and conjuncts serving to organize ideas being the only consistently used features.

Text Sample 66. Author 46, psychology essay – below M (text id: 10893), Dimension 1 score: -1.10

The dual-systems model (Steinberg, 2008) **attempts** to explain heightened risk-taking in adolescence through a discordance between the “early-maturing” socioemotional system and the protracted development of the cognitive control system. Adaptations such as the maturational imbalance model (Casey et al., 2011), and the driven dual-systems model (Luna & Wright, 2016), **also recognise** a hypersensitivity towards socioemotional stimuli and reward, without the regulatory inhibitory control process of the cognitive control system (Casey et al., 2008; Shulman et al., 2016). The structural development imbalance proposed by the dual-systems model **is** characterised by the reduction of cortical grey matter. [...] **Furthermore**, discrepancies in research findings of non-linear

developmental trajectories could be ascribed to variation of surface-based and volumetric approaches to measure brain regions (Vijayakumar et al., 2018; Winkler et al. 2009). Neurodevelopmental studies should not disregard the reliability of voxel-based morphometry, **rather** future research should focus on the use of more consistent measures. The prefrontal cortex (PFC), **is** associated with regulating affective processing in the amygdala, **thus** it's protracted development may lead to decreased impulse control and a propensity for risky behaviour.

Cognitive psychology. 'Moderately Removed' Group *below* the Register Mean: 0.3-0.6

The trend observed in the 'Peripheral' texts above is confirmed by psychology essays in the 'Moderately Removed' group, which add to the topical range in the peripheral group, but illustrate the same trend. Text Sample 67 below overviews existing theories of dyslexia and the causes for the condition proposed by each. References to these theories (*deficit theories; phonological deficit hypothesis; phonological deficit disorder; temporal processing hypothesis*) necessitate frequent nominal references. While features of elaboration are used to explicate the theories (*which says; it does not consider; it attributes*), this explication itself then entails more nominal references to phenomena comprising the essence of the theories (*risk factors; reading impairment, grapheme-phoneme correspondence for alphabet systems; processing speeds; underlying temporal processing deficit*). The focus of the texts is therefore still on the scientific matter, and the minimal elaboration that is present serves the goal of adding substance to the discussed theories rather than the goal of the author's own analysis, which consistently led to increased elaboration in sociology essays.

Text Sample 67. Author 32, psychology essay – below *M* (text id: 10873), Dimension 1 score: – 0.69

There **are** several core deficit theories that can be classed as risk factors for developing dyslexia. The most **widely** developed theory would be the phonological deficit hypothesis, which **says** a necessary symptom must be present for a positive diagnosis of developmental dyslexia. This hypothesis **states** that developmental dyslexia **is** characterized by an inability to learn and apply letter to sound, which in turn **causes** the reading impairment in people with dyslexia **as** they would not be able to learn the grapheme-phoneme correspondence for alphabet systems (Snowling, 1981). [...]

However, the issue with the phonological deficit disorder is that **it does** not consider other aspects of dyslexia, like the phenomenon of motor and sensory deficits [...]. Another theory that **tries** to explain developmental dyslexia is the temporal processing hypothesis, which **says** that developmental dyslexia is caused by deficits in underlying processing speeds and perceiving auditory signals that change rapidly (Tallal, 1980; De Martino et al, 2001). [...] Contrasting the phonological deficit hypothesis, **it attributes** manifestations of phonological deficits as an indicator of an underlying temporal processing deficit (Tallal, 1984).

Finally, Text Sample 68 from the 'Moderately Removed' group is a research report detailing a research procedure and presenting findings on word reading automaticity. As in the essays on technical topics discussed above, nominal features convey a prominent focus on the constructs under analysis (*Stroop effect; response time; (mean) reaction time*) and research process and results (*volunteer sampling; ranks test; font colours*). As before, the rare features of oral elaboration are involved in reporting results (*slower reaction time to incongruent words than congruent because the Stroop test creates a conflict...; highly statistically significant; Data shows that participants took longer to correctly identify the font colours*) rather than offering the author's subjective interpretation.

Text Sample 68. Author 81, psychology essay – below M (text id: 10963), Dimension 1 score: -0.66

The Stroop effect was introduced by John Ridley Stroop in 1935, the task **works** by examining the response time of the participant to name colors of words presented to them. **Most likely**, participants would have a slower reaction time to incongruent words than congruent **because** the Stroop test creates a conflict between an incongruent color and word. [...] An experiment was carried out on 59 participants and the results supported Stroop (1935), **it** was found that there was a delay in participants' reaction time when they were presented with incongruent words. There were **also** many factors which may have affected the results, some of the factors **are** the use of volunteer sampling and there were **also** extraneous variables such as time of day. [...] Table 1. Shows that mean reaction times were longer in the colour incongruent (CI) than the colour congruent (CC) condition. (M= 966.8, SD= 405.1). A Wilcoxon signed ranks test showed that this difference was **highly statistically significant** (T= 403, n= 59, p=< .001). Data **shows** that participants took longer to **correctly** identify the font colours when they were incongruent with the colour of the words written.

6.1.2 Variation within Emails

There are even more frequent instances of extreme deviations from the central tendency in the register of emails, as shown by the number of texts in the two ‘Peripheral’ groups (0.6-0.9 and > 0.9, Table 6.5). This section examines the texts in these groups with regard to direction of deviation – above the register mean, showing an inclination to increased oral elaboration, and below the mean, showing a preference for information density.

Table 6.5. Number of emails per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Emails ($M = 0$; $SD = 0.38$)	Above M	10	23	26	2	
	Below M	13	13	17	6	1

Emails are mainly written by the Aston students within the general domain of school and addressed to their professors or other authorities within the academic context; however, there are some exceptions (e.g., emails written to classmates, landlords, or unspecified entities). Since these exceptions are quite rare, the main focus of the situational analysis is on emails addressed to addressees within the university. The addressee in this case is typically a single person, the relationship is defined by the academic context, such as classes taken by the students or supervision provided to the students, and the status of the author is lower than that of the addressee. The level of interactiveness in emails, which directly address and involve the addressee, is high despite the texts’ monologic nature. The amount of knowledge shared by the participants is medium to high, and its degree and nature (specialist or general) largely depend on the content of the emails. For example, emails that share personal circumstances affecting the

student's academic progress aim to inform the addressee of something they do not yet know, while emails referring to the latest assignment and asking for advice on the appropriate sources assume a large amount of background knowledge, which in this case is partly specialist in nature. Emails that refer to the submission deadline of the latest assignment and describe the technical difficulty the student had with it rely on some amount of shared knowledge (that of the assignment at hand), but do not assume any prior knowledge about the technical difficulty. The content of the emails and their goals thus appears to determine the type and the amount of shared knowledge expected of the addressee. While emails are instances of professional communication in the academic context, this communication represents an exchange between two interactants, typically not shared with other entities, and in that is private rather than public. Time and space are not shared by the participants; the mode is written; the specific medium is permanent electronic; the production circumstances allow the possibility of planning, revision, and editing, although the frequency and speed of email communication in professional contexts often means that no or minimal revision and editing actually occurs.

While these parameters may be considered constant for all texts of the register, as in the case of essays, there are several situational parameters that vary across texts. Even the emails written to figures of authority in the academic context and thus representing a much more restricted set than emails in general, vary in their communicative purpose and topic. It is again not possible to identify a general purpose that would be set in all texts. Texts vary widely with regard to their purposes, some requesting information, others requesting help or guidance, still others providing updates, among numerous other possibilities. Finally, topic, closely associated with communicative purpose, is another situational parameter the value of which cannot be identified for all texts. Emails appear to be bound by very few restrictions in terms of what topics

can be covered, and even the emails of the study, restricted by the academic context, feature a wide variety of topics, discussed in detail further. Purpose and topic, the variable situational parameters of the register, will be the focus of the analysis in this section.

It has been stated, however, that not all the other situational parameters are held constant in emails, and while the situational analysis here has mainly focused on the most common category of texts, namely written by students to professors or other authorities within the academic context, there are other texts, for which the values of several parameters would look different. For example, several emails are addressed by the authors to their classmates. In these emails, they either offer advice and guidance to students in other cohorts or ask questions regarding course materials or address other school-related topics. The relationship between participants in these cases is still defined by the academic environment; however, their relative status is equal, and there is no social distance or power differentials associated with it. In another set of emails, students may write to authorities outside of the academic context. These exceptions include an email to a landlord describing the state of the accommodation and placing a maintenance order, an email to prospective employers following an interview to express appreciation and clarify potential misunderstanding, an email of complaint sent to a local business describing unacceptable behavior of their employee, or an email to a company official describing a negative experience during an internship for that company and demanding a resolution.

In these emails, apart from purpose and topic, the situational factors whose values vary include the relationship with the addressee, the relative status of the participants, and the amount of shared knowledge. In many cases, the relationship is unclear, as the authors may not know the addressee well or at all, or if they do, the length of the relationship is unknown (the email to the

landlord). Relative status may also vary – while the author still seems to be in a lower position than the addressee in some of these emails (the email to a prospective employer after the interview), in others the author’s role as a client, a tenant, or even an employee who did not receive appropriate treatment or services (the two emails of complaint and the email to the landlord) may give them authority in the matter and position them on an equal footing with the addressee, who would otherwise appear to be of a higher standing. The nature of shared knowledge varies in all these possibilities and is heavily dependent on the topic, the length of the relationship, and any previous interactions. For these emails, the variable situational parameters are not defined (Table 6.2). This suggests that these situational factors also can contribute to the variation found across emails. While this additional variability is important to acknowledge, the study mainly focuses on purpose and topic as the variable parameters for all types of emails – addressed to classmates, professors, or authorities outside the academic context. As purpose and topic vary in all these cases and are closely associated with the addressee and the author’s relationship with them, it is assumed that purpose and topic encompass these considerations (e.g., the topic of maintenance and the purpose of requesting assistance occur in an email to the landlord).

Thus, texts in the ‘Central’ group, showing the general register tendency, texts in the two ‘Peripheral’ groups, and texts in the ‘Moderately Removed’ group were again analyzed in terms of the variable situational parameters identified through the application of the framework for situational register analysis (Biber, 1994; Biber & Conrad, 2019). Again, topic was chosen as the basis for the following analyses as the situational parameter that yielded the most informative and interpretable patterns. Table 6.6 presents the specific topics of each analyzed email and the

general topic types identified by the analysis: personal topics, official matters, and class content-related specialized topics.

Table 6.6. Email topics within each distance and direction group. Dimension 1: Oral Elaboration vs. Information Density

Distance group	Above <i>M</i>	Below <i>M</i>
Central tendency (<0.1)	<p>1. Personal topics – 9 emails:</p> <ul style="list-style-type: none"> • Maintenance/ living conditions • Questions re. Master’s program • Personal circumstances causing a delay in assignment submission • Student’s CV & ways to incorporate their marketing experience • Remote vs. in-person work format during Covid-19 (personal & work-related circumstances) • Complaint about group work • Issues with assignment submission • Questions re. choice of format & student’s placement • Struggles with finding evidence/ references (student’s thought process) <p>2. (Class) content-related topics + student’s thought process – 1 email:</p> <ul style="list-style-type: none"> • Choice of essay topic: linguistic & cultural perceptions 	<p>1. Official matters – 5 emails:</p> <ul style="list-style-type: none"> • Company placement & job interview outcome • Course repetition deferral • Course enrollment (dropping out); incorrect submission • Job offer • Assessment submission deadline <p>2. Class content-related topics – 5 emails:</p> <ul style="list-style-type: none"> • Teamwork plan • Addressing criticism on the student’s literary analysis • Effect sizes • Student progress on specific stages of the project • Course materials and readings <p>3. Personal topics – 2 emails:</p>

Distance group	Above <i>M</i>	Below <i>M</i>
Moderately removed (0.3-0.6)	<p>1. Personal topics – 17 emails:</p> <ul style="list-style-type: none"> • Personal circumstances affecting module selection & course progress • Failed module & concerns about the future in the university • Dissatisfaction with the placement experience • Personal experience: success in job interviews and with applications • Student’s financial situation • Module change due to a lack of interest & understanding • Personal circumstances around a new position and a format change (in-person/remote) • Questions to other students re. placement experience • Issues with plagiarism & retaking the module • Personal introduction (to a mentor) • Late submission/ personal problems • Student’s plans re. placement + previous work experience • Student’s thoughts re. the right career path • Student’s concerns re. placement 	<ul style="list-style-type: none"> • Application guidance/ lack of help from the career team • Resubmission & associated circumstances • Deadlines and exam participation <p>1. Official matters – 10 emails:</p> <ul style="list-style-type: none"> • Agenda of student-faculty meeting/ feedback from students to faculty • Policy brief draft (ahead of the meeting re. working for a company; contractual agreement, etc.) • Placement year authorization • Mentee assignment • Placement teams; student work during placement year • STEM event (+ content) • Customer service • Information needed for a formal letter (type of letter unclear) • Exam deferral • Student accommodation reservation

Distance group	Above <i>M</i>	Below <i>M</i>
	<ul style="list-style-type: none"> • Student’s plans for placement • Student’s decision re. placement • Questions re. placement experience <p>2. Official matters – 3 emails</p> <ul style="list-style-type: none"> • Attendance correction in the system • System error/ student’s potential wrong identification in the system • Accepting a position; notice re. attendance in the future <p>3. (Class) content-related topics – 6 emails</p> <ul style="list-style-type: none"> • Structure of the lit review of an essay • Question re. an experiment • Statements re. gender positions in the workplace • Steps of an experiment • Interviewing participants & starting the analysis (*+personal circumstances preventing the start) • Essay topic: personal finances & healthcare 	<p>2.(Class) content-related topics – 4 emails:</p> <ul style="list-style-type: none"> • Absolute freedom in Liberalism (a proposed dissertation topic) • Feedback webinar – recorded presentation • Keynote presentation-related technology issues • Themes/headings in a course assignment <p>3.Personal topics – 3 emails:</p> <ul style="list-style-type: none"> • Assessor feedback & travel plans • Internet, accommodation, Covid-19, and student’s finance-related issues as reasons for a late/failed submission • Resolution of an error with groupwork submission
Peripheral 1 (0.6 – 0.9)	<p>1. Personal topics – 2 emails:</p> <ul style="list-style-type: none"> • Personal experience with coursework and placement • Frustration with unacceptable management/ delayed payment/ lack of guidance (*distinct from the complaint - #1 in official matters) 	<p>1. Official matters – 4 emails:</p> <ul style="list-style-type: none"> • Unacceptable behavior by a staff member & unsatisfactory customer service • Student card loss & ensuing attendance issue • Company placement paperwork • Discount deal with a local business <p>2. Content-related topics – 2 emails:</p>

Distance group	Above <i>M</i>	Below <i>M</i>
Peripheral 2 (>0.9)		<ul style="list-style-type: none"> • Class & gender in fashion as an essay topic • Scientific experiment <p>1. Content-related topics – 1 email:</p> <ul style="list-style-type: none"> • Marketing strategies & addressing consumer needs as an essay topic

Table 6.6 shows a persistent pattern of personal topics consistently showing a tendency towards oral elaboration and official matters and topics related to class content clearly prioritizing information density as follows from the number of emails in each group. This pattern is reminiscent of the trend observed in essays and discussed in the previous section. While the topics addressed by emails are naturally different from those discussed in essays, it appears that the grouping of these topics has the same general basis. That is, the general distinction between the topics in both registers that explained the tendency of some texts towards oral elaboration and others towards information density is that between more or less personal topics. In essays, such relatively more personal topics were found in essays in sociology, law, and marketing and business, and the technical, much less personal topics were addressed in essays in psychology.

In emails, this contrast is perhaps even more stark as authors describe personal circumstances as they request academic accommodation, explain the nature of the problems they are experiencing and ask for help, share thoughts or ask for guidance with regard to a career opportunity or professional training, complain about their negative experiences with group work or company placement, express gratitude for a course they found beneficial, among other possibilities. On the other hand, emails on official matters do not raise personal issues and are

concerned with university policies, such as attendance, student card loss, enrollment, exam deferral, or applications for company placement and interviews. In these emails, informational features often comprise the very reason for writing the email and denote the objects or entities that constitute the matter (the student card lost, the attendance team that would need to get involved, the job description in the documentation that needs to be submitted).

Neither are personal matters the subject of class-content-related emails, which integrate a variety of topics covered in the students' classes – fashion and gender, data collection for an experiment, origins of Liberalism and the current trends, or marketing strategies in addressing consumer needs, among others. In class content-related emails, the nominal sequences are the result of a transfer of the topic of a different project into emails. The frequency of these informational features, which may be expected in the register whose topic they reflect (e.g., essay, thesis, or dissertation), is uncommon for emails, which results in the peripheral position of class content-related texts. It is thus the topic uncharacteristic for emails that accounts for this deviation. In that sense, content-related emails appear to be an interesting hybrid, on the one hand representing email communication between the author and the addressee (typically the professor, whose class the content is from), but on the other hand representing a transfer of content from a different register.

As noted above, the direction of this trend in emails is the same as essays, with personal topics showing a clear preference for oral elaboration and the other two groups, official matters and class content, being more informational. In the case of emails, however, this trend appears to be even more pronounced, with texts even in the 'Central' distance group following the same pattern and lending additional support to the trend (Table 6.6).

In summary, distinctions in topic proved to be highly informative in interpreting the textual variation within emails and identifying the basis for the linguistic deviation from the register central tendency. Emails primarily devoted to explaining personal circumstances contain extensive oral elaboration as the authors are involved and show willingness to discuss the personal matters at hand. Emails on official matters and class content-related emails show a different trend. While some elements of elaboration are present in these emails, their topics clearly require informational features.

It is important to note that while the role of topic has been found to be major in variation across emails, topic should not be interpreted as the sole factor accountable for this variation. As discussed previously, the situational parameters included in the situational framework applied in this study do not represent an exhaustive set. Specifically, variation *within* registers may require and, as shown by research in this vein, typically leads to identification of additional, previously unaccounted for situational factors. It is thus to be expected that such additional situational factors may reveal additional patterns in the linguistic variation in the register, not yet highlighted by topic. For example, while purpose was not seen to be as informative as topic in this analysis, a systematic coding for purpose may be necessary in order to identify such additional patterns. Similarly, a systematic account for the identity of the addressee to the extent possible (e.g., the nature of the relationship with the author) may reveal systematic patterns. As follows from the situational analysis of the register presented at the beginning of this section, the values of several situational parameters are undefined for emails. This large number of undefined parameters implies a wide scope of variation allowed across its texts. Considerations of topic encompass many of those undefined parameters (such as the addressee), as emails on a certain topic are associated with the type of audience they are typically written to. Accounting for those

factors separately as well as identifying other situational distinctions, however, may still be informative. The results presented here are only the first step towards such investigations.

The next section presents text analysis of emails from each group, which exemplifies these trends and discusses their features in more detail.

Text Analysis: Central and Peripheral Emails above and below the Register Mean

Central Tendency

Like essays, emails in the ‘Central’ distance group above and below the mean have not been analyzed separately, as the deviation from the mean is minimal. These emails deviating in both directions incorporate both features of elaboration and information density, which jointly serve the goals of the register. However, as Table 6.6 shows and as noted above, even the emails with this minimal deviation conform to the same pattern of topical differences. The excerpt in Text Sample 69 is an email on a personal topic, in which the author is writing to their landlord with a maintenance request, detailing the problems. The excerpt contains extensive elaboration, most salient features including the present tense, used to report the issues, conjuncts to organize them, adverbs of manner contributing to the specific descriptions of the malfunction, causative adverbial clauses providing reasons for them, among others (demonstrative pronouns, the pronoun ‘it’, and contractions). The author’s involvement is apparent through quite emphatic stance adverbs (*it has always been difficult; is extremely unhappy*). However, informational features also contribute in important ways, naming objects and showing partative relations (*cupboard (design) door; metal wall section; bottom part*), content (*maintenance requests*), or referring to people and locations (*top floor tenant*) (Biber et al, 2021, p. 585). Emails as a register integrate both sets of features, which is reflected in their intermediate position between

the positive and negative ranges of the dimension ($M = 0$; $SD = 0.38$). This email thus exemplifies this trend, not showing any stark preferences towards one end or the other.

Text Sample 69. Author 9, email – central (text id: 10143), Dimension 1 score: 0.07

I'm writing this email on behalf of us all in regards to the maintenance of the house. **Firstly**, the fridge **has** been a recurring problem. **This is because** from day one **it has always** been difficult to **fully** close **it** – **partly** due to the design of the cupboard door and the fact that the freezer **isn't** working **properly and** as a result the ice **builds** up and **doesn't** allow **it** to close. **This further prevents** the fridge from closing **properly** in addition to the cupboard design door. [...] The last of the issues in the kitchen **is** the metal wall section behind the sink. The bottom part of **it has** opened and **essentially** began to peel off. **This also needs** to be fixed **because** if **its** allowed to continue peeling the wall behind **it** will start to get wet from the sink and cause further damage to the property. In line with maintenance requests, **it has also** been brought to my attention that the top floor tenant **is extremely** unhappy with the mattress she **is** having to use **as** she **says** all she can feel **is** the springs within **it**.

Peripheral Texts above the Register Mean

Personal Topics (0.6-0.9)

Text Samples 70 and 71 contain emails in the 'Peripheral' group above the register mean, and both exemplify topics that appear to require extensive oral elaboration. In Text Sample 70, the author offers guidance to a fellow student and shares their experience with placement, professional training students undergo within their program of study, and coursework. The topic is highly personal for the author as they reflect on their first-hand experiences, express their feelings and thoughts about them at the time of the described events and in retrospect, share their judgment formed on the basis of the described experiences, and offer personal advice. The email is replete with adverbs of stance, especially emphatics (*Honestly, it's scary at first but just so amazing especially because...; was genuinely improving; I just found that I was benefitting a lot from the course; and genuinely I was just really enjoying being in a very multicultural environment*), which reveal the author's personal investment. Hedges are used to make

suggestions (*Maybe you could...*) or mitigate the author's subjective judgements (*get any placement not necessarily a good one*). The email is highly detailed – clausal coordination and causative adverbial clauses are used to justify the author's opinions. Numerous instances of features of inexplicit discourse – the pronoun 'it' and contractions – suggest that the email approximates oral communication, where shared context between the interactants allows lack of explicitness. Overall, the personal topic appears to result in the author's high degree of involvement and willingness to elaborate.

Text Sample 70. Author 2, email – personal above *M* (text id: 5291), Dimension 1 score: 0.70

Hi, **Honestly, it's** scary at first **but just so** amazing **especially because** (as silly as this sounds) I **don't think** we **are** exposed **enough** to different cultures **here. Yes! To be honest, I never** considered **it** either until contacted me asking me to consider **as** my grammar was **quite** bad in second year. **Maybe** you could look into a study placement for the language you **feel** weaker **with (stranded)?** Answering your questions: I **just** found that I was benefitting **a lot** from the course I was doing and was noticing that my was **genuinely** improving. I was **also just** thinking about what was more beneficial for me in the future (as in for final year) **and genuinely** I was **just really** enjoying being in a **very** multicultural environment **so didn't** want to leave. **Also** my second placement that I was supposed to be on was an English speaking placement in **so again** didn't feel like I wouldn't benefit as **much** from **it**. **So** I did have a second placement lined up for my second semester as a for a company called in (**also, don't** listen to everything the placement team **tells** you **because** they **just want** you to get any placement not **necessarily** a good one).

The email in Text Sample 71 shows a different relationship between the participants. The email is addressed to a company official who may have the authority to resolve the problems the author has experienced during their work for the company. In spite of this formal relationship, the topic of the email has been considered personal as the text is heavily focused on the author's negative experiences and presents a detailed and rather emphatic account of these experiences. The author expresses their frustration and is again highly involved. Present tense and time adverbs convey immediacy and concreteness as the author describes their work for the company

(Since **then**, I **have** logged on to every single shift **punctually**; I've **contacted** numerous people regarding **this now**). Many of the adverbs express stance and contribute to the emphatic tone of the email (I brought a laptop **especially** for this job; I've **never** experienced anything like **this**; There **has been** no support or assistance, **especially** given that I'm a new worker). Again, it may be the involved and personal nature of the email that results in the use of features of oral communication – demonstrative pronouns, contractions, and clause-final prepositions reveal a lack of concern with revision and editing to meet the expectations of careful writing. Thus, despite the difference in the setting of communication and the relationship between participants, this email too demonstrates the role of personal topics in the degree of oral elaboration.

Text Sample 71. Author 103, email – personal above M (text id: 4081), Dimension 1 score: 0.87

How can **that** be the only response? I **have** raised technical issues with my TM. She gave me my login on the of , which I **have** proof of in the attached images. Since **then**, I **have** logged on to every single shift **punctually**. **has** been off sick for the past week, **and** I was neither informed of this or given an alternative team manager to contact. The whole system **has been** a nightmare - any issues I've had, I've been forced to deal **with (stranded)** and find a solution **for (stranded)** myself. When I did try to contact , I wasn't given a response. There **has been** no support or assistance, **especially** given that I'm a new worker. I've **contacted** numerous people regarding **this now**, none being of any help. I brought a laptop **especially** for this job! I've **never** experienced anything like **this**. I'm having to chase you up, **just** to get paid **correctly**. If I **don't** get paid within the next day, I'm going to have to take **this further** and involve a third-party. Thank you,

Peripheral Texts below the Register Mean: Official Matters and Class Content-related

Emails

The emails below the register mean showing a tendency towards less oral elaboration and increased information density, also differ in topic. Specifically, the relative lack of personal focus or a shift in focus to other objects or entities constitutes the main situational distinction observed in this 'Peripheral' group. As could be seen from Table 6.6, two broad categories of

topics can be identified in this group of less personal emails: emails addressed to a university official or an authority outside of the university with regard to matters of documentation, university policies, accommodation, customer service, job offers, course enrollment, and others and class content-related emails, in which the author typically seeks guidance and approval of an essay or dissertation topic, thus integrating substantial amounts of the content associated with the proposed topic into the body of the email. The text excerpts below present emails from each group.

Text Samples 72-73 exemplify emails concerned with official matters. In the case of Text Sample 72 the author addresses a university official with regard to student card loss and its repercussions for attendance. While the circumstances of the loss are provided as the necessary background information and the email does contain features of oral elaboration, the nominal sequences in the email (*student card; attendance team*) are the necessary references that convey the essence of the problem, the entities involved, and the main reason for writing. This email, therefore, appears to present a balance between some elements of oral elaboration in the personal narrative of events and several instances of informational features, which integrate references key for the purpose of the text.

Official Matters (0.6-0.9)

Text Sample 72. Author 66, email – official below *M* (text id: 8671), Dimension 1 score: -0.77

Dear Sir/Madam, On the I lost my student card **and** as a result of external financial circumstances I was unable to buy a new one **as** I had **already** a previous one. I spoke to 2 of my lecturers and let them know **and** they said **it** would not be an issue to speak with the attendance team if **it needs** to happen. **Today I am finally** able to buy a new student card. I **have** been attending all my lectures except this week I **have** not attended any. I **hope** you **are** able to understand my situation **and** I would **really** not like my attendance to be **negatively** affected. Thank you

The email in Text Sample 73 clarifies the paperwork requirement for the author's company placement. Again, while self-focus in this email is apparent, the frequent references to the job description, the notion central for the official matter raised in the email, result in several instances of nominal sequences. It should be acknowledged that both emails presented here are considerably shorter than the personal emails discussed above, and text length is definitely a factor in the negative score these texts received. That is, features of elaboration could be more frequent in these texts if the texts were longer. Nevertheless, features of information density, largely absent in the oral elaborated emails discussed above, are quite prominent in these short texts and convey important notions that directly represent the subject matter. There thus seems to be a direct relationship between the function of the noun phrases and the official nature of this communication. This in turn suggests that it is the official topics of these emails that can be viewed as the distinct situational characteristic of this group of texts, and it is variation in this situational characteristic that explains the observed linguistic differences.

Text Sample 73. Author 101, email – official below *M* (text id: 3927), Dimension 1 score: -0.61

Good Morning, **Hope you are well.** I wanted to enquire about the paperwork I **need** to upload for my placement to be approved in the handout **it** said I **need** a document with the Job description and my signature provided, **but I only have** a document which **has** the job description listed and the tasks that I would have to do which doesn't have a signature (which **is** in the attachment attached). **In addition**, the paperwork I had to fill in for my placement **is** agreements and DBS checks would you like to upload one of them **also** to confirm that I have a placement with them or **just** the job description.

Content-related Emails (0.6-0.9 & > 0.9)

Similar to the emails focused on official matters, content-related emails gravitate towards information density. However, the nominal sequences in these texts perform a different function, and rather than denoting the objects of an official matter, the noun-noun combinations in

content-related emails refer to a variety of complex notions borrowed from their respective fields and brought into email communication to serve a specific communicative need. This transfer is illustrated in Text Samples 74 and 75. In Text Sample 74, the author shares their plan for an essay topic – celebrity involvement in advertising campaigns. Apart from the salutation, the whole body of the email is devoted to the discussion of this topic, which is not a characteristic topic for email communication. Rather, celebrity involvement in advertising campaigns is the topic of the essay the author plans to write. The nominal sequences in the text (*Christmas atmosphere, esteem needs, consumer behavior, and football stars*) all denote concepts central for that essay, and the mention of all these concepts united by the same topic would otherwise be unlikely in email communication.

Text Sample 74. Author 58, email – content-related below M (text id: 8107), Dimension 1 score: –1.24

Hi, In brief, I **have** mentioned how the celebrities used **are** popular and **have** good reputation with being nominated for awards in their own fields, **also** how the campaign **consists** of celebrities enjoying the Christmas atmosphere **together** whilst wearing products, sharing positivity fulfilling the psychological needs of belongingness and esteem needs. **Also**, the campaign **consists** of them coming to life and showing off their professions in clothing **has** a influence on consumer behaviour. My thoughts for recommendations were **maybe** they could use celebrities of a wider background to show diversity and influence a wider audience, and **maybe** use celebrities that **are** not only football stars and singers.

Likewise, the email in Text Sample 75 requests assistance with an essay outline. The body of the email presents the key ideas for an essay on gender and social class in fashion. The nominal sequences in the email directly reflect that topic (*middle classes, fashion trends, Instagram influencers, key thinkers*) and refer to its source (*essay plan*). This topic is again uncharacteristic for email communication and is transferred into it from the register of essays (or other registers where it may be covered). Thus, it appears that the substitution of the value of that situational parameter (topic typical for an essay rather than an email) results in the consistent

tendency of class content-related emails toward information density and explains the peripheral status of these texts.

**Text Sample 75. Author 28, email – content-related below *M* (text id: 5928),
Dimension 1 score: –0.79**

Hi, **Hope** you're well! Please could you check my essay plan for the question ' **This is** just an outline and will have a detailed plan done by .Paragraph one will discuss the relevance of upper/middle classes in fashion **today**.

Paragraph two will explore what **stimulates** fashion trends in modernity.

Paragraph three will question how gender **is** associated with fashion.

Paragraph four will suggest Simmel's view on Instagram influencers and their role on fashion **today**.

My key thinkers will be Simmel, Park, Burgess and Wirth – each paragraph will include Simmel and for paragraph two, three and four, I will evaluate with the other three thinkers. Thank you

6.1.3 Variation within Image Descriptions

Image descriptions are among the registers where internal variation is quite surprising. The situational analysis of image descriptions shows that the values of each situational parameter are defined for all texts of the register (Table 6.2). The relationship between the participants is not clear, but is likely to be that between the participant and the researcher. There is no direct interaction between the participants as the description is recorded. Image descriptions are monologic, the authors' speech is produced in real time, with no opportunity for planning; it is recorded and transcribed for the addressee to read after the production is over. The time and space are thus not shared by the participants. The general communicative purpose of image descriptions is to describe three visuals presented to the research participants. This task, highly specific in itself and constrained by the content of the visuals at hand, does not seem to entail more specific communicative purposes. The topic of image descriptions is the content of the described visuals. In view of this restrictiveness of the topic, no shared background knowledge is required between the participants. It follows from this analysis that there is hardly any scope for

variation within the register. Yet, Table 6.7 shows that peripheral texts in descriptions are not at all uncommon.

Table 6.7. Number of descriptions per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Descriptions ($M = 0.44$; $SD = 0.51$)	Above M	7	14	19	10	5
	Below M	11	11	20	15	1

These results point to the existence of a different reason for such extensive variation. Exploratory text analysis was therefore conducted on the texts in the ‘Central’, ‘Moderately Removed’, and the two ‘Peripheral’ distance groups in order to identify the additional situational characteristic that corresponds to the linguistic deviation. The process of iterative analysis, i.e., identifying situational basis for the already observed specific linguistic differences being discovered after the fact, is not a novel approach and is described by Biber and Conrad (2019) with regard to between-register analysis. That is, while typically registers in the corpus are initially analyzed for their situational characteristics and the linguistic differences found are then attributed to the situational factors that differ between the registers, often new or unexpected linguistic trends emerge. Such outcomes warrant a reverse approach, namely an analysis of these linguistic results and the identification of their possible basis. In the context of this study, this post-hoc analysis of situation is conducted within a register, with the focus on the three groups of texts that are of interest (‘Central’ and the two ‘Peripheral’ distance groups).

This analysis revealed a pattern in the extent to which the authors relied on the content of the visual or enhanced that content by their subjective commentary. That is, while it may appear

that the task of description is in itself quite restrictive and does not allow the authors any latitude, the contrasts observed in the texts on either side of the register mean demonstrate that approaches to this task may vary. Specifically, some texts focus solely on the content of the visual and state only what is depicted. Such descriptions tend to focus on the ambience, interior design, people's appearance, objects, and other details immediately present in the visual. As a rule, they do not add to what can be seen in the visual, do not offer interpretation, or enhance the content in any other way. While the register as a whole gravitates towards oral elaboration ($M = 0.44$; $SD = 0.51$), these descriptions occur far below the register mean. Their strict focus on the content of the visual results in the noticeably lower rates of occurrence of features of oral elaboration. The level of specificity of some of these texts also leads to informational features denoting specific items of clothes, technology, and other highly specific elements of the image. The informational focus thus reflects this level of detail and engagement with the visual.

In contrast, texts that show a deviation from the mean in the opposite direction and are rich in features of oral elaboration substantially enhance the content through their subjective interpretations. These interpretations involve additional content as authors hypothesize about the nature of the relationship between the people in the visual, the nature of the tasks these people are engaged in, reasons for their behavior (e.g., laughter), and other aspects of the content. Often, while the visual provides some minimal basis for these judgments, they are entirely the product of the authors' imagination. Such commentary naturally calls for increased oral elaboration. The detail observed in the descriptions far below the mean, on the other hand, is not the focus of these texts, which are more concerned with interpretation and analysis than the accuracy of the description.

In essence, the additional situational parameter that accounts for variation in image descriptions – namely the focus on the content of the visual as opposed to the additional content incorporated as a result of the authors’ subjective interpretation – may be construed as variation in topic on a highly specific level. This variation in micro-topic seems to also entail variation in purpose, as the addition of content not present in the visual in the peripheral descriptions above the register mean involves the goal of interpreting the visual. On the other hand, the absence of that situational factor results in the purely descriptive goal of the text. Overall, this analysis confirms that variation across image descriptions, a highly specific, narrow register, fully defined in terms of all situational parameters of the applied framework, is explained by an additional, highly granular situational parameter. The presence or absence of that parameter appears to be determined by a particular approach to the task – the communicative choice to focus on the visual or interpret it, integrating a personal perspective. These contrasting approaches to the task are exemplified in the excerpts below.

Text Analysis: Central and Peripheral Image Descriptions above and below the Register Mean

Central Tendency

A typical text in image descriptions, a register generally oral and elaborated ($M = 0.44$; $SD = 0.51$), is shown in Text Sample 76. This text shows that both a concrete description tied to the content of the visual, i.e., mentions of objects, traits of appearance, clothes, or other characteristics of the people in the visual, and the author’s subjective commentary have a place in the register and both contribute to its overall goal. Oral elaboration clearly conveys the author’s engagement as they make their judgement through such features as the present tense (here of numerous stance verbs) and emphatic adverbs (*Erm, just very young*). Nominal features,

overall relatively infrequent, on the other hand, usually convey the concrete visual-related detail (*seated at a computer desk*). However, the focus on the content of the visual is not limited to just the informational features associated with the dimension and can be found in other parts of the text (*student who's sitting; girl who's sitting; casual attire*). Although these references do not contain the exact informational features comprising the dimension, they still illustrate that this text combines the focus on the visual with the subjective commentary about it.

Text Sample 76. Author 12, image description – central (text id: 3491), Dimension 1 score: 0.50

Okay so I can see erm, a, **it appears** to be a teacher or a lecturer or a guide, er, erm, someone who's guiding someone, **and it looks** like she **is** answering the question for a student who's sitting. **And it seems** that she's **very** surprised that erm, for whatever reply she **actually** received. **And** they both **seem** to be **very** young, er, **but** the person, the girl who's sitting, **seems** to look like she's in her teen years. **So** she **looks** like a teenager, or **maybe** in her early twenties. Erm, **just very** young. **And then** the teacher, or the lecturer, or whatever, **is, looks** like guiding her, er, erm, **looks** like she's in **maybe** her mid to late twenties. Erm, the student **is, looks** like she might **possibly** be seated at a computer desk, and that **is** what she's being helped **with (stranded)**, **and** she **looks very** surprised. **And they** both **seem** to be in **very** casual attire.

Peripheral Texts below the Register Mean (>0.9)

Text Samples 77 and 78 illustrate sparsity of oral elaboration in descriptions below the mean.

Text Sample 77 is an example of the most extreme deviation from central tendency (> 0.9), with minimal use of any Dimension 1 features. The text illustrates that the description is tied to locations, physical properties of the objects in the image, and the concrete activities of the people in it. The author never offers their subjective judgment. The present tense is the only Dimension 1 feature used consistently throughout the description, and minimal elaboration is present through a single occurrence of a causative adverbial clause (*as in the background there **are** sparkles...*). This lack of elaboration is in contrast with the central tendency text above, and while the text in Text Sample 77 still accomplishes the task by providing a clear and

comprehensive description of the image, this text represents a noticeable difference in the communicative means employed by the author.

Text Sample 77. Author 106, image description – focus on the visual below *M* (text id: 3478), Dimension 1 score: -0.52

she **is** showing the other one something on a screen, something on a screen. One of the girls **has** ombré hair, whilst the other **has** braids. For picture s- number two, **it looks** like there **are** two men, one who **is** instructing or tutoring the other. The background **looks** like a library or a classroom. They **have** a desk and on the desk **is** a MacBook. One of the men, with erm, the deeper complexion, **is** looking at the screen. He **is** wearing a chequered shirt with a white tee shirt, with a slight smile. The other guy, who **is** fairer, **is** wearing a colourful tee shirt and **is** looking at the first guy. Picture number three. This picture **has** more of an informal background, **as** in the background there **are** sparkles and a funky design. In front of this background there **are** two people, one a man and one a woman. The man **is** holding some sort of book. Judging by the thickness of **it** I assume **it's** an educational book. The woman **is** not holding anything, **and** they **are** both looking to somebody else whilst laughing.

Text Sample 78, also found below the mean but in a less extreme range, illustrates the focus on visual-related detail, namely physical objects (*hair style , denim jacket, Apple MacBook Pro, leather jacket, combat trousers*), the ambience (*schooling environment*), and the people in the visual (*Caucasian gentleman, around about the same age group*). Some elaboration is present; however, a closer look at the oral features in the text reveals that they do not offer the author's subjective analysis of the depicted content, but are still used in reference to what is immediately obvious from the visual (*however her hair is ombréd; potentially braids; which is mainly covered by a grey jacket*). These adverbial features, therefore, may even seem incongruous with the content of the proposition as there is no apparent reason for hedging and a lack of certainty or emphasis with regard to a concrete physical object (*mainly covered by a grey jacket; potentially braids*). As a result, these oral features hardly contribute in meaningful ways to the description so heavily focused on the concrete.

Text Sample 78. Author 30, image description – focus on the visual below M (text id: 3515), Dimension 1 score: –0.24

I can see two females, **potentially** in a work or schooling environment. One of the females **has** blonde straight hair, **however** her hair **is** ombréd. She **has** more of a pale skin. She's wearing a stripey black and white top with jeans. She has a protective hair style, erm, **potentially** braids. She **is** wearing a blue denim jacket. [...] I can see two, **potentially** students, around about the ages of eighteen to potentially twenty-five, sitting down at a table in front of an Apple MacBook Pro. The first gentleman [...] **has** black hair and **is** wearing a white t-shirt with a red and black chequered shirt. Seated to the right of him **is** a **potentially** Caucasian gentleman, around about the same age group. [...] The female **has** long black hair, she **is** a person of colour, she **is** wearing a black top with a black leather jacket, and burgundy leggings or trousers with leather brown boots. **And** the gentleman **is** wearing a green top, which **is mainly** covered by a grey jacket and beige combat trousers.

Peripheral Texts above the Register Mean (>0.9)

In stark contrast, Text Sample 79 represents the peripheral texts positioned above the mean and is an example of extreme deviation (> 0.9) and a strong tendency to oral elaboration. This text, while addressing the task, is in fact minimally focused on the specific elements of the image and only references them if they lend support to the narrative constructed around the people in the visual, judgments about their attitudes, thoughts, and reactions (e.g., the recurring reference to the laptop these reactions are related to). This description discusses the level of interest and engagement the characters in the visual show in the conversation, interprets facial expressions and the nature of the work done. Prominent features of oral elaboration include the present tense, causative adverbial clauses, clausal coordination, the pronoun 'it', contractions, and adverbs. Adverbs appear to be the main feature that contributes to these texts' particularly high score on the dimension. The functions of these adverbs, specifically the hedging function (frequent use of *maybe, possibly*), are directly related to the subjective focus of these peripheral texts. Hedging is necessary in such interpretive texts as the author cannot be confident in their statements, which are all conjectures and speculation rather than facts. In other cases, however, statements are

intensified (*he's very interested*), suggesting the author's involvement as this interpretation is constructed.

Text Sample 79. Author 32, image description – subjective commentary above M (text id: 3518), Dimension 1 score: 1.48

So, they're sitting **together** at a table with a laptop in front of them. **But**, erm, the, the guy looking at the laptop **looks** like he's **very** interested in what's going on. **But** he, **it's** like he **looks** like he's multitasking, **maybe**. Erm, he's looking at the laptop **but** he's **maybe** listening to what the other guy **is** saying, which **is** why may, he, he may be looking at him. Erm, **but** he **looks** like he's doing both at the same time. '**Cause** he **looks** like he's about to smile. **So**, unless, **maybe**, whatever **it is** could be **possibly** making him smile would be on the laptop rather than the person that's speaking to him. **But yeah**, the other guy **just looks** like he's speaking, he's, he's, he's waiting, **maybe**, for a response **or** he's looking. **Or** he's **just** waiting for him to say something **but** he's looking at him like he's expecting him to say something **and then he's just** on his laptop. **Or** he's, **maybe**, listening to him **as** he's speaking **as well**. **So, this** could be in, **possibly**, in a university setting or, **maybe**, in a, **just** in a lounge area of some kind. Erm, **maybe** they're doing work **together** or something like **that**.

6.1.4 Variation within Business Memos

While business memos generally may vary in topic, addressee, and setting, the business memos in this corpus are highly constrained. The situational characteristics of business memos are largely defined by the fictional scenario: business memos are written texts produced by a single addressor for a single specified addressee in a work context, where the relationship between participants is that of a supervisor and a subordinate. The task implies general shared knowledge between participants, such as the details of the trip, the CEO's plans, and personal preferences, but no specialist shared knowledge is required. The texts are written, planned, and the production circumstances allow careful preparation, revision, and editing; the space and time of production are not shared by the writer and their audience. The general purpose of the document is to inform the recipient, summarizing the results of the search; the specific communicative purposes may include to address the specific requests and expectations of the CEO, evaluate the various

available options on offer, and make plans on behalf of the CEO, but these may vary across texts (variable parameter in Table 6.2). The topic is narrowly specified – the CEO’s trip to Helsinki or, more specifically, accommodation, scenery, food, souvenir shopping, and nightlife.

As the task clearly specified not only the goal of the memos, but every specific topic (accommodation, food, museums, gift shops, etc.) within the general topic (the trip to Helsinki), the values of almost all the situational parameters in the applied framework were defined for all texts of the register. The parameter left undefined was the specific purpose of the text. However, in view of this restrictiveness of the task, additional variation in purpose appears unlikely. In spite of this situational specificity, however, business memos are among the most linguistically varied registers in the corpus (Table 6.8). As in the case of image descriptions, to identify the possible factors corresponding to the observed linguistic differences, texts in the ‘Central’, the two ‘Peripheral’, and the ‘Moderately Removed’ groups were examined.

Table 6.8. Number of business memos per distance and direction group. Dimension 1: Oral Elaboration vs. Information Density

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Business memos ($M = 1.01$; $SD = 0.79$)	Above M	2	10	23	22	9
	Below M	3	6	8	7	22

While business memos are generally an informational register, the deviation from this tendency in the peripheral business memos positioned below the mean presents a rather striking contrast. Such texts adopt a minimalist approach to the task, only including the information that is strictly necessary (many of them even omitting some search items). This parsimony results in lists of places found by the author in response to the prompt and no or minimal explanation,

justification, or detail about the items. Such lists contain numerous nominal sequences of references to locations and activities on offer, many of which are proper names. In fact, in many business memos of this kind, nominal sequences are the only linguistic feature constructing the texts.

While the overall informational focus of the register, further reinforced in the peripheral texts below the register mean, is overall in line with the situational demands of the register (generally informational), the increased oral elaboration in business memos may seem incongruent with some situational characteristics, such as the goal of the text. However, a closer look at the features of oral elaboration reveals that while their presence reflects author involvement as in the case of all the other registers, this involvement is not at odds with the situation of the register. Rather, author involvement in business memos is manifested through detailed information about each search item, supported by expert opinions and reviews, and consistent links between the proposed options and the addressee's personal needs and preferences. Elaboration is thus primarily due to integration of expert opinions in the texts and addressee-focus.

It has thus been confirmed in this analysis that variation in business memos is due to additional, highly specific situational parameters. Arguably, these parameters could be viewed as specific purposes – the variable parameter left undefined (Table 6.2). On the other hand, supporting statements with expert opinions and relating them to the addressee's needs do not appear to be intentionally set communicative goals, and the 'purpose' label may not be appropriate. Regardless of the exact label, the presence of these situational parameters in the texts entails increased oral elaboration. In such texts, the authors are more involved and more clearly portray the relationship with the addressee. The presence or absence of the identified

communicative factors – addressee focus and reliance on expert opinions – or rather the degree to which they are present thus reflects varying approaches to the task at hand. As was discussed in the case of image descriptions, where personal interpretation of the visual was not offered in all texts, business memos also illustrate different communicative strategies, which in turn results in linguistic differences on the dimension. The next section illustrates these varying approaches.

Text Analysis: Central and Peripheral Business Memos above and below the Register

Mean

Central Tendency

Central exemplars of business memos provide basic information about the items listed in the prompt, such as accommodation, cafés, museums, souvenir shops, and entertainment. The references to the selected options, their locations, and their names require the use of nominal sequences, and since these references are indispensable in view of the register's informational goal, they are present in all texts and determine the overall tendency of the register towards information presentation. However, some additional information, which results in features of oral elaboration, is necessary and often provides useful detail about the selected options, thus contributing to the informational goal in important ways, preempting questions and enhancing clarity. Text Sample 80 exemplifies this general trend in business memos:

Text Sample 80. Author 26, business memo – central (text id: 27096), Dimension 1 score: –1.04

Radisson Blu Hotel, this hotel **has** been rated on trip advisor an 8.3 for work trips. **it is** close to the water which and 0.6 miles away from shopping facilities, **so** the manager can purchase gifts, the hotel **includes** a fitness centre, **so** she can stay active whilst she's away and a well rated hotel bar and restaurant, that **provides** a good quality dinner and breakfast. The Radisson **is not far** for public transport which **makes** airport access easy and simple. In total this hotel for the 2-4th of September 2021 **is** within budget at a price of £180 for a two night stay.

Peripheral Texts below the Register Mean (>0.9)

Text Sample 81 is an example of a minimalist approach to the task in the form of a list of items, with sparse comments from the author accompanying the suggestions. The register of business memos as a whole gravitates towards the negative range of the dimension, as the informational focus is required by their general goal, and it was shown above that informational features were clearly the driving force of the central tendency text. However, by giving an even higher priority to informational features, this deviation reveals an almost complete lack of author involvement, making the result of the search technical and impersonal in nature.

Text Sample 81. Author 23, business memo – lack of addressee focus & expert opinions (list format) below *M* (text id: 27093), Dimension 1 score: –2.00

Helsinki 2 Day conference from 2-4 September 1) Hotel options: Clarion Hotel, Radisson Blu seaside hotel - both **have** spectacular water views and under £200 2) Café options: Cafetoria roastery, Café Regatta, Fazer café city centre - all **serve** breakfast and close to the centre 3) Museum: Helsinki City museum, National museum of Finland - learn about city's history 4) Restaurant: Kosmos - **serves** Poronkaristys which **is** sautéed reindeer a national Finland dish 5) Souvenir shop: Kanfurin tupa 6) Evening activities: Apollo live club - comedy acts and live concerts, dining included

Peripheral Texts above the Register Mean (>0.9)

Text Samples 82-83 below illustrate deviation in the opposite direction and are examples of extensive oral elaboration in the register. Text Sample 82 illustrates the use of features of elaboration to accompany the suggestions with a detailed account of their advantages and consistently relate the choices made to the addressee's needs and expectations (*These **are** what I was able to find with the list you **have** given me; which **satisfies** your budget; seafront **is absolutely** stunning so I **think** you'll enjoy it*). Author involvement is conveyed through adverbs or other features of stance, such as *I think* or *absolutely stunning*, and intensifiers, such as *right*

from the Stone Age. This again shows that often increased oral elaboration accompanies the expression of personal views and perspectives. Involvement in business memos seems to reflect variation in the relationship between the author and the addressee. While that relationship is unlikely to have been specified by the prompt, business memos construct the fictional identity of the addressee and portray the relationship with them differently, possibly taking into account the length of their collaboration and the level of familiarity. It is these varying interpretations that appear to have led to wide variation across the texts.

Text Sample 82. Author 65, business memo – addressee focus above *M* (text id: 27139), Dimension 1 score: 0.09

Dear CEO, These **are** what I was able to find with the list you **have** given me: I found a central hotel called Radisson Blu seaside hotel, which **costs** £91 per night and £181 for 2 nights which **satisfies** your budget conditions of £200 and location conditions of being close to water. The seafront **is absolutely** stunning **so I think** you'll enjoy **it**. The hotel **has** a restaurant that **serves** breakfast **too**. There's a National Museum of Finland that **shows** Finland's history starting **right** from the Stone Age, which **is** a 7 minutes drive from the hotel. The typical national dish **is** Sauteed Reindeer **and** the restaurant serving **it** or versions of **it is** called; Kaarna Baari Britto, which **is** about 6 minutes drive from the hotel. **It specialises** in local dishes. Kankurin Tupa **is** a gift shop where you can buy souvenirs for your children and **it's** about 7 minutes from the hotel. **Finally**, there's the Apollo live club that **hosts** events such as comedy and live music concerts (not opera), about 6 minutes drive from the hotel. **All in all**, a short but packed trip with everything you **want** to do close by, I **hope** you **have** a great time **there**. Kind regards, PA

Text Sample 83 illustrates the role of expert opinions in the peripheral texts above the mean.

These expert opinions include the ratings the proposed options received from previous customers on review websites such as TripAdvisor (***It is** rated 4-5 stars which **is** a good option; **this is** a traditional restaurant with 4-5 star ratings; **This is** based in the city centre and **has** received 4-5 star ratings from TripAdvisor*). The author of this text is quite emphatic in their recommendations (*i **have** chosen for my CEO to visit*); however, the text always provides evidence for such strong statements.

Text Sample 83. Author 36, business memo – expert opinions above M (text id: 27107), Dimension 1 score: –0.01

Secondly, in case my CEO wanted to travel to a café to eat breakfast, I **have** selected the Fazer Café Kluuvikatu 3, café which **is also** based in the city centre and **offers** breakfast. **It is** rated 4-5 stars which **is** a good option. A restaurant i shortlisted was: Rosemariini, **this includes** vegetarian and non-vegetarian dishes which **is** good for diversity of food and options. **It offers** different payment methods but **currently offers** a take away service. the Second restaurant i chose which **is** better than the first option **is**: Lappi Ravintola. **this is** a traditional restaurant with 4-5 star ratings and **offers** a 4 course meal, table booking and different payment methods i **strongly recommend** my CEO to visit this restaurant. **Thirdly**, the souvenir shop i **have** chosen for my CEO to visit **is** called: Moomin Shop Forum. **This is** based in the city centre and **has** received 4-5 star ratings from TripAdvisor. This gift shop **contains** gifts such as books, toys, cups and clothes which **is highly** ideal and beneficial for my CEO to buy gifts for herself and children **as** they may like books and toys.

6.2 COMMUNICATIVE VARIATION ACROSS TEXTS: EVIDENCE-BASED STANCE

Section 6.1 identified the situational parameters accountable for linguistic variation within registers on Dimension 1 – ‘Oral Elaboration vs. Information Density’ – and showed that in the registers where the values of topic and purpose were undefined, it is topic or a combination of topic and purpose that explains linguistic variation across texts. In the more narrow, situationally well-defined registers, other situational parameters were identified, unaccounted for by the initially applied framework. It was shown that with regard to Dimension 1 features, the more personal nature of the topic and increased author involvement correlated with oral elaborated discourse, while less personal, more technical topics required information density. While these results may seem intuitive on a dimension of variation often associated with involvement (e.g., Biber, 1988, where the dimension is labeled ‘Involved vs. Informational Production’), the goal of the present chapter is to test these conclusions with regard to a different functional pattern. To that end, analogous analyses were performed on Dimension 4 – Evidence-based Stance.

Dimension 4 is comprised of clausal features of stance, such as stance verbs followed by complement clauses, and does not include nominal features, which may be more immediately associated with topic differences. Apart from the distinction between the lexical verb classes (e.g., public and private), the features associated with Dimension 4 are purely grammatical, and in that the dimension represents a contrast to Dimension 1, on which nominal sequences convey content-related information. As mentioned previously in Chapter 3, another reason for selecting Dimension 4 lies in the fact that the dimension is one of the three ‘universal’ dimensions of variation commonly identified by multidimensional studies – the oral/ literate divide, the narrative/ nonnarrative dimension, and stance (Biber, 2004). The following section therefore reproduces the analyses performed on Dimension 1 and reports them in a more concise manner, addressing the same research question with regard to Dimension 4 features:

RQ 5: To what extent do communicative differences across texts of a register explain linguistic variation within registers?

Table 6.9 below presents the number of texts positioned above or below the register mean within each distance group in each register. It can be seen from the Table that emails, image descriptions, and business memos are the registers that allow the most internal variation on the dimension, with the largest number of texts in the extreme ‘Peripheral’ group. These registers are followed by essays, where the number of texts in the ‘Peripheral 2’ (> 0.9) group is lower. The direction of this deviation in image descriptions and essays appears to be toward more stance expression, while in emails and business memos a relatively equal number of texts appear in the ‘Peripheral’ groups above and below the mean. Evaluations show more substantial variation than on Dimension 1, with several texts in the ‘Peripheral 1’ group (0.6-0.9). The direction of this

extreme deviation in evaluations is also toward more prominent stance expression. Interviews, on the other hand, are again the most uniform of the registers, with only three texts in the ‘Peripheral 1’ group above the mean. An inspection of these three texts revealed that they contained frequent repetitions of discourse markers featuring private and public verbs (such as *you know*; *I’d say*; *you know what I think*). These three texts do not yield substantive interpretable patterns for this study, and interviews are therefore not analyzed in detail. The following sub-sections turn their attention to texts from the other five registers.

Table 6.9. Number of texts per distance and direction group within each register. Dimension 4: Evidence-based Stance

Register	Direction of deviation	Distance group				
		<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Essays ($M = -0.02$; $SD = 0.43$)	Above M	13	19	10	9	3
	Below M	6	24	21	6	
Emails ($M = 0.39$; $SD = 0.65$)	Above M	4	16	13	11	8
	Below M	4	17	18	12	8
Interviews ($M = -0.59$; $SD = 0.26$)	Above M	19	19	11	3	
	Below M	20	22	13		
Descriptions ($M = 0.28$; $SD = 0.58$)	Above M	4	15	10	9	9
	Below M	7	17	29	7	3
Evaluations ($M = -0.28$; $SD = 0.37$)	Above M	8	6	8	6	
	Below M	8	13	13	2	
Business memos ($M = -0.85$; $SD = 0.45$)	Above M	10	22	16	7	3
	Below M	5	18	19	10	1

6.2.1 Variation within Essays

In view of their situational characteristics, namely, their academic topics and goals to present information, synthesize findings, report results, stance is not expected in the register of essays. However, as discussed earlier, these situational characteristics vary across texts, and some academic topics in combination with certain purposes open much more scope for interpretation and analysis than others. It is again these situational characteristics that may lead to a tendency towards more stance expression and result in substantial variation in the register ($M = -0.02$; $SD = 0.43$; Table 6.10).

Table 6.10. Number of essays per distance and direction group. Dimension 4: Evidence-based Stance

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	> 0.9
Essays ($M = -0.02$; $SD = 0.43$)	Above M	13	19	10	9	3
	Below M	6	24	21	6	

Again, texts in the ‘Central’, ‘Moderately Removed’, and ‘Peripheral’ groups were analyzed, and topic and purpose were recorded for each essay. However, while essay topics were reported in the analysis on Dimension 1, the most clear patterns of textual variation on Dimension 4 were revealed by the grouping of texts according to purpose. These purposes are thus the focus of this section, and their grouping is presented in Table 6.11 below with the number of essays in each group. The four purpose groups that contribute to the observed pattern are bolded. While Table 6.11 shows that other purposes were identified in some texts, these four purposes best explain the differences in stance expression. The Table also shows that this pattern holds true even in the ‘Central’ group, where deviation from the central tendency is minimal. As

in the case of Dimension 1, however, the ‘Central’ group is not analyzed with respect to directionality. Nevertheless, the fact that essays in this group contribute to the same pattern lends additional support to the role of purpose in the linguistic variation observed in the register.

Table 6.11. Essay purposes within each distance and direction group. Dimension 4: Evidence-based Stance

Distance Group	Above <i>M</i>	Below <i>M</i>
Central tendency (<0.1)	<ol style="list-style-type: none"> 1. Overview & discuss existing theories/ present stance of the key thinkers in the field/ existing body of research on a topic (in support of an argument) (*sometimes accompanied by own analysis of a specific case) – 8 essays 2. Present results of a qualitative analysis of interview data – 1 essay 3. case study + proposing recommendations – 2 essays 4. state facts, evaluate existing laws, make a claim – 2 essays 	<ol style="list-style-type: none"> 1. State the facts, present & describe (social phenomenon, company performance, efficiency of a business model, etc.) – 3 essays 2. Overview & discuss existing theories/ present stance of the key thinkers in the field – 2 essays 3. case study + proposing recommendations – 1 essay
Moderately removed (0.3-0.6)	<ol style="list-style-type: none"> 1. Present results of a qualitative analysis of interview data – 4 essays 	<ol style="list-style-type: none"> 1. State the facts, present & describe (social phenomenon, company performance, efficiency of a business model, etc.) – 13 essays

Distance Group	Above <i>M</i>	Below <i>M</i>
	<p>2. Overview & discuss existing theories/ present stance of the key thinkers in the field/ existing body of research on a topic (in support of an argument) – 2 essays</p> <p>3. Explain existing laws, assess law enforceability, discuss specific cases – 3 essays</p> <p>4. Case study – analyze a patient’s medical record – 1 essay</p>	<p>2. Report empirical results (usually quantitative) – 4 essays</p> <p>3. Overview & discuss existing theories/ present stance of the key thinkers in the field/ existing body of research on a topic (in support of an argument) – 4 essays</p>
Peripheral 1 (0.6-0.9)	<p>1. Overview & discuss existing theories/ present stance of the key thinkers in the field/ existing body of research on a topic (in support of an argument) – 5 essays</p> <p>2. Explain existing laws, assess law enforceability, discuss specific cases – 3 essays</p> <p>3. Present results of a qualitative analysis of interview data – 1 essay</p>	<p>1. State the facts, present & describe (social phenomenon, company performance, efficiency of a business model, etc.) – 5 essays</p> <p>2. Report empirical results (usually quantitative) – 1 essay</p>
Peripheral 2 (>0.9)	<p>1. Overview & discuss existing theories/ present stance of the key thinkers in the</p>	

Distance Group	Above <i>M</i>	Below <i>M</i>
	<p>field/ existing body of research on a topic (in support of an argument) – 2 essays</p> <p>2. Present results of a qualitative analysis of interview data – 1 essay</p>	

The following trends were identified in the analysis of purpose groups in essays. Stance is expressed to a greater extent when established theories are presented and discussed in order to synthesize the state of the art on a particular topic or in order to argue a point and use the state of the art as a support for the author’s own argument – a goal which entails a heavy reliance on expert opinions and findings of previous research efforts. Alternatively, stance is prominent when qualitative analysis, typically of interview data, is presented, and the thought process of study participants becomes the focus of the essays. Thus, these two goals combine expert stance and the author’s own stance (in essays reporting qualitative analysis of interview data), with the former being more common.

In contrast, essays lower in stance report facts and statistics, or describe a phenomenon (i.e., analysis of the market segment, efficiency of a business model). Such descriptions and evaluations rely on the apparent facts, and the author can usually perform such evaluations without the support of previous observations (e.g., an evaluation of a company’s activity requires the author’s knowledge about the company’s operation, but does not entail references to other authorities who carried out similar analyses and reached similar conclusions). These essays do not share a thought process and do not relay opinions, but rather present facts in a straightforward manner. Another category of essays that commonly do not rely on stance

involves texts that present stages of an experiment and (usually quantitative) empirical findings. While reliance on previous research may seem relevant in such texts, these essays are primarily examples of procedural writing and do not contain extensive literature reviews or discussions, where such references may be expected. At the same time, they do not lend themselves to an expression of the author's own stance as it is objective information presentation that is prioritized in such texts. As a result, texts with these two goals do not use features of speaker-attributed personal stance associated with the dimension.

These patterns are illustrated through text analysis of essays in each group in the next section.

Text Analysis: Central and Peripheral Essays above and below the Register Mean

Central Tendency

Text Sample 84 shows a characteristic use of stance in the register of essays. This text is a combination of straightforward statements of facts which are not the product of the author's thought process and do not contain explicit references to other authorities. While such references may occur in the form of in-text citations, they are not the main focus of the text and are only used to support previously made statements. As a result, they are not integrated in the text itself. On the contrary, when a key thinker and their ideas are central to the discussion, such as the reference to Descartes in this passage, stance becomes explicit and is conveyed through multiple clausal structures. This 'central' text shows that speaker-attributed stance features in this academic register are typically used to denote the stance of experts in the field (rather than that of the author of the essay), whose positions are analyzed and serve as evidence for any subsequent arguments.

Text Sample 84. Author 89, essay – central (text id: 10985), Dimension 4 score: –0.01

Psychology is commonly **defined** as the “Scientific study of the human mind and its functions, especially those affecting behaviour in a given context” (Oxford dictionary, 2004). It’s timeline spans across centuries: from a psychological experiment of language in the 7th century BC by King Psamtik I of Egypt (Hunt, 2007) to modern psychology today. The story of psychology would cease to exist without key thinkers. Take Descartes for example. A French philosopher, and scientist **known** for the phrase “**I think, therefore I am.**” (Descartes,2000). Descartes **can** be **described** as history’s best **known** dualist – **arguing that** the mind and body should be **considered** as separate entities. He **introduced** several core assumptions of Psychology: nativism, dualism and interactionism amongst others. In this essay, the key features and core assumptions will be **discussed** by comparing 2 approaches: The Biological approach and The Psychoanalytical approach. A key feature of both approaches is that a form of structure is important in affecting our behaviour. In the psychoanalytical approach, the tripartite structure of personality governs our behaviour whilst in the biological approach; the structure of the brain is crucial to the outcome of behaviour.

While the essay in the ‘central’ group presents a balance of statements of facts, not always attributed to a particular source, and a discussion of eminent thinkers’ key theories, the peripheral exemplars above and below the register mean represent these two extremes and show that abundance of stance expression and relative lack thereof tend to be associated with a particular communicative purpose prioritized by the essay.

Peripheral Texts above the Register Mean (> 0.9)

The two text excerpts below exemplify the two communicative purposes conducive to evidence-based stance. Text Sample 85 is an essay whose goal is to overview and discuss the existing body of work on a topic to establish the state of the art or prove a point. The excerpt is replete with personal, speaker attributed stance structures, as the stance of the authors of each included study is expressed in the form of their findings and contribution. The past tense, the feature of the dimension associated with evidence, is used in the essay not only to report stance, but to detail the outcome of the experiments on the basis of which the scholars’ conclusions were formed (*participants **found** it easier to recall more details; self-administered interview (SAI)*

*combated the caveats for context reinstatement ; autistic children still **did** not have higher accuracy or detail but they **did** make significantly fewer errors, etc.).* These past facts thus serve as evidence for the expressed stance (i.e., the authors of the cited studies normally suggest or argue something on the basis of their tests and results). This example shows that the features associated with the dimension are essential for this essay's goal to synthesize the existing research on a topic.

Text Sample 85. Author 55, essay above M – overview state of the art (text id: 10914), Dimension 4 score: 0.91

However, Solomon et al. (2016) **suggested (that del)** it **may** not be a relational binding problem but rather the excess use of cognitive functions. Solomon's study into memory **involved** autistic adolescents compared to Bowler's adults which **could suggest** age-specific differences regarding relational binding. Maras et al (2014) **argued that** the self-administered interview (SAI) **combated** the caveats for context reinstatement and the social aspect of the CI. [...] [O]ne aspect of the SAI that they **found to** be useful **was** the Sketch component as autistic participants **found** it easier to recall more details with drawing in relation to the event. Maras et al. **suggested that** this aspect of the SAI may reduce the cognitive demands on AI. [...] Mattison et al. (2015) [...] **found that** autistic children in the Sketch-RC condition (which **consisted** of drawing everything they **could** remember from the earlier video they **watched**) **did** significantly better than the autistic children in the MRC condition. In this study by Mattison et al. the results only **supported** free recall so a further study (Mattison et al., 2018) **assessed** the same conditions also in a probed setting (cued recall) but **found that** autistic children still **did** not have higher accuracy or detail but they **did** make significantly fewer errors than autistic children in the MRC condition.

The second goal that entails personal, speaker-attributed stance expression is presenting results of a qualitative analysis of interview data. Text Sample 86 is an excerpt from such an analysis, which investigates the appeal of the Korean culture to the members of the Korean culture club. Reporting the stance of the interviewee, whose perspectives are being analyzed, is central to this essay, and results in numerous expressions of stance in the passage. In essence, such texts still report the stance of experts on a particular topic: it is the interviewee sharing their first-hand experiences who is now viewed as the expert and the source of knowledge. Author

stance is still rare, although the subjective nature of the analysis in these essays occasionally allows expressions of the author's personal interpretation (*we can infer that; I am assuming that*). Reporting the interviewee's statements and their interpretations is essential for the essays that present this kind of qualitative analysis. This communicative purpose thus appeared to determine the linguistic tendency shown by texts in this peripheral status.

Text Sample 86. Author 65, essay above *M* – report interview analysis (text id: 10933), Dimension 4 score: 0.92

Faith **feels that** Korean culture is respectful, and group orientated, she **states that** “with the whole South Korean culture like they are all like respectful to each other it's a culturist culturistic culture so like I really **enjoy that**”. From this we can **infer that** Faith is drawn to the group orientated mindset, however, I am **assuming that** when she **means** ‘culturistic’ she **meant** collectivist, I did check it by her to be sure. She **feels** like South Korean social culture is like South African culture in terms of mutual respect for the people around you and the collectivism shown amongst people, she **mentions that** “I **love** working together” and “in South Africa you **know** your neighbours, you **know** everyone, here you don't [**know** your neighbours] you don't **know** your flatmates” this **suggests that** she can **relate** to Korean culture more than Western culture as they share similar cross-cultural values that she **understands**, thus **supporting** the effect of relatability on her interest in South Korean culture.

Peripheral Texts below the Register Mean (0.6-0.9)

Texts representing the other extreme within essays, namely making minimal use of clausal stance features, are again associated with two communicative purposes. Text Sample 87 below is an excerpt from an essay whose goal is to state facts, present, and describe – specifically, the potential market for a company expansion. This goal requires knowledge of the company's activity as well as of particular markets, but it does not entail references to specific authority figures who provide it. Rather, this information is presented without being attributed to any particular source, and it is these facts that are more critical for the goal of the essay than the experts providing them. Speaker-attributed stance is therefore rare in such texts.

Text Sample 87. Author 29, essay below M – describe, evaluate (text id: 10869), Dimension 4 score: –0.86

This report will aim to **present** an effective marketing analytical report including a strategy for the esports company ‘ReKTGlobal’, allowing them the opportunity to transform their revenue, as well as evolve company operations on a global scale. [...] South Korea currently accounts for 255m USD of the global share of 1.1B USD. China (with Taiwan also contributing an undisclosed amount’, brings in 250m USD, with Japan accounting for 77M. The reasons Japan the “spiritual home of gaming” brings in such a low amount in comparison despite its godfather status, is due to its market value (in Esports rather than gaming) being much more localised than its counterparts. Japan’s Esports market, focuses on broadcasting events of ‘Street Fighter’, rather than the more globalised juggernauts ‘League of Legends’, ‘Overwatch’ alongside some others. Although SK ranks highest, most of its revenue is actually attributed to Ad revenue, as well as its jumpstart on the industry whereby “South Korea’s Culture Ministry **sought** to get ahead of the curve in 2000 with the creation of the Korea e-Sports Association (KeSPA)” In spite of this, not only is China hot on their heels, but although Taiwan does not have isolated statistics available for their chunk on esports revenue, they are still huge underrated power players. Despite the vastly inferior population of 23m, Taiwan rakes in 1.3B USD through the gaming industry.

Text Sample 88 illustrates the other communicative purpose, which does not involve stance expression – reporting empirical results, usually of quantitative studies. It was noted above that these essays are examples of procedural writing, highly focused on the steps of the research process and its outcome. Such accounts of the methods, procedures, and results contain minimal, if any, references to the stance of others, and are almost certain to not contain any instances of the author’s own stance.

Text Sample 88. Author 27, essay below M – report empirical results (text id: 10862), Dimension 4 score: –0.65

According to table 1, extraversion and unhealthy eating habit has the weakest negative correlation ($r = -.013$, $n = 195$, $p = .856$). There is a weak positive correlation between neuroticism and unhealthy eating behaviour ($r = .024$, $n = 195$, $p = .739$). There **was** a negative and statistically significant relationship between conscientiousness and unhealthy eating habit ($r = -.168$, $n = 195$, $p < 0.05$). According to table 2, the overall model is not significant ($R^2 = .033$, $F = 1.305$, $df = 5/194$, $p = .264$). Neuroticism factor and Extraversion factor **got** a positive Beta 0.476 and 0.603, whereas openness factor and agreeableness factor have a negative Beta -0.791 and -1.294 . However, only conscientiousness factor is significantly **predict** a negative unhealthy eating behaviour ($B = -3.527$, $p = 0.044$) which is smaller than 0.05. According to AVOVA result,

conscientiousness is a significant effect of unhealthy eating habit ($R^2 = .028$, $F = 5.614$, $df = 1/194$, $p = .019$). In conclusion, the factor of conscientiousness has significant negative effect to unhealthy eating habit, this **support** the hypothesis of there is a negative correlation between conscientiousness score and unhealthy eating habits.

6.2.2 Variation within Emails

Emails are a register with the highest mean score on the dimension, but also the highest level of variation across texts ($M = 0.39$; $SD = 0.65$; Table 6.12).

Table 6.12. *Number of emails per distance and direction group. Dimension 4: Evidence-based Stance*

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Emails ($M = 0.39$; $SD = 0.65$)	Above M	4	16	13	11	8
	Below M	4	17	18	12	8

As the topics and purposes of the emails in the ‘Central’, ‘Moderately Removed’, and the ‘Peripheral’ groups were analyzed as the two variable situational parameters, it was found that topic offered the most clear insight into the variation in the register, and the same groups of topics as previously identified with regard to Dimension 1 features – personal topics, official matters, and class content-related topics – explain the observed linguistic deviation from the register central tendency. Specifically, emails on personal topics tend to occur above the register mean and incorporate substantial amounts of speaker-attributed stance, while emails on official matters or on content-related topics use stance to a considerably lower extent. While these trends are the most obvious in peripheral exemplars of the register, the pattern is clearly supported even by texts with a minimal deviation in the ‘Central’ group. Table 6.13 below presents these topics and the corresponding number of emails in each topic group above and below the register mean.

Table 6.13. Email topics within each distance and direction group. Dimension 4: Evidence-based Stance

Distance Group	Above <i>M</i>	Below <i>M</i>
Central tendency (<0.1)	<p>1. Personal topics – 4 emails</p> <ul style="list-style-type: none"> • Concerns/ questions re. presentation quality • Success in interviews & job applications • Application guidance/ lack of help from the career team • Student’s CV & the place of digital marketing experience in it 	<p>1. Class content-related topics – 2 emails</p> <ul style="list-style-type: none"> • Childbirth & women’s social status in African and Asian societies as an essay topic • Plan for the study, interviewing participants, data analysis <p>2. Official matters – 1 email</p> <ul style="list-style-type: none"> • Customer service <p>3. Personal topics – 1 email</p> <ul style="list-style-type: none"> • Dissatisfaction with the placement experience
Moderately removed (0.3-0.6)	<p>1. Personal topics – 9 emails:</p> <ul style="list-style-type: none"> • Personal circumstances around module selection & course progress • Unsatisfactory performance by a group member • Personal circumstances & thought process re. coursework & placement • Circumstances around assignment resubmission 	<p>1. Class content-related topics – 7 emails:</p> <ul style="list-style-type: none"> • Transition from a Marxist-Leninist dictatorship to democratic federal republic as an essay topic • Diaspora engagement in supporting democracy & justice as a dissertation topic • Structure of the literature review of an essay • Feedback webinar – recorded presentation

Distance Group	Above <i>M</i>	Below <i>M</i>
	<ul style="list-style-type: none"> • Praise of the module; expression of appreciation • Position/ format transfer due to a change in circumstances • Accommodation for a coursework deadline • Student's thoughts re. the right career path • Assessment submission deadline <p>2. Official matters (*with personal circumstances) – 2 emails:</p> <ul style="list-style-type: none"> • Unsatisfactory taxi service • Student's experience/ motivation to teach <p>3. Official matters – 1 email:</p> <ul style="list-style-type: none"> • Information needed for an official letter (type unclear) 	<ul style="list-style-type: none"> • Student books & readings • Effects of cannabis on the brain as a research topic • KAAAnE essays structure; academic reference search <p>2. Official matters – 6 emails:</p> <ul style="list-style-type: none"> • Student training • Quality of education during Covid-19 – value for money (to Members of the Parliament) • Agenda of student-faculty meeting/ feedback from students to faculty • Fundraiser – money for a volunteering opportunity • Job application • Accepting a position, notice about future attendance <p>3. Personal topics – 5 emails:</p> <ul style="list-style-type: none"> • Advice about company placement & coursework • Accommodation, Covid-19, student's finances, and internet-related issues • Student's financial situation

Distance Group	Above <i>M</i>	Below <i>M</i>
Peripheral 1 (0.6-0.9)	<p>4. Class content-related topics – 1 email:</p> <ul style="list-style-type: none"> • Criticism on student’s literary essay 	<ul style="list-style-type: none"> • Personal introduction • Personal decision re. placement
	<p>1. Personal topics – 6 emails:</p> <ul style="list-style-type: none"> • Remote vs. in-person work during Covid-19 • Placement termination, job loss • Received guidance, changed attitude to the subject, well-wishes • Decision re. course assignment • Concerns re. relevant literature for an assignment • Late submission, personal problems 	<p>1. Class content-related topics – 4 emails:</p> <ul style="list-style-type: none"> • China: democratization, tech/ industrial revolution as an essay topic • Class and gender in fashion as an essay topic • Steps of an experiment • Marketing choices addressing consumer needs as an essay topic
	<p>2. Official matters – 4 emails:</p> <ul style="list-style-type: none"> • Course enrollment (removal from roster) • Exam deferral • Student accommodation reservation • Student card loss & attendance 	<p>2. Official matters – 4 emails:</p> <ul style="list-style-type: none"> • Maintenance request/ living conditions • Attendance correction in the system • Student feedback to faculty re. placement teams & work during placement year • Reference letter request
	<p>3. Class content-related topics – 1 email:</p> <ul style="list-style-type: none"> • Keynote presentation-related technology issues 	<p>3. Personal topics – 3 emails:</p> <ul style="list-style-type: none"> • Master’s degree • Student’s concerns re. placement • Student’s plans re. placement

Distance Group	Above <i>M</i>	Below <i>M</i>
Peripheral 2 (>0.9)	<p>1. Personal topics – 5 emails:</p> <ul style="list-style-type: none"> • Assessor feedback; travel plans • Circumstances around a late submission • Circumstances around group work submission • Personal circumstances causing a delay in assignment submission • Assignment submission-related problems <p>2. Official matters/personal – 2 emails:</p> <ul style="list-style-type: none"> • Company placement & job interview outcome (*although includes personal attitude towards the job) • Unacceptable behavior by a staff member & unsatisfactory customer service (* although includes feelings regarding the experience) <p>3. Class content-related topics – 1 email:</p> <ul style="list-style-type: none"> • gender positions in the workplace (checking understanding) 	<p>1. Class content-related topics – 4 emails:</p> <ul style="list-style-type: none"> • Policy brief draft ahead of a meeting: contractual agreements, GDPR legislation, storage limitation/ data retention, high security – 2-step verification • Steps of an experiment • Effect sizes • Personal finances & healthcare as an essay topic <p>2. Official matters – 4 emails:</p> <ul style="list-style-type: none"> • Discount deal with a local business • STEM event promotion • Library system, library deadlines, book return policy • Unsatisfactory company performance towards employee

Apart from the fact that these patterns explain linguistic deviation in the register, an important finding that emerges from these patterns concerns the relationship between the situational parameter of topic and grammatical variation. It was stated earlier that the choice of Dimension 4 for this analysis was motivated by the predominantly grammatical composition of the dimension and, specifically, its clausal structure. That is, apart from the verb class distinction, such as private or public verbs, the dimension does not contain lexical features; moreover, the dimension does not contain nominal features that may be thought to immediately reflect topic. Grammatical variation has not been commonly associated with topical differences (e.g., Biber & Conrad, 2019, p. 48). The outcome of this analysis, however, shows that topical differences are in fact reflected in grammar. The next section presents text analysis illustrating this trend.

Text Analysis: Central and Peripheral Emails above and below the Register Mean

Central Tendency

It was noted earlier that speaker-attributed stance is characteristic for the register of emails on the whole ($M = 0.39$; $SD = 0.65$). Text Sample 89 illustrates this typical use of Dimension 4 features in an email to a professor regarding progress on a project. Unlike essays, stance in emails is usually directly attributed to the author of the email, as is illustrated by the use of most of the private verbs in this email. The author is always a participant of the events described in the emails, the agent of the steps taken or the thoughts being shared. It is most likely this fact that accounts for the personal nature of stance in the register and the high use of stance features in the register overall.

Text Sample 89. Author 55, email – central (text id: 7938), Dimension 4 score: 0.34

Hi, **Hope (that del)** you **had** lovely holidays! I just **wanted to** update you, I've done 4 of my friend interviews now, aside from one 50 minute one, the rest are between 1-1.5hrs, so I don't **know** whether that **means (that del)** we should keep going or try to do the analysis first. But that being **said**, I've put my study up on anyway but no one has signed up! If they don't it's fine I **can** still send my advert to friends I **guess** but I just **wanted to** ask what to do from here (aside from start analysing). I would **like to** go through them with you (I've transcribed two) but my exams open this and I'm kind of swamped with revision so it'll be hard for me but hopefully when it's over we **can** do that, but I **can** still conduct interviews if I **need to**. **Hope to** hear from you soon, Final Year

Peripheral Texts above the Register Mean. Personal Topics (> 0.9)

As was shown by Table 6.13, emails on personal topics are consistently associated with increased use of stance in peripheral texts above the register mean. Text Sample 90, for example, is an email detailing the circumstances surrounding the author's inability to meet the submission deadline and a request for an extension. Stance features comprise this explanation and the speech acts of apology and requests (*apologize, ask*), convey the author's thoughts (*think/don't think + that deletion*), and inform the addressee (*wanted to let you know*), but also report the stance of the immediately involved participants (*My sister said that*) and the impersonal stance of other unspecified entities (*Due to being announced as moving into tier 3 yesterday*). In addition to features of stance, the past tense is used to provide evidence of the author's position (*my parents came to collect me abruptly and I ended up leaving my laptop*). It is apparent from this example that stance is the driving force of emails on such personal topics and is essential to the fulfillment of their major goals.

Text Sample 90. Author 85, email – personal above *M* (text id: 9833), Dimension 4 score: 1.78

Dear ,

I **apologise** for having to **ask** this, but is it possible for me to upload my assignment on the ? Due to being **announced** as moving into tier 3 yesterday, my parents **came** to collect me abruptly and I **ended** up leaving my laptop. My sister **said that** she will post my laptop, but I don't **think (that del)** I would be able to get it until next week now. I normally keep my work on my memory stick but I **think (that del)** I **left** it in my laptop.

Once again I'm sorry, I just **wanted to** let you **know** my situation and if it is possible to have a later deadline. Thank you, (2nd Year)

Peripheral Texts below the Register Mean. Official Matters and Class Content-related

Topics (> 0.9)

In contrast, the two emails below (Text Samples 91-92) exemplify communication on official matters and class content-related topics. Text Sample 91 is written in a professional setting and aims to describe a potential opportunity to an unspecified addressee. This description contains relevant facts necessary for the addressee's understanding of the deal. Author stance does not aid that understanding and is therefore not required by this goal, which results in a limited functional range of the stance expressions that do occur. Stance in this email is limited to clichés that represent email conventions to request action (*were **wondering** whether you would*), make a suggestion in a tentative indirect way (*we would **love** it to be something like 10-15%*), and express interest in a response (*I will be **looking forward to hearing** from you soon*) rather than the author's genuine attitude or beliefs as is typically the case in emails on personal topics.

Text Sample 91. Author 67, email – official below M (text id: 8722), Dimension 4 score: -0.55

To whom it may concern, My name is and I am part of the team at . We recently **came** to your , branch and **were** referred to Head Office in regards to a possible discount for our members. At the moment we have many local businesses signed up and were **wondering** whether you would also be interested. In terms of how it would work, our members will receive a card with your logo on as well as the logos of other businesses in the area and would have to present it in exchange for a discount at your branch. In regards to the discount, it is all up to yourselves but we would **love** it to be something like 10-15% off of the customer's order. Thank you for your time and I will be **looking forward to hearing** from you soon,

Stance is similarly not required in emails rich in class content. The email in Text Sample 92 is on a content-related and quite technical topic – stages of a scientific experiment. There are only a few occurrences of personal stance (*do I **need to**; Also **want to say**...*) and references to

the author's past actions (*I sent you*) in this rather long text, most of which is focused on clarifying the steps of the experiment.

Text Sample 92. Author 52, email – content-related below *M* (text id: 7681), Dimension 4 score: -0.76

Hi , Not sure what you **mean** by a blocked order. Does this **mean (that del)** all no load trials presented and then all positive trials presented and then all negative trials presented and this would be counter balanced so some participants would get all negative trials first and then all positive trials and then all no load conditions. And this would be based on the B-G list you **gave** me? And the powerpoint I **sent** you is version A and would that be the mixed condition? So half pps would get no load, positive, negative, no load, positive, negative. Or do I **need to** change this to also be mixed so for example no load trial, positive trial, negative trial and then positive trial, negative trial and no load trial. Or do I just let half pps do version A that I **sent** you and leave that how it is and the other half just do the B-G blocked condition which will be randomised. Also **want to say** about the amount of pps as contact hours are max 30 so this would **mean** only able to get 40 pps instead of planned 60. Thank you for your help again!! Kind regards,

6.2.3 Variation within Image Descriptions

Image descriptions ($M = 0.28$; $SD = 0.58$) is the second register after emails most reliant on stance expressions. However, descriptions also show extensive variation across texts (Table 6.14), which cannot be attributed to the situational parameters of the applied framework, all the values of .which are defined for all texts of the register. The communicative factors accounting for linguistic variation were thus again identified through exploratory text analysis.

Table 6.14. Number of descriptions per distance and direction group. Dimension 4: Evidence-based Stance

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Descriptions ($M = 0.28$; $SD = 0.58$)	Above M	4	15	10	9	9
	Below M	7	17	29	7	3

Text analysis revealed that one situational distinction that explained a highly prominent use of stance in some peripheral texts and a lack of stance features in others was the same distinction identified on Dimension 1. That distinction reflected varying approaches to the task: focus on the content of the visual or the choice to accompany the content with extensive subjective commentary. Expectedly, texts offering a subjective commentary abound in author-attributed stance, while texts that only cover the content of the visual do not use stance features. However, that distinction cannot fully account for the observed variation with regard to stance, as linguistic differences were also identified *within* one of these groups – namely, the texts that provide an interpretation of the visual. This further distinction concerns the nature of the expressed stance. All Dimension 4 features – different types of verbs and their various complements – represent personal, author-attributed stance. These features are naturally associated with texts that provide detailed subjective commentary. However, many of the texts accomplish *the same goal* through quite different linguistic means, which also convey stance but are not associated with the dimension (e.g., hedging devices ‘*sort of/ kind of*’ or adverbial features of (un)certainty – *maybe, possibly*). This set of features directly contributes to the authors’ interpretation of the visual. However, as they do not comprise the dimension, texts rich in these features but lacking in Dimension 4 features receive low scores and constitute the peripheral group below the register mean. The stance expressed through these alternative linguistic devices is quite different from that expressed through Dimension 4 features – specifically, such mostly adverbial expressions of stance tend to be less personal as stance is not attributed directly to the speaker (a distinction also identified by Egbert & Biber, 2023). This highly specific situational difference in the type of stance conveyed is illustrated in the texts below.

Text Analysis: Central and Peripheral Image Descriptions above and below the Register Mean

Central Tendency

As was shown earlier, image descriptions generally maintain a balance between the content of the visual and an interpretation of that content. In view of the subjective nature of the task, author-attributed stance is to be expected in the register. A text illustrating this balance is presented in Text Sample 93. Stance in this passage is not only that of the author (*I guess; I believe*), but in their interpretation the author attributes stance to the general unspecified addressee (*you can't see in the picture, but you can see...*) as well as the characters of the visual (*she's just shown her something that she finds funny; she doesn't really know like why the other woman is laughing*).

Text Sample 93. Author 2, description – central (text id: 3514), Dimension 4 score:0.23

And she is sitting down, whereas the other white woman is standing up, sort of leaning over her. And the Black woman is laughing. Erm, it **could** be that she's just shown her something that she **finds** funny. Erm, because sh- er, the Black lady is laughing. Erm, the white lady look, she's got like a half smile. So she's not like fully smiling but erm, it's like she's only just sort of, she doesn't really **know** like why the other woman is laughing **I guess**. [...] So the woman on the left is a Black lady, and the man on the right is South Asian, **I believe**, erm, and he's holding a book. And again they're both quite, er, smiley. They're not looking at each other. It **appears that** they're looking at a third person who you **can't see** in the picture, but you **can see** their hand on the left. The man on the right, he is wear- he has short hair and a large beard, and he's wearing a grey jacket with a green t-shirt, and erm, like camel trousers. And the lady on the left, she has long black hair.

Peripheral Texts above the Register Mean (>0.9)

Text Sample 94 is rich in author-attributed stance. While not all the private verbs have a personal pronoun subject (such as *I guess (that del); I feel like*), the author still takes agency of the proposition in almost all instances of 'seem' (*it may seem to me; To me it seems as if; To me it*

seems like; And it seems to me). As noted above, due to their personal nature, features associated with the dimension also serve to attribute stance to the characters of the visual (*postgraduates who are just enjoying their break; She seems to be engaging in a conversation; he seems to be enjoying the conversation*).

Text Sample 94. Author 28, description – interpretation of the visual above M (text id: 3511), Dimension 4 score: 1.46

And she **seems to** be laughing particularly because it may **seem to** me that she's never **seen** the picture before. And I **guess (that del)** they are uni students, or maybe postgraduates who are just **enjoying** their break. The second picture is of two guys. One **seems to** be an Asian guy, and another may be a white British guy. The Asian guy is wearing a red tartan print jumper or jacket, and he **seems to** be watching something on his laptop. To me it **seems** as if he may be studying and he's trying to get his work done, but his friend is distracting him. And his friend is wearing a blue floral type [...] and he **seems to** be distracting his friend from his studies, maybe making jokes... blue shirt... and is making jokes. The third p-person or the third picture is of two ethnic minority background students. To me it **seems** like the first one is a Black woman and the second is an Asian male. The Black woman is wearing a black leather jacket and burgundy leggings, and b-brown boots. And it **seems** to me that she is in the social community area, maybe in a lounge for students. She **seems to be engaging** in a conversation, but I **feel** like she's a bit shy because of the way her hands are held together and she's clutching tightly to her phone. And the second guy, the Asian guy, he **seems to be enjoying** the conversation.

Overall, the passage illustrates that an interpretation of the visual through features of personal, author-attributed stance allows multiple functional possibilities, all of which are lacking in texts that solely focus on the content of the visual. This group of texts was illustrated in the previous section with regard to a lack of oral elaboration and will therefore not be repeated here (although the features of interest are different, the texts illustrating their absence are the same). Rather, the peripheral text below the register mean that follows exemplifies a different contrast and contains features of impersonal stance, not directly attributed to the author.

Peripheral Texts below the Register Mean (>0.9)

Text Sample 95, a quite lengthy description, contains rare instances of Dimension 4 features. A closer look at the text shows that other hedging devices (*kind of, maybe*) are extremely prominent and serve the same goal of interpreting the visual. However, they differ from the features associated with the dimension in that they do not attribute stance to any particular entity and express a general uncertainty. Thus, while both sets of features serve the same goal, namely to offer an interpretation of the content of the visual, in the peripheral texts below the mean this goal is served by a different kind of stance –impersonal stance primarily conveyed through adverbial features. While the exact reasons for the choice of impersonal stance over speaker-attributed stance are not obvious, the distinction in the level of personal agency expressed through stance features accounts for the peripheral status of some image descriptions below the register mean.

Text Sample 95. Author 64, description – impersonal stance below *M* (text id: 3553), Dimension 4 score: -0.74

And it, again, **looks** like, maybe, they're working on, maybe not helping each other som-, this time but, maybe, working on a joint piece of work together, maybe like some group work. Erm, they both **look** relatively happy, not really laughing as much as the other two but just smiling. Erm, and yeah, kind of like they're, maybe, discussing something as well because they're ki-, erm, they're kind of looking what, er, one of the guys is looking into, er, at the other one in kind of like agreement or to, maybe, he's done something to see if that's okay with the piece of work. Erm, again, they're dressed quite casually, just in a shirt and a tee shirt. And a iMac, erm, and that's that for that one. Looking at the last image, erm, I get kind of, er, focus group kind of a vibe. So, erm, maybe that they're in a discussion with somebody that's holding a focus group. Erm, I getting more of a thing where they, these two don't **know** each other because they **look** a little bit more awkward, erm, even though they're both smiling. Erm, it's one boy, one Asian guy and one, er, African girl. Erm, and they **look** a little bit more awkward or, maybe, like they don't **know** each other as well. Erm, but that's something people in focus groups generally are a little bit more awkward, I **think**, when they have to kind of discuss something.

6.2.4 Variation within Business Memos

Business memos overall show a tendency towards a lack of author-attributed stance, as can be seen from the register mean ($M = -0.85$; $SD = 0.45$). The register does, however, show substantial variation across its texts (Table 6.15).

Table 6.15. Number of business memos per distance and direction group. Dimension 4: Evidence-based Stance

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Business memos ($M = -0.85$; $SD = 0.45$)	Above M	10	22	16	7	3
	Below M	5	18	19	10	1

As discussed previously, this variation is likely due to additional factors, not accounted for by the situational framework applied by the study, as the values for almost all the included situational parameters can be defined for all business memos in the corpus. One parameter – specific purpose – was left undefined; it is thus either highly specific purposes or additional situational parameters that account for the observed linguistic variation. As in the case of Dimension 1, these parameters were identified as texts from ‘Central’, ‘Moderately Removed’, and ‘Peripheral’ groups were examined.

This analysis revealed that variation in the use of stance features could be attributed to the same situational factors as tendencies toward oral elaboration and information density. These factors include references to expert opinions and relating the proposed choices to the needs and expectations of the addressee. In addition, the authors often emphasize their active role in the task and show involvement through explicit references to the work they did (*I suggested/recommended*). The presence of these situational factors corresponds to increased use of stance

features. It thus appears that both sets of features, namely the increased use of features of oral elaboration and the clausal features of author-attributed personal stance can be explained through these situational factors. As noted in the discussion of Dimension 1 results, these situational factors – expert opinions, addressee-focus, and involvement – are not viewed in this study as specific purposes, but rather as distinct situational factors accounting for variation.

While essays relied predominantly on expert stance and it was the author's own stance that was expressed in emails, peripheral business memos that stand out in their unusual reliance on stance features incorporate both types of stance expressions, using them to convey both the author's position and that of an expert authority. The next section presents analysis of business memos from the 'Central' and 'Peripheral' groups above and below the register mean.

Text Analysis: Central and Peripheral Business Memos above and below the Register

Mean

Central Tendency

It was noted above that business memos as a register do not tend to rely on stance, which appears in line with its informational goal. This is reflected in the text below (Text Sample 96), which offers quite a detailed description of the researched options, but does so in a neutral, rather impersonal manner. This business memo does, however, relate the suggestions to the addressee's preferences through the use of private verbs, although these references are singular occurrences in an otherwise straightforward objective description (*if you **wish to** tell you about the history of the city; However, if you **wish** for an alternative museum*).

Text Sample 96. Author 21, business memo – central (text id: 27091), Dimension 4 score: -0.78

The hotel offers a 1 bedroom suite for £378 for 2 nights with a complimentary breakfast for each morning of your stay in their restaurant the Bistro Gimmis. However, for your

breakfast you also have the option of the Cafeteria Roastery. This is a local café a short car journey from the hotel where they have 4.5 star reviews for their breakfast foods and are a highly recommended place to eat in the city. During one of the days you will attend the Luomus, which is the Finnish Museum of Natural History where you **can** view the exhibits and have a guided tour if you **wish to** tell you about the history of the city. However, if you **wish** for an alternative museum there is also the Arteneum Art Museum which holds art from the 19th century onwards **telling** the history of Finland through art.

Peripheral Texts above the Register Mean (>0.9)

In contrast to this general tendency, the author's personal stance gains prominence in peripheral texts above the register mean. In Text Sample 97, the author directly refers to their recommendations several times throughout the text (*I have **suggested** for them to stay in the Hilton Helsinki strand [...] Secondly, a café I have **recommended**; I have **stated that***). The text also refers to the addressee and offers guidance as to how these recommendations can be followed (*it is a beautiful setting to **observe** the art pieces on show; you **can** eat buffet which really provides a lot of options*). Finally, recommendations are consistently supported through expert opinions in the forms of reviewer comments (*reviews have **suggested that**; The reviews say **that***). Features of stance thus serve several concrete functions in such texts.

Text Sample 97. Author 54, business memo – expert opinions & addressee focus above *M* (text id: 27127), Dimension 4 score: 0.18

For the CEO's day at Helsinki I have **suggested** for them to stay in the Hilton Helsinki strand [...] Secondly, a café I have **recommended** is the Fazer Café Kluuvikata 3. [...] It is a little pricey but the reviews have **suggested that** it is great, especially with the timing from 8am to 7pm, the CEO will be able to attend the café from morning to evening. A museum to attend is the museum of contemporary art museum, which is located in the central of Helsinki. It costs \$0-15, but with the small price to pay it is a beautiful setting to **observe** the art pieces on show. For a restaurant meal I have **stated that** the Restaurant Saaga as it serves the national dish **called** the Poronkäristys. The reviews **say that** the place has great ambiance. It also serves a vegetarian option as well as an all you **can** eat buffet which really provides a lot of options.

Peripheral Texts below the Register Mean (>0.9)

The other extreme, namely peripheral texts that do not incorporate any stance features at all, is comprised of business memos that adopt the minimalist approach to the task and are written in the form of lists. Text Sample 98 is an example of a business memo, in which the author's position, the addressee's expectations and preferences, or expert opinions about the proposed options are entirely absent.

Text Sample 98. Author 112, business memo – lack of addressee focus & expert opinions (list format) below *M* (text id: 27080), Dimension 4 score: –1.95

Hotel- 'Arctic DeLuxe.' Early bird rate- 378 euros for 2 nights. Cafe- Address- Toolonlahdenkatu 2, 00100, Helsinki Helsinki City Museum-Address- Aleksanterinkatu 16, 00170 (opens 11am). Restaurant Konstan Moija- Address- Hietalahdenkatu 14, 00180. Souvenir shop- Annensoppi- Address- Fredrikinkatu 68, 00100. Live music- Storyville- Address- Museokatu 8, 00100.

6.2.5 Variation within Evaluations

Linguistic variation in evaluations (not observed with regard to Dimension 1 features) is unexpected considering the situational restrictions of the register. Evaluations are defined through all the situational parameters in the applied framework (Table 6.2). The texts target an unspecified audience; the domain is defined as 'research project' as the texts were elicited for the purposes of corpus compilation, and they are most likely to have been produced in a public setting. Text are written, permanently recorded in an electronic form, which allows authors opportunities for revision and editing. Additionally, evaluations, mini-essays about the Aston campus facilities, appear heavily controlled in topic and purpose. The topic is held constant for all texts, and the general communicative purpose is to describe these facilities and express personal attitudes. This goal as well as the topic of these mini-essays are highly specific. However, additional, more specific communicative purposes may be present, and the parameter is considered variable (Table 6.2). No shared knowledge is necessary on the addressee's part as

the texts provide a description of the Aston facilities. Overall, this situational analysis illustrates that evaluations are situationally well-defined and are a quite specific register. The amount of internal variation to be expected in a register like this is thus quite low.

Yet, the register does show some internal variation with regard to stance features (Table 6.16). The basis for this variation is again identified through exploratory text analysis.

Table 6.16 Number of evaluations per distance and direction group. Dimension 4: Evidence-based Stance

Register	Direction of deviation	<0.1	0.1-0.3	0.3-0.6	0.6-0.9	>0.9
Evaluations ($M = -0.28$; $SD = 0.37$)	Above M	8	6	8	6	
	Below M	8	13	13	2	

While generally evaluations are quite neutral descriptions of the facilities with only occasional use of personal stance through statements of likes and dislikes, the deviation towards more stance presents a much more analytical approach to the task, in which the author not only lists the objects of the likes and dislikes, but discusses the origins of the problems, the reasons behind them, proposes a course of action, and explains why the changes would be an improvement. Proposals for future renovation and change in particular are a common feature of texts which abound in personal stance and may be viewed as a highly specific additional purpose. While the presence or absence of this specific purpose was not seen to correlate with higher stance use in all peripheral texts above the mean, this goal is commonly set by this group of texts and is worth noting. This characteristic as well as the overall higher level of personal stance appears to indicate heightened author involvement. It usually shows that the author has

given careful thought to the issues they are raising in the text, cares about their resolution, and uses the mini-essay as an opportunity to effect change.

On the other extreme, namely texts whose peripheral status is due to a complete lack of author-attributed stance, are texts mostly descriptive and factual. While the two peripheral groups in the case of this register may not present as stark a contrast as these extremes did in the other registers, author personal involvement appears to be the distinguishing factor between the peripheral texts above and below the mean. Alternatively, the difference between texts on the opposite extremes may be interpreted as a difference in highly specific communicative purposes present or absent in texts – for example, the purpose to propose solutions, to support suggestions through statements of personal opinions and references to negative past experiences. The presence of these specific purposes, however, still appears to signal author's heightened involvement in the task. The text samples below present the analysis of the central and peripheral exemplars of the register and illustrate the role of this parameter in linguistic variation.

Text Analysis: Central and Peripheral Evaluations above and below the Register Mean

Central Tendency

Generally, evaluations ($M = -0.28$; $SD = 0.37$) follow business memos in their lack of reliance on author-attributed stance. Text Sample 99 illustrates an evaluation representative of this general register trend. Contrary to what may be expected in view of the task, texts are factual rather than opinionated, and the most commonly used stance features are the private verbs *like* and *dislike*, which most likely repeat the wording of the task. Otherwise, authors' evaluations mainly focus on the reasons for the likes and dislikes and thus describe the positive and negative features of the campus rather than express personal stance toward them.

Text Sample 99. Author 31, evaluation – central (text id: 92), Dimension 4 score: – 0.34

Secondly, the library is another thing I **like** about the Aston Campus. It includes a massive openspace for loud study, where you **can** also eat, including multiple floors for quiet study. The colour scheme in the library is also very pretty and one of the main things that attracted me to the library. The members of staff are always friendly and approachable and with it being in the middle of campus, it is easily accessible for all students. Lastly, the student union is a main facility that I **like** about Aston campus. The drinks are cheap and the atmosphere there is vibrant and exciting. It attracts all the students and has become a main place for students to socialise and have fun. It includes a pool table, a bar, a restaurant downstairs, a seating area and plugs for when you want to study. This is a place I have come to socialise and chill with friends, as well as study many times and it's also a place I would definitely **recommend** first year students to come to make new friends.

Peripheral Texts above the Register Mean (0.6-0.9)

Some evaluations, however, are exceptions from this trend and demonstrate increased author involvement, conveyed through frequent use of Dimension 4 stance features and past tense to recount first-hand experiences as evidence. Text Sample 100 demonstrates numerous instances of public and private verbs and their complements (*I would have to **say that**; I believe that; I feel that*) conveying the author's views on the state of Aston facilities. Opinions are accompanied by detailed explanations, and the passage also contains proposals of future improvements (*I think (that del) it may **need** some renovating; hopefully they **can** be reviewed and changed*), which rely on Dimension 4 features and demonstrate the author's involvement in the task as they interpret it as a chance to suggest changes.

Text Sample 100. Author 5, evaluation – involvement above *M* (text id: 122), Dimension 4 score: 0.40

I would have to **say that** the new Student Union is beautiful to look at, but the purpose of being there has not been fulfilled. I **believe that** it is due to COVID-19 partially that many of Astons students have not been able to utilise this facility. However, I **feel that** Aston should be able to allow us to utilise this building well, as we helped fund it. [...] I have not been able to use this facility [...], which has been quite disappointing to **say** the least. I also **believe that** the main building is quite disappointing, respectfully. Although it has been there for ages, I quite **dislike** the outside building. [...] With this being **said**, I **dislike** it purely due to the fact that it **looks** old, and I **think (that del)** it may **need** some

renovating to keep up with the times. Lastly I **dislike** the food canteen facility in the main building. This is simply because the food is not varied and there is only a limited range that I do not **like**. Also, it would be nice if they **could** cater to those with severe dietary needs, such as gluten. [...] All in all, these are the three things I **dislike** the most in Aston and hopefully they **can** be **reviewed** and changed.

Peripheral Texts below the Register Mean (0.6-0.9)

In contrast, Text Sample 101 appears to differ in tone and directness due to a relative lack of personal stance features. This text is mostly descriptive, focused on positive and negative aspects of the facilities, and these statements are not enhanced through analysis or proposals of improvements.

Text Sample 101. Author 82, evaluation – lack of involvement below *M* (text id: 178), Dimension 4 score: –0.96

It is very small and **can** get too crowded at times, especially during the summer, where everyone tends to sit outside on the grass for picnics etc. Also a lot of the time events are taking place on campus in tents etc. and with the already tight space, it **can** be a bit of a squeeze as there's not much room to begin with. It would be nice if the campus was a bit bigger especially since there's such a high number of students that attend Aston University; it just becomes hard to accommodate enough space for students to manoeuvre across from, in summer, due to the crowdedness on the campus. Another dislike is the library. It is far too small for the amount of students attending Aston University. There's 3 different floors, however ground floor is fully occupied by 10 am and majority of the time people aren't really studying, it is more of a get together and meet up spot.

6.3 CONCLUSION

The chapter has examined communicative variation among texts of several registers with regard to two functional patterns – ‘Oral Elaboration vs. Information Density’ and ‘Evidence-based Stance’. It was shown that the registers in the corpus vary in the degree to which they are situationally well-defined (Biber, 1994) and in the amount of linguistic variation they allow.

While some situational parameters can be specified for all texts of a register, others are variable.

It was then shown that it is the specific situational parameters that vary among texts of a register that account for linguistic variation within that register. The situational parameter whose role in

explaining linguistic variation within registers has been shown to be major is the type of topic – specifically, the study made a distinction between personal and impersonal topics. Texts on personal topics were shown to lend themselves to increased oral elaboration, while impersonal (technical, scientific, or official) topics required increased information density. In some registers, the study then uncovers patterns of association of certain topics with particular communicative goals.

The study observes linguistic variation even within quite narrow, well-defined registers, where such extensive variation is unexpected in view of the restrictiveness of the task. In these registers, the study identifies additional, highly granular situational parameters that account for linguistic variation. Overall, it is conclusively shown that linguistic differences among texts of a register are always explained by specific situational parameters variable among texts of that register.

CHAPTER 7. INDIVIDUAL VARIATION ACROSS REGISTERS

RQ 6: How do authors vary across registers?

Chapter 6 analyzed communicative variation across texts within registers. The analysis of this register-internal variation revealed important distinctions in the communicative approaches to the same tasks, such as the highly specific tasks of business memos or image descriptions. In other cases, situational factors, such as the varying topics of emails, were found to have a major influence on the linguistic composition of the texts. Likewise, in essays, considerations of topic and the closely associated communicative purpose determined the linguistic features pervasive in the texts.

That analysis yielded the following major findings. First, registers vary with regard to the extent of internal linguistic variation they allow. This extent of communicative freedom permitted within a register may on the one hand be determined by the number of situational parameters, the values of which are specified for a register. On the other hand, even registers with all previously accounted for situational parameters defined for their texts exhibit linguistic variation. In these cases, linguistic variation within these registers is attributable to some additional communicative distinctions within registers. Thus, a new parameter is identified that explains linguistic variation. Other registers, however, are quite restrictive, and their situational characteristics do not allow much internal variation. As a result, all their texts are quite uniform in their communicative characteristics, meeting the same situational demands.

This chapter builds on these findings and applies them to an analysis of individual variation across registers. Each author in the corpus produced a text within each of the registers. In Chapter 6, each text within the registers was characterized with regard to its place relative to the register mean and the direction of deviation. This approach thus resulted in each author's

texts being annotated for their place in their respective registers. The study now examines if authors vary in their interpretation of different registers. That is, some authors may consistently produce texts that are among the central exemplars of every register. These authors may therefore show a high level of awareness of the conventions of each register and its situational constraints. As a result, this group of authors may produce texts that are always in line with these situational demands.

On the other extreme, some authors may consistently show deviation from what is commonly accepted in the registers of the study, and that deviation may consistently be in the same direction (i.e., more informational texts or more oral elaborated ones across registers). Their reasons for such consistent deviations may be lack of awareness of the situational make-up of the registers as a result of a lack of experience and exposure, their lack of sensitivity to conventions, potential transfer of the conventions of other registers, or their personal choice to diverge from these conventions.

Finally, some authors may not represent either trend and may make unique decisions about the extent of oral elaboration or information density or the amount of stance necessary in their texts in each register. This chapter explores individual variation with regard to such cross-register patterns and presents case studies of these groups of authors – authors consistently conforming to the central tendency, authors prioritizing a particular dimension function in at least half of their texts, and authors showing a range of approaches to register. While only the registers with the widest range of internal variation were included in the analysis in the previous chapter, all six registers are of interest in this examination of cross-register patterns, as the authors' closeness to the central tendency in the registers of interviews, for example, is part of

their cross-register pattern.⁵This investigation is again carried out on two dimensions – ‘Oral Elaboration vs. Information Density’ (Section 7.1) and ‘Evidence-based Stance’ (Section 7.2).

7.1 INDIVIDUAL VARIATION ON DIMENSION 1: ORAL ELABORATION VS. INFORMATION DENSITY

The case studies presented in the next sections yield the following general insights. First, it is shown that authors who are consistently ‘central’ across registers appear to reveal a keen awareness of register conventions, different across situations. That is, in each case these authors closely adhere to these varying conventions, representing the central tendency of such different registers as image descriptions (generally oral and elaborated) as well as essays or business memos (generally informational). Second, some authors show systematic preferences for highly informational or highly elaborated texts in most registers, thus demonstrating an individual preference across situational contexts. A more noteworthy trend, however, lies in the fact that even these authors do not always produce texts in line with their preference, and in one or two registers may deviate in the opposite direction; this divergence from their linguistic preference, however, is not caused by register restrictions. For example, if an author has consistently produced highly informational texts across registers, one of their texts (e.g., a business memo) may still be highly elaborated. We cannot conclude, however, that this difference is due to the situational restrictions of business memos. Chapter 6 showed that the register of business memos allows extensive internal variation, which should have made it possible for the author to produce

⁵ As described in Section 3.1, each speaker is represented by a single text per register, which renders comparisons of group means (i.e., of each author’s texts) across registers impossible. While the design of this study does not permit such an investigation, it should be acknowledged that this analysis is necessary for a more complete profile of each author’s language use across registers – namely, an examination of whether all the works of a given author are positioned in a certain way with respect to the register mean or whether these positions vary.

a text in line with their cross-register tendency. Yet, this author's deviation neither aligns with their cross-register tendency (i.e., they did not produce an informational text), nor is it caused by the restrictiveness of the register. This result suggests that authors may be guided by considerations other than their individual preferences or register. These considerations may involve effectiveness, i.e., what the authors believe to be effective in accomplishing the goals of a certain task, or appropriacy, i.e., what the authors believe to be expected in a certain context, among others.

While some authors showed tendencies persisting across registers, the most frequent pattern in the corpus is that of authors who are highly inconsistent in their positions in different registers. These authors show not only varying degrees of conformity to the register norms, but different pattern of direction of deviation from that norm when their texts are quite removed from the central tendency. In the case of the author analyzed below (Author 52), for example, while compliance with the general trend was shown in evaluations, interviews, and essays, a pattern of deviation towards a more oral and elaborated extreme emerged in the author's email, and this pattern was considerably more pronounced in their business memo. On the other hand, the same author's description is entirely lacking in elaboration and focuses only on the image at hand. This shows that the author chooses a variety of different strategies as they produce texts of different registers. While the norms of essays, evaluations, and interviews may have been interpreted by the author as relatively fixed and unambiguous, they may have felt that these norms were less rigid in emails, business memos, and descriptions. It may have been this perceived latitude that resulted in the author fulfilling the expectations of the registers in their own unique way.

The varying strategies observed in these three registers may further be the result of differences in the nature of the tasks. The register of emails is naturally occurring, and the author's email is an example of a real communicative need in an important university project. Elaboration in this text as a result of the author's involvement is therefore not surprising. Heightened involvement in the business memo, on the other hand, cannot be the result of the same factors due to the fictional nature of the scenario; however, the presence of an interlocutor, even if a fictional one, may have led the author to adopt a personal tone and make the suggestions relevant to a hypothetical addressee. Alternatively, previous experience or an interest in the business context may have resulted in this heightened engagement. The opposite trend in image descriptions may in that case reveal their lack of interest in the task due to a lack of a real interactant or a real purpose.

While the exact reasons for such choices are impossible to identify, the fact that these authors are the most representative of the corpus (rather than the authors with a clear trend toward elaboration or information presentation) suggests that authors not only adapt their language depending on the situation of use, but also authors may show variation with regard to their perceptions of the strictness of register norms. Certain registers may be interpreted as more situationally restrictive than others, and if a register is perceived to allow variation, authors choose varying communicative strategies in meeting the demands of such registers (e.g., focus on the content vs. subjective interpretation in image descriptions). An additional consideration is that authors differ from each other with regard to the registers in which they show deviations from the central trends. That is, some authors may consider evaluations restrictive and only see one way of meeting the demands of this register, while others may see possibilities for variation in that register, but view the norms of other registers, such as essays, as fixed.

The direction of deviation offers an even more detailed insight. For example, authors who deviate in the direction of the register's general trend (i.e., if a register is generally informational, their deviation towards greater information density is then in line with this trend) differ from authors whose deviations go against the general trend (i.e., if a register is generally informational, their deviation towards greater oral elaboration could be viewed as more extreme). Identifying reasons for such deviations appears to be a more feasible task if the author's identity is known and the observed tendencies can be supported through the existing research on their language use. At a minimum, the analysis here is intended to show that analysis of individual variation can be enhanced through the additional considerations of the author's perception of register, and specifically, their approaches to a variety of diverse contexts situationally defined to different degrees and allowing varying degrees of linguistic freedom.

7.1.1 Case Study 1: Authors 'Central' across Registers

Author 12 was selected for this case study as an author with a majority of texts (five out of six) positioned close to the register means. This author's image description and evaluation are found in the 'Central' distance group (< 0.1), and their email, essay, and interview are in the 'Minimally Removed' group (0.1-0.3), also considered a representation of the central tendency. This author's business memo, however, is at some distance from the register mean in the 'Moderately Removed' group (0.3-0.6). Nevertheless, the overall inclination shown by this author in most cases appears to be towards texts that are prototypical exemplars of a register rather than exceptions from the general trend. This author's texts in the registers where they represent the central tendency are shown below.

'Central' Texts (< 0.1): Image Description and Evaluation

Descriptions ($M = 0.44$; $SD = 0.51$)

Text Sample 102 is an excerpt from the author's image description. Image descriptions generally gravitate towards oral elaboration as can be seen from the descriptive statistics for the register. However, as noted previously, the register's position on the dimension should be interpreted in relative terms rather than represent an absolute preference towards one dimension function or the other. Informational features were shown to contribute to the register in important ways, commonly referring to objects or other notions inferable from the visual. Descriptions thus tend to incorporate both sets of features as both functions – oral elaboration and information packaging – are performed by the register. This author's image description represents this tendency. Elaboration is more prominent as the author accompanies the description with an interpretive commentary, whereas some informational features serve the purpose of concrete references to an object (*computer desk, glass table*) or judgments based on the appearance of the people in the visual (*teen years*). Elements of the two contrasting approaches to the task of an image description – the focus on the content of the visual and subjective commentary – observed earlier in the peripheral texts are combined in the text of this author, both present in moderate rather than extreme amounts. Rather than giving precedence to one communicative strategy, this author alternates between the two, focusing on what is immediately present in the visual (the mention of the furniture items and the participants' positions, their clothes, and their appearance) and what the author infers from this content (the fact that one person is guiding the other, the emotion of surprise, or the fact that they may be waiting for someone). Both strategies effectively combine to serve the goals of the register, and while the personal commentary may not be required by the task of description, it enhances the description and adds substance to it,

thus contributing to the same end. The author's choice of intertwining the two strategies may demonstrate their awareness of this effectiveness.

Text Sample 102. Dimension 1 score: 0.50

And they both **seem** to be **very** young, er, **but** the person, the girl who's sitting, **seems** to look like she's in her teen years. **So** she **looks** like a teenager, or **maybe** in her early twenties. Erm, **just very** young. **And then** the teacher, or the lecturer, or whatever, **is**, **looks** like guiding her, er, erm, **looks** like she's in **maybe** her mid to late twenties. Erm, the student **is**, **looks** like she might **possibly** be seated at a computer desk, and that **is** what she's being helped **with (stranded)**, **and** she **looks very** surprised. **And they** both **seem** to be in **very** casual attire. [...] the background **seems** to be **very** casual **once again**, **it seems** to be a beaded erm, circ- circular thing on the background. **It looks very** nice. **And** there **seems** to be a table, a glass table, in front of them. **So it** could **just simply** be a place where you can't come together to study, but **just** to catch up and speak, **and maybe** they're waiting on someone.

Evaluations ($M = 0.15$; $SD = 0.24$)

Evaluations as a register also presents a balance of oral and informational features; however, the role of informational features is more prominent than in descriptions, as can be seen from the descriptive statistics for the register. Again, Author 12 maintains that balance in their texts, using a substantial number of nominal sequences to name the campus facilities and campus features (Aston campus; Aston students Union; prayer room (facility), lecture halls; water fountain), which are the objects of the evaluation, and student activities these facilities are meant for (group work meetings; revision time; university experience). Features of elaboration, on the other hand, are used to express the author's personal evaluation of these facilities (**really useful; always very clean; as it is large and very beautifully designed; is also very beautiful during the day and night**). Again, both sets of features work together to fulfil the goal of the register, and neither can be dispensed with without compromising the efficiency of the text. This strict need for both informational and oral features for the core goal of evaluations appears to explain the lack of variation within the register. Author 12, alongside most other authors, is therefore quite restricted

in their linguistic choices, as a certain balance of oral and informational features is required by the situational constraints of the register – namely, the communicative purpose of evaluating Aston campus resources (which first need to be named with the help of informational features).

Text Sample 103. Dimension 1 score: 0.19

Aston campus **has** many different things that made my university experience convenient. **Firstly, it has** a lot of open and close space to study alone or with a group throughout the main building, library, and the Aston students Union, **and** these spaces were **really** useful during group work meetings and revision time and a place to rewind watching shows and eating snacks, **moreover** all of the rooms were **always very** clean and have a warm pleasant atmosphere in them. **Furthermore**, as a Muslim student I **really appreciate** the new prayer room facility **as it is** large and **very beautifully** designed, **and** the prayer room as well as the washroom for ablution **is** located in a **really** nice convenient place. The lecture halls **are also always** clean and tidy, **it is** spacious and **has** place in front in some rooms to charge your laptop if needed. The green scenery and water fountain of the university **is also very** beautiful during the day and night.

‘Minimally Removed’ Texts (0.1-0.3): Email, Essay, Interview

Emails ($M = 0$; $SD = 0.38$)

Emails are still lower on the dimension than image descriptions and evaluations. The register thus again integrates features of both extremes. The shortness of emails, however, may have resulted in the register’s lower position on the dimension than the other two registers, and thus the lower scores are to an extent the result of both sets of features’ failure to occur. The central exemplars of the register, such as Text Sample 104 below, may thus not seem as rich in Dimension 1 features overall as the previous examples from the other two register. Nevertheless, both oral elaboration and information density fulfil important functions in the register. In Author 12’s email below, features of information density convey notions central for the essay topic the student is discussing with the professor. The minimal deviation from the central tendency in this text therefore reflects the class content-related topic of the email. Elaboration, however, is necessary as the author shares their thought process and reasoning as they propose the topic and ask for the addressee’s feedback, and the goal of the email would not have been achieved

without either set of features. Author 12's email is therefore a representative example of the way oral elaboration and information density co-construct email communication.

Text Sample 104. Dimension 1 score: -0.25

Good evening,

Okay, I understand I should focus on the societal expectations of motherhood **since this is** the sociological understanding and not the psychological. Would **it** be okay if I made an essay about ' ' Where I will be writing about Japan low birth rate being blamed on women working and choosing to have voluntary infertility, which **is** being said **it's because they're** imitating men, Hindu culture that **states** Pregnancy **is** stated as the start of womans adult identity and Uganda women not being accepted to society until they have at least 2 or more children.

Is it okay if I **don't** focus on the UK?

Yes, I would like to book an appointment please,

Thank you

Regards

Essays ($M = -0.24$; $SD = 0.27$)

Similarly, in essays, Author 12 represents the general register trend, namely serving the informational as well as the discursive goals of the register, with clear precedence given to the former. The essay below relies quite heavily on informational features, which denote the core phenomena related to Buchanan's presidency. The essay is primarily focused on historical facts, and for the most part the author does not offer their interpretation – only a few features of elaboration are present (present tense, conjunct, adverb). It is apparent from this passage that elaboration of the historical events is, in fact, present in the text in the form of nonfinite adverbial clauses, not associated with the dimension but providing additional information about Buchanan's actions; however, the function of these clauses is different from the oral elaboration which is associated with the dimension and has been shown to reflect the authors' subjective analysis. While this subjectivity was observed in peripheral exemplars considerably above the register mean, it is generally not characteristic for the academic register of essays. Author 12's

text illustrates that and shows that while some features of elaboration are involved, the register prioritizes information density and typically does not leave much scope for subjective expression.

Text Sample 105. Dimension score: -0.50

‘The Kansas-Nebraska Act in 1854 reopened the stipulated query over state-wide slavery (Mometrix, 2021). [...] The faith of Kansas located in the mid-west becoming either a freed state or a slave state **has** reached boiling point (Mometrix, 2021). Border ruffians helped to secure a proslavery legislature in Kansas, which drafted a proslavery constitution known as the Lecompton Constitution. **Meanwhile**, anti-slavery activists established an extra-legal regime of their own based in Topeka. The political and ideological disputes over the Kansas- Nebraska Act opened bloodshed for 5 year and polluted the endorsers' Pierce and Douglas Democratic candidate campaign, allowing a perfect campaign run for Buchanan’ (Khan Academy, 2020). ‘Buchanan took a conventional presidential campaigning approach: making no public appearances and **very** few press statements, choosing to leave his campaign to his supporters named ‘Buchaneers’; Fremont did the same whereas budding Republican Lincoln campaigned **heavily**’ (Cooper, 2020).

Interviews ($M = 1.62$; $SD = 0.21$)

Conversely, interviews are heavily subjective, which is reflected in a high concentration of features of oral elaboration in all texts of the register. The register shows minimal internal variation, as oral elaboration serves its essential goals – to provide detailed responses on a personal topic. Author 12’s interview represents that trend, with numerous adverbial features expressing stance and clausal coordination contributing to that goal most notably.

Text Sample 106. Dimension 1 score: 1.79

So, erm, I went **downstairs**, had pancake **and then**, erm, **yeah, and then, I just, and then, yeah, and then** I did, **and then, obviously**, you **know**, I did my own things. **And then** by, erm, lunch, by the time they were fini she'd I **think** around three, four, erm, my mum made pasta-not pasta, rice and **this** a, a **very** traditional dish, rice and like meat and, erm, a salad and, erm, what, **it's** like sauces kind of. Erm, **and then it has** a banana **as well**, we **actually** we add to **it**. Erm, **and yeah, basically** I had **that** and a drink, **obviously, as well**. I had, erm, Pepsi, **and then**, erm, we, **yeah, then I just** did my own things **again and then**, erm, my sister **actually**, my younger sister is, erm, **really** interested, I **don't** know why **but** during quarantine she's been **really** interested in cooking **so** she made us these brownie kind of things.

This sensitivity to register conventions is consistently revealed by the author across five registers, all quite distinct in their situational characteristics and cutting across a variety of settings. Written academic registers like class essays, school email communication, specific tasks like evaluation and image description, and an informal interview all require understanding of quite different situational contexts. Author 12 varies their language across these distinct contexts and shows a tendency to conform to their expectations in each case.

‘Moderately Removed’ Texts (0.6-0.9): Business Memo

Business Memos ($M = -1.01$; $SD = 0.79$)

However, Author 12 shows some divergence from this trend in the register of business memos, where their text is found in the ‘Moderately Removed’ group. While the informational focus of the register is still clearly maintained, Author 12 introduces elaboration through multiple references to review ratings justifying their recommendations. Adverbial features also express personal stance (*very nice café; option is the closest at **only** being 2 minutes away; **highly** rated shops*), which is supported through expert opinions. While some deviation in this register is apparent, it is clear that in most of their texts Author 12 shows a tendency toward conforming to the register trend. It may also be argued that this deviation does not demonstrate a lack of understanding of the register conventions, but conversely, enhances them and results in a highly effective business memo that meets the goals of the register.

Text Sample 107. Dimension 1 score: -0.63

For the 2 day conference from 2-4 September 2021. Here **are** the following itinerary items I **have** found for your trip: Hotel Hotel- Clariton Hotel Helsinki which **is** a £129 night stay and **has** been rated a 4.4. Hotel- Hotel Indigo Helsinki Boulevard which **is** a £103 night stay and **has** been rated a 4.5. Hotel- Hilton Helsinki strand which **is** a £174 which **has** been rated a 4.4. Cafe For the cafe options all these hotels **are** located near the market square of which the second option **is** the closest at **only** being 2 minutes away **and**

there **is** a **very** nice cafe called Restaurant Haven located in the famous market square. Souvenir shops There **are also** several souvenir shops located in the market square, the two I **recommend are** Liopto and Finnska souvenirs **as** they **are** the largest and **highly** rated shops to buy traditional Finnish souvenir items.

7.1.2 Case Study 2: Authors ‘Peripheral’ across Registers

While no author in the corpus was found to be at the periphery of a register at all times, in the case of some authors, a trend towards peripheral texts with a clear inclination towards one dimension function or the other across registers can be traced, as at least half of their texts are found in the peripheral groups above or below the mean. This section presents two contrasting authors, who gravitate towards information density or oral elaboration across registers.

Tendency towards Information Density

Author 66 is an author with almost all texts (5 out of 6) positioned below the register means (with the exception of their essay, which is at the mean). The author thus shows a consistent tendency towards information density, and in three registers – image descriptions, emails, and business memos – this tendency is much more pronounced than in others. The Text Samples below illustrate this trend.

‘Peripheral 1’ Texts (0.6-0.9): Image Description and Email

Descriptions ($M = 0.44$; $SD = 0.51$)

Author 66 is considerably below the mean in image descriptions, which on the whole are oral and elaborated. Author 66’s description is quite short (Text Sample 108 contains the whole text), which contributes to the low text score. The text still contains features of oral elaboration, but these features are not nearly as prominent as in the typical descriptions in the register, and the description is overall quite limited in both the amount of covered content and the author’s own

interpretation of it. The description contains three nominal sequences, which are repeated and in essence convey the author's main focus (*learning environment* and *university backgrounds*), as the author reiterated throughout the description that the students were learning, not adding much additional detail. This text thus exemplifies a noticeable lack of elaboration. This tendency to concision may have also led the author to use nominal sequences, which in this text convey the author's main idea in a highly compressed manner in place of lengthy elaborated structures. A combination of these two trends results in the author's peripheral position in the register. It is important to note again that all texts of a register, i.e., texts that conform to its conventions to a greater or lesser extent, are still perceived to be instantiations of that register. That is, the text produced by Author 66 differs substantially, linguistically and situationally, from the 'central' image descriptions. The author does not cover all the required content and does not provide extensive commentary. Yet, this text represents their individual interpretation of the task, their choice of how to approach it, and may be the result of a combination of individual and communicative factors, such as a lack of interest and engagement, lack of focus, lack of experience with imaginative and interpretive tasks, or an intentional choice to state what is minimally sufficient. Regardless of the exact reasons, the text represents the author's attempt to produce a text in response to the same task that was addressed by the other authors. The individual approach adopted by this and other peripheral authors is thus a valid example of the register and contributes to variation within its bounds.

Text Sample 108. Dimension 1 score: -0.27

So the first image **looks** like a student being helped. A s-student learning **even**. Erm, **it looks** like a teacher assisting her. Erm, the second image **looks** like **almost** the same thing but **instead** two students helping each other. Erm, both **look** like... yeah, **just looks** like college, like uni. Like **just** lear- a learning environment. Erm, third one, kind of **looks** like they're both learning instead of like the two images where they were, they had

the other one helping each other. **It** kind of **looks** like these two are both learning from someone. Erm, all three images kind of **look** like **really** relaxed learning environments. Erm, **yeah**, they **look** like college or university backgrounds. Mm.

Emails ($M = 0$; $SD = 0.38$)

Author 66's email, analyzed earlier as communication on official matters, is again quite short, which contributes to the low dimension score. Some oral elaboration is present in the passage as an explanation of the personal circumstances surrounding the card loss is necessary to fulfil the goals of the text. Nominal sequences, however, are essential to this email and contain crucial references. While in the case of image descriptions, this lack of elaboration and informational focus could be viewed as an individual choice, this appears less likely in the author's email, where the nominal sequences are required by a communicative consideration, such as topic. Nevertheless, this email contributes to the pattern of the author's consistent tendency, intentional or otherwise, to rely on features of information density and not use extensive elaboration, partly due to brevity.

Text Sample 109. Dimension 1 score: -0.77

Dear Sir/Madam, On the I lost my student card **and** as a result of external financial circumstances I was unable to buy a new one **as** I had **already** a previous one. I spoke to 2 of my lecturers and let them know **and** they said **it** would not be an issue to speak with the attendance team if **it needs** to happen. **Today I am finally** able to buy a new student card. I **have** been attending all my lectures except this week I **have** not attended any. I **hope** you **are** able to understand my situation **and** I would **really** not like my attendance to be **negatively** affected. Thank you

'Peripheral 2' Texts (> 0.9): Business Memos

Author 66's business memo is the most extreme example of their peripheral status. Text Sample 110 illustrates that elaboration is almost completely lacking in this text, which is driven solely by nouns. Even the features not associated with the dimension are nominal, and no clausal features are present, making it impossible for prominent features of elaboration to appear (such as the present tense). Again, for individual or communicative reasons, the author shows a tendency to

brevity and presents minimally sufficient information, only covering the search items required by the task.

Business Memos ($M = -1.01$; $SD = 0.79$)

Text Sample 110. Dimension 1 score: -2.13

Radisson Blu Seaside hotel - £187 for 2 nights Cargo breakfast cafe, 1 min drive from the hotel Konstan Molja, traditional food restaurant with lots of variety not **just** seafood, 1 min drive from the hotel National museum of Finland, good reviews, most popular for history of Finland, 7 min drive from hotel Annensoppi souvenir shop, good souvenirs for kids and adults, 7 mins drive from hotel Apollo live club, live music and no opera, 6 min drive

While in two other registers – essays and interviews – Author 66 is in the ‘Central’ and ‘Minimally Removed’ distance groups, thus representing the central tendency of those registers, the author shows a clear preference for information density in the three texts analyzed here. Factors affecting these cross-register differences may vary, and while some possible reasons for the peripheral status and specifically minimal elaboration in descriptions, emails, and business memos were mentioned above, factors that preclude deviation in the other registers may include the highly restrictive nature of the essay topic and goal which required a greater elaboration than the author would have produced under different circumstances, while the spoken mode and casual setting of the interview results in oral elaboration in the case of all authors. The registers in which Author 66 does deviate from the central tendency, on the other hand, may have been free of such restrictions outside of the author’s control (such as topic) and may thus have allowed the author a greater opportunity for self-expression.

Tendency towards Oral Elaboration

While the extent of deviation is not always extreme – four of the registers produced by Author 45 are in the ‘Moderately Removed’ distance group and one is in ‘Peripheral 2’ (> 0.9) – author

45 shows a pattern of a consistent preference for deviation across registers, and in four of these texts the preference is for oral elaboration. The Texts Samples below demonstrate this tendency in the author's image descriptions, emails, essays, and interviews.

'Moderately Removed' Texts (0.3-0.6): Image Description, Email, Essay, Interview

Descriptions ($M = 0.44$; $SD = 0.51$)

It has been shown that oral elaboration is consistently used in descriptions to express the author's interpretation of the content of the visual or add a personal perspective. The image description in Text Sample 111 discusses the level of engagement of the participants, adding background information about the purpose of their meeting and the formality of the setting. The author hedges their statement throughout the text (*seems/ looks like; perhaps*) as almost the entire text is the product of the author's interpretation.

Text Sample 111. Dimension 1 score: 0.84

Erm, **it seems** as if, erm, there's a dark haired boy **and then** there's a blonde, erm, brunette haired boy. **And** the darker haired boy **seems** like he **is**, erm, having a look, erm, at something on the laptop while the brunette boy **is** kind of chatting to him. They **look** like they're friends. They, they **don't** look like they're in a formal setting **at all**. Erm, they kin-, erm, the black haired boy, erm, he's kind of, erm, more engaged in, er, like, erm, whatever's he's looking at in the laptop, erm. **And** the one, the other his friend, he **looks** like he's, erm, more engaged in the conversation rather than, erm, you **know**, the laptop **anyway**. Erm, the third image **looks**, erm, like there's two people. They **look** like they're **perhaps** meeting some friends **and it's just** them two that the picture **is of (stranded)**.

Emails ($M = 0$; $SD = 0.38$)

The author's email on a class content-related topic represents an exception from the trend discussed above, namely that of content-related topics gravitating towards information density. This email illustrates that the text is, in fact, not content-heavy, although it does use nominal sequences related to the content and in that supports the trend. However, the email is dominated

by the author's thoughts about the topic rather than information about gender representation in various jobs. The email clearly shows the author's interest and engagement with the issue, which leads them to raise the question with the professor, explaining the reasons for their interest and describing the search they did. Elaboration is thus quite prominent. The fact that the author incorporates such extensive elaboration even in an email that is written on an academic topic may suggest that elaboration is their preferred communicative strategy even in tasks that involve information presentation.

Text Sample 112. Dimension 1 score: 0.51

Hi, I was **just** reading a education topic essay that you **have** given us couple weeks **back**, **its** from tutor2u. **It mentions** that more primary school teachers **are** female (I **know that for sure because** we **have** learnt about it) **but then goes** on to say "heads and deputies with ultimate responsibilities for the school **are** likely to be men". I **have** looked **online** to see if i can find any evidence **but** nothing I've come across **supports it**. I **think** if it is correct, **its** a **very** good point to have for gender differences, for boys and girls. **So**, I'm **just** wondering if you **think this is** correct and something worth remembering?

Essays ($M = -0.24$; $SD = 0.27$)

Author 45's essay is a discourse analysis of interview data on the topic of students' Covid-19 experiences. It was shown earlier that such essays tend to rely heavily on oral elaboration – they raise questions quite personal to the interviewee, who shares their experiences and is thus highly involved. This involvement is then conveyed by the author who presents their responses and their analysis of them, identifying the prominent themes. It is therefore not clear if this extensive elaboration, which appears to be the result of the topic, can be attributed to the author's individual preference. The author may have been restricted by the topic and the analytical purpose of the essay, which determined their language use; alternatively, the author may have chosen the topic and the task that aligned best with their preference and natural tendencies.

While these reasons can only be investigated if more texts written by this author on a variety of

topics are examined, in the data available for this study this essay contributes to the cross-register tendency to oral elaboration shown by this author.

Text Sample 113. Dimension score: 0.19

Natasha **draws** up on the fact that the university experience she **is currently** having **is very** different to what she had prior to the COVID-19 restrictions in her first year. Although, she **highlights** that she had been able to enjoy some of the perks of being a student in her first year of study, she **emphasises** the lack of **it now** and **realises** that she had taken for granted some of the opportunities she had in first year, **as** she **states** - “I **definitely** should have taken those erm (.) opportunities and **just** spend **it** with the people I **like** being around”. **This** may highlight the idea that Natasha **feels** regret **because** certain experiences **are** not as **readily** available as they were **before**.

Interviews ($M = 1.62$; $SD = 0.21$)

Finally, the high dimension score in interviews, the most oral and elaborated register overall, is not surprising. The personal topic of the interview and the casual setting lead the author to explain their routine and food choices in detail, expressing stance and showing involvement throughout the interview.

Text Sample 114. Dimension 1 score: 2.12

Erm, **so normally**, every single day I **have** around two full meals. **So**, that would be breakfast and dinner. **So**, for breakfast, I **normally have** either baked oats, erm, **and** that will be with different toppings **so, for example**, be like golden syrup, white chocolate. **And then** I can have like a cup of tea or a hot chocolate with **it**. Other times, I might be having scrambled eggs, toast and, erm, tea. I **don't really** do **too much** breakfast **because** I **normally just** kind of **make it** to the bus **because** I'm rushing for work. Erm, **and then, o-obviously**, I'm working most of the time during the day **so I don't really** have lunch **often**.

Business Memos ($M = -1.01$; $SD = 0.79$)

In quite stark contrast with this consistent trend is the author's business memo, where the elaboration that the author produces in other registers is entirely absent. Text Sample 115 contains a complete text of the author's business memo and shows that the author does not fully address the task. As discussed earlier, reasons such as a lack of familiarity with the business context or inability or unwillingness to construct an imaginary relationship with a fictional

superior in the text may have influenced the author's choices. These reasons, on the other hand, may include external circumstances, such as a lack of time, as it is apparent from the author's other texts that a more oral and elaborated business memo in their case could be quite likely. In the previous chapter, the register of business memos was shown to permit quite extensive variation. This suggests that the author should not have been restricted by the situational demands of the register and could have made the choice in line with their tendency to increased oral elaboration illustrated in the other registers. Yet, in their interpretation of what constitutes a business memo, the author appears to be guided by considerations other than their individual linguistic preferences or register.

Text Sample 115. Dimension 1 score: -2.42

Clarion Hotel Helsinki - £108 per night (under budget) Early Bird Restaurant - for breakfast (near the hotel) Helsinki City Museum - for city history Lappi Ravitola - **serves** Poronkarystys (national dish - sauteed reindeer) Annensoppi - souvenir shop Apollo live club - **has** live gigs

7.1.3 Case Study 3: Authors of Different Distance Groups and Directionality across

Registers

Finally, while some authors showed a preference for oral elaboration or information density, with at least half of their texts gravitating towards a particular dimension function, the majority of authors in the corpus are less consistent and show varying tendencies across registers. Such authors show much more complex and versatile patterns in their varying interpretations of register conventions. Author 52 is representative of this large group of authors whose texts are found in different distance groups, deviating from the mean in different directions.

'Central' Texts (< 0.1): Evaluation

Evaluations ($M = 0.15$; $SD = 0.24$)

Author 52's evaluation is a text that best represents the central tendency, being positioned extremely close to the register mean. The text incorporates both the necessary references to the Aston facilities through nominal sequences, which is characteristic for the register, and elaboration, which conveys the author's subjective judgment and its justification, essential for the goal of evaluations as well. The author conforms to these expectations and maintains this balance without favoring one extreme or the other.

Text Sample 116. Dimension 1 score: 0.09

Aston campus is home to thousands of students and **has** many good qualities, ideal for a range of things. One of them being **its** location. When I first visited Aston on an open day, I **absolutely** loved how close **it** was to city centre. Birmingham city centre **has** amazing establishments for students to socialise **at (stranded) and** knowing **this** was a stone's throw away **makes it** perfect for any student. When I had lectures on campus, I was able to get food or take a break somewhere with my friends **and** the convenience of the location made **it** possible to do **so** without having to worry about timing, which I **know** in other universities, they **do** not have this luxury. **Now**, my favourite thing about Aston campus, **is** the university library.

'Minimally Removed' Texts (0.1-0.3): Essay and Interview

Essays and Interviews are again quite close to their respective register means, which suggests that the author shows awareness of their commonly accepted norms. The author's essay (Text Sample 117) employs a combination of oral and informational features as it discusses symptoms, diagnoses, and test results. Nominal sequences in the passage denote the tasks implemented in these tests, while elaboration is used for presentation clarity and organization (conjuncts) and additional explanations (present tense). Again, the author does not favor either set of features to a substantial degree, and both contribute equally to the goals of the text, overall reflecting the intermediate status of essays on the dimension.

Essays ($M = -0.24$; $SD = 0.27$)

Text Sample 117. Dimension score: 0.03

A diagnosis of ASD according to the DSM-IV-TR **entails** a minimum of two social impairments, one communication and one in restricted and repetitive behaviours (RRB) (American Psychiatric association, 1994). [...] ToM **is** the ability to attribute mental states to oneself and others (Tager-Flusberg, 2006). 80% of children with ASD, failed the Sally Anne false belief task (which **investigates** ToM) (Baren-Cohen & Leslie; Frith, 1985) compared to **typically** developing (TD) children. **However**, 20% of children **still** passed. **This doesn't** account for performances by higher-functioning individuals with autism (HFAI), as Happe (1995) **states** passing false-belief tasks **are** predicted by verbal mental age. **Furthermore**, Happe's 1994 strange stories test found individuals with autism (IWA) showed difficulties with irony, sarcasm, and metaphors.

Interviews ($M = 1.62$; $SD = 0.21$)

As noted above, a lack of deviation from the mean in interviews is to be expected: the register has been shown to be quite uniform, with all its texts found in the oral and elaborated range. Text Sample 118 shows that Author 52 provides a detailed account of their day, constructed most notably through adverbs, clausal coordination, and the present tense.

Text Sample 118. Dimension 1 score: 1.81

I **wake up, probably**, in the afternoon time about one or **generally** about **now**. Erm, I **always drink** a cup of tea, I've **always** drunk a cup of tea since Year Seven in school, erm, I **need** that tea to **fully** wake me up. **And** if I **don't** have **it**, **then** I can **really** feel the caffeine effects. Erm, I **only have** the one cup, **though**, erm, my mum **normally has** like a cup an hour. Erm, **but anyways**, after **that**, erm, **occasionally**, I'll have some eggs. Eggs **are** my favourite thing to have for breakfast. Erm, I **love** fried eggs.

'Moderately Removed' Texts (0.3-0.6): Email

Emails ($M = 0$; $SD = 0.38$)

Some noticeable deviation from the central tendency towards oral elaboration is observed in the author's email. This email, whose topic is categorized as class content-related, requests clarification on the stages of an experiment. Yet, Author 52 uses a substantial number of features of elaboration, such as discourse-organizing adverbs, demonstrative pronouns, clausal coordination, and the present tense, as they share their understanding of the relationship between the stages and their order. This detail-orientedness suggests involvement, which leads the author to integrate features of oral elaboration even in a highly technical text.

Text Sample 119. Dimension 1 score: 0.34

Hi, Not sure what you **mean** by a blocked order. **Does this** mean all no load trials presented and **then** all positive trials presented and **then** all negative trials presented and **this** would be counter balanced **so** some participants would get all negative trials first and **then** all positive trials and **then** all no load conditions. **And this** would be based on the B-G list you gave me? **And** the powerpoint I sent you **is** version A **and** would **that** be the mixed condition? **So** half pps would get no load, positive, negative, no load, positive, negative. **Or do** I need to change **this** to **also** be mixed **so for example** no load trial, positive trial, negative trial and **then** positive trial, negative trial and no load trial. **Or do** I **just** let half pps do version A that I sent you and leave that how **it is and** the other half **just** do the B-G blocked condition which will be randomised. **Also want** to say about the amount of pps **as contact hours are** max 30 **so this** would mean **only** able to get 40 pps instead of planned 60. Thank you for your help **again!!** Kind regards,

‘Peripheral 1’ Texts (0.6-0.9): Image Description and Business Memo

Finally, the author’s image descriptions and business memos demonstrate the opposite trend from that of conformity to the register norms found in interviews, essays, and evaluations and to a lesser extent in emails. Conversely, the image description and business memo below show a substantial deviation from their respective means. This deviation, however, is in opposite directions, the image description showing a pronounced tendency towards information density and a lack of elaboration and the business memo supplying extensive detail and relating the suggestions to the addressee’s expectations, gravitating towards more elaboration than is characteristic for the register.

Descriptions ($M = 0.44$; $SD = 0.51$)

The image description in Text Sample 120 illustrates that the author is highly focused on the detail of the visual, which results in several nominal references (*denim jacket; Apple MacBook; furniture thing; polka dots*) and elaboration is only present in the form of the present tense. No detail is added to the content of the visual, which results in a lack of specificity usually created

through adverbial features. The author also does not offer any interpretation of this content, and the text is almost entirely lacking in their subjective stance.

Text Sample 120. Dimension 1 score: -0.43

She **is** wearing a blue denim jacket. She **has** braids in her hair **and** she **is** looking in another direction. The woman standing up **appears** to be helping her with something. The woman standing up **is** white and the woman sitting down, erm, is Black. In the second image, there **are** two students. One **is** of Asian descent and the other one **is** white. They **have** a Apple MacBook in front of them. There **is** a green furniture thing behind them with green and dark green polka dots. The m-, the guy, the Asian guy **is** on the left **and** the white male **is** on the right. The white male **has** a black earring, he **is** wearing a grey floral top with pink writing and **has** brown hair.

Business memos ($M = -1.01$; $SD = 0.79$)

On the contrary, the author's perspective is evident in their business memo as every suggestion is explained and justified through the author's knowledge of the addressee's preferences and expert opinions. While the text contains the necessary references to locations, conveyed through nominal sequences, the author's tendency to oral elaboration in this text is quite pronounced.

Text Sample 121. Dimension 1 score: -0.31

Hi Sophie, I **have** compiled a itinerary of all the things you would want to do on your trip to Helsinki for your two-day conference on the 2nd-4th September. **Firstly**, I **have** found a hotel called the Radisson Blue seaside hotel **as i know** you **like** the water and would want to be near **it**, this hotel **is only** 1 mile away from the Hetrikani beach and **has** stunning ocean views from the room. **it is** £77 a night within your budget and **has** many features such as A/C, bar/lounge, a gym in case you would want to blow off any steam from the conference. **It has** many positive reviews and **is** deemed one of the best in Helsinki!

7.2 INDIVIDUAL VARIATION ON DIMENSION 4: EVIDENCE-BASED STANCE

In this section, the authors cross-register tendencies are analyzed with regard to evidence-based stance (Dimension 4). Again, three types of authors are identified: authors who conform to the register trends, i.e., their texts represent the register central tendency in at least half of the registers; authors who consistently show creativity with regard to register, i.e., their texts occur

on the periphery of at least half of the registers and show a relatively stable preference towards an increasingly frequent use of stance or lack thereof; authors of different distance groups and directionality across registers, i.e., their texts represent contrasting tendencies with regard to stance use across registers, with texts ranging from compliance to the central tendency in some registers to considerable deviation from it in others.

The study observes the same pattern with regard to authors who show a systematic deviation towards a particular dimension end. First, these authors clearly show a trend towards more or less stance expression across several registers. Second, this trend does not apply to all registers, and there are usually exceptions from this pattern. Even in cases of quite extreme deviation in a particular direction in most of their texts, the same author may show an equally extreme deviation in the *opposite direction* in another one or two registers. This suggests that even in cases of such cross-register preferences, these trends are not absolute. More importantly, these exceptions from the authors' individual trends were not motivated by the situational restrictions of the register. That is, the registers in which authors diverged from their preferences allow internal variation (as demonstrated in Chapter 6). This means that the authors should not have felt constrained by their situation and could have produced a text consistent with their preferences. Yet, these authors appear to have been guided by considerations other than these preferences or the restrictions imposed by the register. This finding seems to point to the existence of additional factors influencing authors' decisions, such as their own considerations of appropriacy and accepted norms (not necessarily in line with the actual demands of the register) or considerations of effectiveness – i.e., communicative strategies and approaches they believe to be better suited for certain tasks. Importantly, these considerations, which lead the authors to

diverge from their preferences, appear to reflect individual beliefs rather than actual register constraints.

As on Dimension 1, the most common pattern was observed in the case of authors who represent the central tendency in some registers as well as deviate from it in either direction in others. Again, these results show that authors appear to perceive the conventions of different registers as more or less restrictive and thus allowing more or less scope for variation. In the case of the author analyzed below (Author 26), for example, adherence to the norms of essays and business memos may indicate the author's perceptions of these registers' conventions as strictly specified by their formal contexts. On the other hand, interviews, evaluations, emails, and image descriptions may be perceived as less formal contexts, and, as a result, associated with less strict conventions or a lack of clear, accepted norms. A greater degree of latitude may therefore be thought to be possible in contexts with no rigid rules. That is, authors' approach to register seems to entail not only knowledge of the situational differences, but also awareness of the degree of the constraints imposed by the situation. These varying ideas of constraints then lead authors to attempt varying degrees of deviations from the register trends. These deviations again appear to point to individual choices of communicative strategies that authors believe to be more effective for particular registers (such as enhancing a description with an interpretation or addressee-focus and references to expert opinions in business memos).

The following sections present case studies of authors who consistently represent the central tendency, show systematic preference for an increased amount of stance or a lack of stance, and authors whose texts occupy a variety of positions across registers.

7.2.1 Case Study 1: Authors ‘Central’ across Registers

Author 50 is an extreme example of compliance with register conventions in all six registers, with five texts (essays, business memos, image descriptions, evaluations, and interviews) in the ‘Central’ group (< 0.1) and one text (emails) in the ‘Minimally Removed’ distance group (0.1-0.3). It was shown that the registers of the study rely on stance to different degrees, with some registers like emails and image descriptions integrating author-attributed stance, others, such as essays and evaluations, much less so, and interviews and business memos being generally low in stance. The author’s adherence to these general trends, unique in the case of each register, indicates the author’s sensitivity to situational distinctions and expertise in modifying language accordingly. Excerpts from this author’s texts are presented below.

‘Central’ Texts (< 0.1): Essay, Business Memo, Image Description, Evaluation, Interview

Essays ($M = -0.02$; $SD = 0.43$)

As discussed before, essays as a register occupy an intermediate position on the dimension and while their focus on objective information presentation is not associated with large numbers of stance features, some stance in the form of expert opinions or analysis of interview data is found in these texts. The excerpt below from Author 50’s essay shows private verbs denoting the goals of the text (*it will be scrutinised and critiqued*) and the analytical steps it will take (*will show*). The past tense contributes to these goal as it is used to report facts about the models comprising the topic of the text and thus lay the foundation for the evaluation of these models. As noted before, the author’s personal stance is not characteristic for the register, and it is apparent from this text that Author 50 contributes to that trend.

Text Sample 122. Dimension 4 score: -0.10

There are a variety of new and emerging models as well as older models and it will be **scrutinised** and **critiqued** on the bases of how effective or ineffective, they **were** in terms of helping the team in the process of creating a fully functional e-business. There will also be several real-world examples that will **show** success stories of each model and how it **was** used effectively. The notion in the essay will **show** the newer models are particularly effective for creating modern new e-businesses and older models while having their place and **can** be effective are becoming outdated. The B2C model is the fundamental core of our e-business. The model **was** originated by Michael Aldrich in 1979 who **used** TV as a medium to reach out to consumers (Kenton, 2019). The model is based on offering products to be sold directly from businesses to consumers. The model suits us as it **allowed** us to have complete control over the prices we charge for our products. This is in the hope that our profit margins are higher as we don't deal with retailers or intermediaries who charge for their services.

Business Memos ($M = -0.85$; $SD = 0.45$)

Although the author's business memo is positioned close to the mean, the text resembles peripheral exemplars written in the form of lists that show a tendency toward a lack of stance features. This text does, however, contain instances of private verbs, one of which makes the text somewhat more addressee-oriented (*enjoy your trip*). Overall, the minimal use of author-attributed stance in this text is in line with the general tendency shown by the register, whose goals typically do not entail stance expression.

Text Sample 123. Dimension 4 score: -0.92

ITINERARY FOR HELSINKI (2-4 SEPTEMBER) PLACE OF STAY: HILTON HELSINKI STRAND- JOHN STERBERGUN RANTA 4, HELSINKI. 2 NIGHT STAY. PLACES TO EAT: (CAFE; RESTAURANT) GREEN HIPPO PUNAVVORI, PUNAVUONENKATU 2, HELSINKI CAPPI, ANNANKATU 22, HELSINKI (**OFFERS** TRADITIONAL MEAT BASES DISHES) PLACES OF INTEREST: (MUSEUM, SOUVENIR SHOP, LIVE MUSIC) NATIONAL MUSEUM OF HELSINKI, MANNERHMINNTE 34, HELSINKI KANKURINBTUPA, POHJOISESPLANADI 35, HELSINKI NOSTURI, TELAKKAKATU 8, HELSIKI **ENJOY** YOUR TRIP! YOURS SINCERELY, LEE

Image Descriptions ($M = 0.28$; $SD = 0.58$)

On the contrary, image descriptions are conducive to personal stance, and Author 50's text again reflects this trend. The private verbs in the text denote both the author's own stance towards the content (*it looks like*) and stance attributed by the author to the characters of the visual (*he's sort*

of *understanding* it). As discussed before, this tendency is characteristic for descriptions, where the quantitative results – namely, the relatively high rates of occurrence of stance features – correspond to a functional diversity in the use of these features in accordance with the notion of foregrounding – statistical deviance and the subsequent psychological salience (Mahlberg & Wiegand, 2018). This trend naturally emerges in Author 50's text as they adopt a balanced approach to descriptions integrating both factual information and their interpretation of it.

Text Sample 124. Dimension 4 score: 0.31

Er, this image that I'm looking at, it **looks** like like a friend... er, there are two people in the picture, and one of the people is **showing**, er, the other person a funny sort of video or picture. And the other person is sort of laughing, erm, and holding her right arm to her chest, er, and laughing er, hysterically. Erm, and I'm now going to **describe** the second image. Erm, er, there are two people in this image and one person **seems to be describing** something that he's **seeing** on the computer that is in front of them, to the other person. Erm, and yeah, the other person **looks** like he's sort of **understanding** it. Erm, yeah, they're just sitting down and are- they're both in front of a computer.

Evaluations ($M = -0.28$; $SD = 0.37$)

In contrast, evaluations are lower in stance, although when stance features are present, they signal the author's engagement with the task and are often accompanied by proposals of future improvements in the sections of evaluations focusing on the negative characteristics. The evaluation written by Author 50 is an example of such use. The text is mostly factual, presenting concrete descriptive detail, and author-attributed stance is integrated to express the author's personal as well as other students' dissatisfaction and a hope for future action (*It's something that I believe should be addressed*).

Text Sample 125. Dimension 4 score: -0.35

It would be nice if there was another building close by to the main building which also had lecture halls and classrooms to free some of the pressure and reduce some of the crowding that occurs in the main building at times, especially when it's busy during the morning and end of the day. Secondly, there are constant issues with the gates for Astons halls residences. I **know (that del)** this may **seem** minor or trivial but it's an issue that I

feel is well known and many students **find** a problem while moving around the university. The key cards given to us often don't work while exiting the halls of residence. There **can** be times where after ten or more tries of placing your key card on the area to rescanned for exit or entry into halls. It's something that I **believe** should be addressed as it **can** cause delays at the gate as this is the only known and provided ways of exiting your halls to get to university campus. Hopefully, this issue is resolved in the near future.

Interviews ($M = -0.59$; $SD = 0.26$)

Author 50 represents the tendency of interviews to minimally rely on the type of clausal stance features associated with the dimension. Complex clausal structures are rarely found in interviews, where stance tends to be expressed through other features, characteristic for oral discourse, such as adverbs and modal verbs. Text Sample 126 shows that the stance features associated with the dimension are limited to rare private verbs, mostly denoting individual food preferences and thus also being limited in their functional range.

Text Sample 126. Dimension 4 score: -0.66

Erm, yeah, sometimes I **like** doing like potato hash with eggs, erm, maybe bacon sometimes, but I try and shy away from that. Erm, and like moving on to like snacks throughout the day, because I don't really eat until later, erm, yeah, I probably have like a cereal bar or some fruit. Erm, yeah, I've tried to like sort of keep healthy, because my diet before wasn't great. But it is what it is, you know. Erm, yeah, and then like, I like sort of erm, sometimes I have like two dinners, so like when I finish work at like five or six, I might eat like some pasta, but then I might get hungry again during like, around eight or nine. So it **depends**. Erm, I'm not really sure what I would exactly eat. Maybe some chicken and some rice actually. Erm, or some-some fish. So I do **like**, like fried, fried rice and prawns really. I **like** sort of like Asian sort of cuisine.

'Minimally Removed' Texts (0.1-0.3): Email

Emails ($M = 0.39$; $SD = 0.65$)

The author's email is minimally removed from the register mean, and overall represents the register tendency towards author-attributed stance. As discussed before, it is always the author's personal stance that is expressed through the dimension features. Author 50's email is an example of a text on a personal topic, offering a personal introduction and establishing rapport

with the addressee. Several of the private verbs and their complements convey the author's wishes and aspirations, while others denote their thought process or are used as part of email conventions. Therefore, overall the email is quite representative of the trend in the register and shows a characteristic use of stance features in personal contexts.

Text Sample 127. Dimension 4 score: 0.18

Hello, First of all **excuse** my lateness in responding to you regarding this. I was sort of overwhelmed this term with my university work and completely **forgot (that del)** I applied to this! My name is and I am delighted to have been allocated as your mentee. I study at and I am currently in my second year of study. I would **like to find out** more about you, what your job entails and why you became an as I too **want to** get into being a . Is it possible for you to do so through email? as I am currently on christmas break. I will be able to meet you face to face when I return in early and will send another email to **confirm** when and where I will be able to do so. **Looking forward to hearing** from you Kind regards

7.2.2 Case Study 2: Authors 'Peripheral' across Registers

Tendency towards Stance

In contrast to Author 50, who showed consistent conformity to the general trend of each register, Author 73 shows deviation from the register trend in four of their texts, and in three of them this deviation is quite extreme, placing the author in the 'Peripheral' distance groups. In these three texts – essays, evaluations, and image descriptions – the author shows a marked preference for stance features. This increased use of stance features in three registers, two of which do not generally favor stance expression (essays and evaluations), suggests that the author's interpretation of the situational characteristics of these registers is distinct from the general trend. Examples of this marked use are presented below.

'Peripheral 1' Texts (0.6-0.9): Essay and Evaluation

Essays ($M = -0.02$; $SD = 0.43$)

The author's essay fulfills the communicative goal of presenting a qualitative analysis of interview data. The high rates of occurrence of Dimension 4 features in this text thus support the pattern discussed earlier – namely, peripheral texts with this goal being prone to stance expression, as both the interviewee's stance and the author's analysis of the discourse involve author-attributed stance. The text contains a variety of Dimension 4 features, including private and public verbs and their complements reporting attitudes as well as the speech of the interviewee. The past tense serves to support the author's analysis with the evidence of the interviewee's statements (*was made evident through the repetition of "unfortunate" which was mentioned three times during the interview*).

Text Sample 128. Dimension 4 score: 0.68

The sense of the pandemic being a difficult circumstance **was** made evident through the repetition of "unfortunate" which **was mentioned** three times during the interview. This repetition **resonates** with an individual thus **spotlighting** how bad her experience is. Mia is unable to shake the ill feelings off and due to this, it is likely that Mia **cannot** help but **call** the circumstance unfortunate as it **seems** like it is the only key word that can truly capture what the feeling is like. Social interactions due to government lockdowns has **meant that** Mia has been unable to socially interact with others [...] Mia has likely come to terms with it and **accepted** it through the "unfortunate" being paired with "...it is what it is". It **suggests (that del)** her experience serves some kind of ambiguity and later **elaborates** the effects of this ambiguity by **saying (that del)** the situation "is out of my control", **rendering** her useless in a sense. Mia's emphasis on the word "much" **denotes** a significant emotional toll on her, regardless of anything else that may be happening in her life.

Evaluations ($M = -0.28$; $SD = 0.37$)

Prominent use of stance features is even less characteristic for evaluations, which tend to be factual and descriptive rather than personal. However, this author's evaluation exemplifies the heightened personal involvement discussed as the basis for the linguistic differences in the peripheral texts above the register mean. This author uses personal stance rather than general statements regarding the Aston facilities to construct the evaluation. Again, as is commonly the

case in texts that openly express the author's stance, the author makes suggestions for future improvements (*I really do **think (that del)** this can be improved for ease and better signage for us*). The author uses the past tense to recount their experiences with the criticized facilities, and these experiences are used as evidence of the need for change. It appears that this author (and others who share their thoughts on improvements) view the task as a possibility to effect change, and stance features contributing to that goal add directness and urgency to the texts.

Text Sample 129. Dimension 4 score: 0.60

I also **dislike** how there is not much emphasis on the facilities they have to offer in terms of sporting. [...] and I **think (that del)** that would help greatly towards students' well-being as I am **assuming that** it is not always easy to **speak** to someone; sometimes, releasing stress through physical activity works better for some. They do not utilise this facility well which **seems** like a waste to me. Also, the fact that they have not **advertised that** there is a gym behind Aston that offers discounted rates for students attending Aston. Getting through some of its facilities on campus **can** be a little bit difficult to manoeuvre. For example, in the main building, certain lifts only go to certain areas of different floors which **can** be really confusing. It **was** confusing in first year and is still confusing now that I am in 2nd year, about to go in my 3rd. I really do **think (that del)** this can be improved for ease and better signage for us.

'Peripheral 2' Texts (> 0.9): Image Description

Image Descriptions ($M = 0.28$; $SD = 0.58$)

The same engagement with the task can be seen in the author's image description. The author offers an extensive commentary regarding every described object or person in the visual, which suggests that expressing personal stance and showing involvement is a common communicative strategy for this author when the situational characteristics of the texts allow this freedom of expression. Specifically, in the three texts discussed so far, this freedom seems allowed by the personal topic of the essay (Covid-19 experiences) and its goal to analyze the discourse surrounding this topic; in evaluations, the author appears to have seen a chance to voice their opinion and effect change in the university facilities, which directly concern the author's daily life; in image descriptions, the author does not feel restricted by the content of the visual, but

rather appears to see its fictional nature as a possibility for interpretation. Text Sample 130 shows how the reality of the visual is created through the author's interpretation.

Text Sample 130. Dimension 4 score: 1.61

So, in this image, we have two ladies who **appear to** be looking at something on a monitor that is very, very entertaining. We have a person of colour who is wearing a denim jacket, a blue denim jacket, who **seems to** have her makeup done very, very nicely. [...] Erm, and the lady in the denim jacket **seems to be enjoying** what is on the monitor more than the lady in the black and white striped shirt. They **seem to be** in a formal setting as **indicated** by the white background. So, I would probably **think that** they are in a university or something like that. And yeah. As for image two, we have two boys, two university looking boys, who **appear that** they are in a social setting due to the pillow in the background. It's **looking** very funky. And they **seem to** be engaging in some serious content as **indicated** by the one in a chequered shirt.

This tendency towards the peripheral status also persists in the author's email. However, the direction of deviation in this case is towards a lack of stance, indicating that although the author shows a cross-register pattern towards stance expression in many of their works, this pattern should not be generalized to their language use across contexts.

'Moderately Removed' Texts (0.3-0.6): Email

Emails ($M = 0.39$; $SD = 0.65$)

The author's email is found in the 'Moderately Removed' group, which shows that the deviation in the direction of a lack of stance is less prominent, and even this deviation still results in a considerable number of stance features in this email on a personal topic as the author offers guidance to a fellow student. Private verbs serve an important function in the text as the author shares experience and gives advice. Thus, it appears that despite this text representing a deviation in the opposite direction from the register's central trend, this deviation is not extreme, and the situational characteristics of this text (the familiar addressee of an equal status and a personal

topic) are conducive to stance (as the register of emails generally is), providing ground for the author's predilection for stance expression.

Text Sample 131. Dimension 4 score: -0.15

I'm good thank u. You'll soon get the hang of , the only place you **need to know** is the "my courses" tab and that's it rilly; maybe stuff to do with the library etc. I'm in 2nd year now but good luck with your exam, I'm sure you'll do good! In terms of the reading, the list isn't compulsory as you **said** but I strongly **recommend** reading the essential reading. Normally for each module, there's an "essential" and maybe a couple "further" reading books if you **want**, but definitely read the essential if you **want to** get an edge over the rest as most don't do it. And no, I don't **know** anyone selling their books unfortunately but I'll **ask** around. we all get by with the library as its really good in terms of finding sources and more time you'll find when writing essays that whatever you're looking for is on the internet also like copies of books; articles and stuff. **Hope (that del)** that helps
x

Tendency towards a Lack of Stance

In contrast, Author 59 shows a tendency towards a lack of stance features in four of their texts, two of which are moderately removed from the mean (0.3-0.6) and two of which are peripheral exemplars (>0.9). The analysis of these four texts revealed that the relative lack of stance in all of them appears to be the result of concreteness in the author's language use. An exception from this preference for a relative lack of stance observed across registers is the author's image description, which shows a moderate (0.3-0.6) deviation in the opposite direction. Nevertheless, lack of stance is more common across the author's works and may be indicative of a tendency pervasive across registers.

'Moderately Removed' Texts (0.3-0.6): Evaluation and Interview

Evaluations ($M = -0.28$; $SD = 0.37$)

The concreteness mentioned above is exemplified in the register of evaluations. It was stated earlier that stance features in texts scoring low on Dimension 4 are often limited to the private

verbs *like* and *dislike* which most likely mirror the wording of the task. This minimal expression of stance is observed in this author's text. The focus of the text is on such specific facts as the food options available in the neighborhood, the means of transport, other facilities, such as shopping and banking, and specific jobs students can hold on campus. It is these facts that comprise the evaluation rather than the author's views of them.

Text Sample 132. Dimension 4 score: -0.70

One of the biggest things I **like** about Aston Campus is its prime location. It is situated on the edge of central Birmingham, is neighbour to the express highway and is within a 10-minute walk to every central bus route and 3 different train stations. The campus's accessibility doesn't stop there, it is within walking distance to any cuisine you **could** possibly crave from a light chip shop to high end restaurants such as Pushkars on the high street. This extends to grocery shopping, banking and any retail shopping. Another aspect of the campus I **like** is its ability to give students a voice. Participation is encouraged at all levels of communication and interaction within campus. Being a student, you will **see** fellow students working as tour guides on open days, part of IT teams, mentoring groups and even student Hub offices.

Interviews ($M = -0.59$; $SD = 0.26$)

As discussed before, interviews resort to other means of stance expression and do not employ complex and elaborate clausal structures. The author's interview is an extreme example of that tendency. Again, this text is a highly concrete account of the dishes and the times of the meals. It is this account that is the essence of the text, and the author's thoughts about it are not necessary in response to the concrete questions of the interview.

Text Sample 133. Dimension 4 score: -0.90

So, currently I'll eat when it's in the, er, in the day. So, that'll be at like eight p.m. and I'll pray and then I'll eat, well, as much as I **can**. But usually, it's a lot of fast food. So, and I'll have like samosas or fries and then a lot of water cause we don't get to drink in the day either. And, and then I'll eat again in a c-, a couple of hours later at like one a.m. And that'll just basically be any leftovers from what I ate earlier. And then we have to stop eating before prayer time, which is at like four o'clock in the morning, it's like half four. So, before then, I'll eat again. I'll probably have like porridge or more leftovers and I'll try cook a bit of chicken.

‘Peripheral 1’ Texts (0.6-0.9): Business memo and Email

Business memos ($M = -0.85$; $SD = 0.45$)

Concreteness is again a feature of the author’s business memo, where the focus is exclusively on the suggested places and activities. The excerpt shows that value judgments occur in the form of phrases (e.g., expensive), but rarely as clausal complement structures. The single occurrence of a private verb with its complement is the only instance of relating the suggestions to the addressee’s preferences in a rather long text.

Text Sample 134. Dimension 4 score: -1.50

Souvenir shop inside too aswell as a bar -Near both hotels Restaurants: (Reserve online available, and dine in available) Restaurant Savoy, 4.6 stars, Finnish food, in between both hotels, expensive. Restaurant Finnjavel Salonki/ Sali, 4.5 stars, quite far out, west of hotel Kattanjoka, expensive Helsinki music centre- 2nd september, 7pm-9pm. La passion de simone- classical piece 19-28 euros, book in advance. near Kattanjoka Bear park near to the cafes and many restaurants in eye distance from first restaurant incase ceo **wants to** eat something other than Finnish (many asian cuisines)

Emails ($M = 0.39$; $SD = 0.65$)

Finally, the concrete report in the excerpt from the author’s email contributes to the same trend and, considering the overall reliance of emails on stance (contrary to the other registers discussed here), the deviation in the author’s text is even more striking. That is, the general trend in business memos, interviews, and evaluations is towards a lack of stance, and the deviation in the author’s texts is therefore in the same direction. In contrast, emails rely heavily on personal stance. A substantial deviation in the opposite direction therefore appears to contribute even more meaningfully to the cross-register trend of a lack of stance revealed by the author.

Text Sample 135. Dimension 4 score: -0.46

Hi, Unfortunately i will not be able to attend as i am working on placement and **cant** take time off to attend rep meetings. I am a rep for placement year . Feedback i have received from students:

- Students have been struggling with the communication with placement team in emails, students have to chase up a lot and the team does not **report** back any changes in any processing.
- There is also a confusion regarding the required 150 days of work, does it account for annual leave given by employers and how is it tracked, will it be signed off on by employers?
- Some placement tutors have been hard to get in touch with and sort out a date with for the visit required. Others such as my own (Dr) have been really good and visits have been very well communicated.

‘Moderately Removed’ Texts (0.3-0.6): Image Description

Image Descriptions ($M = 0.28$; $SD = 0.58$)

However, as noted before, no matter how persistent a pattern of individual preference is, exceptions from it seem to suggest that no pattern gives credence to generalizations about individual use. Text Sample 136 is such an exception and shows quite extensive use of stance as the author accompanies the description with commentary. It has been shown that descriptions, despite being a situationally well-defined narrow register, allow authors a variety of approaches to the task, and some authors opt for a highly specific focus on the content and a lack of personal stance. Author 59, therefore, could have chosen that approach, which would have been more in line with the trend they reveal in the other registers. The author, however, makes a choice that runs counter to their individual preference even when a choice in line with that trend is possible. This again poses questions about the nature of individual expression and the motivation for individual linguistic choices. While this motivation may lie in part in the situational restrictions of the register and in part in authors’ unique individual preferences, the example of Author 59 seems to suggest that there are cases when no restrictions are imposed by the situation, but individual preferences are also not observed. As discussed earlier, it may be that even in the absence of situational restrictions, some authors may have preconceived ideas regarding the communicative strategies that best accomplish the tasks of certain registers. As a result, Author 59’s description reflects and even enhances the register trend rather than contributes to their own

preference. These varying beliefs of what is appropriate in a register seem to be distinct from the author's cross-register tendency or what is required by the situation and may therefore constitute an additional consideration in investigations of individual language use.

Text Sample 136. Dimension 4 score: 0.87

But the, they both **seem to** be having a good time. Er, yeah, they're also dressed very casual, erm, way too laid back for some formal, er, workwear. Maybe in college, even, they're quite young, probably like early twenties. Er, then image three, I **think, I think (that del)** they're also in university, although the background's a bit weird, It-it's a bit, it doesn't look like a university like area but it's, I don't **know**, well, we do have two g-, two people that are working together. The, er, there's a, there's a whitish Latino guy and there's a Black girl and they're, they're w-, the guy is holding a book so I'm **assuming (that del)** it's in university. Er, they're sitting down near a glass table. . They're also wearing like kind of casual clothes. Well, it is kind of cas-, really casual. But they don't **seem to** be close at all. I'd probably **think (that del)** they're strangers, er, like it's j-, it's just the way that they are together, like they're not overl-, they're sitting next to each other but they're not overly close or they're not looking at each other, they're looking away.

7.2.3 Case Study 3: Authors of Different Distance Groups and Directionality across

Registers

This section presents an author highly diverse in their approach to register. Author 26 is represented by texts in each distance groups, with deviation mostly towards increased stance use. That is, the author does not seem to have a clear preference with regard to stance features that cuts across situational contexts. Rather, depending on the register, the author makes choices to conform to its situational norms more or less closely.

'Central' Texts (< 0.1): Essay

Compliance with the register general tendency is the most obvious in essays. The text below shows attribution of stance to experts, whose views are used by the author to construct the argument of the essay. This reliance on expert opinions, conveyed through private verbs and

their complements, is directly associated with the goal of overviewing the existing body of work on a topic, common in the texts of the register.

Essays ($M = -0.02$; $SD = 0.43$)

Text Sample 137. Dimension 4 score: -0.001

Gender and sexual power are topics that are arguably very subjective, definitions of both often vary depending on social groups or individuals. For example, for some, sexual power may refer to conventional and/or conservative notions that derive from religion or cultural belief systems, whereas others may **attribute** sexual power to individual preferences and beliefs. According to Tredway (2014) gender shares the same variability according to different cultures. Gender in 2020 **seems to** be a term that is increasingly tip toed around due to its now extensive characterisations. Judith Butler, perhaps one of the most relevant theorists in gender exploration, **suggests that** gender is something that is not a choice and that is socially constructed and placed upon an individual at birth, who then practises and “performatively produces” it, Butler (1990). She **explains that** through the repetition of particular acts such as dressing, acting and living a certain way is what brings ‘gender’ into physicality. This is what she **calls** ‘iterability’, as she **describes** gender as something that is fictitious and imposed onto people.

‘Minimally Removed’ Texts (0.1-0.3): Business Memo

Business Memos ($M = -0.85$; $SD = 0.45$)

Likewise, the author’s business memo is generally representative of the register trend. Stance is minimal in the register overall and in the text below. The author’s text prioritizes facts over opinions, and while modal verbs of possibility associated with the dimension do occur in the text to enhance the suggestions, overall features of author attributed stance do not contribute to this informational focus.

Text Sample 138. Dimension 4 score: -0.65

Radisson Blu Hotel, this hotel has been **rated** on trip advisor an 8.3 for work trips. it is close to the water which and 0.6 miles way from shopping facilities, so the manager **can** purchase gifts, the hotel includes a fitness centre, so she **can** stay active whilst she’s away and a well rated hotel bar and restaurant, that provides a good quality dinner and breakfast. The Radisson is not far for public transport which makes airport access easy and simple. In total this hotel for the 2-4th of September 2021 is within budget at a price of £180 for a two night stay.

‘Moderately Removed’ Texts (0.3-0.6): Evaluation and Interview

Evaluations ($M = -0.28$; $SD = 0.37$)

The author shows a noticeable deviation from the general trend of relative lack of stance in evaluations. Their text shows a clear emphasis on proposed solutions to the outstanding problems rather than simply presents a list of the negative features of the university facilities. Features of author-attributed stance play a major role in these proposals and are, therefore, quite prominent in the author's text. This integration of solutions and future improvements is not required by the task and thus may represent the author's communicative need or ideas of strategies that contribute to the efficiency of the text. These ideas, however, are contrary to the overall trend of the register, namely that of a relative lack of stance, and the author deviates from the register trend in a direction opposite to the overall trend.

Text Sample 139. Dimension 4 score: 0.26

The first and perhaps the most inconvenient to me throughout the years, is the lack of parking facilities on campus. Whilst yes, it **can** be tricky to find parking in a lot of places in Birmingham, I **think (that del)** the least that a university **can** do for its paying students is provide good, secure and cheap parking facilities. Due to the small number of spaces available there have been instances where I have missed or been late for lectures. I **know (that del)** many others who have experienced the same thing. The next facility that I **dislike** at Aston University is the library. I **think that** it is outdated in terms of technology and decor. I **think (that del)** [illegible] should be [illegible] put in place so that students have a more modern place to study in. I **think that** systems should be introduced so that students **can see** whether study spaces are free. I **think that** in study areas students should be able to control the temperature of the room so that they are comfortable.

Interviews ($M = -0.59$; $SD = 0.26$)

A deviation of the same extent but in the opposite direction is found in the author's interview.

Unlike evaluations, where the overall register trend is the lack of stance, in interviews a lack of stance is the norm, and a deviation in the same direction is less surprising. Again, the emphasis on concrete detail rather than the author's thought process does not call for features of personal stance.

Text Sample 140. Dimension 4 score: -0.99

I recently have been on like a real health tip anyway so I've been trying to like incorporate a lot of, erm, just good stuff into my diet. So, I start off with, erm, a fresh juice, usually, celery juice. Erm, and then, like apple cider vinegar mixed with water. And then, erm, porridge with like blueberries, flax seeds, nut seeds, erm, you **know** good stuff. I'm trying to stay away from sugar so I don't put any sugar in it, cinnamon, stuff like that. Erm, so yeah, that's like the first half of the day. And then, if I **need** a snack in between, like lunchtime it's usually, probably like a piece of fruit.

'Peripheral 1' Texts (0.6-0.9): Email

Emails ($M = 0.39$; $SD = 0.65$)

The author's email shows an extreme deviation towards increased use of stance features. The highly personal topic is conducive to expression of personal stance, and the focus on communication with third-parties as well as the author's thoughts and inquiries makes Dimension 4 features central to the text.

Text Sample 141. Dimension 4 score: 1.15

hello, I **hope (that del)** this email finds you well. Im emailing regarding the conditions of my placement assignments that are due in the later year. Since arriving back from in the nick of time some two weeks ago, I have received an email by my placement organisation **stating that** they have terminated the contract for my placement. is worsening at a rapid rate and having **spoken** to some of my colleagues it **seemed** like the best option to stay in England regardless. Having **said** all of this, I was **wondering** if there could be some conditions changed regarding the terms of my assignments. My placement has been shortened by 4 months and I have been laid off, giving me no further contact with the school until I visit to **say** goodbye when is once again safe. So in this regard, I **feel that** this puts me at a slight disadvantage compared to those who may still be working at home. Is there a way that the assignments I have due **can** be adjusted to my experience? without jeopardising my grades. Thank you I **hope (that del)** you **can understand** my position and **can** come to a middle ground. Kind regards,

'Peripheral 2' Texts (> 0.9): Image Description

Finally, an even more extreme deviation is found in the author's image description, where the author's stance is again prominent in the inferences they make about the visual. Frequent repetitions of stance verbs, such as 'seem' and 'appear' result in an especially high text score on the dimension.

Image description ($M = 0.28$; $SD = 0.58$)

Text Sample 142. Dimension 4 score: 1.46

Okay so in the first picture I **can see** two girls that **appear to** be friends. Erm, they **seem to** be quite relaxed in front of each other. And the one that is standing up **seems to** be showing the one that is sitting down, the lady that's sitting down, something funny, or she, the lady that's sitting down, **appears to** be laughing at something. Erm, but yeah all-overall they **seem to** be quite relaxed. And the background doesn't really **indicate** where they are. Erm, but regardless they **seem** pretty comfortable with each other. Erm, in the second picture, so similar situation, where I **think (that del)** they would be friends. And it **looks** like they're in like, in maybe a library or a café.

7.3 CONCLUSION

This chapter has examined individual differences in the interpretation of register space – namely, authors' perceptions of the degree of linguistic freedom allowed by registers situationally defined to various extents. It showed that authors systematically distinguish between registers, but also reveal different trends with regard to how they position themselves in these registers. One persistent pattern emerging from this analysis points to the fact that authors are guided by a variety of considerations as they navigate registers – some of these considerations may be consistent individual preferences of a particular expression; others may involve register constraints, but still others appear to be a distinct group of considerations that may reflect people's ideas of appropriacy or efficiency of certain communicative strategies in certain registers that lead them to make choices at odds with their own prominent cross-register trends and not required by situational restrictions. A general outcome of this analysis is the fact that a cross-register perspective can offer a layer of new, crucially important nuance for a study of individual language use.

CHAPTER 8. DISCUSSION AND CONCLUSIONS

The study identifies patterns of functional linguistic variation in the corpus – linguistic dimensions of variation (Chapter 4) and examines the role of several predictors of linguistic variation across texts, including characteristics of users, such as age, gender, and the authors' individual expression, and characteristics of texts – the register category the text belongs to. The study observes a major effect of register, explaining 76-79% of variation on three dimensions and 38% on one dimension. The other predictors, age, gender, and author identity, do not contribute to the overall variance on the dimensions, explaining 1% of variance at most. The size of this effect of register may be attributed to the extreme variation in the registers in the corpus, which in turn may be a by-product of the goal set in the process of corpus compilation. That is, the goal in collecting the Corpus of 100 Idiolects was to obtain one text per author from a range of registers. While this goal resulted in a rich resource representing each author through a variety of contexts, a principled selection of these contexts was not the goal of this process, as any context would have contributed to this variety. As a result, the first major difference between the registers in the corpus is the distinction between naturally-occurring and elicited registers. While the first group – essays, emails, and text messages – represents registers with a variety of real-world topics and purposes, the second group – business memos, evaluations, interviews, and image descriptions – are registers produced in response to a specific task in a research setting. A further analysis of the specific goals and other situational characteristics of the registers in both groups shows that the registers within the naturally occurring and elicited groups also display major differences with respect to these considerations. In view of such strong situational distinctions, the paramount role of register in linguistic variation is not surprising. The study carries out a register analysis (Chapter 4) on each dimension and analyzes the identified

functional patterns in texts of each register. When each age and gender group is investigated across registers, the same patterns of register variation are observed within each social group (Chapter 5).

The study then turns its attention to intra-register variation observed on the dimensions and examines various sources of this variation. First, the study tests the possibility that this internal variation is accounted for by the social groups of the study – age and gender (Chapter 5). Second, communicative differences among texts are examined as the source of internal variation within registers (Chapter 6). Third, the study then presents patterns of author differences across registers with regard to the positions of their texts in a range of registers (Chapter 7). The following subsections discuss the major outcomes of these three steps.

8.1 ROLE OF SOCIAL GROUP IN LINGUISTIC VARIATION ACROSS TEXTS

The present study examined the effect of social group, age and gender, on functional linguistic variation. The overall result of this investigation points to an absence of such effects of either age or gender, with groups not showing significant differences on any of the four dimensions. The study also shows that the observed register differences are equally prominent in each of the groups, and generally groups follow the same patterns of register variation, finding the same registers distinct to a significant or nonsignificant degree. As the study examines register-internal variation observed on all dimensions, it reveals that social group is not the reason for this extensive variation. Instead, it becomes apparent that each of the age and gender groups vary extensively within registers. This pattern, repeatedly observed in the case of each social group pointed to a different basis for this variation – a possibility explored by the study further through analysis of texts within registers.

The lack of differences among social groups on the dimensions generally and within registers is worth discussing in order to reconcile it with previous research. Age is represented by five groups of 19-23-year-old students. These groups were viewed as years of undergraduate study, which could show if the length of immersion in the academic context resulted in differences in functional linguistic patterns in the academic registers of the corpus. It should be noted again, however, that the corpus metadata did not contain evidence that the age reported by the participants corresponded to years of study. Thus, while this correspondence may exist, the results are interpreted with caution as levels of study may be represented by a range of ages and the same age group may include different levels of study. To an extent, this possibility may explain the lack of effect of age this study observes in the register of essays. This result is particularly surprising on Dimension 1 – the contrast between clausal and phrasal linguistic features – in view of the findings of earlier studies of complexity development across university levels (Staples et al., 2016). Staples et al. see a clear progression from clausal to phrasal complexity from lower to higher levels of university study. One especially notable pattern that emerges from their study is the development of pre-modifying nouns over the course of the students' university careers – a feature more strongly associated with academic prose than any other kind of informational writing (e.g., Biber & Gray, 2016, cited in Staples et al., 2016, p. 163). Pre-modifying nouns are also the feature most strongly associated with Dimension 1 of this study, i.e., its informational pole. Yet, the older groups of this study do not show a trend towards the informational end of the dimension in essays (which could be seen as an increased use of that feature).

While the lack of correspondence between age and year in the program may certainly be the reason for this outcome, perhaps another reason lies in the nature of the texts in the two

corpora under analysis – specifically, the kind of internal communicative variation that exists in these texts. The trend observed by this study in essays consists in the facts that i. age groups do not differ from each other; ii. all age groups show considerable variation within themselves; iii. as communicative variation is examined later in the study (Chapter 6), it is attributed to differences in topic. Specifically, these topical differences are presented as the contrast between personal and technical or scientific (impersonal) topics. It is then shown that while disciplinary distinctions are highly informative (sociology vs. psychology essays), what really explains the polar opposite positions of the disciplines in the register is this topical contrast. Essays in sociology, business and marketing, and law all allow the author’s subjective take on a topic and a discussion of a subject matter on which they have some general knowledge. These essays were in stark contrast with psychology essays, which reported on empirical procedures and involved scientific content. Thus, while phrasal complexity is essential to technical topics, the non-technical personal topics do not require the same informational focus (the higher emphasis on phrasal features in science writing compared to humanities is confirmed by Staples et al., 2016 and Biber & Gray, 2016). Therefore, it appears that this contrast among the texts of the register is so stark that it has to be consistently observed by and is equally apparent to students of all levels, and this inherent variability in essays results in the lack of mean differences among the age groups. On the other hand, Staples et al.’s study is based on the four disciplines of the BAWE corpus – Arts and Humanities, Social Sciences, Life Sciences, and Physical Sciences, and texts from Life and Physical Sciences were grouped to represent science writing (Staples et al., 2016, p. 155). The study reports an increased use of phrasal features in all disciplines. However, it is not clear whether the included disciplines revealed the same topical contrast. That is, while the same contrast between science and humanities writing may be represented in Staples et al.’s

study, the topics within Humanities (such as topics in history exemplified by the study with specialized content and a lack of the author's subjective take) may not necessarily lend themselves to the same oral and elaborated (clausal) expression that the personal topics in this study did (cf. experiences during Covid-19 or leadership qualities vs. Postmodernist theory). It may be hypothesized then that while in the case of topics such as the topics in history a trajectory towards phrasal complexity is in line with the disciplinary norms, in the case of personal topics such as the ones seen in this study the disciplinary norms may in fact not require the development of phrasal complexity in the first place (or at least not to the same extent). If that is the case, while in Staples et al.'s study this development was uniform across all the examined disciplines, in the case of this study there may be a certain tension between some essays requiring increased phrasal complexity over time (psychology reports) and others not calling for phrasal expression at all at any stage (personal topics in sociology). As noted earlier, in this scenario, this extreme variation among texts would result in a lack of age group differences, as all groups would have to recognize that contrast. While these explanations are only potential reasons for the observed trends, a future empirical comparison of developmental patterns in registers of varying degrees of internal variability may shed light on the seemingly contradictory findings of these studies.

The approach the present study takes to investigating gender differs from those adopted in previous studies in several ways. First, the study does not focus on individual linguistic features previously found to reflect gender differences. One criticism of this focus on linguistic features, identified as characteristic for men's or women's discourse, is that the functions of those features may have been different in the register in which they differentiated between men and women than they are in the registers of all future studies based on them. It is therefore

problematic to base assumptions about gender on the same features, setting out to find differences (Aries, 1996). Second, while other studies of gender have relied on combinations of features and holistic measures like indices of “women’s language” based on these combinations, it has been pointed out that such measures combine functionally different features not related to each other in discourse, which makes the meaning of such indices questionable (Aries, 1996). The goal behind identifying such measures seems clear: similar to the individual-feature approach, these studies rely on features that should reveal gender differences. Aries (1996) notes, however, that combining unrelated features because they are expected to represent women’s or men’s language has been motivated by the wrong assumption that these features share a function and proposes instead investigations of communicative strategies in discourse, which would focus on the features that contribute to a given strategy. The approach taken by this study seems in line with this recommendation. This study identifies patterns of functional linguistic variation that exist in the domain and *then* tests whether genders differ in ways of navigating these functional patterns. The present study does include features previously found indicative of gender differences, but in contrast to previous studies focusing on gender, it does not limit the set of included features to the ones that are expected to reflect gender differences. The identified patterns, based on *functional feature co-occurrence*, are thus in stark contrast to any of the previous feature combinations used in studies of gender.

The observed result suggests that functional linguistic patterns captured by the four dimensions of this study, used to an equal extent and in the same way by both genders, are fundamental to communication rather than reflective of different socialization practices (as linguistic variation based on gender is viewed by sociolinguists, e.g., Wardhaugh, 1993). That is, while specific linguistic features, many of which occur on the dimensions, have been previously

shown to be used to different extents by men and women, genders do not differ with respect to general functional patterns (i.e., dimensions of functional linguistic variation). The amount of talk, hedges and emphatics, coordination and subordination, discourse markers, and personal pronouns are some of the features previously found to reflect gender differences. These features, while fundamental to the dimensions, comprise their functional patterns only in part. Their individual functions are not equivalent to the functional pattern at the core of each dimension, as it is patterns of co-occurrence of several features co-constructing discourse that comprise the complex functional pattern of the dimension. The key difference between the results of this study and previous findings of investigations of gender thus lies in the fact that while individual features may have been used differently by men and women, showing a difference in their individual functions (e.g., hedging), a complex communicative pattern, such as the extent of oral elaboration or information density, is the kind of linguistic resource employed equally by both genders, and variation along this cline is equally characteristic for both. Similarly, the clines from abstract to concrete discourse, from others-oriented descriptive to self-oriented or interactive discourse, and from more to less prominent expression of stance, all comprised of several interrelated linguistic features, are used by both genders to an equal extent.

It is important to note that this outcome does not denounce any of the previous findings of gender differences in individual features in some contexts. Rather, these patterns point to the fact that linguistic differences in the speech of men and women are quite specific and concern a particular aspect of communication in particular registers, such as an expression of certainty or directness (e.g., hedging or lack thereof), expression of emotion or involvement (e.g., boosters), and conversation strategies (minimal responses, tag questions). Major communicative patterns,

such as identified by the dimensions of this study, however, were not found to reflect gender differences.

Importantly, most of the previously identified differences were observed in conversation (or more generally, oral interactive registers) as can be seen from these typically oral features and as pointed out in the literature (Friginal & Hardy, 2014). Most of the registers of this study, in contrast, represent very different domains. Oral interviews are the only interactive register in the corpus, but even in the case of interviews communication is quite restricted by topic and the research setting in which it was conducted. The other spoken register in the corpus, image descriptions, is monologic, and this lack of a real interlocutor makes the register non-interactive and thus less likely to reflect gender differences. More informational contexts, however, containing objective statements about self and the world, were shown to not reveal gender differences (e.g., Soskin & John, 1963, cited in Aries, 1996). Thus, it may also be the nature of the registers of the corpus of this study that results in the non-significance of gender. Alongside several previous studies investigating combinations of features in different contexts (e.g., Mullaac et al., 1985; Mullaac & Lundell, 1986; Soskin & John, 1963), which reported an overall minor effect of gender or gender differences dependent on the situation of use, the present study observes a major effect of register in functional variation and consistent adaptations to the demands of the situation shown by both men and women. Further, within a single register, both men and women show communicative differences across their texts. These findings therefore add to the body of research that suggests that i. it is informative to investigate gender with relation to the situation of use; ii. results may differ depending on the linguistic unit of analysis. While differences may lie in highly specific linguistic features in some contexts, they will not exist in general functional patterns characterizing communication at large.

8.2 COMMUNICATIVE VARIATION ACROSS TEXTS WITHIN REGISTERS:

FUNCTIONAL CORRESPONDENCE BETWEEN SITUATION AND LANGUAGE

As linguistic variation within registers is examined, the study finds that this linguistic variation is accounted for by specific situational distinctions among texts. These results reflect the theoretical grounding of register research, namely the systematic relationship or the functional correspondence between situation and language, specifically its recent reconceptualization (Biber & Egbert, 2023), which aims to account for the possibility of situational and linguistic variation at all levels of analysis. Biber and Egbert write that while the relationship between situation and language is well established between registers, register research to date has hinged on the assumption that registers can be defined in terms of their situational characteristics. The values of these situational characteristics are therefore assumed to hold true for all texts of a register. This assumption would not allow the possibility of register-internal variation, as registers in this approach are accepted to be situationally uniform. This assumption, however, has been contradicted by empirical findings pointing to the fact that registers are not linguistically or situationally homogenous (e.g., Gray, 2015; Biber et al., 2020; Biber et al., 2021; Goulart et al., 2022; Egbert & Gracheva, 2023). On the basis of these persistent trends, Biber and Egbert (2023) propose that it is the same systematic functional relationship between situation and language that should account for variation within registers and that it is possible to measure both linguistic and situational factors within a register on a continuous scale and compare texts in terms of these measures.

In part, the present study implements this approach by developing a scale from texts best representing a register to texts that are far on its periphery and comparing these texts, grouped on the basis of their linguistic differences, in terms of the situational basis corresponding to these

differences. Importantly, this study moves away from situationally defining each register as a whole, and as the framework for situational register analysis is applied, it is illustrated that at least several situational parameters are variable across texts. It is these variable situational parameters – purpose and topic – that were shown to correspond to the linguistic differences in the registers where these parameters were undefined. Alternatively, additional highly granular situational factors were shown to account for linguistic variation among texts of quite narrow registers, defined through all major situational parameters of the applied framework. Thus, while the study relies on the widely accepted traditional definition of registers as named, culturally recognized categories of texts (Biber et al., 2020), these results confirm the fact that such categories are not uniform and members of a culture appear to have varying ideas of what constitutes a named category of texts, which can as a result gain quite different linguistic representation.

8.2.1 The Role of Topic and Purpose in Variation across Texts

While communicative purpose has received considerably more attention as a predictor of linguistic variation in register research (Biber et al. (2021) in conversation; Goulart et al. (2022) in student academic writing; Wood (forthcoming) in statutory law; Egbert & Gracheva (2023) in introductory university textbooks) and topic has remained on the sidelines of such investigations, the present study highlights the paramount role of topic in explaining linguistic variation within registers. Topical differences, however, are not independent from purpose, and an important pattern observed and discussed by this study is the interdependence of the two. This close association between these two textual situational factors is, in fact, apparent if we examine the definitions of the two constructs and some consistent patterns emerging from previous research on granular intra-register situational distinctions. Chandler and Munday (2016, cited in Goulart

et al., 2022) define communicative purpose as “the primary goal or intention of anyone involved in an act of communication on a given occasion.” Topic, on the other hand, defined in very similar ways by a number of dictionaries, is “a matter dealt with in a text, discourse, or conversation”, from Greek *ta topika* (‘matters concerning commonplaces’ by Aristotle) from *topos* ‘a place, a commonplace’ (Thompson, 1995, p.1470).

Taken together, these definitions seem to suggest that topic may be seen as the “given occasion” in the definition of purpose put forth by Chandler and Munday, and while purpose is the general “goal or intention”, topic is the specific matter that goal is directed towards. Indeed, although topic has been less targeted as the subject of research, the coding schemes applied by studies focusing on purpose reveal that it is quite challenging to isolate topic from the highly specific purpose labels they propose. That is, often the communicative purposes used in such coding schemes also contain information about the topic of the texts where the purpose is commonly found, and the more specific the purpose label, the higher the likelihood of it containing topic information. For example, in Goulart et al.’s (2022) study, examples include ‘to describe a tangible object’, ‘to narrate a personal event’, ‘to narrate a past event’, with the additional third narrative purpose ‘to give a procedural account’, which limits narration to “methods for an experiment or study.” Social purposes of genre families identified by Nesi and Gardner (2012) involve several examples, such as ‘to develop awareness of motives and/or behavior in individual (including self) or organization’ or ‘demonstrate the ability to design a product or procedure that could be manufactured or implemented.’ This convergence of topic and purpose seems to be also observed in the domain of legal language as Wood (forthcoming) develops a detailed framework for analysis of purpose in US laws. Many of the identified purposes, again due to their specificity, either contain the scenario in which they would be

applied, the scope of application, or the specific element of the description: e.g., ‘prescribes a criminal offense for an action’, ‘prescribes a mandatory duty or responsibility to a human agent or government entity’, ‘provides general directives that do not regulate human action’, ‘describes a process in a specific situation/context, including description or the role(s) of one or more humans/entities’. The underlined portions of these purpose labels seem to contain concrete references to what could be construed as the topical domain of the law (e.g., criminal offences, duties and responsibilities). In the framework for situational analysis (Table 2.1), Biber and Conrad (2019) list general purposes (narrate / report, describe, inform / explain / interpret, persuade, how-to / procedural, entertain, edify, reveal self) and specific purposes (summarize information from numerous sources, describe methods, present new research findings, teach moral through personal story), and it is immediately apparent that the specificity in purpose leads to the inclusion of topic. The approach to purpose in the present study (essays on Dimension 4) is also not an exception: the purposes ‘to overview state-of-the-art/ expert opinions’, ‘to present qualitative analysis of interview data’, ‘to evaluate or analyze (typically company activity or market segment)’, and ‘to report empirical (quantitative) findings’ are clearly not independent of topic. These examples illustrate that while the purposes are meant to denote “goals and intentions”, these highly specific labels appear to also contain the subject matter these goals are applied to.

Further examples come from Nesi and Gardner’s BAWE corpus, where such distinctions in topic within the purpose labels seem to serve as the linchpin in distinguishing between quite specific varieties of academic writing: for example, Explanations and Critiques have the purpose ‘to demonstrate/ develop an understanding of the object of study’, while Case Studies set the

purpose ‘to demonstrate/ develop an understanding of professional practice through the analysis of a single exemplar’ (cited in Staples et al., 2016, p. 157).

Interestingly, it is attempts to develop frameworks with highly specific communicative purposes, typically to account for variation among texts of a single register, that seem to inevitably integrate topical information into the purpose labels (where topics appear to be the concrete “occasions” or the concrete “subject matter” of the general communicative purposes ‘to demonstrate an understanding of’, ‘to describe’, ‘to prescribe’, etc.). It appears then that specificity in purpose (by default a general intention or goal) entails information about topic (i.e., the concrete subject matter the purpose applies to).

Prominence of topic in explaining intra-register linguistic differences, especially with regard to Dimension 1 – ‘Oral elaboration vs. Information density’, is among the main outcomes of this study. Topical distinctions yield clear patterns of difference with regard to oral and informational features both in essays and in emails, and this pattern is worth examining further and more closely in other registers. It seems especially useful to test the effect of topic while controlling for the possible related variables, such as discipline. While the effect of discipline is acknowledged by this study, it is demonstrated to the extent possible with the data available that disciplinary distinctions do not fully account for the observed pattern. While psychology and sociology are Life and Social sciences and both investigate human subjects, it was the specific topics that shed light on the linguistic differences between them in essays. The study additionally suggests that the disciplinary distinctions and the varying degrees of focus on personal vs. technical, scientific matters in them may be associated with the varying possibilities for empirical research in the disciplines, with some disciplines lending themselves to qualitative analysis, evaluation, and discussion (as in the case of more personal topics within sociology in

this study) and others more reliant on concrete scientific evidence (scientific investigations in cognitive psychology essays). It was further shown that this pattern of linguistic and topical differences is upheld by essays in Marketing and Business and Law, the specific topics of which were personal enough to lend themselves to a subjective speculation and discussion and thus required oral elaboration. Thus, it is the distinction between more personal as opposed to more technical subject matter that appears to account for the linguistic differences rather than simply discipline labels. The results observed in emails, i.e., the effect of topic in the absence of disciplinary distinctions, give credence to this interpretation.

It is worth mentioning, however, that these clearly interpretable patterns emerged from ‘topic types’, namely coherent groups of topics, rather than the specific topics of the texts. That is, topics were grouped into personal and technical (or impersonal) in essays, and the same broad categories were identified in emails, official matters and class content-related topics representing the impersonal category. Topic types thus represent a grouping of texts, whose specific topics can be unified under a more general topical domain. In the case of this study the central distinction between the identified topic types that proved informative was the degree to which the topics were personal or technical. In texts on topics labeled ‘personal’, the author could typically rely on and integrate their own judgments on issues that constitute their daily experiences even if these experiences are not first-hand (gender dynamics, Covid-19, leadership qualities, conflict in the workplace, etc.). Conversely, the impersonal, technical topic type exhibits a heavy informational focus: in essays, this informational focus is found in texts that present empirical results; in emails, the information presented pertains to official or class-content related matters. While in this study these general topical domains received ‘personal’ and ‘technical/ impersonal/ informational’ labels, it cannot yet be argued that these same topical

distinctions will emerge from other corpora. What is shown, however, is that such topical differences can be major in accounting for linguistic variation.

Having identified general ‘topic types’, the study uncovers another interesting pattern. It becomes apparent that certain purposes are systematically associated with particular topic types. Specifically, this topic type-purpose cooccurrence was discussed in essays as personal topics, more closely connected to daily human experience, are found to be associated with the goals of analysis, discussion, and evaluation. In contrast, essays on technical, scientific topics consistently set the goals of reporting research procedures and summarizing findings, but technical topic types do not lend themselves to a discussion or speculation (e.g., on the alpha-inhibition hypothesis). It is useful at this point to return to the discussion of the relationship between topic and purpose at the beginning of this section, where it was suggested that topics are the concrete subject matter that communicative purpose (more general in nature) is applied to in specific instances of communication. If purpose is a general intention or goal that gains a concrete representation through a concrete topic, the correspondence observed by this study between this general intention and a general topic-type seems to align with that hypothesis.

Examining this connection more closely may confirm that certain topics lend themselves to certain goals, and certain communicative goals are typically set or typically possible with regard to certain topics. However, this pattern may look different in different registers and with regard to different linguistic variables. In this study, for example, this topic-purpose correspondence was observed in essays, but the same pattern did not seem to emerge from emails, where topic stood out as the main predictor.

A difference with regard to the linguistic variables was also already observed in this study. While the role of topic as the most informative situational factor explaining variation in

features of oral elaboration and information density was quite clear, the results observed on Dimension 4 made the picture more complex. Part of the motivation in selecting Dimension 4 to replicate the analyses conducted on Dimension 1 consisted in testing the importance of topic on this dimension, comprised mostly of grammatical features of stance. The only lexical component present on the dimension is the lexical classes of verbs, such as private and public. It may therefore seem unlikely that topic will play a part in predicting predominantly grammatical features. In contrast, Dimension 1 contained nominal sequences, which may be anticipated to reflect topic. However, the same grouping of email topics, namely personal topics, official matters, and class content-related topics, helped explain the linguistic differences across texts on Dimension 4, yielding an equally clear pattern. This finding is interesting for several reasons. First, it conclusively shows that topic plays a paramount role in explaining several different patterns of linguistic variation at least in the register of emails –comprised of features of oral elaboration and information density and author-attributed stance. Second, this result shows that distinctions in topic are, somewhat surprisingly, clearly reflected in grammatical features (Biber and Conrad (2019, p. 48), for example, write that “topical differences are less influential for determining grammatical differences”). While distinctions in purpose may be more readily associated with grammar, as the role of purpose has now been consistently highlighted (by studies cited above), the fact that different grammatical structures contribute to more or less personal topics has not been emphasized.

This result also shows that although the two dimensions (Dimension 1 and Dimension 4) are comprised of different features and capture unique functional patterns, these functional patterns are related and, specifically, variation on both appears to be predicted by the degree of personal involvement. Dimension 1 is a cline from oral elaboration to information density, and

the dimension structure is not limited to features of involvement, such as emphatics, also containing features associated with distinctions in mode (e.g., contractions), amount of shared knowledge between the participants (e.g., demonstrative pronouns), or the immediate and situated nature of the discourse (present tense, adverbs). Dimension 1 also contains clausal coordination and causative adverbial clauses, the features associated with elaboration, which played a major part in the dimension interpretation. Most of these features as well as the analysis of texts from different registers and their place on the dimension (i.e., between-register analysis in Chapter 4) did not suggest that involvement was the central component of the functional pattern, although involvement has been commonly associated with more oral discourse in other studies (e.g., Biber, 1988). Its role became obvious, however, through the situational parameter of topic when texts within registers were assigned topic labels and grouped into personal and technical 'topic types'. Dimension 4, on the other hand, is a continuum from more to less author-attributed stance and in that is immediately associated with involvement. Since the major distinction between the identified topics both on Dimension 1 and Dimension 4 and both in emails and essays is between more or less personal topics, the degree of personal involvement has thus emerged as the common denominator determining the place of the texts of these registers on both dimensions.

An exception from this pattern, however, are essays on Dimension 4. In the case of these texts, the topic or discipline labels did not prove informative, and it was purpose that explained the differences in the use of stance features across texts. While this result appears puzzling at first, the fact that in all other cases it seems to be the possibility of personal involvement that has formed the basis for topic type differences offers some insight. The stance features comprising Dimension 4 have been referred to as 'author-attributed' stance due to their grammatical

characteristics: that is, the verbs and their complements occur with a subject, which usually denotes the holder of the stance. The holder of the stance in most texts in the corpus is the author of the texts, and indeed in most texts the author's prominent expression of stance through Dimension 4 features corresponds to author involvement. This is true for emails, where the degree of involvement was reflected in topic differences (personal vs. official matters and class content); this is also true in the registers of evaluation of Aston facilities, image descriptions, interviews, and business memos, as in all these cases the author of the text and the holder of stance are the same. This is not the case in essays, however. It was emphasized before that it is in fact rare for essay authors to express their own stance. Instead, stance features in this register express the stance of experts (prominent thinkers, scientists, authors of cited research) or interview subjects whose perspective is presented and analyzed. Thus, in these cases, these grammatical structures still express author-attributed stance, but since the stance is in most cases not the author's, it is not associated with involvement. Therefore, there is no reason to expect in the case of this dimension in this register that a difference in more or less personal topics will explain linguistic variation. This linguistic variation simply does not reflect the personal or impersonal nature of the discourse in these essays. Rather, it is the communicative purpose or, perhaps, the focus on opinions vs. facts that requires higher or lower rates of occurrence of expert stance: specifically, the essays that present results of qualitative analyses of interview data and overview and discuss theories or the state of the art in the field (focus on opinions) are likely to rely on expert stance more, while essays that report either empirical results or offer an objective evaluation (e.g., of company activity, market segment) (focus on facts) do not require features of expert stance. This example of essays on Dimension 4 thus illustrates a divergence of

topic and purpose, showing that their interrelationship looks different depending on the register and the linguistic function in question.

In sum, these results have shown that the situational parameters of topic and purpose are not independent, and this complex interrelationship, specifically the possibility that certain topic types are systematically associated with certain purposes, may be worth examining with regard to other linguistic patterns (such as the other two dimensions of this study – ‘Abstract vs. Concrete discourse’ and ‘Others-oriented descriptive vs. Self-oriented or interactive discourse’) in a range of registers. While it has been mostly purpose that has been credited with explaining variation, understanding the nature of the relationship between topic and purpose may contribute important insights into our conceptualization of intra-register linguistic differences.

8.2.2 The Role of Additional Granular Situational Factors

While topic and purpose are the variable parameters in the registers of essays and emails, the study further demonstrates that even when all the values for all the situational parameters in the existing framework are defined, texts may exhibit substantial linguistic variation. According to the hypothesis of functional correspondence on all levels of analysis, the existence of linguistic variation within highly narrow registers should again be related to some situational variable. The grouping of texts based on the observed linguistic differences and a comparison of the groups on the opposite ends of the scale from ‘central’ to ‘peripheral’ exemplars revealed clearly interpretable patterns and led directly to the identification of such additional situational variables. While these factors may be specific to the registers of the study (e.g., the distinction between focus on the content of the visual and extending it through additional content and subjective interpretation is highly specific to the task of image descriptions), these identified characteristics may still serve as a basis for future register analyses, especially if these analyses target register-

internal linguistic variation. For example, if the specific situational parameters identified in image descriptions, business memos, and evaluations are thought of more generally, they may be construed as distinctions in degree of (or presence or absence of) subjective judgment, degree of reliance on (or presence or absence of) expert opinions, and degree of (or presence or absence of) addressee focus. As follows from all the analyses and discussion of these findings in the previous chapters, however, the underlying reason for these situational distinctions again appears to be the degree of involvement. Heightened author involvement was discussed in evaluations on Dimension 4 as some authors chose to explicitly express their stance towards the state of Aston facilities and proposed a course of action and specific improvements; author involvement was observed in business memos as some authors referred to their actions, the search they performed, the results they obtained, and consistently related these results to the needs and expectations of the addressee, supporting them through expert opinions, i.e., reviews and ratings; finally, in image descriptions involvement could be viewed as the reason for the different approach to the task by some authors, who accompanied their description with an extended interpretive commentary. Thus, while the specific situational parameters varied from one register to another, these different parameters seem to have revealed author involvement as the underlying basis for the linguistic differences on both dimensions and in almost all registers (with the exception of essays discussed above).

The results of this study have conclusively demonstrated in a range of registers and with regard to two linguistic patterns ('Oral Elaboration vs. Information Density' and 'Evidence-based Stance') that registers are not uniform: rather, there is substantial variation across texts as shown by the existence of central and peripheral linguistic exemplars within each. These differences have been systematically associated with a difference in highly specific situational

factors. These specific situational parameters thus capture patterns of variation that extend beyond the differences between registers. As a result, the linguistic account of each specific register is enriched through a more fine-grained analysis, in which the focus is not on the mean (as a measure of central tendency of a register in comparison to other registers), but on each particular text and its place in the register space. This analysis of each register much more accurately reflects a complex reality, in which an entire variety cannot be characterized either by an average measure of its linguistic composition or a global account of its situation. Instead, the characteristics of texts in a register comprise a range of situational and linguistic possibilities. While some of these possibilities may be among its 'central' exemplars, representing the more common or conventional interpretations of the situation, others deviate from this tendency. These linguistic deviations have been shown to be the result of variation in a particular situational parameter in the texts (e.g., topic, communicative purpose, reliance on expert opinions, addressee-focus, subjective interpretation of content).

8.2.3 Situational and Linguistic Granularity in Broader Context

These findings are further supported by and should be discussed in light of the view of registers as mental categories (Keller, 2021) rooted in the prototype theory (Rosch, 1976), which distinguishes between prototypical (or central) and less representative members of a category. Keller (2021) proposes that speakers of a language have a certain mental representation of a register, formed through associative links between situational cues and linguistic structure acquired in the process of language use. Once these links are developed, speakers then identify similar situations in which corresponding linguistic structure is activated by the same or similar situational cues. These mental models include knowledge of the speakers' relationship between participants, how information is likely to be decoded, and other situational characteristics. That

is, a language user has “mental representations of a situation” and “the relevant knowledge about what is and is not appropriate for the situation”, which results in their ability to select the linguistic resources based on these considerations (Keller, 2021, p. 52). However, atypical situational cues may occur alongside situational characteristics that are part of such a mental representation of a register. When that happens, these atypical situational cues activate linguistic features associated with them, thus resulting in a linguistic deviation from the language commonly associated with the register. At the same time, situational cues that are commonly associated with the register may not be relevant, which leads to their suppression and the suppression of the corresponding linguistic features in favor of the new features that now correspond to the new situational parameter (Keller, 2021, p. 57). Furthermore, situational cues from different register categories may simultaneously activate linguistic features associated with several register categories, which would explain the existence of hybrid texts (such as those identified by Biber & Egbert, 2018), i.e., texts containing situational and linguistic markers of several registers. It is this activation of uncharacteristic linguistic features corresponding to an atypical (or new, additional, or different) situational cues that results in intra-register variation.

This view therefore provides an explanation for the phenomenon of variation across texts within a register from the perspective of authors’ mental representation of registers, while the present study, in turn, supports the existence of new, additional situational characteristics brought into a register or, conversely, disregarded by the author (or suppressed in their mental model of the register). One illustrative example of such competing situational cues can be found in the register of emails in the present study – specifically, class content-related emails. These emails quite noticeably involve (situational and linguistic) features of another register – the academic essays, dissertations, or other projects carried out by the authors in other, markedly

different situational contexts. The activation of linguistic features of those contexts (in this case associated with the original project and its topic and goals) was obvious in class content emails, which incorporated features of information density characteristic for the situational make-up of the original register whose situational characteristics were now intertwined with those of emails. Another example of a 'hybrid', although less stark, may be the essays presenting the results of qualitative analysis of interview data. On both dimensions, these essays stood out as more elaborated (on Dimension 1) as the topics were consistently deemed more personal and more rich in stance (on Dimension 4) as a result of their speculative, analytical goal. In essence, the reason for these trends is the elements of conversation (i.e., the author's conversation with the interviewee) permeating the texture of the essay. The authors cannot help being elaborated in such texts as they analyze excerpts from the conversation with the interviewee and the discourse themes they identify in them. Similarly, these conversational elements called for features of stance. The situational cues associated with the register of conversation thus led to the activation of linguistic features characteristic for that register. At the same time, the situational cues directly associated with essay writing must have also caused the activation of the linguistic resources necessary for that task. These examples here and the analyses of central and peripheral texts presented in the previous chapters illustrate that the described deviations can be viewed through the lens of mental register representation.

The results of this study and the fact that authors have mental models of what constitutes a register, however, raise questions about situational parameters with regard to their restrictiveness, on the one hand, and the authors' ability to exercise personal choice in observing or disregarding some of these situational considerations on the other. The topic and purpose of a text (an essay, an email, or any other text), for example, are the kinds of factors that inevitably

have to be considered in text production. Any text is produced on a certain topic, and any text has to fulfil particular communicative goals. On the other hand, the situational characteristics identified in the other registers, such as the degree of reliance on expert opinions and addressee-focus in business memos, the interpretation of the visual in image descriptions, and the choice to propose future directions in evaluations, appear to represent ‘optional’ choices under the authors’ control. This interpretive freedom became apparent when each of these registers (business memos, image descriptions, evaluations) showed exemplars that relied on a particular situational factor (e.g., expert opinions) and texts that disregarded that consideration (while no text can lack a topic or purpose). Thus, there seem to be inherent differences in the situational parameters with regard to the degree of agency a language user can have in introducing their own interpretation of a register or disregarding one or several of its expectations. The highly specific situational parameters accounting for linguistic variation in highly narrow, situationally well-defined registers appear to represent (at least in this corpus) the type of characteristics amenable to choice. Conversely, it appears that topic and purpose may be the kind of considerations imposed on authors by the communicative context to a greater extent than the more granular situational parameters, more open to interpretation.

It is worth asking at this point what may account for this addition or disregard for a particular situational parameter and what prompts the language users in this communicative choice? And could such disregard for a situational factor or an addition of a factor incongruous with the situational make-up of a register lead to a text no longer representing that register? If so, what kinds of situational parameters could draw that line and determine the status of a text within or outside a register? While these are all questions that arise from this research rather than are answered by it, they warrant some discussion. The very idea of ‘central’ as well as ‘peripheral’

exemplars of a register precludes the possibility that texts may no longer represent the register due to the extreme degree of their deviation from its norm. The two facts would be mutually exclusive. Indeed, when Biber and Egbert (2023) state that registers should not be *defined* in terms of their pre-determined situational characteristics because such an approach does not allow for variation within a register, but rather registers can and should be *described* in terms of their situational and linguistic characteristics, it is implied that the peripheral texts are an integral part of this internal complexity. The authors argue that registers are not a scientific construct that can be defined through a number of characteristic features, and a scientific category does not entail variation in its members with respect to these defining characteristic features. A register, in contrast, is a cultural construct. Speakers of a language have a mental representation of such a category, but as research discussed above illustrates, these mental models are flexible. Therefore, cultural categories, recognized by speakers of a language, in contrast to scientific categories, are not based on a rigid set of criteria of inclusion and allow variation. Additionally, culturally recognized categories, again in contrast to scientific, may feature hybrid texts which do not clearly belong to any of the registers whose features they combine. This would be impossible in a scientific category as such a text would need to be classified into a new category based on a set of its newly discovered defining characteristics (Biber & Egbert, 2023).

While these are the assumptions the current study operates on as it accepts that all business memos, for example, represent the register, the variation observed within the registers of the study inevitably poses questions regarding the existence of register boundaries and, relatedly, the reasons for variation. It has been stated before that variation within registers has stemmed from communicative differences (in topic and purpose) and in other cases was perhaps to a greater extent dictated by individual decisions of the language users. While the exact reasons

for these decisions are not clear, they may be the result of a unique communicative need felt by the author in a situational context that typically does not satisfy these communicative needs (e.g., highly frequent self-references and a heavy interpersonal focus in business memos). Despite certain situational restrictions, an author could pursue this communicative need and produce a text uniquely different from the trend. This possibility makes peripheral texts especially interesting objects for linguistic investigation as they appear to have the potential to gradually widen the scope of the allowed possibilities within a register if enough speakers produce more similar texts over time. However, the example here – a highly elaborated business memo – does not deviate from the register to the extent that it stops fulfilling its core goals and covering its required topic. That is, while it incorporates an additional situational and linguistic component, it was clear from all the performed analyses that such a text successfully achieves the basic goal of informing the addressee on all the required topics. However, the study did include business memos that failed to do so by considerably limiting the expected content and, as a result, did not effectively meet the informational goal. These one- or two-line texts barely satisfied the core situational characteristics of the register. This peripheral exemplar could also gradually become a more representative text of this register if more of these texts are produced; these texts are not a singular occurrence even now in the corpus of the study. The efficiency of such texts, however, is immeasurably lower due to their disregard for the vital situational considerations, such as the specified values of topic (accommodation, food preferences, entertainment, giftshop, etc.) and purpose (to inform about all the above topics). It was noted above that these core considerations could be distinguished from the more granular distinctions that occur within a register, the former being the required considerations and thus more restrictive and the latter being ‘optional’ considerations that could be a matter of choice. While such peripheral examples of business

memos were discussed by this study as texts not reliant on expert opinions and not integrating addressee focus, it could instead be questioned whether such texts in fact meet the core characteristics of the register and satisfy the expectations of topic and purpose. While in the initially proposed framework for situational analysis (Biber, 1994), it was emphasized that the situational parameters are not hierarchically organized, the parameters underlying variation *within* registers may differ from between-register analysis. This observation regarding the nature of situational parameters, some of which may determine the status of a text within a register and others are a matter of communicative choice within a register, calls for further studies in intra-register variation, which may eventually propose that a hierarchical relationship between such higher-order and lower-order situational parameters may exist.

8.3 INDIVIDUAL VARIATION ACROSS REGISTERS: INDIVIDUAL FREEDOM AND IDEAS OF APPROPRIACY

The investigation of authors' texts across registers has shed light on individual views on register conventions. On both examined dimensions, the study has identified authors who have shown quite consistent preferences (as shown by at least half of the authors' texts) towards the central tendency across all registers, or a tendency towards one or the other dimension function – i.e., authors that consistently prefer oral elaboration across registers were counterposed to authors who are consistently more informational, and authors who systematically expressed more personal stance were counterposed to those that tended not to. On the one hand, this consistency may lend support to conclusions regarding the persistence of an author's linguistic choices: if an author's linguistic behavior shows stability across situational contexts, it suggests that they may have an established linguistic preference. On the other hand, there seem to be other unique

insights offered by the cross-register approach taken by this study, namely the investigation of the specific place the author occupies within a register and direction of deviation.

First, this approach reveals an author's interpretation of the situational characteristics of each register. The registers in the study are situationally and linguistically distinct; thus, if an author consistently represents the central tendency of most registers, we learn that this author is sensitive to distinctions in situational contexts, on the one hand, but also, importantly, shows an awareness of the common trend of each particular context. That is, this author has a clear understanding of what typically comprises an academic essay, email communication, or an informal interview, all quite different from each other, and conforms to the general trends of these registers. It is also possible, however, that such authors are not consciously aware of the existence of norms, but are guided by their implicit knowledge acquired through exposure. Regardless of the nature of this understanding, an author's central status across registers tends to be indicative of their general register sensitivity.

Conversely, authors who tend to occur on the periphery of most registers, reveal a lack of compliance with the mainstream trend and seem to always prioritize their distinct communicative needs over the register conventions. However, this lack of convergence with the other authors may not necessarily stem from a lack of awareness of the register conventions, although that may be a possibility. A lack of understanding of, inattention and insensitivity to context could lead an author to repeatedly produce texts strikingly different from the central trend across contexts, especially in contexts that require exposure and even explicit instruction (e.g., academic essays). However, since all texts within a register are viewed by the study as belonging to that register, peripheral status is not interpreted through the lens of appropriacy. Rather, it is viewed as an author's unique and valid interpretation of the situational context. Consistent deviation from the

trend across registers may then be an intentional expression of creativity and distinctiveness. This systematic deviation from the central trend may be the result of authors' belief that register boundaries can be expanded (as discussed in the previous section) and their choice to defy conventions out of considerations of effectiveness. Indeed, it can be argued that sometimes peripheral texts revealed a better understanding of the conventions of the register than the central exemplars, contributed to its goals in important ways, and enhanced effectiveness. Examples of this can be found in the register of business memos, where the central texts were not as detailed and elaborate regarding the proposed options, did not always relate them to the search criteria, or support them through evidence to the same extent as the peripheral texts above the register mean did. In such cases, deviation from the trend definitely does not betray a lack of understanding of the register conventions, but conversely, shows that this understanding may in fact be more refined.

Another layer of information about the peripheral status of authors' texts across registers is added by the direction of deviation. As noted above, a particular direction of deviation across registers lends support to conclusions about an author's persistent linguistic choices. However, it may not be interpreted in this uniform way across registers. For example, while a clear preference for information density in image descriptions and in business memos does mean that an author tends to be consistently informational rather than oral and elaborated, in business memos this deviation is in line with the register trend (since the register is overall informational), but in image descriptions this deviation would be contrary to the register trend. In that sense, the author's deviation towards information density is much more meaningful (or a more creative choice) in image descriptions than it is in business memos.

The reasons for consistent deviations in one direction or the other can be hypothesized about, but based on the data of this study are impossible to establish with certainty. However, the approach applied here could yield a wealth of knowledge into the tendencies of an author whose identity is known, poses interest in itself, and could explain a certain linguistic pattern. For example, a 19th century fiction author occurring in the central tendency of 19th century fiction novels produces texts in line with the 19-century tradition of fiction writing; if another hypothetical author is markedly informational in their fiction, deviation from the central tendency towards information density in other registers could stem from their tendency to bring in linguistic resources of the register that they have expertise in: for example, this author's personal letters, which as a register are more oral and elaborated (Biber, 1988), showing a deviation in the direction of information density, common for their fiction, could be an example of such transfer.

Although consistent patterns – representing the central tendency or deviating towards one dimension function or the other across most registers – were shown to be persistent, they were never uniform. Even the authors who showed such consistency in most of their texts could diverge from their preferred pattern in one or two registers. Furthermore, authors who showed this overall consistency across registers are, in fact, much more rare than the authors who show diversity of patterns across registers. That is, the majority of authors vary widely with regard to their positions in the registers, and their texts may be central exemplars in some registers, occupy a medial position in others, and be on the periphery of still other registers. It appears that it is this pattern of divergence that is more complex and harder to interpret, but at the same time even more informative than the patterns of linguistic consistency.

One example of such complexity is an instance of authors 1) showing a deviation from their preferred cross-register pattern and 2) this deviation *not* being required by register constraints. For example, an author who consistently preferred a lack of personal stance across most registers showed a deviation in the opposite direction in image descriptions. It was shown, however, that the situational characteristics of image descriptions did not restrict other authors with regard to the extent to which stance could be expressed. Instead, descriptions showed quite wide variation, with texts ranging from rich in personal stance and offering subjective interpretation to low in stance if the authors chose not to provide personal commentary. It is therefore *not* the situational restrictions of the register that led this author to make a choice contrary to their preference shown in several other registers when a choice in line with their preference was possible. While the exact reasons are again unclear, such results suggest that considerations other than register or individual linguistic preferences, such as the authors' *ideas* about appropriacy or efficiency of certain approaches to a task, add a layer of complexity to the concept of individual cross-register variation. Based on the observations of this study, these beliefs appear distinct from the demands of the situation of use (i.e., register considerations) *as well as* individual preferences for particular expression. Yet, these ideas appear to guide authors in making choices that they consider most fitting for particular registers. This nuance has been made possible by a combination of the two approaches adopted here – variation within each register, on the one hand, which showed the extent and the kind of variation a register allows, and individual variation across registers.

The register perceptions that may guide authors in their linguistic choices appear to encompass, at least to an extent, considerations of strictness or permissiveness of situational context. That is, academic essays may be seen as a more restrictive register due to the formality

of the academic context and its association with more clear-cut conventions, in which case authors may choose to adhere to the central tendency. In contrast, authors may perceive other registers as less strict due to the overall informality of the context, thus allowing freedom in the form of deviations from the accepted conventions. The results, therefore, suggest that a complete account of individual language use should, if possible, be based on a variety of registers, and the reasons lie not only in the effect of register (Cvrček et al., 2020; Gracheva, under review), but in the varying interpretations authors tend to have of the register conventions.

The cross-register approach to individual variation has overall yielded insights that could be useful in future research in that avenue. More complete insights into individual variation may result from an initial detailed examination of the register, in which the author's language is examined, to establish what is characteristic for the register, how much communicative variation is observed in it, and where the author stands in this variation. Further cross-register investigations would then show if the author's views on register restrictions vary across contexts. While the present study could not systematically investigate each author's language use across registers due to a lack of data for each author, it showed that register mattered for such research efforts in a number of ways. Specifically, the first part of the investigation, namely variation across texts of each register, informed its second part – the investigation of variation among people. All and any of these insights could be applied to investigations of the language of political figures – an avenue that appears to benefit the most readily from these approaches due to the availability of texts of various registers regularly produced by these modern-day authors. These approaches could also be applied to investigations of fiction and nonfiction authors, in whose cases several texts from these registers are available.

It has thus been shown that analysis of register-internal variation and analysis of individual language use across registers yield complementary insights. On the one hand, texts vary within a single domain and reflect communicative distinctions within a register. These communicative distinctions may represent authors' choices of how to address a particular task. These choices reflect unique approaches to the tasks as in image descriptions and business memos, interpretations of the goal of the register, as in evaluations, where some authors proposed solutions instead of simply listing their likes and dislikes. Often this choice of approaches on both dimensions reflected authors' willingness or tendency to reveal self and express subjective judgment. In other registers, considerations, such as topic and purpose, determined linguistic choices. On the other hand, individual variation across registers revealed patterns of differences in individual interpretation of various situational contexts and their boundaries. Importantly, the cross-register perspective showed that authors are often guided by considerations other than register constraints or their individual linguistic preferences. These additional considerations, which may involve ideas of appropriacy or effectiveness of a particular communicative approach to a task, pose new questions for future research into individual language use.

8.4 CONCLUSION: LIMITATIONS AND FUTURE DIRECTIONS

8.4.1 Limitations of the Corpus

Several limitations of the present project need to be acknowledged. One such limitation concerns the corpus of the study and lies in the contrived nature of the registers included in the corpus.

Most of the analyses conducted by the study are based on two naturally occurring registers (essays and emails) and four registers elicited in a research setting (evaluations, business memos,

image descriptions, and interviews). While the elicited registers of the corpus showed an impressive situational and linguistic range, and the findings of the study – the inherent diversity of these registers – are even more likely to be observed in naturally occurring registers, which can be anticipated to be even more varied, these findings and the claims made on their basis need to be tested on other registers.

Although the Corpus of 100 Idiolects proved a valuable resource that yielded rich insights into the questions of the study, a substantial limitation of the corpus is the lack of data for investigations of authors' individual language. In the cross-register investigation of individual language use, this study examined this individual expression with regard to the amount of internal freedom a register allows and authors' positions relative to the registers' central tendency. The dataset, in which each author is represented through only one text, however, precluded an investigation of an author's individual expression that would take into account variation across each author's works. As a result, the findings about an author's standing in each register, restricted to one text, cannot be generalized to what is known about that author's linguistic preferences in that register overall. By extension, conclusions about that author's cross-register patterns should also be made with caution. The outcomes of this study strongly suggest, however, that such a cross-register examination of individual language use relative to the register trends could result in comprehensive and rich individual linguistic profiles.

Finally, while the corpus contains a rich annotation of the speakers' social characteristics, its design simply does not permit conclusions typical for sociolinguistic investigations. In this study, this limitation concerns the variable of age. The study views the age groups as markers of students' stages in their university careers, but on the whole these groups represent a single quite homogenous groups of speakers – young adults. The conclusions reached by the study thus

contribute to the body of research on this population, but they cannot offer insights into the effect of more substantial age differences on functional linguistic variation. Such future investigations based on other corpora, which include a wider age variety, could address questions of age grading with regard to general functional patterns such as the ones identified by the study.

8.4.2 Limitations of the Method

A limitation of the method lies in the choice to rely on the register mean as the benchmark against which deviation from the central tendency was measured. As the mean is not a concrete point of reference, but rather an amalgam of all text scores comprising the register, all distributed differently in different registers, it may not be the best measure of what is typical, characteristic, or common in a register. That is, the mean is computed on the basis of all text scores, including the scores of the peripheral exemplars, which thus comprise the mean alongside all other texts of the register. The texts representing the central tendency of a register were therefore not discussed in these terms (typical, characteristic, or common) for that reason. The mean is only referred to as a measure of the central tendency and not as the ‘norm’ for the register. To address this limitation, future studies could choose other measures of central tendency, perhaps more representative of the register ‘norm’, such as the register median, which would offer a concrete point of reference. Alternatively, the register mean could be computed excluding the peripheral exemplars, which could result in a more accurate representation of the central tendency.

Another limitation of the method used concerns the register of text messages, which could not be included in the key analyses of this study due to considerable limitations in text length. The methods applied by this study therefore need to be further developed and improved to accommodate short texts. While text messages may be limited in length, they are definitely not limited in their situational diversity, allowing an almost limitless range of purposes, topics,

and addressee-associated factors. As in all registers, this situational diversity corresponds to a wide linguistic range. To an extent, this range was demonstrated by this study. However, the shortness of these texts posed a major limitation to the applied quantitative method and resulted in score inflation. Thus, other approaches are needed to investigate the extensive variation within text messages and the basis for this variation.

A natural limitation of the study lies in limiting the interpretation of the results to some situational parameters, such as topic and purpose in some registers. While topic showed highly informative patterns in some registers, such as emails, it is important to acknowledge the possible effect of other situational factors in a highly varied register like emails. It was discussed earlier that while topic and purpose were the variable situational parameters chosen as the basis for the subsequent text analyses, emails vary with regard to other situational factors as well. These other factors include the exact addressee (professors, other university officials or authority figures, classmates or other fellow students, colleagues, or other entities, such as local businesses, landlords, or unspecified entities in letters of complaints), the relationship with the addressee, the relative status, and the amount of shared knowledge. While all these variables could be important predictors of variation across emails, their values are harder to specify (e.g., the relationship between the author and an addressee like their landlord or colleague is not clear), and some of these emails are singular cases rather than groups of texts that could yield a pattern. Additionally, while the situational parameters in the framework are not hierarchical (Biber, 1994), and there is no precedence of some situational factors over others, it appears intuitive that considerations of purpose and topic inevitably involve considerations of audience (e.g., if the purpose is to describe the malfunctioning appliances in the author's home and request maintenance, which is the topic of the email, it is implied that the landlord or a similar figure is

the audience). Nevertheless, acknowledging the existence of other variable parameters is important for future investigations of this highly diverse register with a wide scope of situational possibilities and seemingly few restrictions. While topic was shown to be the factor explaining linguistic variation here, the study does not claim that all internal variation in the register is due to this factor alone.

8.4.3 Future Directions

Situation of Use

Future directions that emerge from the findings of this study include viewing both linguistic features and the associated situational parameters in continuous space. While it was stated above that the linguistic deviations within a register are due to a presence or absence of a particular situational parameter, it would be more accurate to say that these linguistic deviations reflect the degree to which a particular situational parameter is present in the texts. The linguistic characteristics of texts are measured on a continuous scale, namely through dimension scores of each text, which in their turn are based on weighted combinations of the rates of occurrence of the linguistic features comprising the dimensions. The differences between these text scores and the register central trend vary in extent, which is conveyed through the measure of distance of each text from the register mean. However, if linguistic variation has traditionally been accepted to be of continuous nature and linguistic and situational variation are functionally related, situational variation should likewise be measured in continuous terms (Biber & Egbert, 2023). The feasibility and potential of continuous rather than categorical approaches to situational analysis have been previously demonstrated (Biber et al., 2020). Moreover, Egbert, Biber and Keller (in preparation) illustrate a systematic relationship between linguistic and situational dimensions of variation identified in online registers (Biber & Egbert, 2018; Biber et al., 2020).

The study correlates the linguistic and situational variables (linguistic and situational dimensions), both measured continuously, in a random sample of texts from the corpus of online registers (CORE) and identifies patterns of linguistic/ situational correspondence in the domain. That study concludes that it is possible to trace clear relationships between linguistic and situational dimensions. While the present study identified the explanatory situational variables in the process of analysis and did not have predetermined situational parameters that could be measured continuously, the continuous nature of situational variation was observed in the process of text analysis and is important to acknowledge and investigate in a systematic way in the future.

An important new direction is the shift in focus from the central tendency to peripheral texts. This explicit focus on the outliers accentuates the fact that registers are situationally and linguistically complex and facilitates understanding of the full range of texts recognized by a culture as exemplars of a register. The existence and emergence of peripheral texts may also mean that registers are dynamic and flexible categories, allowing new members and redefining themselves. The peripheral exemplars may then be of interest as markers of change in cultural conceptualizations of a register. A promising future extension of this line of research is a diachronic investigation of register development as culturally recognized categories.

Specifically, the transition of a peripheral exemplar of a register into a more central position in that register with time can shed light on what this developmental trajectory may involve. Such register metamorphosis may occur as a) more speakers produce texts initially considered peripheral and a category reconstruction occurs (Keller, 2021), but also b) as other even greater deviations occur within the register. Diachronic accounts of these processes may offer insights into the formation of new registers on the one hand, as register categories may go from more

amorphous and highly varied due to speakers' lack of clarity with regard to their situational make-up to more situationally uniform as such clarity increases. Alternatively, once registers are well-established as a culturally recognized category, they may show a 'funneling' pattern from uniformity to a higher internal diversity in the form of sub-registers, which in this case could be an indicator of the register's status as an established enough category to develop further complexity. Such efforts would offer essential knowledge of the speakers' perceptions of the situational characteristics of a register and the construction of their mental models of a cultural category.

Characteristics of the Users

It was mentioned above that a natural extension of this work for sociolinguistic research could be an analogous investigation of more diverse age groups that could reveal patterns of age grading if age-grading applies to functional linguistic variation. While adolescents and young adults are the most common subject of analysis in age-based variation in sociolinguistics (Barbieri, 2008), adult speakers have been shown to exhibit variation associated with major life events and transition points between life stages (Cheshire, 1987; Seifert, Hoffnung & Hoffnung, 2000, cited in Murphy, 2010). Investigating functional linguistic variation among age groups that represent diverse life stages and experiences appears to be a promising research initiative. Additionally, similar investigations could be carried out with other social variables, such as social class and educational background, as predictors of functional linguistic variation.

With regard to individual characteristics of the users, an important finding in this study that needs to be explored further consists in the additional considerations of context restrictiveness or freedom that appear to guide authors in their choice of communicative

approaches to a text. The study observed patterns of deviation that ran counter to the pattern of author linguistic preferences and were also not imposed by register constraints as the registers in question showed substantial communicative variation. It was proposed that these could be considerations of appropriacy or effectiveness, distinct from actual register demands, but a targeted investigation of what constitutes these additional considerations based on several texts from each author and possibly complemented by interviews with the authors could be a new direction for future research on individual variation.

8.4.4 Conclusion

The present study explored the relationship between distinct predictors of linguistic variation – characteristics of the users or the situation of use. The study was intended to illustrate that these distinct strands can be united in a single investigation of linguistic variation – their common goal – and new insights can be gained into each of the predictors through examining its links to the others. It is hoped that while the present study only took a step in this direction, the recognition of the mutual relationship between characteristics of the users and the situation of use becomes more prominent in future studies of discourse.

References:

- Aijmer, K. (2018). 'That's Well Bad': Some new intensifiers in spoken British English. In V. Brezina, R. Love, & K. Aijmer (Eds.), *Corpus approaches to contemporary British speech. Sociolinguistic studies of the Spoken BNC 2014* (pp. 60-95). Routledge.
- Alim, A. (2004). *You know my steez: An ethnographic and sociolinguistic study of style-shifting in an African-American speech community*. Duke University Press.
- Aries, E. (1996). *Men and women in interaction: Reconsidering the differences*. Oxford University Press.
- Asif A., & Frog (2015). An introduction to registers in communications. In A. Asif & Frog (Eds.), *Registers of communications* (pp. 13-23). Finnish Literary Society.
- Auer, P. (2007). *Style and social identities: Alternative approaches to linguistic heterogeneity*. Mouton de Gruyter.
- Bakan, D. (1966). *The duality of human existence*. Rand McNally.
- Bakhtin, M. (1986). *Speech genres and other late essays*. University of Texas Press.
- Balasubramanian, B. (1982). An objective approach to the linguistic style of press advertisements in English. *Bulletin of the Deccan College Post-Graduate and Research Institute*, 41, 11-15.
- Barber, B. (1957). *Social stratification*. Brace.
- Barbieri, F. (2008). Patterns of age-based linguistic variation in American English. *Journal of Sociolinguistics*, 21(1), 58-88.
- Barbieri, F., & Wizner, S. (2019). Appendix 1. Annotations of major register and genre studies. In D. Biber & S. Conrad, *Register, genre, and style*.
- Bell, A. (1984). Language style as audience design. *Language in Society*, 13, 145-204.
- Bernstein, B. (1971). *Class, codes, and control*. Routledge & Kegan Paul.

- Biber, D. (1985). Investigating macroscopic textual variation through multi-feature/multi-dimensional analyses. *Linguistics*, 23, 337-360.
- Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language*, 62, 384-414.
- Biber, D. (1988). *Variation across speech and writing*. 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 University Press.
- Biber, D. (2004). Conversation text types: A multi-dimensional analysis. In G. Purnelle, C. Fairon, & A. Dister (Eds.), *Le poids des mots: Proceedings of the 7th International Conference on the Statistical Analysis of Textual Data* (pp. 15-34). Presses Universitaires de Louvain.
- Biber, D. (2006). *University language. A Corpus-based study of spoken and written languages*. John Benjamins.
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5, 97-116.
- Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory*, 8(1), 9-37.
- Biber, D., & Burges, J. (2000). Historical change in the language use of women and men. *Journal of English Linguistics*, 28(1), 21-37.
- Biber, D., & Burges, J. (2001). Historical shifts in the language of women and men: Gender differences in dramatic dialogue. In D. Biber, & S. Conrad (Eds.), *Variation in English: Multi-dimensional studies* (pp. 157-170). Longman.
- Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.

- Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly*, 36(1), 9-48.
- Biber, D., & Egbert, J. (2018). *Register variation online*. Cambridge University Press.
- Biber, D., & Egbert, J. (2023). What is a register? Accounting for linguistic and situational variation within – and outside of – textual varieties. *Register Studies*, 5(1), 1-22.
- Biber, D., Egbert, J., & Keller, D. (2020). Reconceptualizing register in a continuous situational space. *Corpus Linguistics and Linguistic Theory*, 16(3), 581-616.
- Biber, D., Egbert, J., Keller, D., & Wizner, S. (2021). Towards a taxonomy of conversational discourse types: An empirical corpus-based analysis. *Journal of Pragmatics*, 171, 20-35.
- Biber, D., & Finegan, E. (1994). Intra-textual variation within medical research articles. In N. Oostdijk & P. de Haan (Eds.), *Corpus-based research into language* (pp. 201-222). Rodopi.
- Biber, D., & Finegan, E. (1994). Multi-dimensional analyses of authors' styles: some case studies from the eighteenth century. *Research in Humanities Computing*, 3, 3-17.
- Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9, 2-20.
- Biber, D., & Gray, B. (2012). The competing demands of popularization vs. economy: Written language in the age of mass literacy. In T. Nevalainen & E. Traugott (Eds.), *The Oxford handbook of the history of English* (pp. 314-328). Oxford University Press.
- Biber, D., & Gray, B. (2016). Grammar emerging in the noun phrase: The influence of written language use. *20th Anniversary Collection, English Language and Linguistics*.
- Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, 37(5), 639-668.

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999/ 2021). *Longman grammar of spoken and written English (LGSWE)*. Longman.
- Seoane, E., & Biber, D. (2021). A corpus-based approach to register variation. In D. Biber, & E. Seoane (Eds.), *Corpus-based approaches to register variation* (pp. 1-17). John Benjamins.
- Blankenship, J. (1974). The influence of mode, submode, and speaker predilection on style. *Speech Monographs*, 41, 85-118.
- Bloomfield, L. (1933). *Language*. Henry Hold & Co.
- Blom, J. & Gumperz, J. (1972). Social meaning in linguistic structures: code-switching in Norway. In J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 407-434). Holt, Rinehart and Winston.
- Brezina, V., Love, R., & Aijmer, K. (Eds.). (2018). *Corpus approaches to contemporary British speech. Sociolinguistic studies of the Spoken BNC 2014*. Routledge.
- Brown, J.D. (2009). Choosing the right type of rotation in PCA and EFA. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 13(3), 20-25.
- Brown, P. (1998). How and why are women more polite: some evidence from a Mayan community. In J. Coates (Ed.), *Language and gender: A reader* (pp. 81-99). Blackwell.
- Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge University Press.
- Carter, R. (2004). *Language and creativity. The art of common talk*. Routledge.
- Carter, R. (2007). Response to special issue of Applied Linguistics devoted to language creativity in everyday contexts. *Applied Linguistics*, 28(4), 597-608.
- Chafe, W.L. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and written language: exploring orality and literacy* (pp. 35-54). Ablex.

- Chafe, W.L. (1985). Linguistic differences produced by differences between speaking and writing. In D.R. Olson, N. Torrance, & A. Hildyard (Eds.), *Literature, language, and learning: The nature and consequences of reading and writing* (pp 105-123). Cambridge University Press.
- Chandler, D., & Munday, R. (2016). *A dictionary of media and communication* (2nd ed.). OUP Oxford.
- Cheshire, J. (1987). Age and generation-specific use of language. In U. Ammond, N. Dittmar, & K. Mattheier (Eds.), *Sociolinguistics: An international handbook of the science of language and society* (pp. 760-767). Walter de Gruyter.
- Cheshire, J. (2005). Syntactic variation and beyond: Gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics*, 9 (4), 479-508.
- Cheung, I. (2017). Plain language to minimize cognitive load: A social justice perspective. *IEEE Transactions of Professional Communication*, 60(4), 448-457.
- Chomsky, N. (1979). Human language and other semiotic systems. *Semiotica*, 25, 31-44.
- Coates, J. (1996). *Women talk: Conversation between women friends*. Blackwell.
- Coates, J. (2004). *Women, men and language: A sociolinguistic account of gender differences in language* (3rd edition). Pearson Longman.
- Coleman E.B., & Blumenfeld, J.P. (1963). Cloze scores of nominalizations and their grammatical transformations using active verbs. *Psychological Reports*, 13(3), 651-654.
- Conrad, S. (1996). Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education*, 8(3), 299-326.
- Coupland, N. (1980). Style-shifting in a Cardiff work-setting. *Language in Society* 9(1), 1-12.
- Coupland, N. (2001). Language, situation, and the relational self: Theorizing dialect-style in sociolinguistics. In P. Eckert & J.R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 185-211). Cambridge University Press.

- Coupland, N. (2007). *Style. Language variation and identity*. Cambridge University Press.
- Coupland, N. (2010). The authentic speaker and the speech community. In C. Llamas & D. Watts (Eds.), *Language and identities* (pp. 99-112). Edinburgh University Press.
- Crawford, M. (1995). *Talking difference: On gender and language*. Sage.
- Crosby, E., & Nyquist, L. (1977). The female register: An empirical study of Lakoff's hypotheses. *Language in Society*, 6, 313-322.
- Crystal, D., & Davy, D. (1969). *Investigating English style*. Longman.
- Cvrček, V., Laubeová, Z., Lukeš, D., Poukarová, P., Řehořková, A., & Zasina, A.J. (2020). Author and register as sources of variation. A corpus-based study using elicited texts. *International Journal of Corpus Linguistics*, 25(4), 461-488.
- Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219-253. DOI io.iO75/lab.2.3.oid
- Drieman, G.H.J. (1962). Differences between written and spoken language. *Acta Psychologica*, 20, 78-100.
- Dubois, B. L., & Crouch, I. (1975). The question of tag questions in women's speech: They don't really use more of them, do they? *Language in Society*, 4, 289-294.
- Durmus, E., & Cardie, C. (2018). Understanding the effect of gender and stance on opinion expression in debates on “abortion”. *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media* (pp 69-75). Association for Computational Linguistics.
- Eckert, P. (2000). *Linguistic variation as social practice*. Blackwell.
- Eckert, P. (2003). Elephants in the room. *Journal of Sociolinguistics*, 7(3), 392-397.

- Eckert, P. (2004). Adolescent language. In E. Finegan & J. Rickford (Eds.), *Language in the USA: Themes for twenty-first century* (pp. 361-374). Cambridge University Press.
- Eckert, P. (2017). Age as a sociolinguistic variable. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 151-167). Blackwell.
- Eckert, P., & McConnell-Ginet, S. (2003). *Language and gender*. Cambridge University Press.
- Egbert, J. (2012). Style in nineteenth century fiction. A multidimensional analysis. *Scientific Study of Literature*, 2(2), 167-198.
- Egbert, J. (2015). Publication type and discipline variation in published academic writing: Investigating statistical interaction in corpus data. *International Journal of Corpus Linguistics*, 20(1), 1-29.
- Egbert, J., & D. Biber, D. (2023). Key feature analysis—A simple, yet powerful method for comparing text varieties. *Corpora* 18(1), 121-133.
- Egbert, J., Biber, D., & Keller, D. (in preparation). Correlating situational context and language use. New methods for an evolving theory of register.
- Egbert, J., & Davies, M. (2019). If olive oil is made of olives, then what's baby oil made of? In J. Egbert & P. Baker (Eds.), *Using corpus methods to triangulate linguistic analysis*. Routledge.
- Egbert, J., & Gracheva, M. (2023). Linguistic variation within registers: granularity in textual units and situational parameters. *Corpus Linguistics and Linguistic Theory*, 19(1), 115-143.
<https://doi.org/10.1515/cllt-2022-0034>
- Egbert, J., & Mahlberg, M. (2020). Fiction – one register or two? *Register Studies*, 2(1), 72-101.
- Esser, J. (1993). Quantitative methods in stylistics and definitions of style. *Neuphilologische Mitteilungen*, 94(3/4), 297-312.

- Finegan, E. (1980). *Attitudes towards English usage: The history of a war of words*. Teachers College Press.
- Finegan, E., & Biber, D. (1994). Register and social dialect variation: An integrated approach. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 315-47). Oxford University Press.
- Finegan, E., & Biber, D. (2001). Register variation and social dialect variation: the register axiom. In P. Eckert & J. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 235-267). Cambridge University Press.
- Firth, J. (1957). *Papers in linguistics*. Oxford University Press.
- Firth, J. (1968). *Selected papers of J. R. Firth (1952-1959)*. (F. R. Palmer, Ed.). Longman.
- Fleishman, S. (1998). Gender, the personal, and the voice of scholarship: A viewpoint. *Signs*, 23(4), 975-1016.
- Francis, W. N. (1983). *Dialectology. An introduction*. Longman.
- Friginal, E. (2009b). *The language of outsourced call centers: A corpus-based study of cross-cultural interaction*. John Benjamins.
- Friginal, E., & Hardy, J. A. (2014). *Corpus-based sociolinguistics*. Routledge.
- Gardner, S., Nesi, H., & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics*, 40(4), 646 - 674. <https://doi.org/10.1093/applin/amy005>
- Giles, H. (1984). *The dynamics of speech accommodation*. Mouton.
- Giles, H., & Powesland, P. (1975). *Speech style and social evaluation*. Academic Press.
- Gorsuch, R.L. (1983). *Factor analysis*. Erlbaum.

- Goulart, L. (2022). *Communicative text types in university writing*. (Publication No. 9798841774686) [Doctoral dissertation, Northern Arizona University]. ProQuest Dissertations Publishing.
- Goulart, L., Biber, D. & Reppen, R. (2022). In this essay, I will...: Examining variation of communicative purpose in university writing. *Journal of English for Academic Purposes*. 59. <https://doi.org/10.1016/j.jeap.2022.101159>
- Gracheva, M. (2022). Style of creative nonfiction. A multidimensional analysis of literary essays. *Scientific Study of Literature*, 12(1). <https://doi.org/10.1075/ssol.22002.gra>
- Gracheva, M. (under review). The role of situation in individual style: A powerful confounding variable or a new dimension? The case of American presidential discourse.
- Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. John Benjamins.
- Gregory, M. (1967). Aspects of varieties differentiation. *Journal of Linguistics*, 3(2), 177-198.
- Grob, L. M., Meyers, R. A., & Schuh, R. (1997). Powerful/powerless language use in group interactions: Sex differences or similarities? *Communication Quarterly*, 45, 282-304. doi:10.1080/01463379709370066
- Gumperz, J. (1964). Linguistic and social interaction in two communities. *American Anthropologist*, 66(6), 137-153.
- Gumperz, J. (1968). The speech community. In *International encyclopedia of the social sciences*. Macmillan.
- Gumperz, J. (1971). The speech community. In A.S. Dill (Ed.), *Language in social groups: Essays by John J. Gumperz*. Stanford University Press.
- Gumperz, J. (1982). *Discourse strategies*. Cambridge University Press.
- Halliday, M.A.K. (1973) *Explorations in the functions of language*. Edward Arnold.

- Halliday, M.A.K. (1978). *Language as a social semiotic*. University Park Press.
- Halliday, M. A.K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- Hardy, J. & Roemer, U. (2013). Revealing disciplinary variation in student writing: A Multidimensional analysis of the Michigan Corpus of upper-level student papers (MICUSP). *Corpora* 8, 183-207.
DOI: 10.3366/ cor.2013.0040.
- Hart, R. P. (1984). *Verbal style and the presidency: A computer-based analysis*. Academic Press.
- Hartman, M. (1976). A descriptive study of the language of men and women born in Maine around 1900 as it reflects the Lakoff hypotheses in "Language and women's place." In B. L. Dubois & I. Crouch (Eds.), *Proceedings of the Conference on the Sociology of the Languages of American Women* (pp. 81-90). Trinity University Press.
- Hasan, R. (2009a). The place of context in Systemic Functional Model. In M.A.K. Halliday & J.J. Webster (Eds.), *Continuum companion in Systemic Functional Linguistics* (pp. 166-189). Continuum.
- Heath, S. B. (1986). Socio cultural contexts of language development. In California State Department of Education (Ed.), *Beyond language: Social and cultural factors in schooling language minority students* (pp. 143-186). Evaluation, Dissemination & Assessment Center.
- Heini, A., Kredens, K., & Pezik, P. (2021). *Corpus of 100 Idiolects*. Aston Institute for Forensic Linguistics.
- Hockett, C.F. (1987). *Refurbishing our foundations: Elementary linguistics from an advanced point of view*. John Benjamins.
- Holmes, J. (1995). *Women, men, and politeness*. Longman.
- Hyland, K. (2002a). Directives: Argument and engagement in academic writing. *Applied Linguistics*, 23, 215-239.

- Hymes, D. (1974). *Foundations in sociolinguistics*. University of Pennsylvania Press.
- Ishikawa, Y. (2011). A corpus-based research on THANKS in British dialogue. *Proceedings of the International Conference on Language, Literature, and Linguistics IPELR*, 26 (pp. 384-389).
- James, D., & Drakich, J. (1993). Understanding gender differences in amount of talk: A critical review of research. In D. Tannen (Ed.), *Gender and conversational interaction* (pp. 281-312). Oxford University Press.
- Jespersen, O. (1922). *Language: Its nature, development, and origin*. George Allen & Unwin.
- Johnson, C. (1994). Gender, legitimate authority, and leader-subordinate conversations. *American Sociological Review*, 59, 122-135.
- Johnstone, B. (1996). *The linguistic individual: Self-expression in language and linguistics*. Oxford University Press.
- Joos, M. (1962). *The five clocks*. Brace.
- Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31–36.
<https://doi.org/10.1007/BF02291575>
- Kay, P. (1977). Language evolution and speech style. In B. G. Blount & M. Sanches (Eds.), *Sociocultural dimensions of language change* (pp. 21-33). Academic Press.
- Keller, D. (2021). *Register representation and processing*. (Publication No. 9798505511756) [Doctoral dissertation, Northern Arizona University]. ProQuest Dissertations Publishing.
- Kroch, A.S. (1978). Toward a theory of social dialect variation. *Language in Society*, 7, 17-36.
- Labov, W. (1966). The effect of social mobility on linguistic behavior. *Social Inquiry*, 36, 186-203.
- Labov, W. (1969). *The study of nonstandard English*. National Council of Teachers of English.

- Labov, W. (1986). The social stratification of (r) in New York city department stores. In H. B. Allen & M.D. Linn (Eds.), *Dialect and language variation* (pp. 304-329). Academic Press.
- Labov, W. (1990). The intersection of sex and social class in the course of linguistics change. *Language Variation and Change*, 2, 205-254.
- Laferriere, M. (1979). Ethnicity in phonological variation and change. *Language*, 66, 603-617.
- Lakoff, R. (1975). *Language and woman's place*. Oxford University Press.
- Lakoff, R. (1990). *Talking power: The politics of language in our lives*. Basic Books.
- Lamb, S. (1991). Acts without agents: An analysis of linguistic avoidance in journal articles on men who batter women. *American Journal of Orthopsychiatry*, 61(2), 250-257.
- Lindenfeld, J. (1969). The Social conditioning of syntactic variation in French. *American Anthropologist*, 71: 890-898.
- Llamas, C. (2006). *Age*. The Routledge Companion to Sociolinguistics.
- LePage, R., Tabouret-Keller, A. (1985). *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge University Press.
- Li, M., & Li, X. (2022). Gender differences and pragmatic identity construction of stance markers in business speeches from the perspective of communicative action theory. *Lecture Notes on Language and Literature*. DOI: 10.23977/langl.2022.050504.
- Matthiessen, C. (1993). Register in the round: Diversity in a unified theory of register analysis. In M. Ghadessy (Ed.), *Register analysis: theory and practice* (pp. 221-292). Pinter.
- Macaulay, R. (1991). *Locating dialect in discourse. The language of honest men and bonnie lassies in Ayr*. Oxford University Press.
- Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist*, 45(4), 513-520.

- Malinowsky, B. (1946). The problem of meaning in primitive languages. In C. K. Ogden & I. A. Richards (Eds.), *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism* (pp. 296-336). Brace & World, Inc.
- Marko, K., Reitbauer, M., & Pickl, G. (2022). Same person, different platform. Challenges and implications for forensic authorship analysis. An exploratory study of Instagram and Twitter users. *Register Studies*, 4(2), 202-231.
- Moerk, E. (1970). Quantitative analysis of writing styles. *Journal of Linguistics*, 6(2), 223-230.
- Moore, S. E, Shaffer, L., Goodsell, D. A., & Baringoldz, G. (1983). Gender or situationally determined spoken language differences? The case of the leadership situation. *International Journal of Women's Studies*, 6(1), 44-53.
- Mulac, A., Incontro, C. R., & James, M. R. (1985). Comparison of the gender-linked language effect and sex role stereotypes. *Journal of Personality and Social Psychology*, 49(4), 1098-1109.
- Mulac, A., & Lundell, T. L. (1986). Linguistic contributors to the gender-linked language effect. *Journal of Language and Social Psychology*, 5(2), 81-101.
- Mulac, A. & Lundell, T. (1994). Effects of gender-linked language differences in adults' written discourse: Multivariate tests of language effects. *Language and Communication*, 14(3), 299-309.
- Mulac, A., Lundell, T. L., & Bradac, J. J. (1986). Male/female language differences and attributional consequences in a public speaking situation: Toward an explanation of the gender-linked language effect. *Communication Monographs*, 53, 115-129.
- Murphy, B. (2010). *Corpus and sociolinguistics: Investigating age and gender in female talk*. John Benjamins.
- Nesi, H., & Gardner, S. (2012). *Genres across the disciplines: Student writing in higher education*. Cambridge University Press.

- O'Barr, W., & Atkins, B. (1980). 'Women's language' or 'powerless language'? In S. McConnell-Ginet, R. Borker, & N. Furman (Eds.), *Women and language in literature and society* (pp. 93-110). Prager.
- Oberlander, J., Gill, A. (2004). Individual differences and implicit language: personality, parts-of-speech and pervasiveness. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26).
- Osgood, C.E. (1960). Some effects of motivation on style of encoding. In T.A. Sebeok (Ed.), *Style in language* (pp. 293-306). MIT Press.
- Pennebaker, J.W., & King, L.A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6), 1296-1312.
- Pennebaker, J. W., & Stone, L.D. (2003). Words of wisdom: Language use over the life span. *Journal of Personality and Social Psychology*, 85, 291-301.
- Pinker, S. (2014). *The sense of style*. Penguin Books.
- Precht, K. (2008). Sex similarities and differences in stance in informal American conversation. *Journal of Sociolinguistics*, 12(1), 89-111. Doi: 10.1111/j.1467-9841.2008.00354.x
- Preisler, B. (1986). *Linguistic sex roles in conversation*. Mouton de Gruyter.
- Preston, D. (1991). Sorting out the variables in sociolinguistic theory. *American Speech*, 66(1), 33-56.
- Preston, D. (2001). Style and the psycholinguistics of sociolinguistics: The logical problem of language variation. In P. Eckert & J. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 279-304). Cambridge University Press.
- Reid, T.B.W. (1956). *Linguistics, structuralism, philology*. Archivum Linguisticum.
- Rickford, J., & Blake, R. (1990). Contraction and deletion of the copula in Barbadian English. *Proceeds of the sixteenth Annual meeting of the Berkley Linguistic Society* (pp. 257-268). Berkley Linguistic Society.

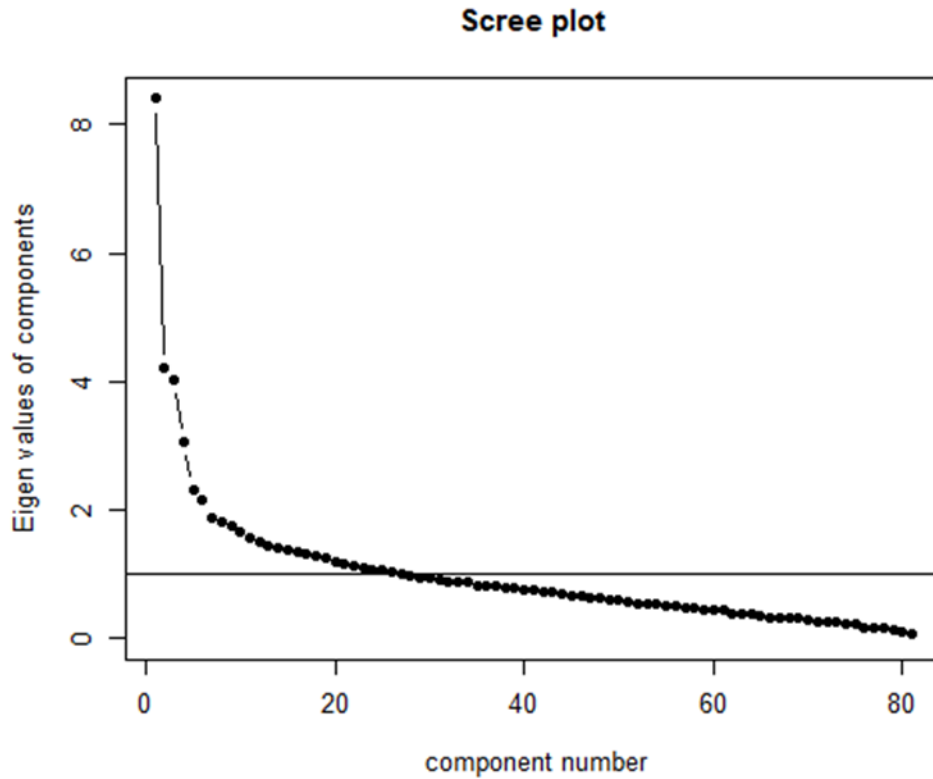
- Rickford, J. R., & McNair-Knox, F. (1994). Addressee- and topic-influenced style shift: a quantitative sociolinguistic study. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 235-276). Oxford University Press.
- Romaine, S. (1980). Stylistic variation and evaluative reactions to speech: problems in the investigation of linguistic attitudes in Scotland. *Language and Speech*, 23(3), 213-232.
- Romaine, S. (1982). What is a speech community? In Romaine, S. (Ed.). *Sociolinguistic variation in speech communities* (pp. 13-24). Edward Arnold.
- Rosch, E. (1976). Classification of real-world objects: Origins and representations in cognition. In S. Erlich & E. Tulving (Eds.), *La memoire semantique*. Bulletin de Psychologie.
- Rubin, D.L., & Greene, K. (1992). Gender-typical style in written language. *Research in the Teaching of English*, 26(1), 7-40.
- Sabino, R. (2018). *Languaging without languages*. Brill.
- Sagrestano, L. (1992). Power strategies in interpersonal relationships. *Psychology of Women Quarterly*, 16, 481-495.
- Sandell, R. (1977). *Linguistic style and persuasion*. Academic Press.
- Saussure, F. (1916). *Course in general linguistics*. Duckworth.
- Schmid H., Würschinger Q., Fischer, S., & Küchenhoff, H. (2021). That's cool. Computational sociolinguistic methods for investigating individual lexico-grammatical variation. *Frontiers in Artificial Intelligence*, 3. doi: 10.3389/frai.2020.547531
- Seifert, K. L., Hoffnung, R. J., & Hoffnung, M. (2000). *Lifespan development* (2nd ed.). Houghton Mifflin.
- Sheffler, T., L. Kern, & Seeman, H. (2022). The medium is not the message: Individual level register variation in blogs vs. tweets. *Register Studies*, 4(2), 171-201. <https://doi.org/10.1075/rs.22009.sch>

- Shore, S. (2015). Register in Systemic Functional Linguistics. In A. Asif & Frog (Eds.). *Registers of communications* (pp. 54-76). Finnish Literary Society.
- Slatcher, R., Chung, C., Pennebaker, J., & Stone, L. (2007). Winning words: Individual differences in linguistic style among U.S. presidential and vice-presidential candidates. *Journal of Research in Personality, 41*, 63-75.
- Soskin, W. F., & John, V. P. (1963). The study of spontaneous talk. In R. Barker (Ed.), *The stream of behavior* (pp. 228-281). Appleton-Century-Crofts.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Harvard University Press.
- Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication, 33*, 149-183.
- Stenström, A., & Jørgensen, A. M. (2009). *Young speak in a multilingual perspective*. John Benjamins.
- Stuart-Smith, J., & Timmins, C. (2010). The role of the individual in language variation and change. In C. Llamas & D. Watts (Eds.), *Language and identities* (pp. 39-54). Edinburgh University Press.
- Swales, J., Ahmad, U., Chang, Y.Y., Chavez, D., Dressen, D., & Seymour, R. (1998). Consider this: The role of imperatives in scholarly writing. *Applied Linguistics, 19*, 97-121.
- Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge University Press.
- Thompson, D. (1995). *The concise Oxford dictionary* (9th ed.). Clarendon Press.
- Trudgill, P. (1974b). *Sociolinguistics*. Penguin Books.
- Trudgill, P. (1986). Social identity and linguistic sex differentiation. In H. B. Allen & M. D. Linn (Eds.), *Dialect and language variation*. Academic Press, Inc.
- Trudgill, P. (1986). *Dialects in contact*. Basil Blackwell.

- Ure, J. (1982). Introduction: approaches to the study of register range. *International Journal of the Sociology of Language*, 35, 5-23.
- Van den Broeck, J. (1977). Class differences in syntactic complexity in the Flemish town of Maaseik. *Language in Society*, 6, 149-181.
- Verhagen, V., & Mos, M. (2016). Stability of familiarity judgments: Individual variation and the invariant bigger. *Cognitive Linguistics*, 27(3), 307-344. DOI 10.1515/cog-2015-0063
- Wolfram, W. (1995). *Sociolinguistics: An introduction to language and society*. Penguin Books.
- Wardhaugh, R. (1993). *An introduction to sociolinguistics*. Blackwell.
- Watt, D. (2010). The identification of the individual through speech. In C. Llamas & D. Watts (Eds.), *Language and identities* (pp. 76-85). Edinburgh University Press.
- Winter, E. (1982). *Towards a contextual grammar of English: the clause and its place in the definition of sentence*. George Allen and Unwin.
- Wolfson, N. (1989). *Perspectives. Sociolinguistics and TESOL*. Heinle & Heinle Publishers.
- Wood, M. (2023). *Communicative function and linguistic variation in State statutory law*. [Doctoral dissertation. Northern Arizona University]. ProQuest Dissertations Publishing.
- Zipf, G.K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.

APPENDICES

Appendix 1. Scree plot of eigenvalues – four-factor solution



Appendix 2. Factor loadings

Linguistic feature	Factor 1	Factor 2	Factor 3	Factor 4
Contractions	0.73	-0.40	0.15	-0.17
Present tense	0.58	0.18	-0.02	0.15
2 nd person pronouns	-0.06	-0.03	-0.37	0.02
Proverb 'do'	0.21	-0.03	-0.24	-0.10
Demonstrative pronouns	0.54	0.29	0.06	-0.15
Emphatics	0.38	-0.19	-0.30	-0.07
1 st person pronouns	0.35	0.08	-0.61	0.25
Pronoun 'it'	0.37	-0.09	-0.23	0.10
Verb 'be'	0.05	0.05	-0.14	0.02
Causative subordination	0.31	-0.03	-0.22	-0.06
Discourse particles	0.01	-0.06	-0.15	0.05
Nominal pronouns	0.24	-0.40	0.24	0.13
Hedges	0.43	-0.34	0.24	-0.08
Amplifiers	0.21	0.01	0.03	-0.02
'Wh' questions	0.05	0.04	-0.09	0.10
Possibility modals	-0.06	-0.03	-0.09	0.28

Clausal coordination	0.54	-0.34	0.25	-0.05
'Wh' clauses (verb complements)	-0.03	0.00	-0.06	0.24
Stranded prepositions	0.45	-0.07	-0.10	-0.11
Attributive adjectives	-0.26	0.26	0.38	-0.21
Past tense	-0.21	0.11	-0.24	0.37
3 rd person pronouns	0.06	-0.44	0.52	0.36
Perfect aspect	0.08	0.29	-0.11	0.13
'Wh' relative clauses on object position	0.12	0.11	0.17	-0.03
'Wh' relative clauses on subject position	-0.03	0.10	0.28	-0.01
Phrasal coordination	-0.02	0.04	0.17	0.06
Nominalizations	-0.32	0.56	0.17	-0.19
Time adverbials	0.13	-0.07	-0.31	0.03
Place adverbials	-0.12	-0.14	-0.21	-0.11
Predictive modals	0.28	0.18	-0.31	0.04
Conditional subordination	0.04	0.01	-0.26	0.11
Necessity modals	-0.04	0.03	-0.15	0.01
Split auxiliaries	0.32	0.14	0.03	0.01
Conjuncts	0.63	0.03	0.09	-0.13
Agentless passive	-0.03	0.34	0.25	-0.01
Agentive passive	-0.07	0.35	0.17	-0.04
(Other) adverbial subordination	0.07	-0.07	0.04	0.09
Downtoners	0.24	0.09	-0.08	-0.16
Predicative adjectives	0.18	0.41	0.05	-0.01
Verb 'have'	0.23	0.00	-0.38	-0.13
Progressive aspect	0.13	-0.43	0.53	0.25
Attitudinal adjectives	0.11	0.16	-0.14	0.04
Epistemic adjectives	0.08	0.20	-0.05	0.07
Premodifying nouns	-0.50	-0.04	0.00	-0.40
'That' verb complement clauses controlled by stance verbs	0.04	0.07	0.19	0.76
'To' verb complement clauses controlled by stance verbs	-0.01	-0.04	0.04	0.44
Abstract nouns	-0.11	0.34	0.34	0.00

Concrete nouns	-0.03	-0.45	0.26	-0.06
Activity verbs	0.10	-0.18	-0.21	0.24
Type-token ratio	0.42	0.51	0.47	-0.30
Word length	-0.46	0.38	0.22	-0.33
Word count	0.59	0.35	0.06	-0.24
Existential 'there'	0.13	-0.12	0.36	-0.10
(Other) adverbs	0.30	-0.02	-0.09	-0.04
Public verbs	-0.12	0.11	-0.02	0.56
Private verbs	0.24	-0.01	-0.06	0.38
Suasive verbs	-0.11	0.19	0.01	0.02
'Seem'/'appear'	0.06	-0.34	0.52	0.18
Verb complement clauses controlled by public verbs	-0.09	0.12	0.08	0.32
Verb complement clauses controlled by private verbs	0.06	0.17	0.08	0.26
'That' deletion controlled by public verbs	-0.15	0.01	-0.02	0.48
'That' deletion controlled by private verbs	0.19	-0.04	-0.08	0.28
Adjective complements	-0.03	0.35	-0.11	0.14
'That' relative clauses on subject position	0.05	0.09	0.17	-0.02
'That' relative clauses on object position	0.10	0.11	0.04	0.13
Concessive subordination	0.00	-0.05	0.02	0.00
'Of' genitive	-0.21	0.48	0.17	-0.18
Prepositional phrases (noun + preposition)	-0.08	0.47	0.11	-0.02
Other prepositions	0.08	-0.02	0.10	-0.18
Adjective + prepositional complement	0.04	0.24	-0.01	-0.01
'Ed' reduced relative clauses	-0.02	0.30	0.15	0.00
'Wh' prepositional complement	0.10	0.00	0.04	0.03
'Ing' verb complement clauses	-0.04	-0.04	-0.14	-0.01
'To' adverbial clauses	-0.09	0.13	0.01	0.12
'Ing' adverbial clauses	-0.01	0.19	0.14	-0.05

'To' relative clauses	0.04	0.06	-0.02	-0.11
'Ing' relative clauses	0.00	-0.06	0.34	0.05
's' genitive	-0.19	-0.05	-0.03	-0.20
'That' noun complement clauses	0.09	0.35	0.13	0.04
'To' noun complement clauses	-0.01	0.41	0.09	0.05
'wh' relative clauses – pied piping	-0.03	0.28	0.06	0.05

Appendix 3. Social groups across registers

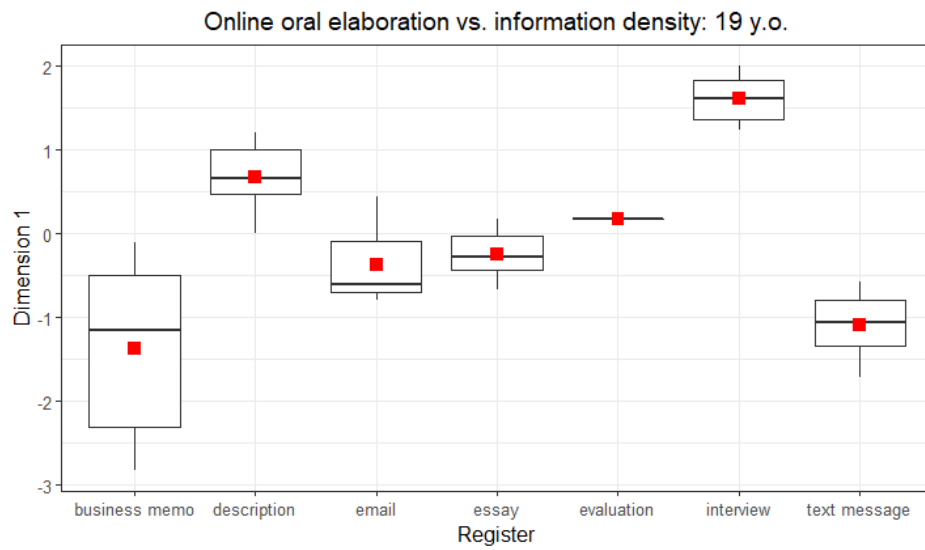
Registers within Age groups on Dimension 1. Descriptive statistics

Register	19 y. o.		20 y. o.		21 y. o.		22 y. o.		23 y. o.	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-1.37	1.15	-1.16	0.83	-0.94	0.82	-0.85	0.63	-1.11	0.74
Description	0.68	0.44	0.42	0.48	0.39	0.53	0.47	0.49	0.4	0.67
Emails	-0.37	0.51	0.07	0.3	0.04	0.36	0.01	0.38	-0.07	0.43
Essays	-0.24	0.32	-0.25	0.29	-0.24	0.24	-0.25	0.31	-0.19	0.15
Evaluations	0.18	0.01	0.28	0.22	0.11	0.25	0.13	0.26	0.12	0.07
Interviews	1.61	0.31	1.62	0.21	1.65	0.22	1.62	0.21	1.57	0.19
Text messages	-1.09	0.43	-0.93	0.61	-0.97	0.76	-0.73	0.6	-1.01	0.7

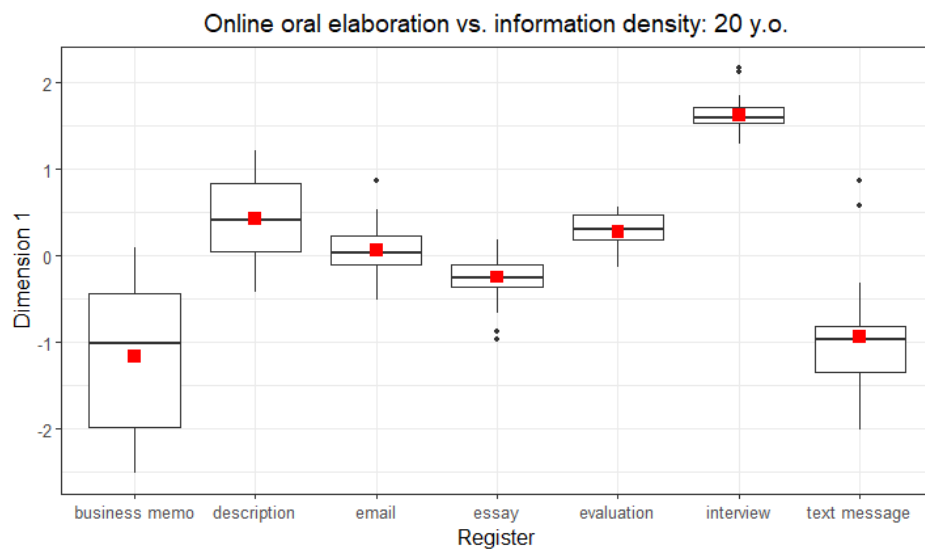
Variance explained by Register within age groups on Dimension 1

Age group	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
19 y. o.	6	17.98	<0.001	77%
20 y. o.	6	100.3	<0.001	77%
21 y. o.	6	109.4	<0.001	73%
22 y. o.	6	109.5	<0.001	76%
23 y. o.	6	23.01	<0.001	76%

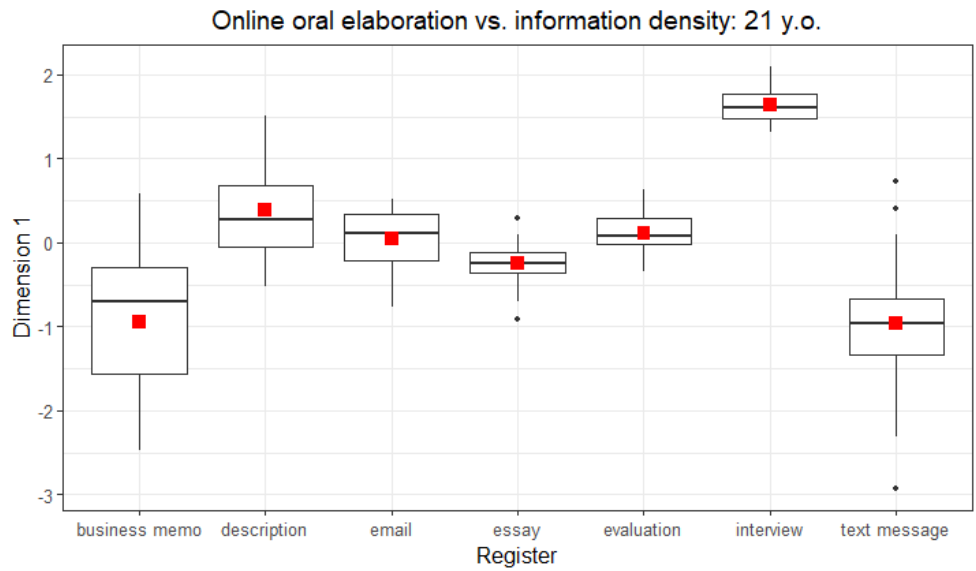
Register differences within the 19-year-old group



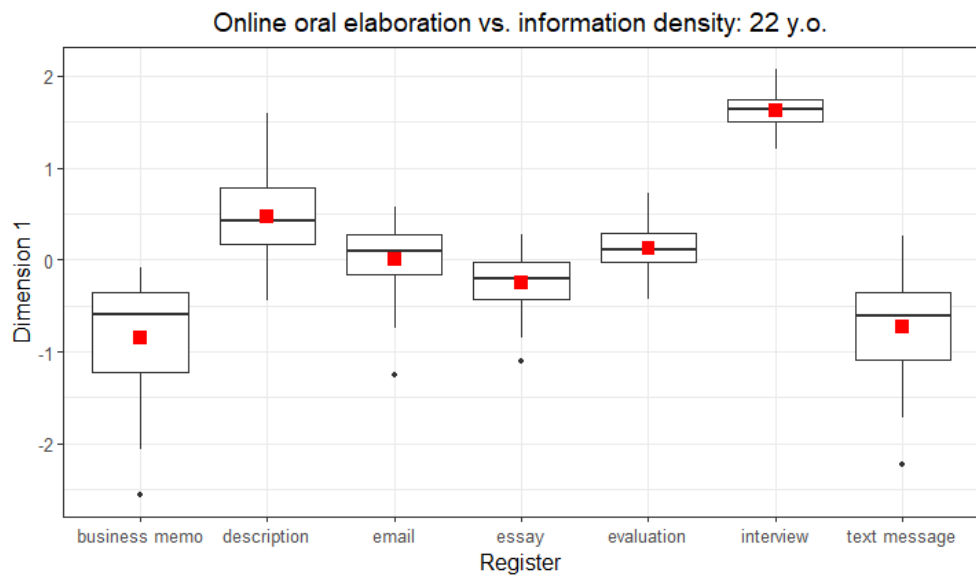
Register differences within the 20-year-old group



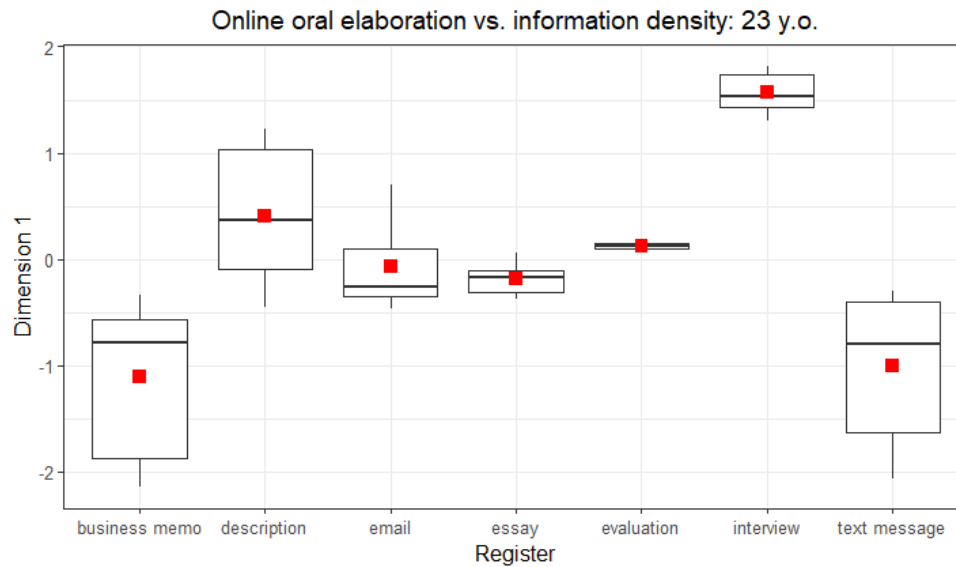
Register differences within the 21-year-old group



Register differences within the 22-year-old group



Register differences within the 23-year-old group



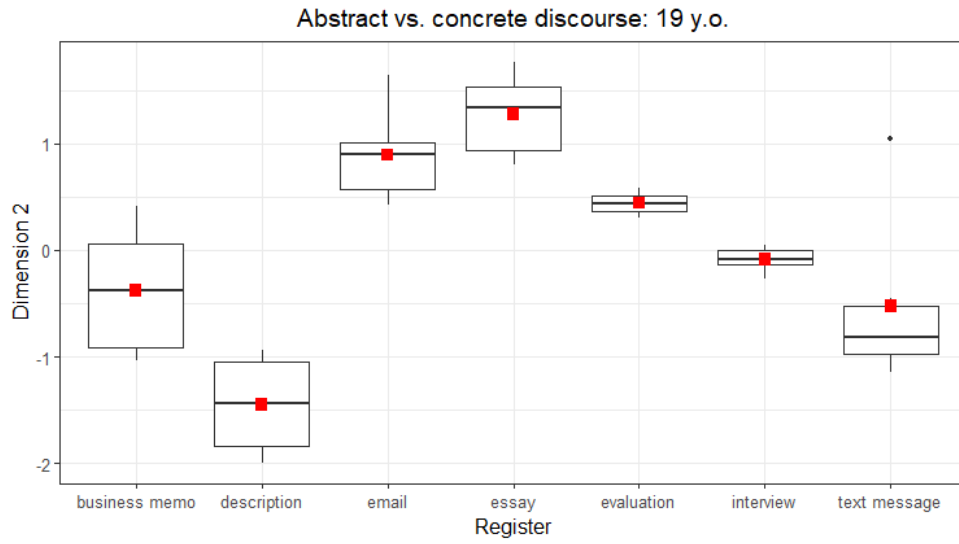
Registers within Age groups on Dimension 2. Descriptive statistics

Register	19 y. o.		20 y. o.		21 y. o.		22 y. o.		23 y. o.	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.38	0.61	-0.4	0.57	-0.25	0.52	-0.21	0.55	-0.34	0.39
Description	-1.45	0.48	-1.26	0.38	-1.32	0.33	-1.32	0.49	-1.26	0.35
Emails	0.89	0.44	0.79	0.51	0.63	0.51	0.66	0.45	0.75	0.47
Essays	1.28	0.39	1.28	0.32	1.35	0.3	1.26	0.22	1.32	0.3
Evaluations	0.44	0.2	0.55	0.35	0.57	0.29	0.59	0.39	0.69	0.27
Interviews	-0.09	0.12	-0.06	0.17	-0.04	0.18	-0.07	0.17	-0.01	0.09
Text messages	-0.53	0.81	-0.97	0.48	-0.6	0.79	-0.52	0.64	-0.6	0.77

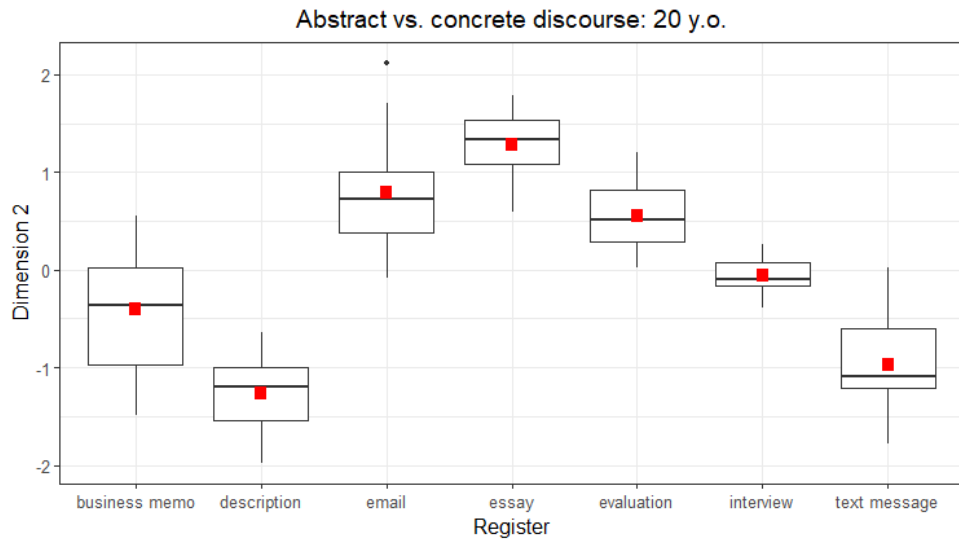
Variance explained by Register within age groups on Dimension 2

Age group	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
19 y. o.	6	19.22	<0.001	79%
20 y. o.	6	133.3	<0.001	82%
21 y. o.	6	131.9	<0.001	76%
22 y. o.	6	113.8	<0.001	77%
23 y. o.	6	30.38	<0.001	81%

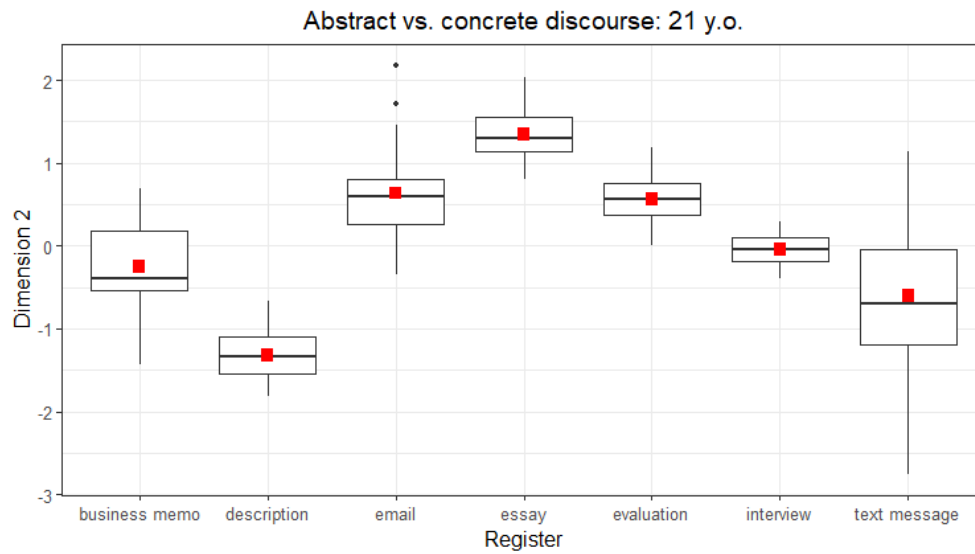
Register differences within the 19-year-old group



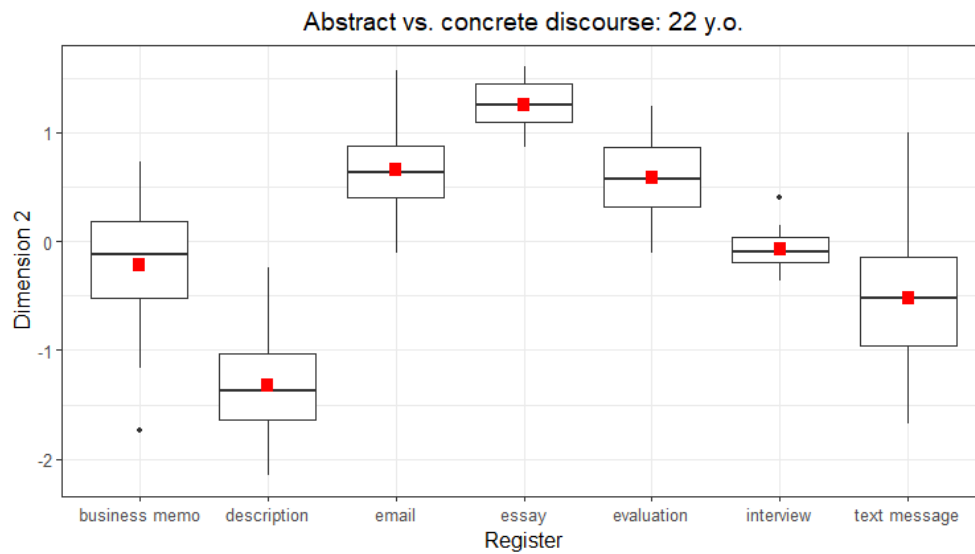
Register differences within the 20-year-old group



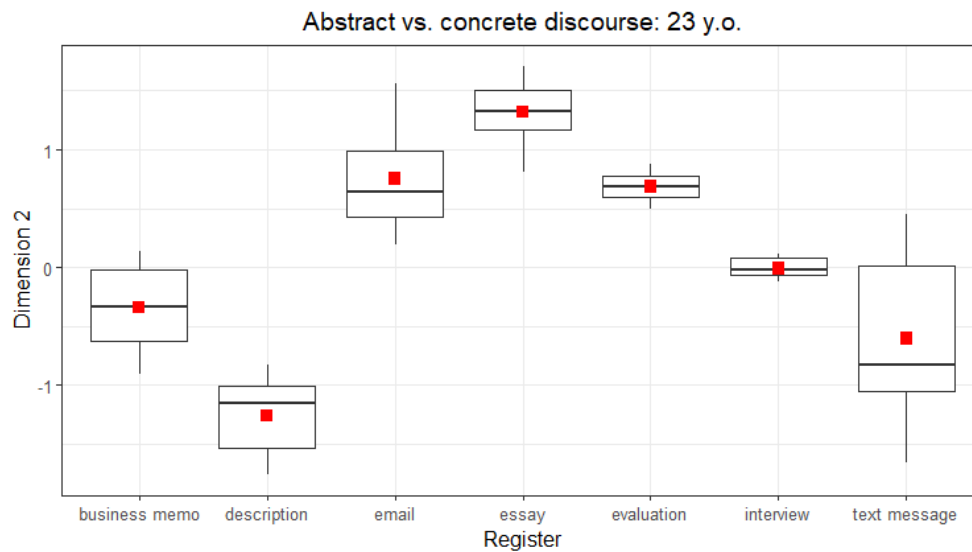
Register differences within the 21-year-old group



Register differences within the 22-year-old group



Register differences within the 23-year-old group



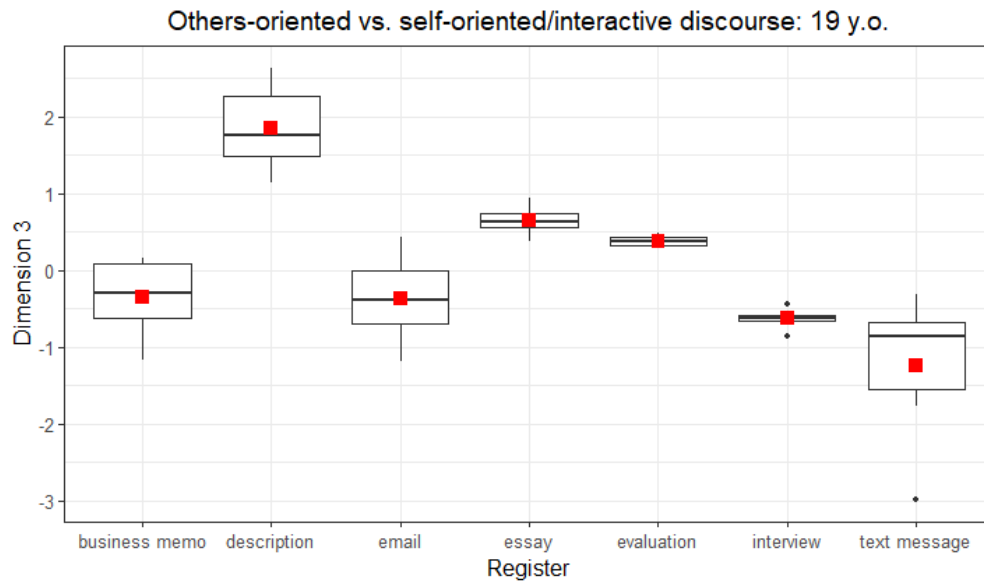
Registers within Age groups on Dimension 3. Descriptive statistics

Register	19 y. o.		20 y. o.		21 y. o.		22 y. o.		23 y. o.	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.36	0.52	-0.28	0.4	-0.17	0.47	-0.12	0.44	-0.09	0.41
Description	1.85	0.57	1.55	0.27	1.46	0.45	1.69	0.42	1.54	0.37
Emails	-0.37	0.58	-0.4	0.52	-0.51	0.44	-0.4	0.43	-0.36	0.22
Essays	0.65	0.2	0.69	0.23	0.7	0.31	0.7	0.22	0.5	0.22
Evaluations	0.38	0.14	0.25	0.25	0.06	0.26	0.05	0.23	0.06	0.07
Interviews	-0.63	0.13	-0.71	0.22	-0.64	0.16	-0.67	0.18	-0.61	0.23
Text messages	-1.24	0.99	-1.05	0.9	-0.96	0.94	-1.15	0.85	-0.68	1.11

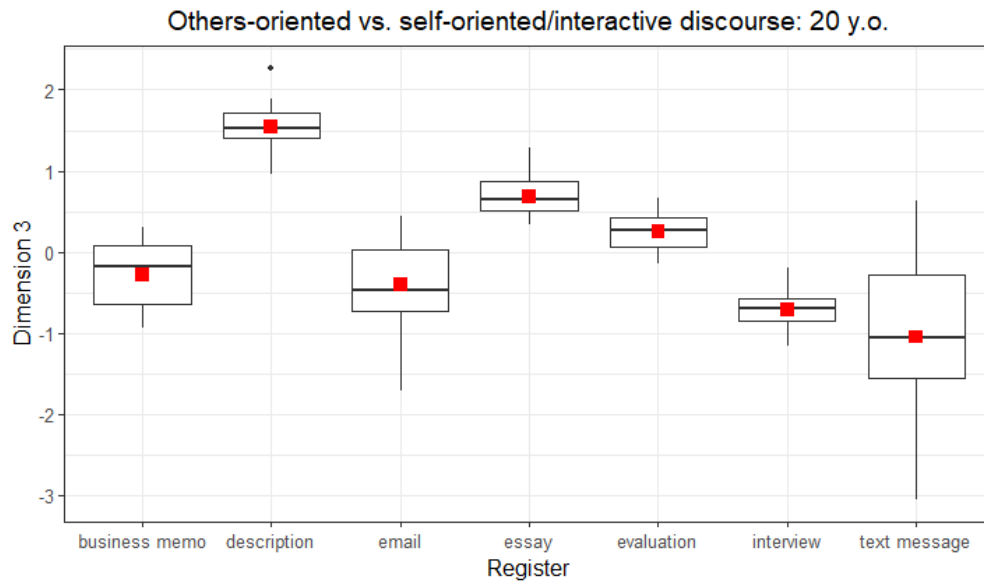
Variance explained by Register within age groups on Dimension 3

Age group	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
19 y. o.	6	19.13	<0.001	79%
20 y. o.	6	99.17	<0.001	77%
21 y. o.	6	105.7	<0.001	72%
22 y. o.	6	132.6	<0.001	79%
23 y. o.	6	17.47	<0.001	71%

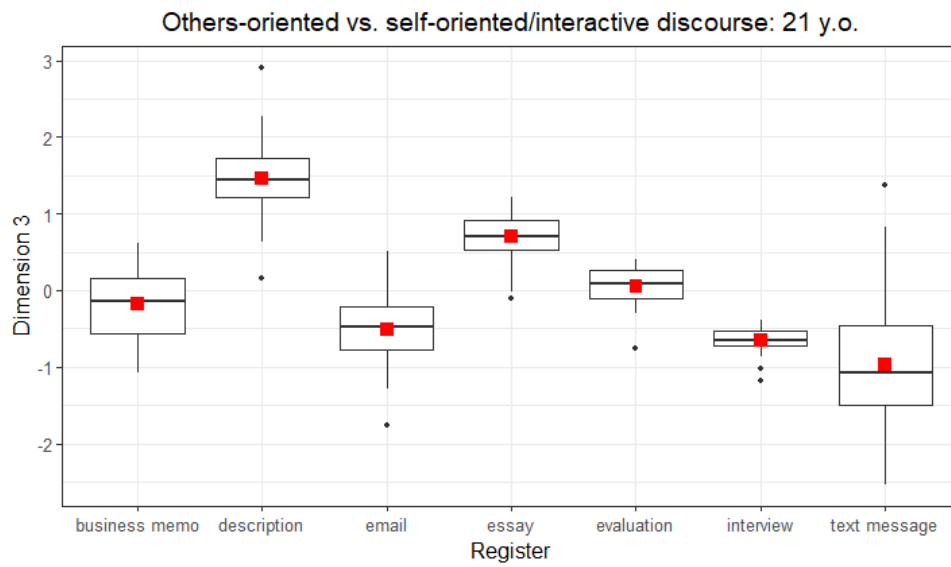
Register differences within the 19-year-old group



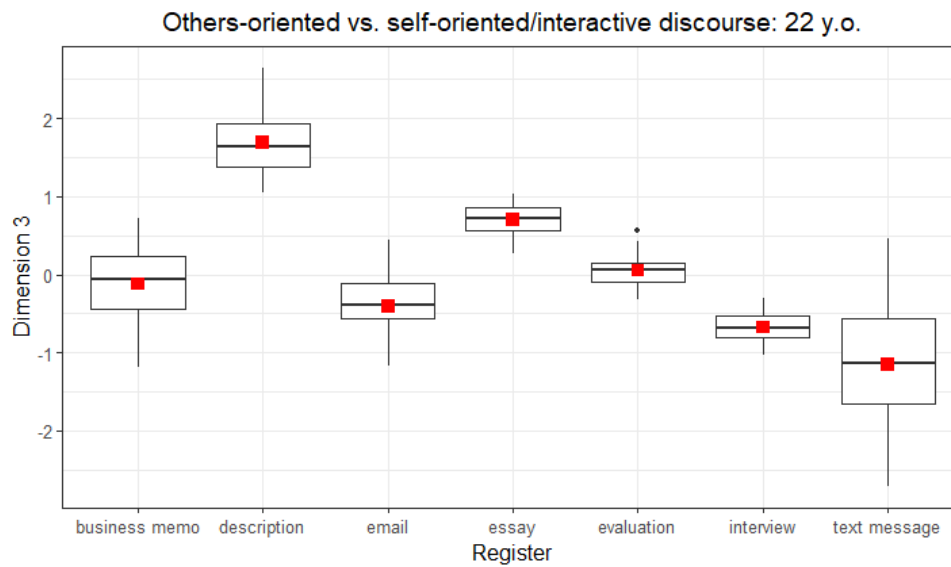
Register differences within the 20-year-old group



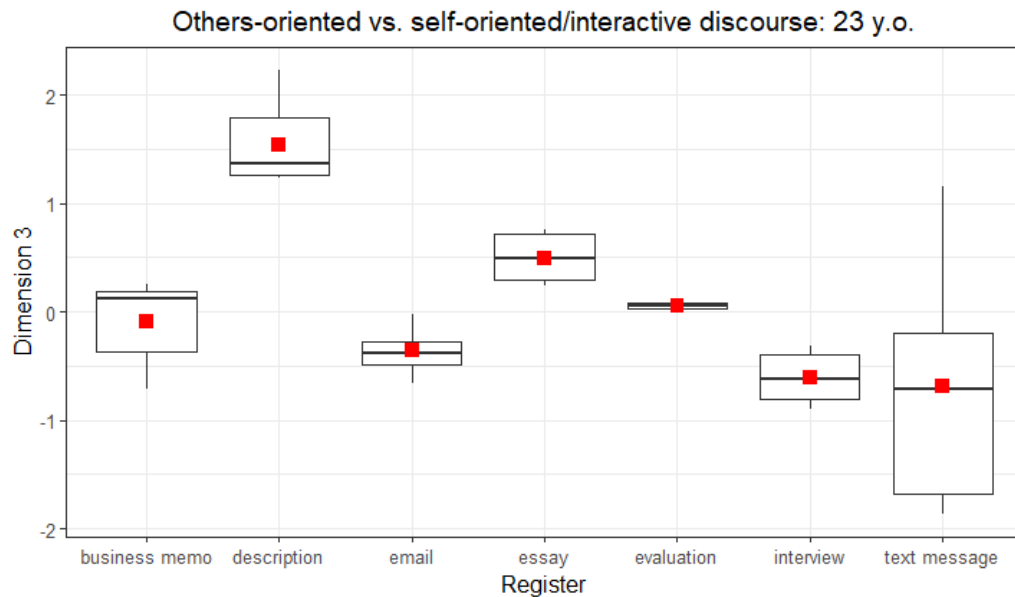
Register differences within the 21-year-old group



Register differences within the 22-year-old group



Register differences within the 23-year-old group



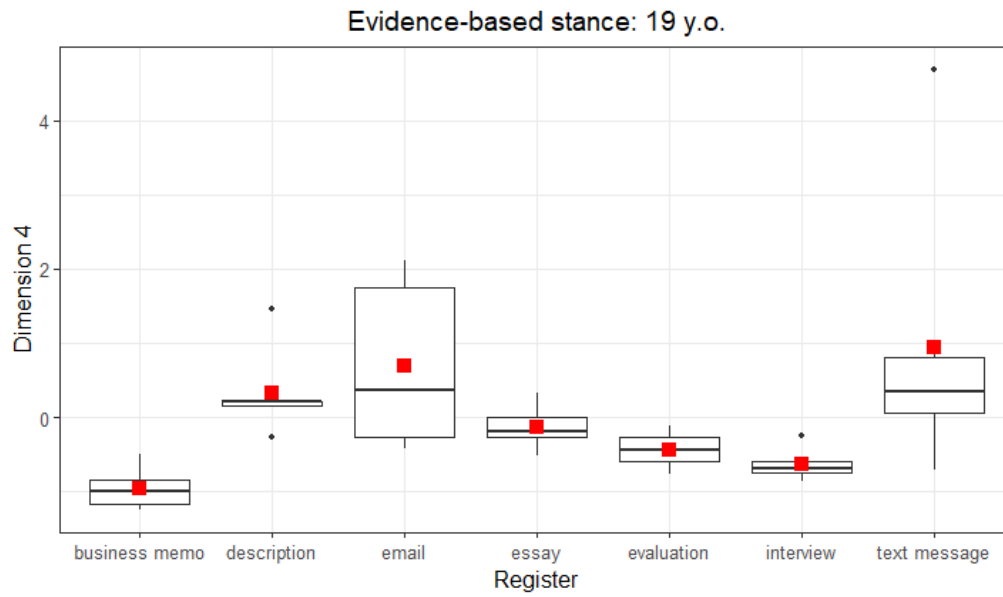
Registers within Age groups on Dimension 4. Descriptive statistics

Register	19 y. o.		20 y. o.		21 y. o.		22 y. o.		23 y. o.	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.96	0.28	-0.97	0.48	-0.81	0.44	-0.73	0.48	-1.04	0.33
Description	0.33	0.59	0.27	0.49	0.17	0.52	0.41	0.67	0.27	0.79
Emails	0.69	1.17	0.49	0.7	0.33	0.5	0.34	0.71	0.2	0.34
Essays	-0.13	0.29	0.1	0.39	0	0.52	-0.11	0.35	-0.2	0.43
Evaluations	-0.44	0.46	-0.38	0.47	-0.19	0.33	-0.33	0.38	-0.23	0.31
Interviews	-0.63	0.21	-0.65	0.24	-0.52	0.29	-0.6	0.25	-0.66	0.21
Text messages	0.94	1.92	0.41	1.08	1.25	1.91	0.93	1.5	1.75	1.97

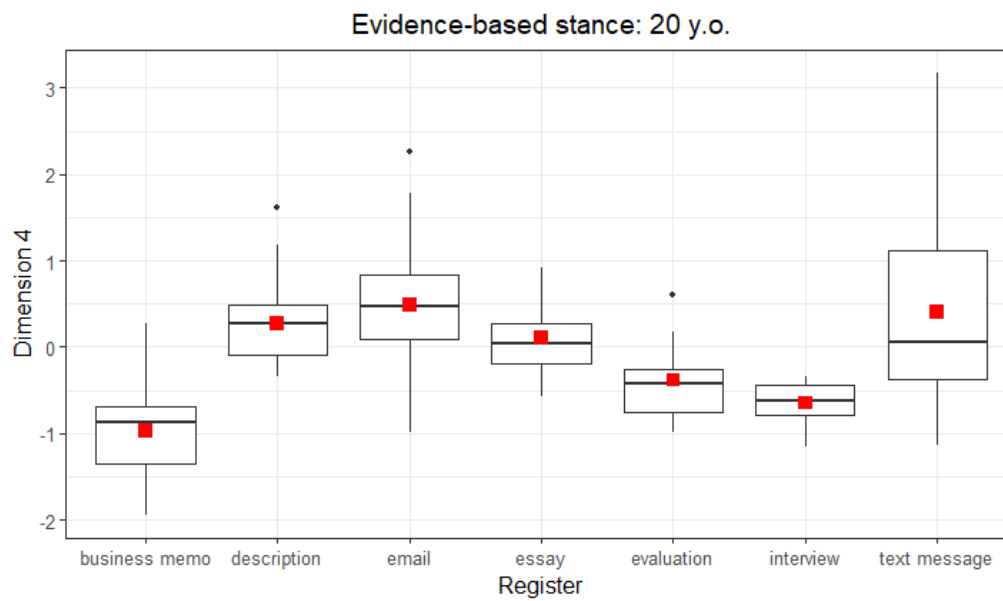
Variance explained by Register within age groups on Dimension 4

Age group	<i>df</i>	<i>F</i>	<i>P</i>	<i>R</i> ²
19 y. o.	6	3.14	>0.001	38%
20 y. o.	6	23.24	<0.001	44%
21 y. o.	6	22.96	<0.001	36%
22 y. o.	6	19.86	<0.001	37%
23 y. o.	6	7.81	<0.001	52%

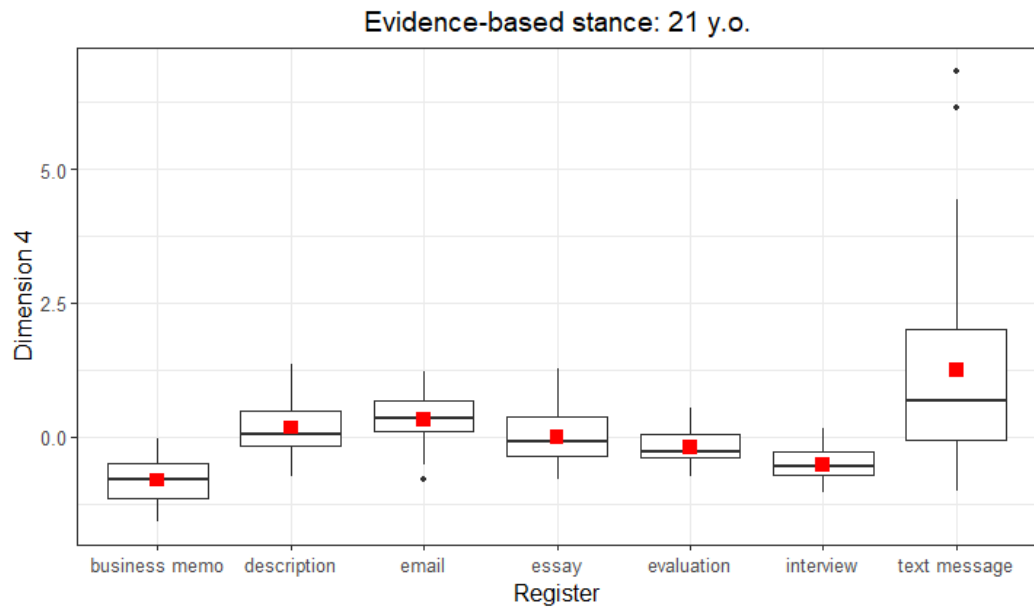
Register differences within the 19-year-old group



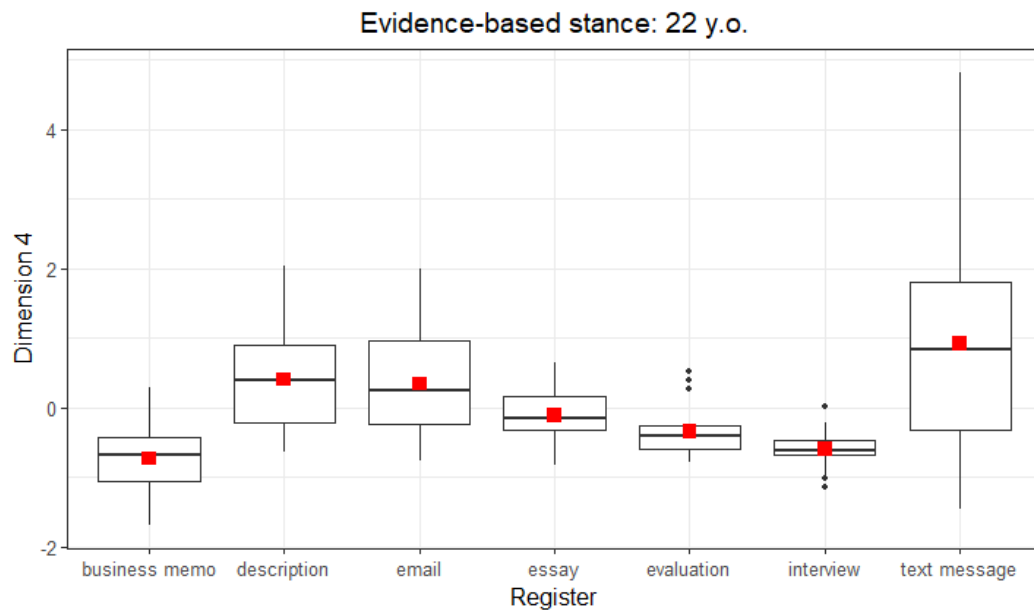
Register differences within the 20-year-old group



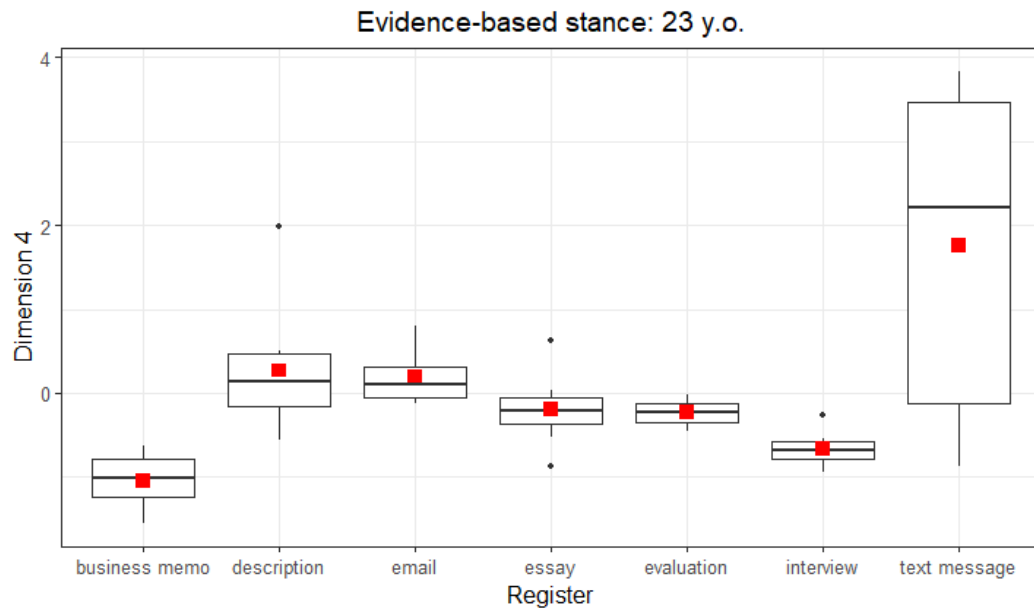
Register differences within the 21-year-old group



Register differences within the 22-year-old group



Register differences within the 23-year-old group



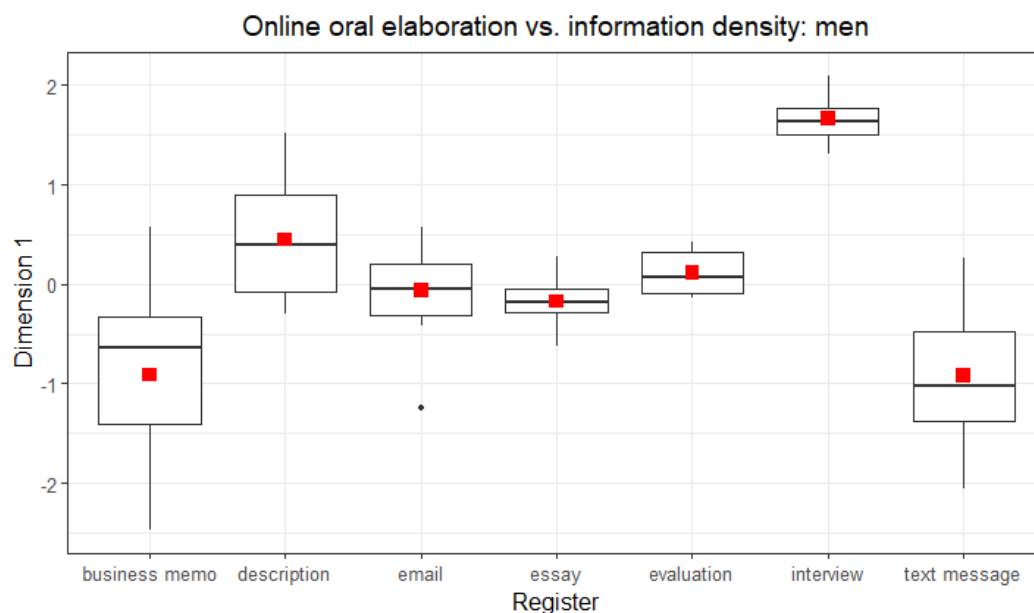
Genders across registers. Dimension 1

Register	Men		Women	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.91	0.77	-1.04	0.79
Description	0.45	0.58	0.43	0.49
Emails	-0.06	0.41	0.02	0.37
Essays	-0.17	0.21	-0.26	0.28
Evaluations	0.12	0.21	0.15	0.25
Interviews	1.66	0.22	1.62	0.22
Text messages	-0.92	0.63	-0.89	0.66

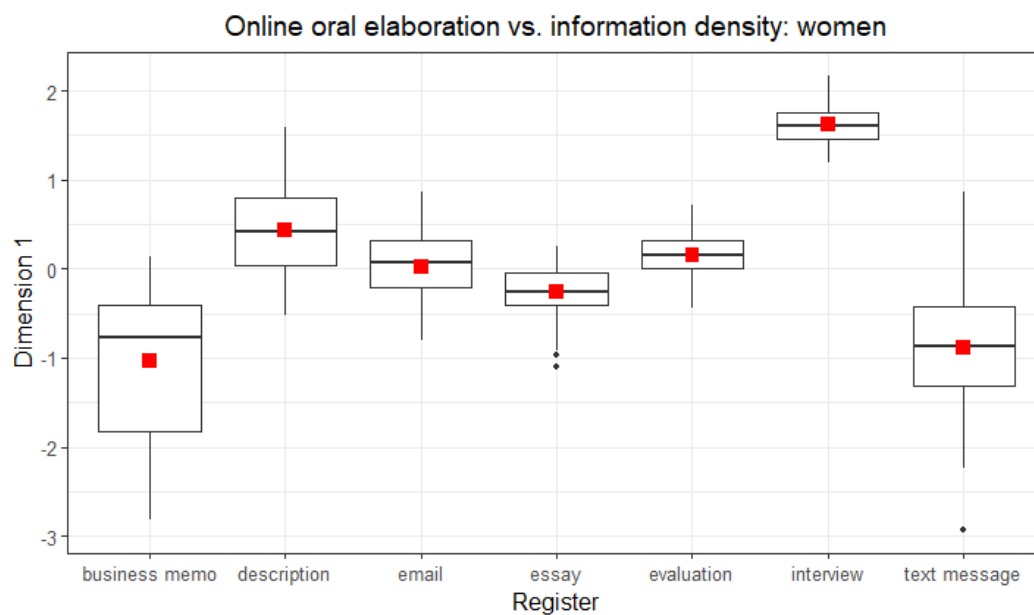
Variance explained by Register within Gender group. Dimension 1

Dimension	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Men	6	63.92	<0.001	75%
Women	6	291.5	<0.001	75%

Register differences within the male group



Register differences within the female group



Genders across registers. Dimension 2

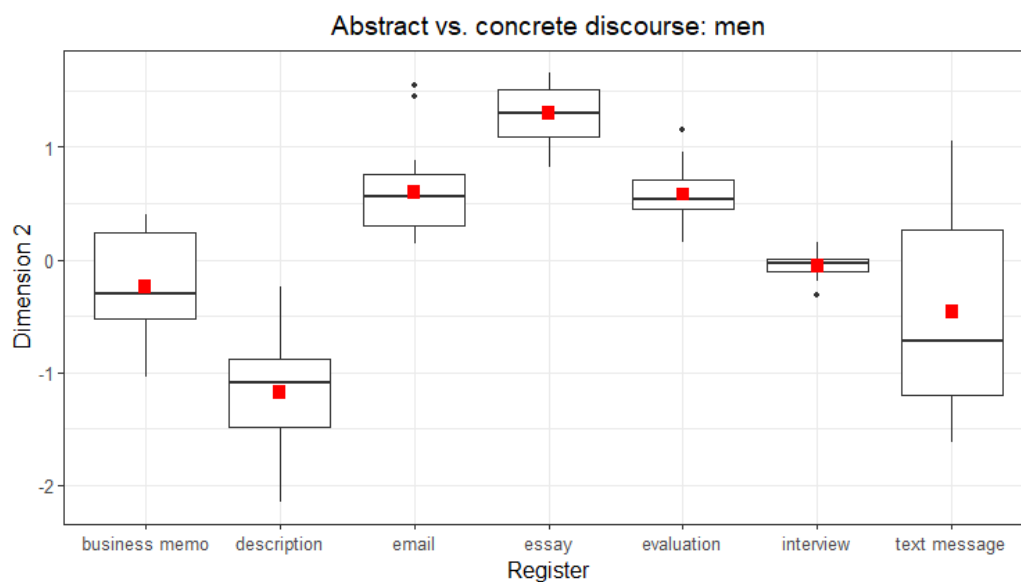
Register	Men		Women	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.24	0.47	-0.31	0.55
Description	-1.18	0.45	-1.34	0.37
Emails	0.6	0.38	0.72	0.5
Essays	1.3	0.26	1.3	0.3

Evaluations	0.58	0.29	0.59	0.32
Interviews	-0.05	0.12	-0.06	0.17
Text messages	-0.46	0.87	-0.72	0.65

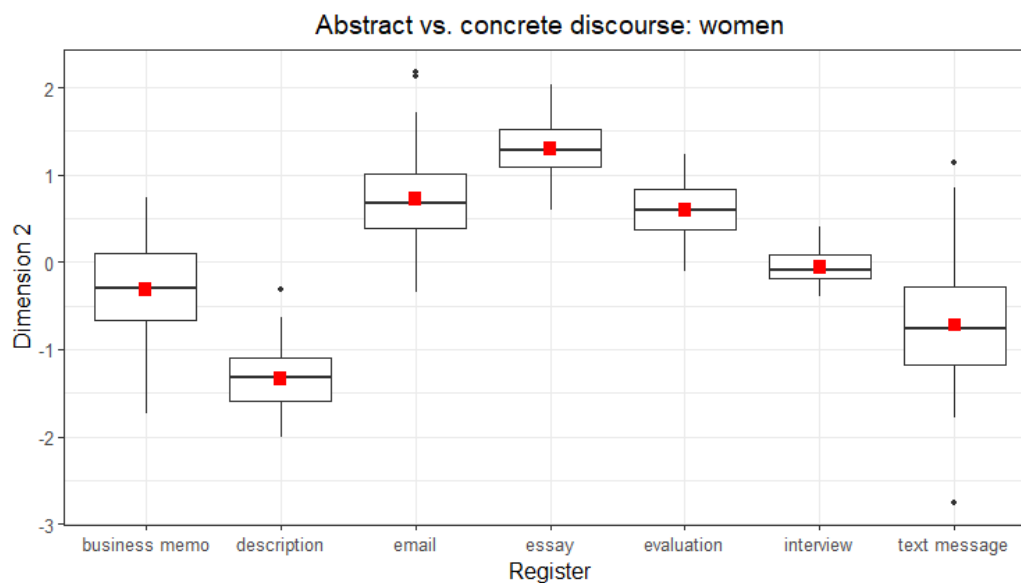
Variance explained by Register within Gender group. Dimension 2

Dimension	df	F	p	R ²
Men	6	58.28	<0.001	73%
Women	6	366.5	<0.001	79%

Register differences within the male group



Register differences within the female group



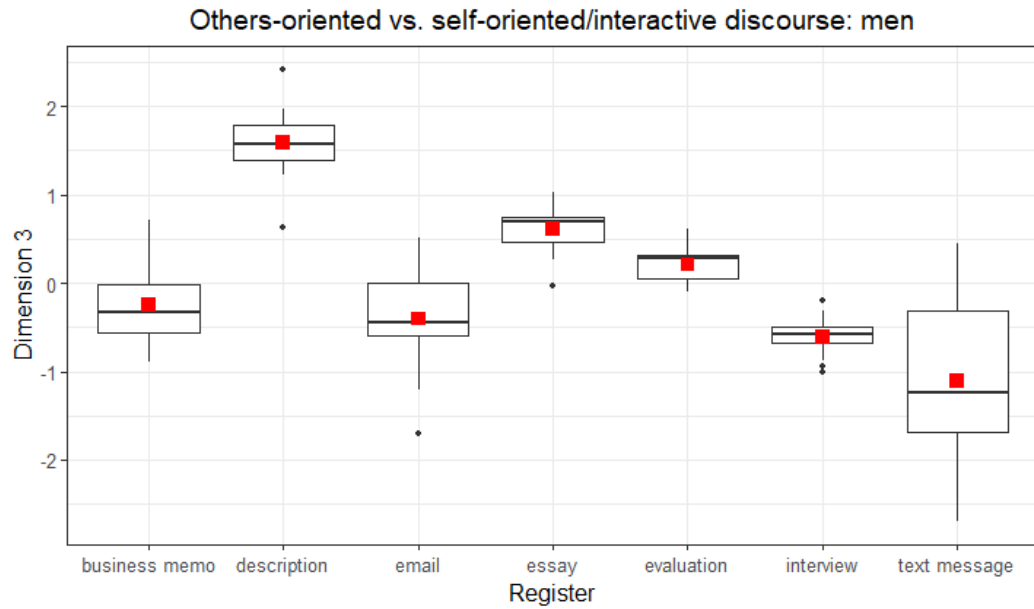
Genders across registers. Dimension 3

Register	Men		Women	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.25	0.41	-0.17	0.45
Description	1.59	0.37	1.57	0.43
Emails	-0.4	0.56	-0.44	0.42
Essays	0.62	0.27	0.69	0.25
Evaluations	0.21	0.2	0.09	0.26
Interviews	-0.61	0.22	-0.68	0.18
Text messages	-1.11	0.94	-1.02	0.91

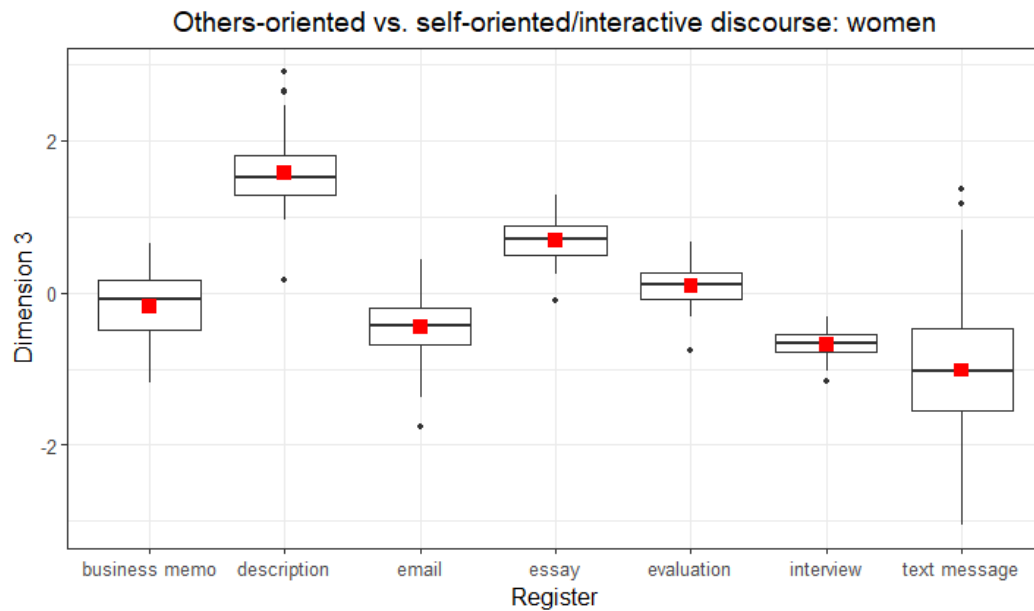
Variance explained by Register within Gender group. Dimension 3

Dimension	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Men	6	62.89	<0.001	75%
Women	6	299.0	<0.001	75%

Register differences within the male group



Register differences within the female group



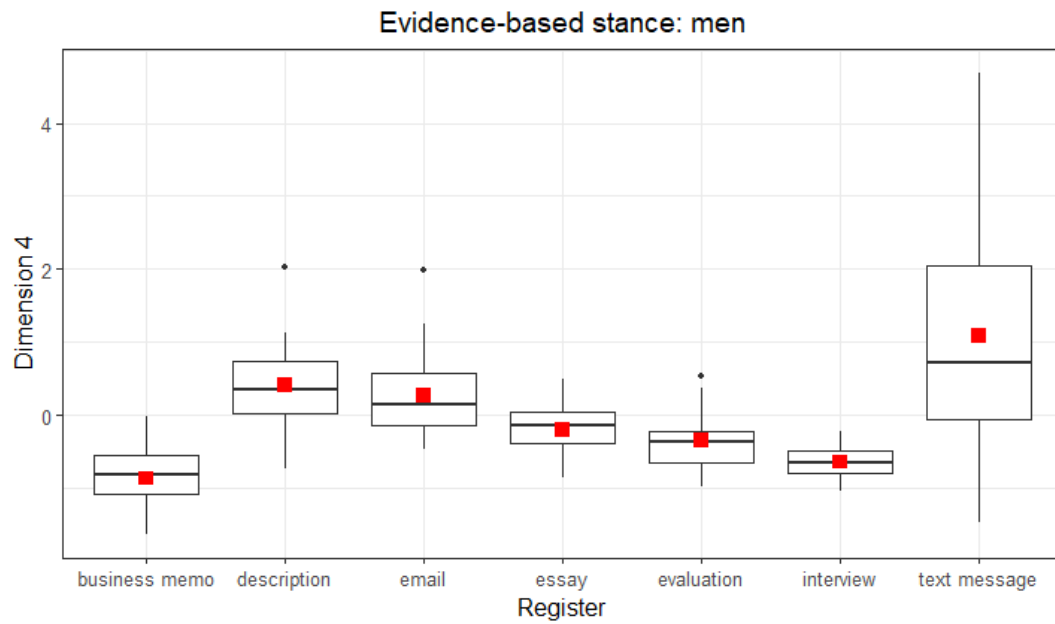
Genders across registers. Dimension 4

Register	Men		Women	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Business memos	-0.87	0.41	-0.84	0.46
Description	0.41	0.6	0.26	0.57
Emails	0.27	0.62	0.42	0.67
Essays	-0.2	0.36	0.02	0.44
Evaluations	-0.34	0.44	-0.25	0.36
Interviews	-0.64	0.24	-0.58	0.26
Text messages	1.09	1.62	0.93	1.66

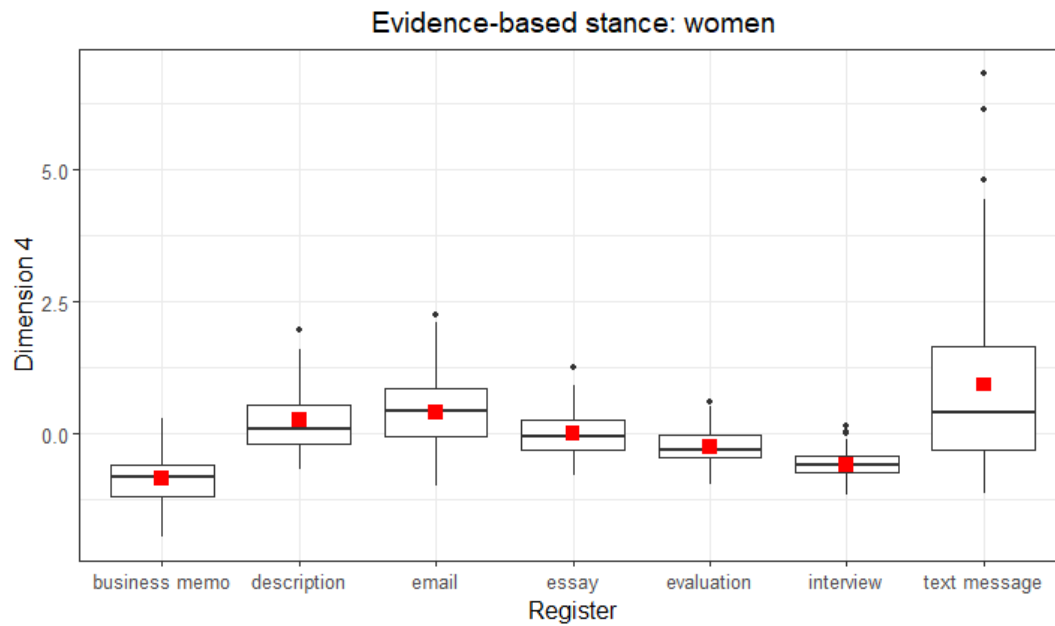
Variance explained by Register within Gender group. Dimension 4

Dimension	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Men	6	15.52	<0.001	42%
Women	6	53.17	<0.001	35%

Register differences within the male group



Register differences within the female group



Appendix 4. Social Groups within Registers

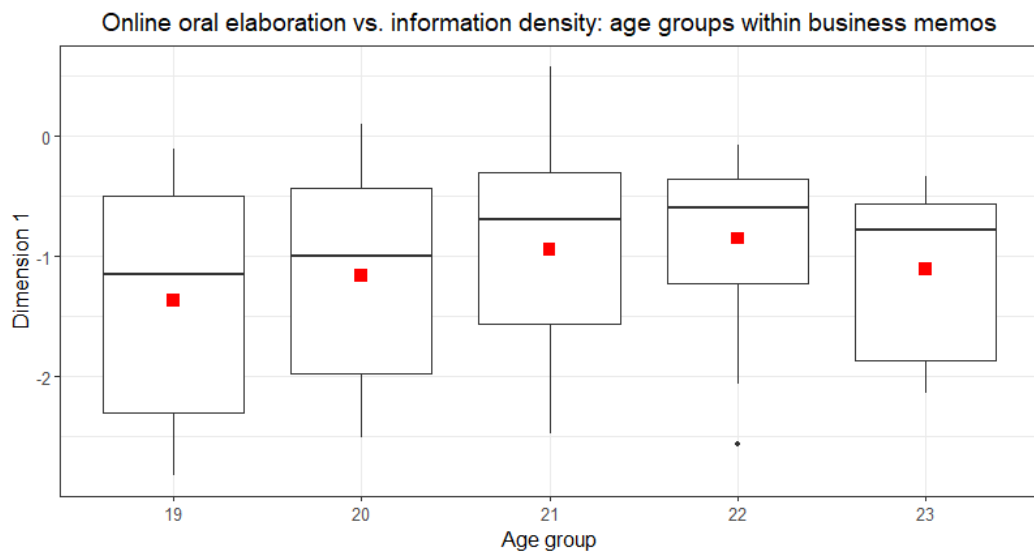
Age groups within registers. Dimension 1

Age	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
19	–	1.15	0.68	0.44	–	0.51	–	0.32	0.18	0.01	1.61	0.31	–	0.43
	1.37				0.37		0.24						1.09	
20	–	0.83	0.42	0.48	0.07	0.3	–	0.29	0.28	0.22	1.62	0.21	–	0.61
	1.16						0.25						0.93	
21	–	0.82	0.39	0.53	0.04	0.36	–	0.24	0.11	0.25	1.65	0.22	–	0.76
	0.94						0.24						0.97	
22	–	0.63	0.47	0.49	0.01	0.38	–	0.31	0.13	0.26	1.62	0.21	–	0.6
	0.85						0.25						0.73	
23	–	0.74	0.4	0.67	–	0.43	–	0.15	0.12	0.07	1.57	0.19	–	0.7
	1.11				0.07		0.19						1.01	

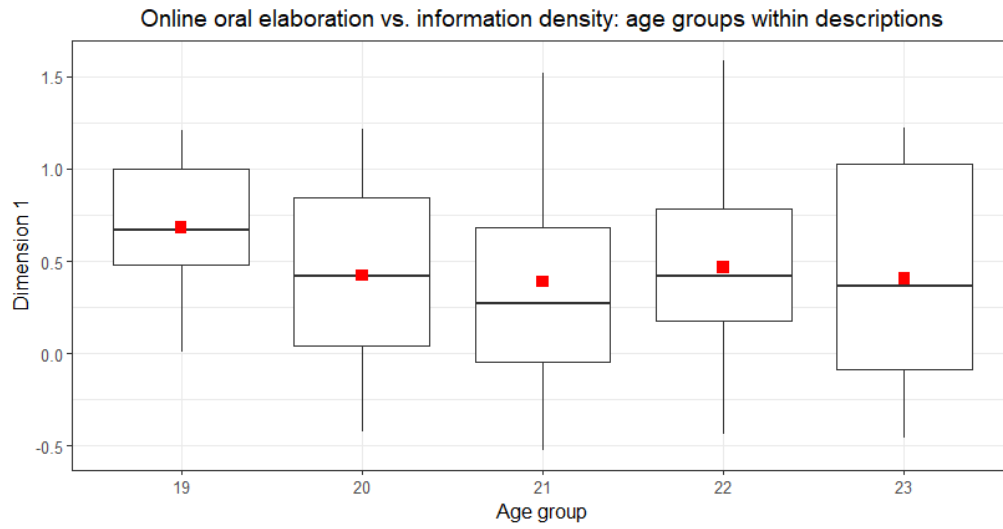
Variance explained by age within registers. Dimension 1

Register	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Business memos	4	0.99	>0.001	3%
Descriptions	4	0.46	>0.001	2%
Emails	4	1.91	>0.001	7%
Essays	4	0.09	>0.001	0.3%
Evaluations	4	1.12	>0.001	7%
Interviews	4	0.23	>0.001	0.8%
Text Messages	4	0.82	>0.001	3%

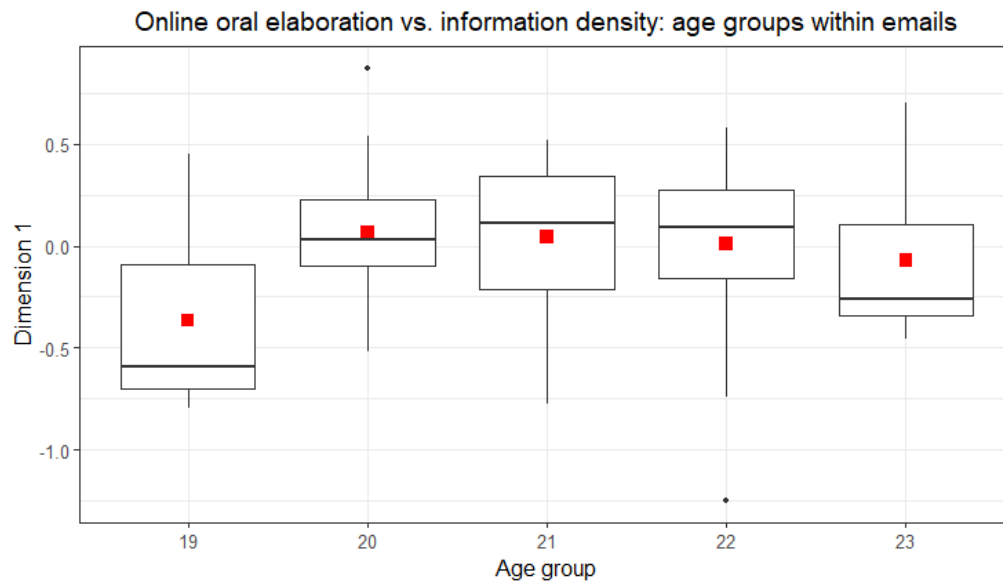
Age groups within business memos. Dimension 1



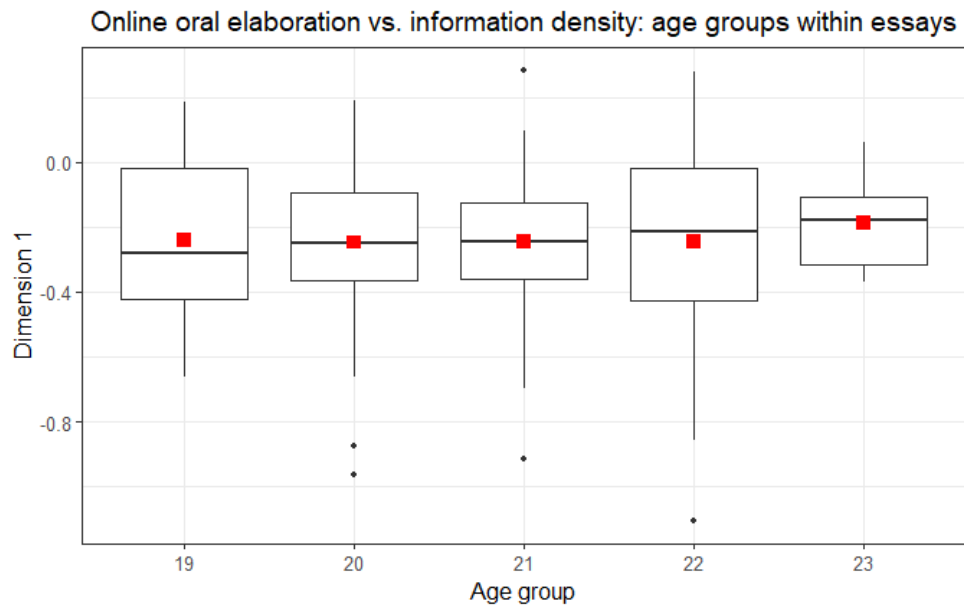
Age groups within descriptions. Dimension 1



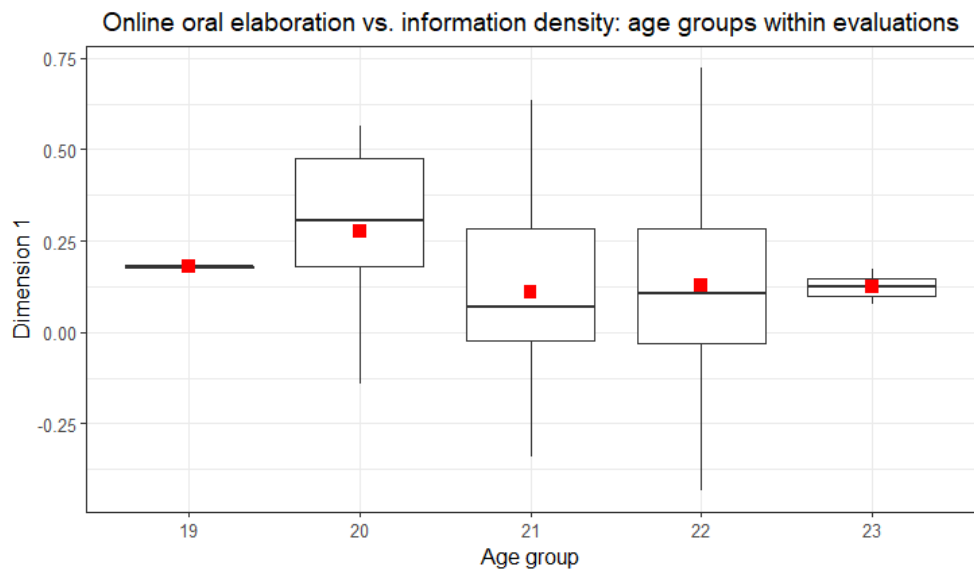
Age groups within emails. Dimension 1



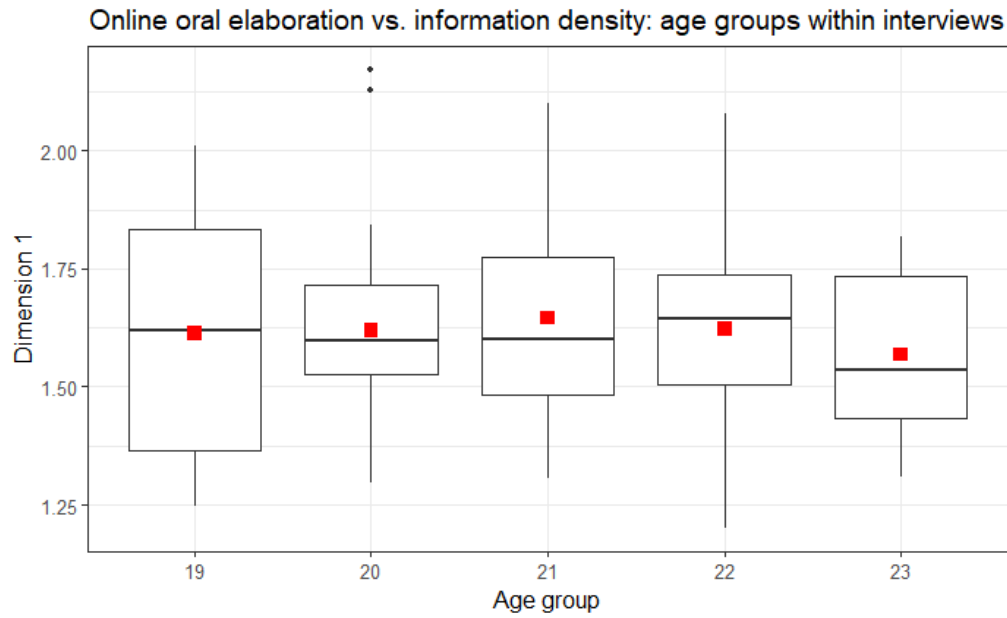
Age groups within essays. Dimension 1



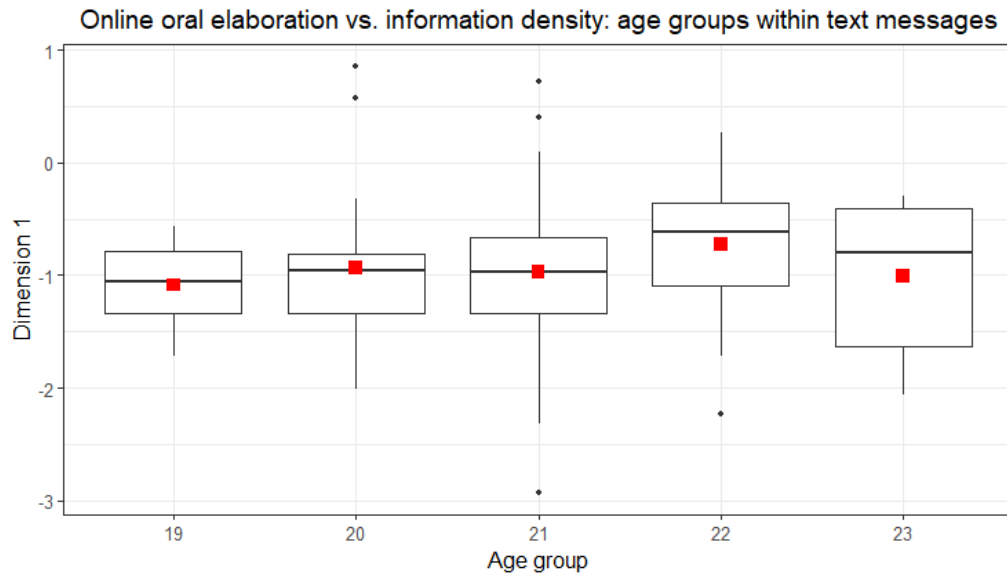
Age groups within evaluations. Dimension 1



Age groups within interviews. Dimension 1



Age groups within text messages. Dimension 1



Age groups within registers. Dimension 2

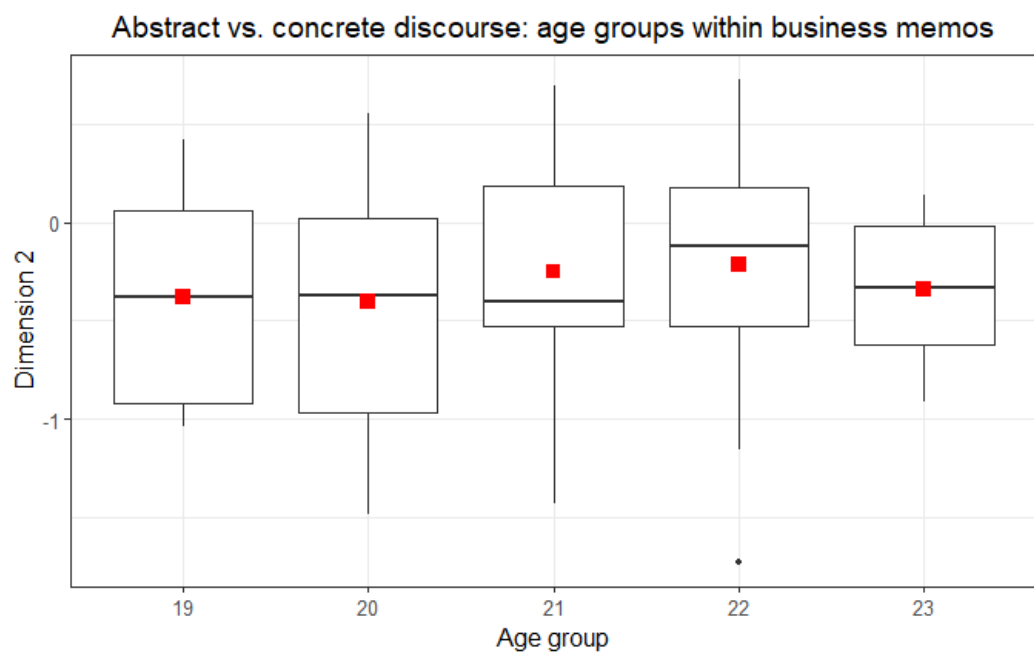
Age	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
19	–	0.61	–	0.48	0.89	0.44	1.28	0.39	0.44	0.2	–	0.12	–	0.81	
20	0.38	0.57	1.45	0.38	0.79	0.51	1.28	0.32	0.55	0.35	–	0.06	0.09	–	0.48
	–0.4		1.26								–		0.17		

21	–	0.52	–	0.33	0.63	0.51	1.35	0.3	0.57	0.29	–	0.18	–0.6	0.79
	0.25		1.32								0.04			
22	–	0.55	–	0.49	0.66	0.45	1.26	0.22	0.59	0.39	–	0.17	–	0.64
	0.21		1.32								0.07		0.52	
23	–	0.39	–	0.35	0.75	0.47	1.32	0.3	0.69	0.27	–	0.09	–0.6	0.77
	0.34		1.26								0.01			

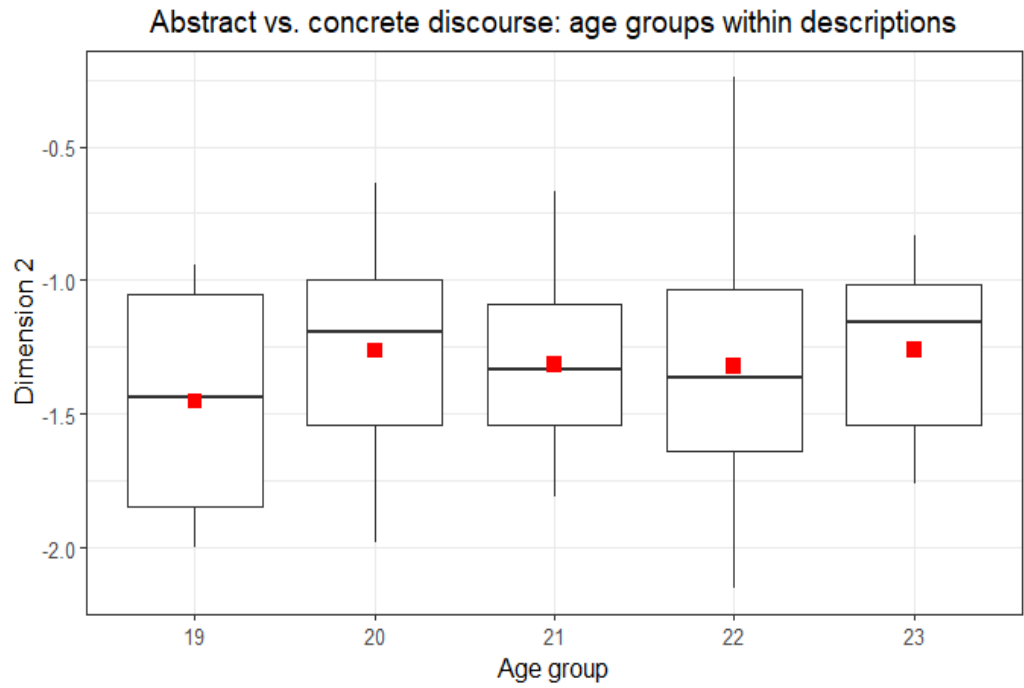
Variance explained by age within registers. Dimension 2

Register	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Business memos	4	0.57	>0.001	2%
Descriptions	4	0.32	>0.001	1.2%
Emails	4	0.72	>0.001	3%
Essays	4	0.48	>0.001	1.8%
Evaluations	4	0.15	>0.001	1%
Interviews	4	0.36	>0.001	1.3%
Text Messages	4	1.94	>0.001	7%

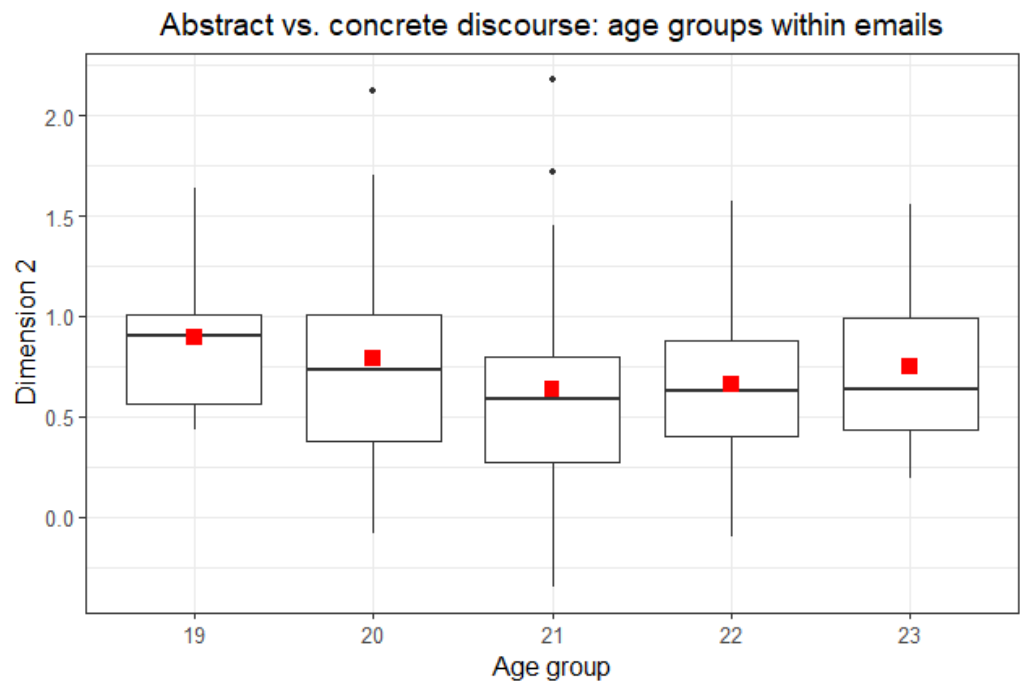
Age groups within business memos. Dimension 2



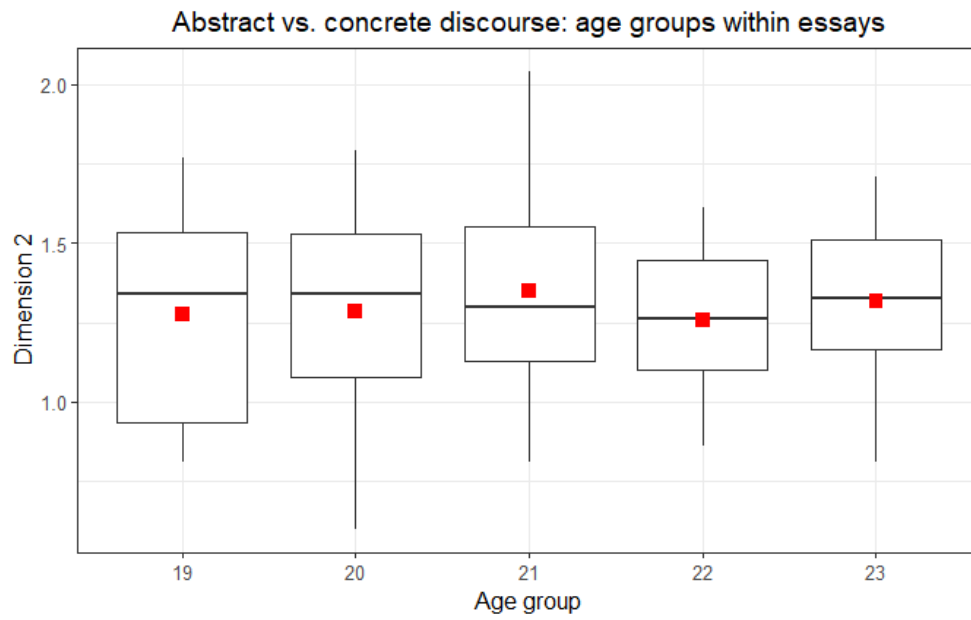
Age groups within descriptions. Dimension 2



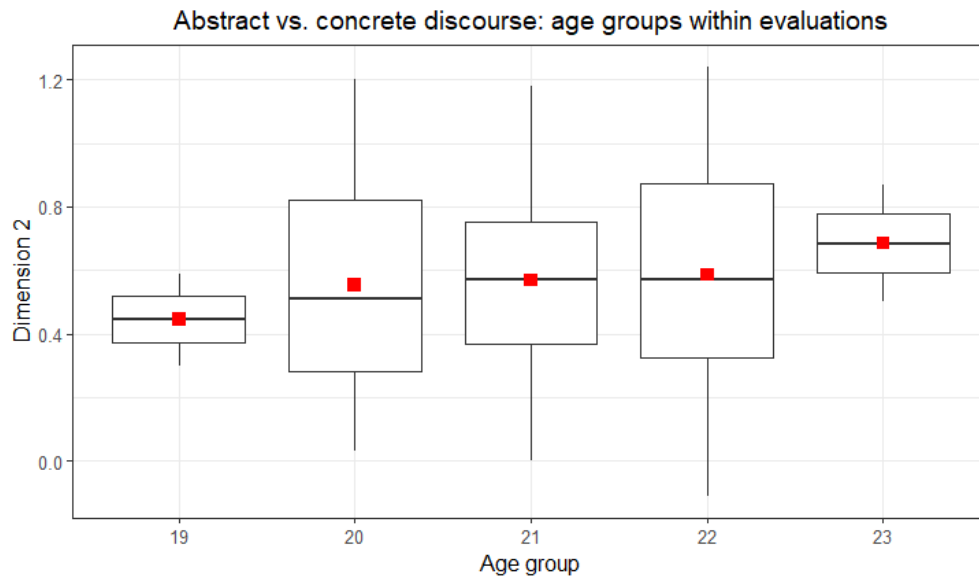
Age groups within emails. Dimension 2



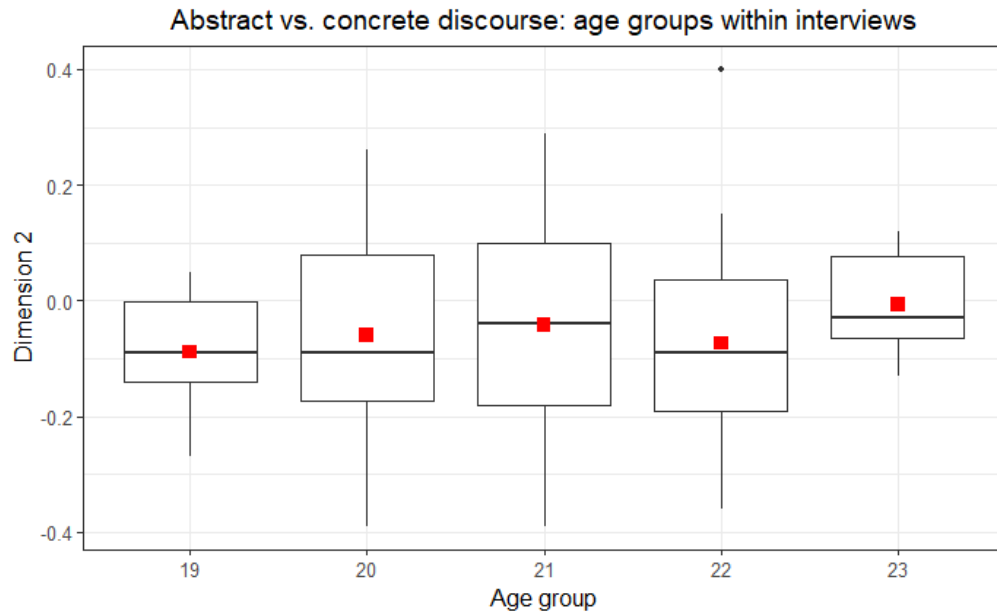
Age groups within essays. Dimension 2



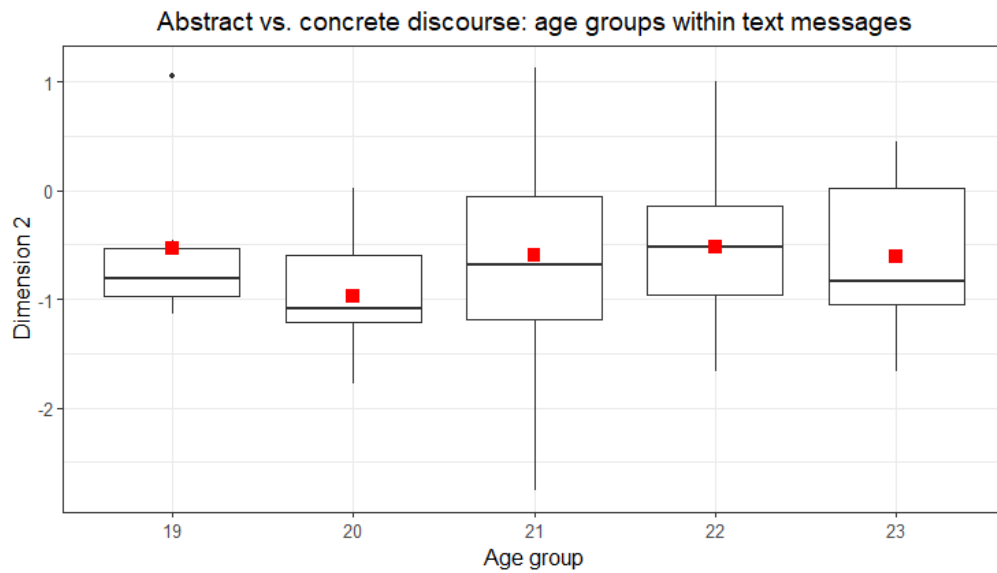
Age groups within evaluations. Dimension 2



Age groups within interviews. Dimension 2



Age groups within text messages. Dimension 2



Age groups within registers. Dimension 3

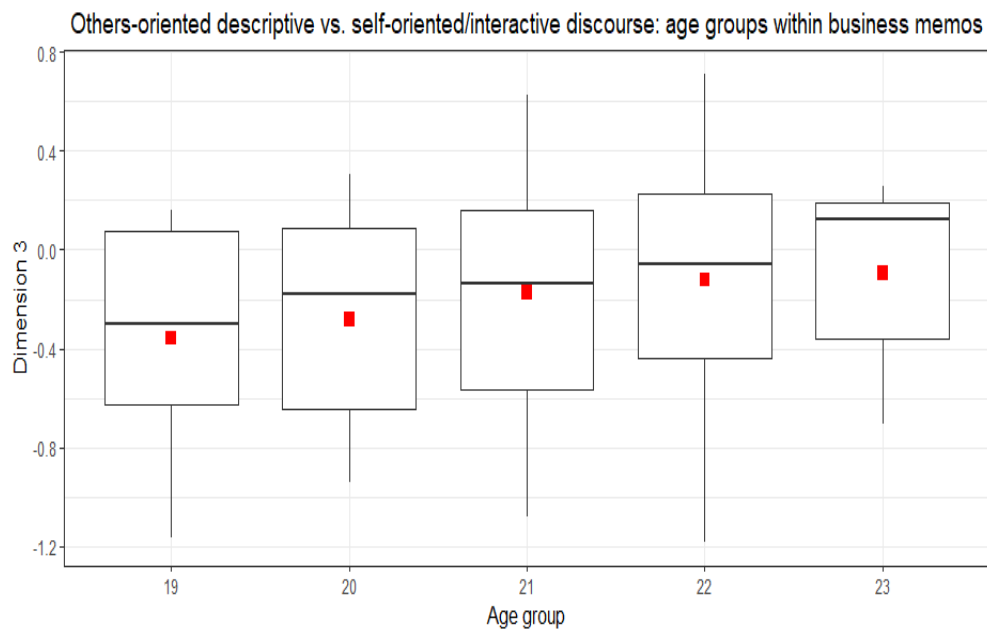
Age	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
19	–	0.52	1.85	0.57	–	0.58	0.65	0.2	0.38	0.14	–	0.13	–	0.99
	0.36				0.37						0.63		1.24	
20	–	0.4	1.55	0.27	–0.4	0.52	0.69	0.23	0.25	0.25	–	0.22	–	0.9
	0.28										0.71		1.05	

21	–	0.47	1.46	0.45	–	0.44	0.7	0.31	0.06	0.26	–	0.16	–	0.94
	0.17				0.51						0.64		0.96	
22	–	0.44	1.69	0.42	–0.4	0.43	0.7	0.22	0.05	0.23	–	0.18	–	0.85
	0.12										0.67		1.15	
23	–	0.41	1.54	0.37	–	0.22	0.5	0.22	0.06	0.07	–	0.23	–	1.11
	0.09				0.36						0.61		0.68	

Variance explained by age within registers. Dimension 3

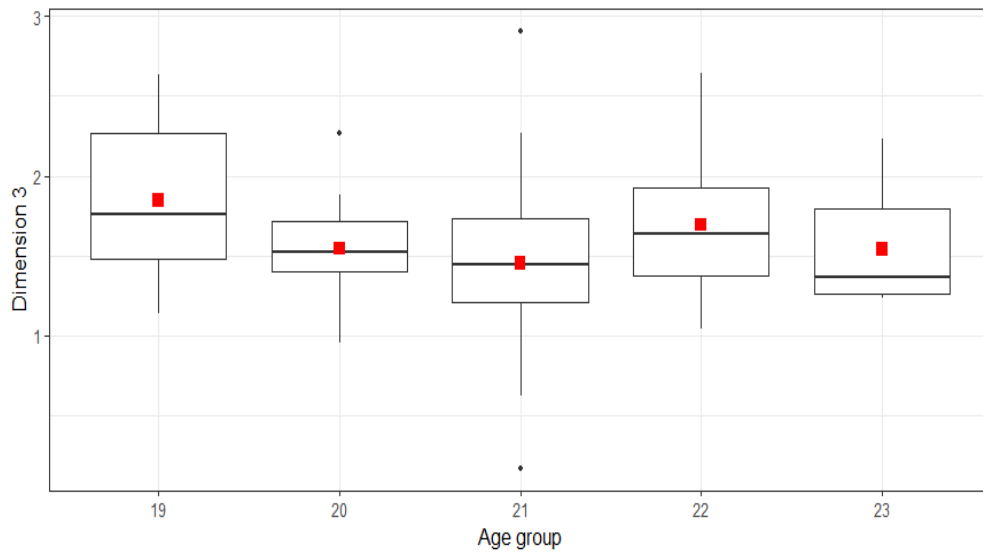
Register	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Business memos	4	0.81	>0.001	3%
Descriptions	4	2.20	>0.001	7.7%
Emails	4	0.40	>0.001	1.5%
Essays	4	1.18	>0.001	4.3%
Evaluations	4	2.31	>0.001	14%
Interviews	4	0.76	>0.001	3%
Text Messages	4	0.55	>0.001	2%

Age groups within business memos. Dimension 3



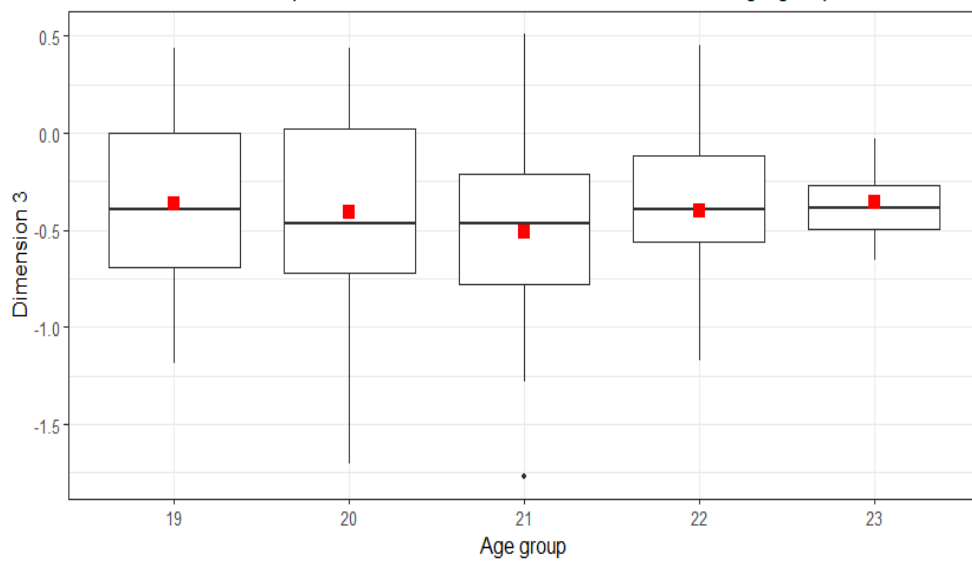
Age groups within descriptions. Dimension 3

Others-oriented descriptive vs. self-oriented/interactive discourse: age groups within descriptions

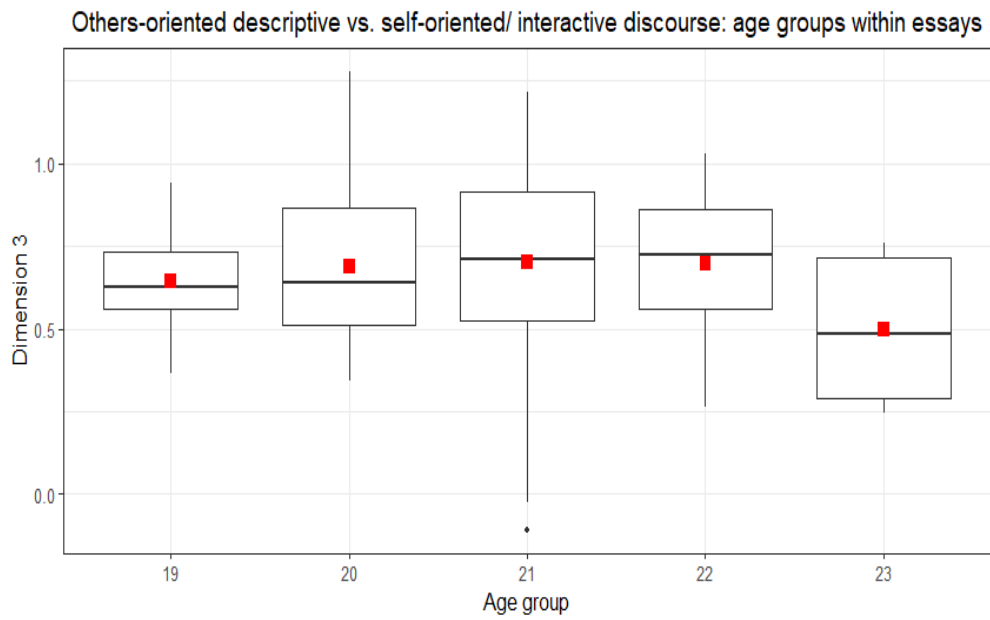


Age groups within emails. Dimension 3

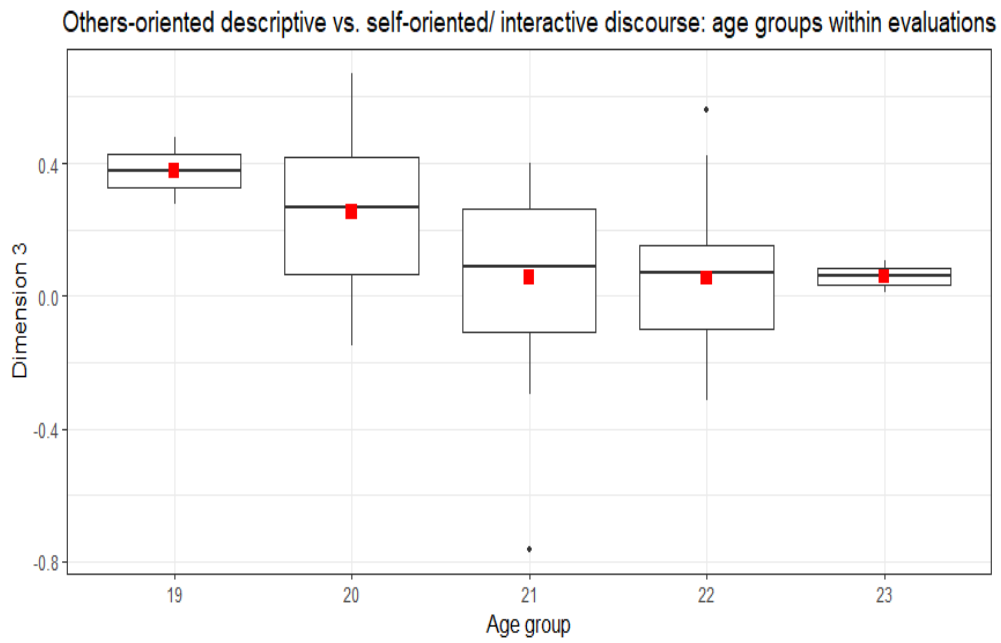
Others-oriented descriptive vs. self-oriented/ interactive discourse: age groups within emails



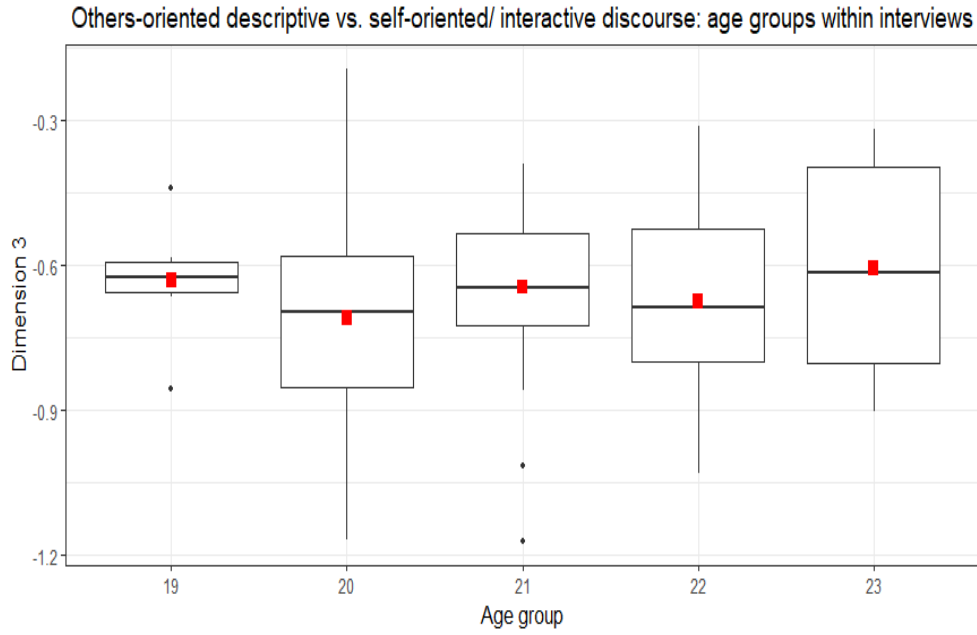
Age groups within essays. Dimension 3



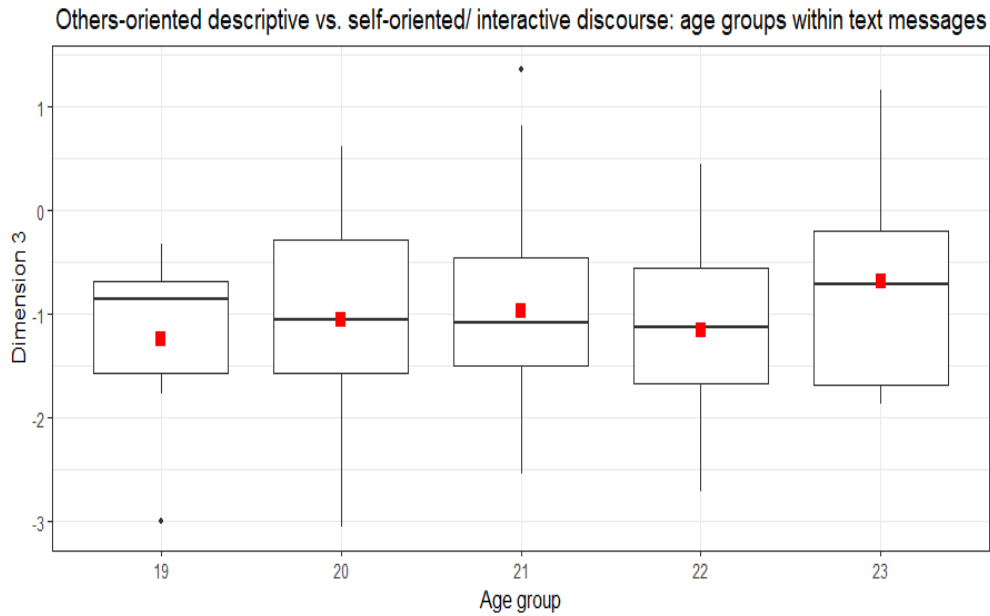
Age groups within evaluations. Dimension 3



Age groups within interviews. Dimension 3



Age groups within text messages. Dimension 3



Age groups within registers. Dimension 4

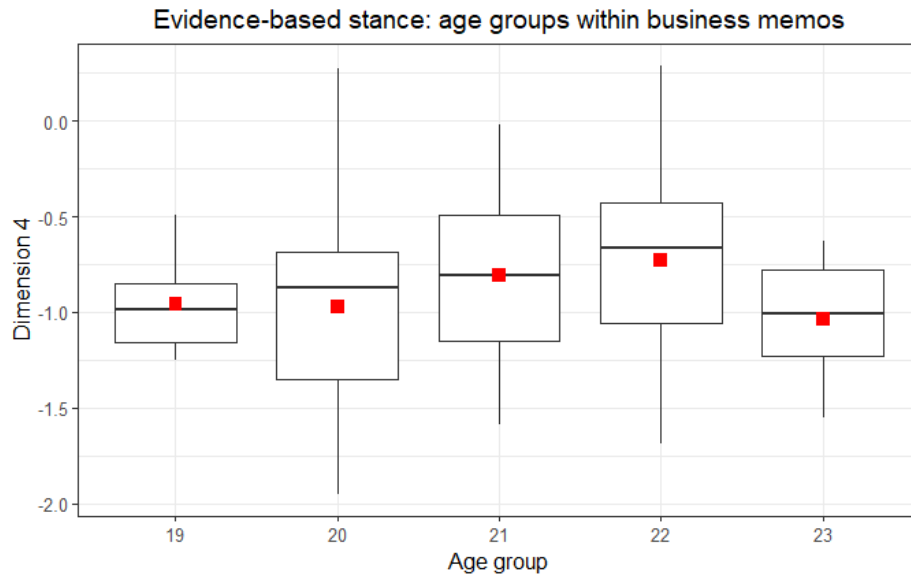
Age	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD

19	–	0.28	0.33	0.59	0.69	1.17	–	0.29	–	0.46	–	0.21	0.94	1.92
	0.96						0.13		0.44		0.63			
20	–	0.48	0.27	0.49	0.49	0.70	0.10	0.39	–	0.47	–	0.24	0.41	1.08
	0.97								0.38		0.65			
21	–	0.44	0.17	0.52	0.33	0.50	0	0.52	–	0.33	–	0.29	1.25	1.91
	0.81								0.19		0.52			
22	–	0.48	0.41	0.67	0.34	0.71	–	0.35	–	0.38	–	0.25	0.93	1.50
	0.73						0.11		0.33		0.60			
23	–	0.33	0.27	0.79	0.2	0.34	–0.2	0.43	–	0.31	–	0.21	1.75	1.97
	1.04								0.23		0.66			

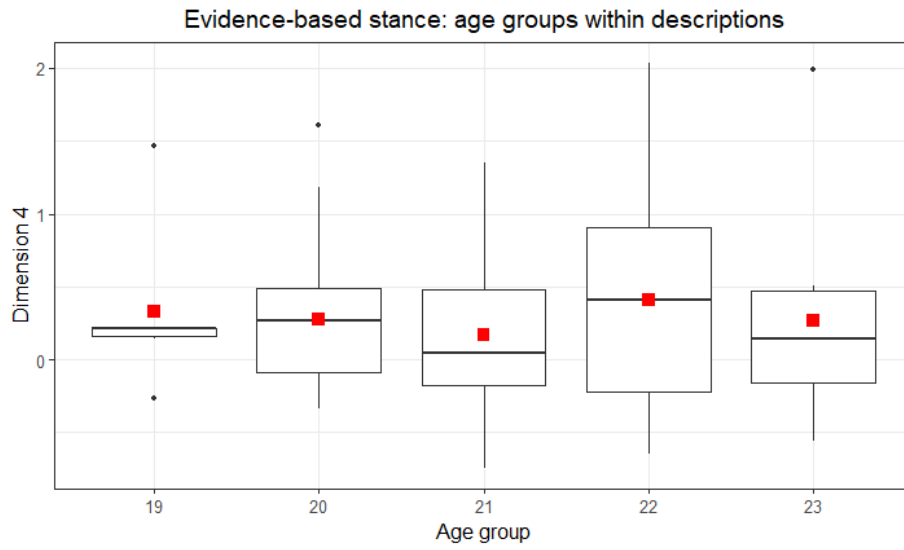
Variance explained by age within registers. Dimension 4

Register	<i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Business memos	4	1.62	>0.001	6%
Descriptions	4	0.73	>0.001	3%
Emails	4	0.75	>0.001	3%
Essays	4	1.40	>0.001	5%
Evaluations	4	0.78	>0.001	5%
Interviews	4	1.20	>0.001	4%
Text Messages	4	1.59	>0.001	5%

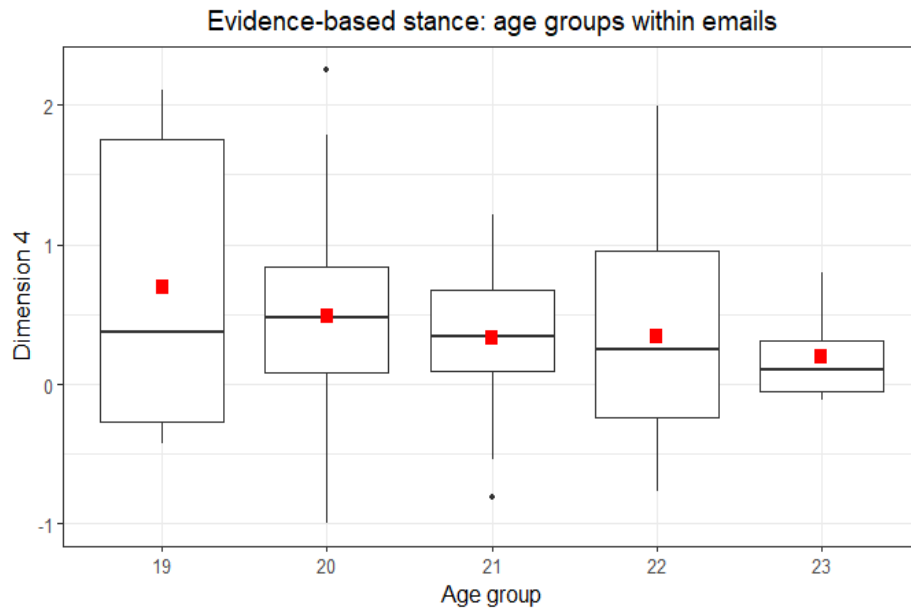
Age groups within business memos. Dimension 4



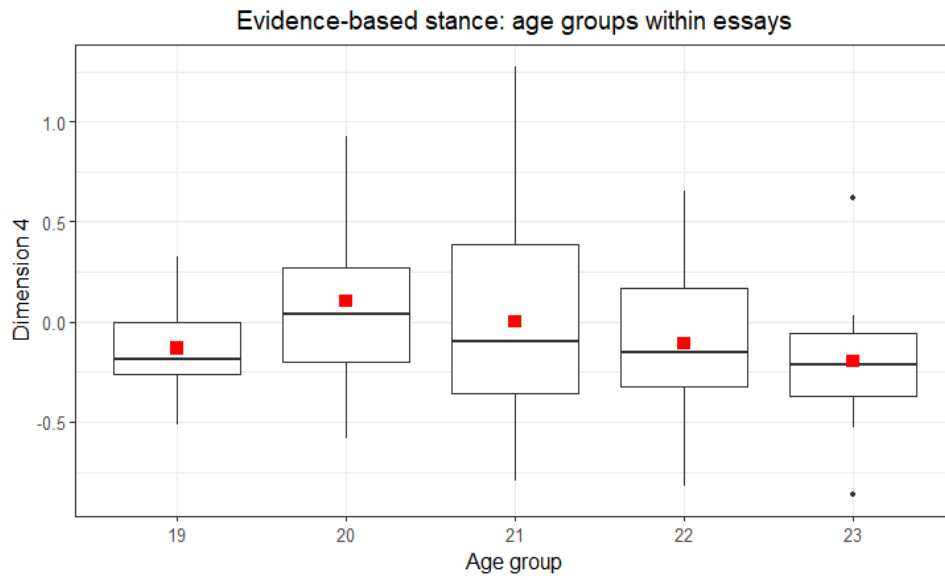
Age groups within descriptions. Dimension 4



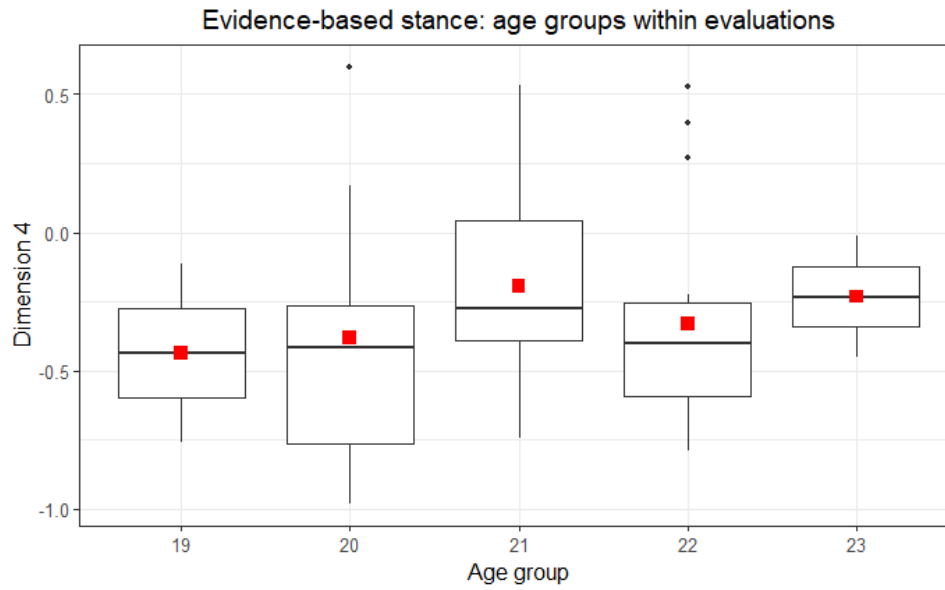
Age groups within emails. Dimension 4



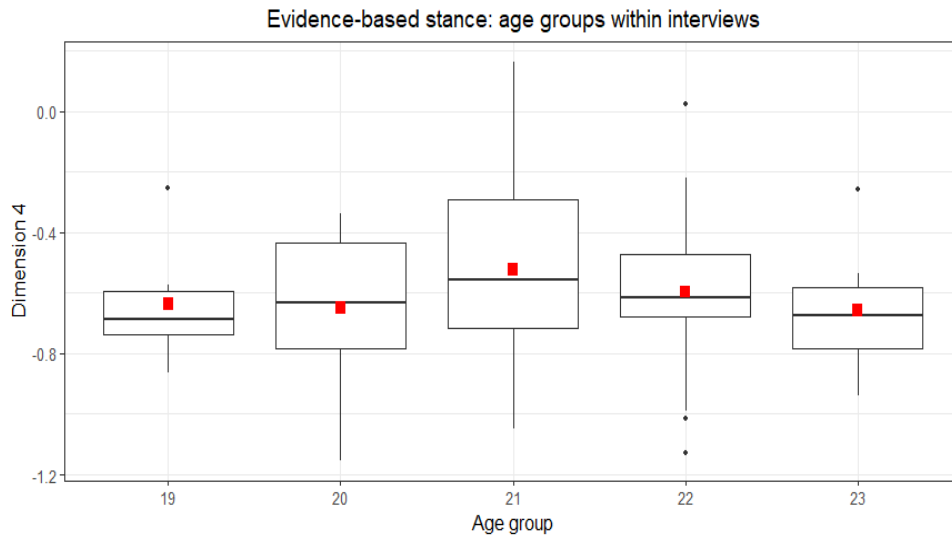
Age groups within essays. Dimension 4



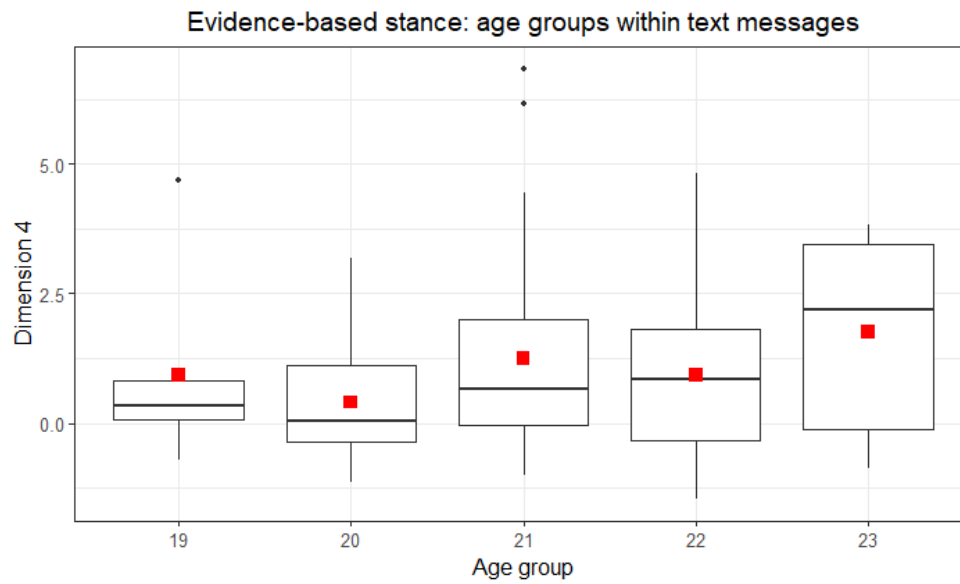
Age groups within evaluations. Dimension 4



Age groups within interviews. Dimension 4



Age groups within text messages. Dimension 4



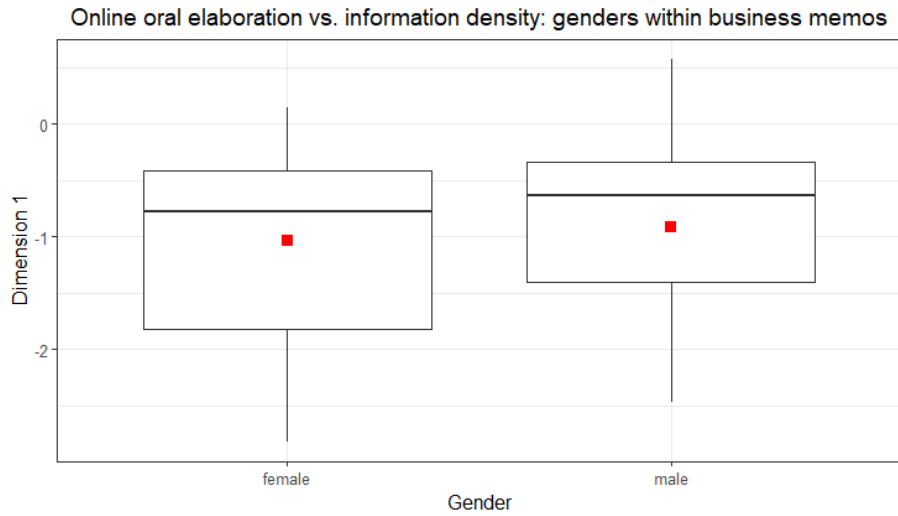
Gender groups within registers. Dimension 1

Gen der	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Men	-	0.77	0.45	0.58	-	0.41	-	0.21	0.12	0.21	1.66	0.22	-	0.63
	0.91				0.06		0.17						0.92	
Wo men	-1.04	0.79	0.43	0.49	0.02	0.37	-	0.28	0.15	0.25	1.62	0.22	-	0.66
							0.26						0.89	

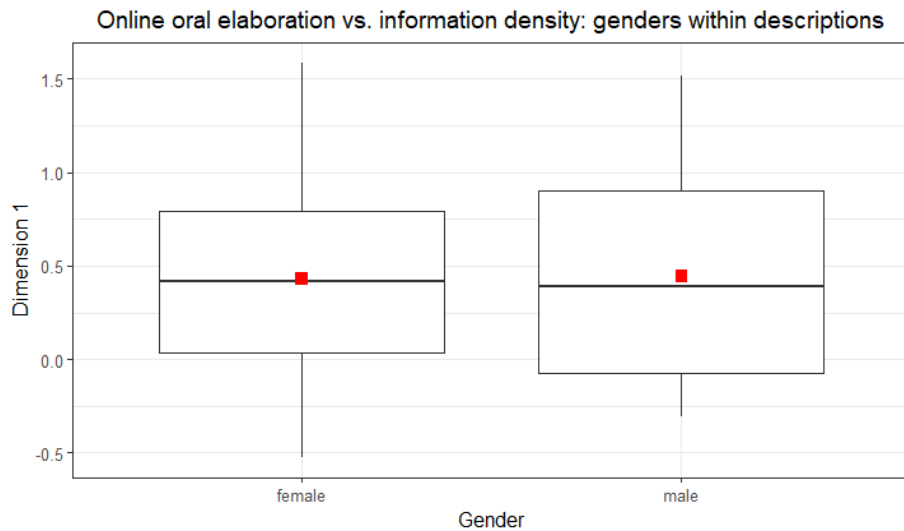
Variance explained by gender within registers. Dimension 1

Register	df	F	p	Cohen's d
Business memos	1	0.39	>0.001	0.2
Descriptions	1	0.01	>0.001	0.04
Emails	1	0.83	>0.001	-0.2
Essays	1	1.76	>0.001	0.3
Evaluations	1	0.16	>0.001	-0.1
Interviews	1	0.67	>0.001	0.2
Text Messages	1	0.03	>0.001	-0.04

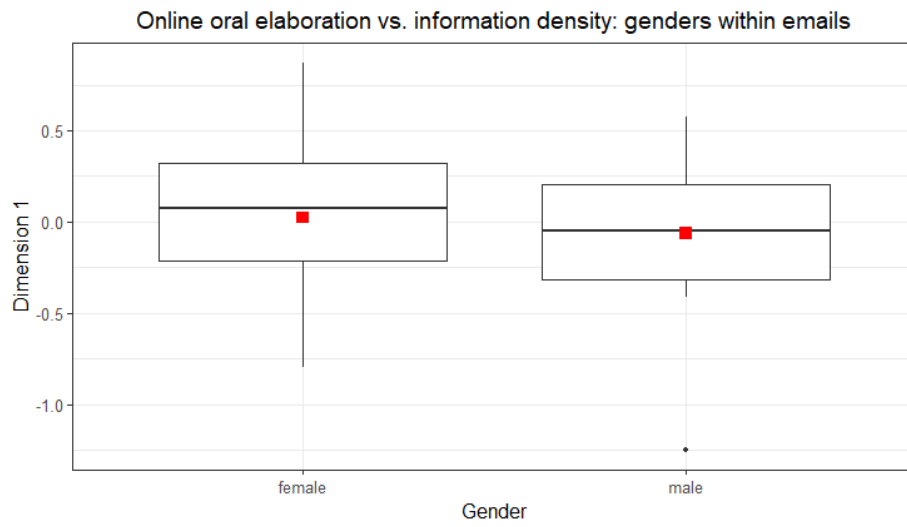
Gender groups within business memos. Dimension 1



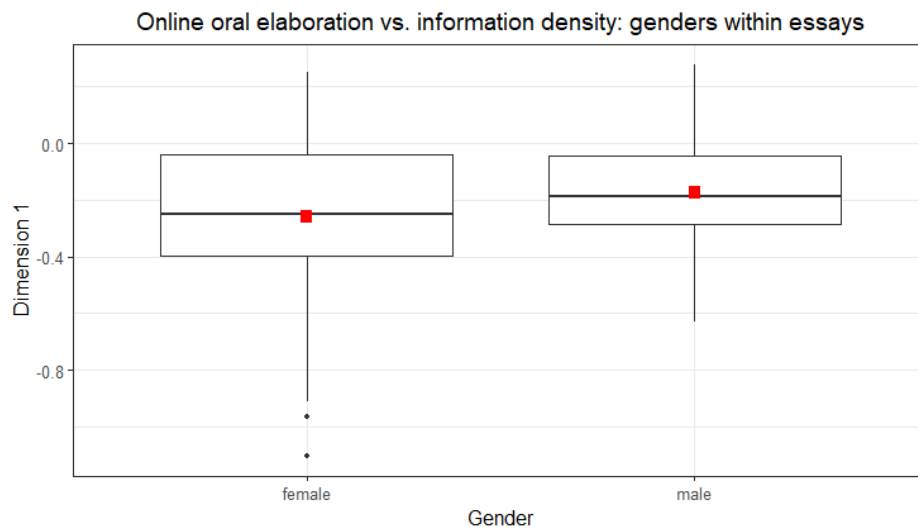
Gender groups within descriptions. Dimension 1



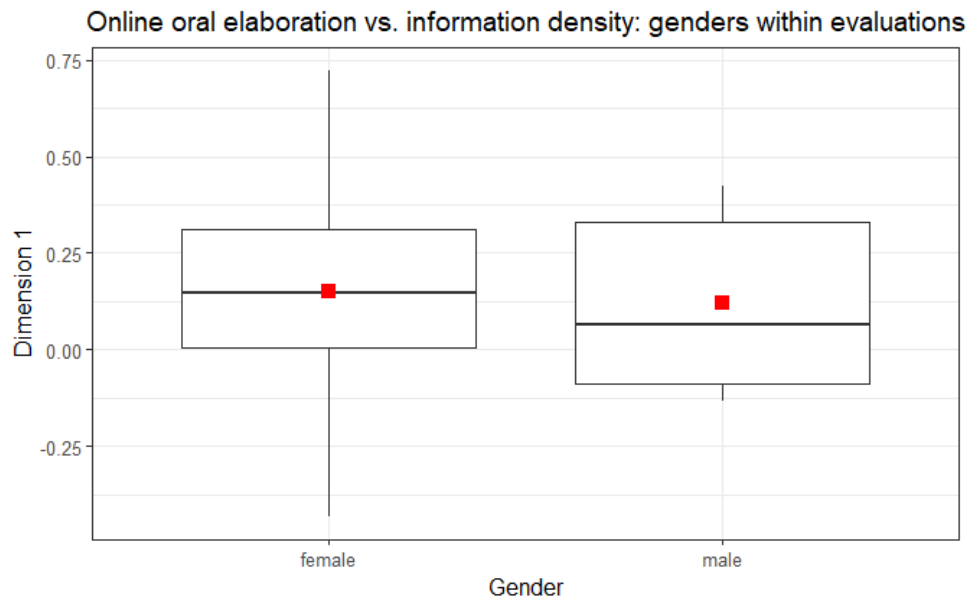
Gender groups within emails. Dimension 1



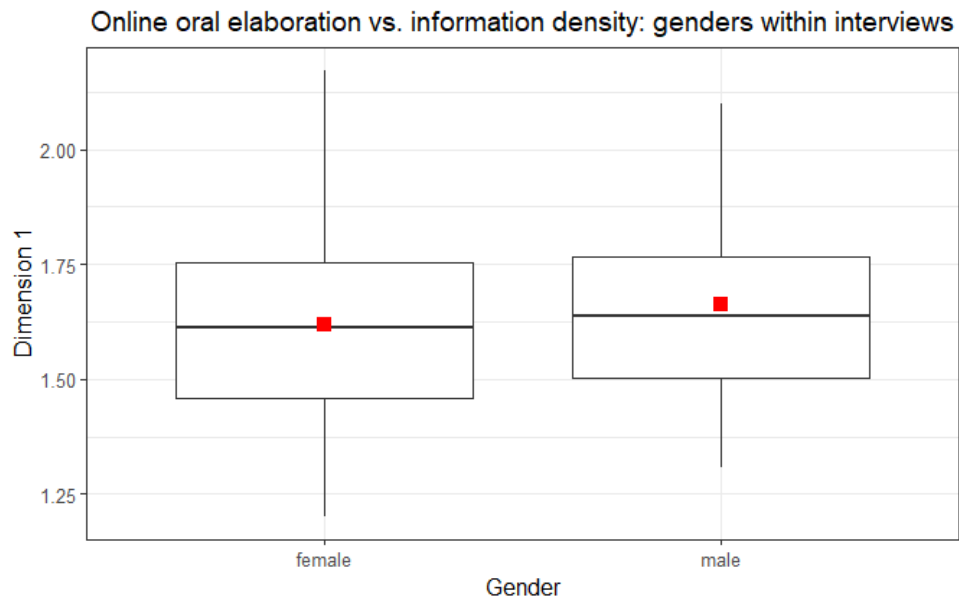
Gender groups within essays. Dimension 1



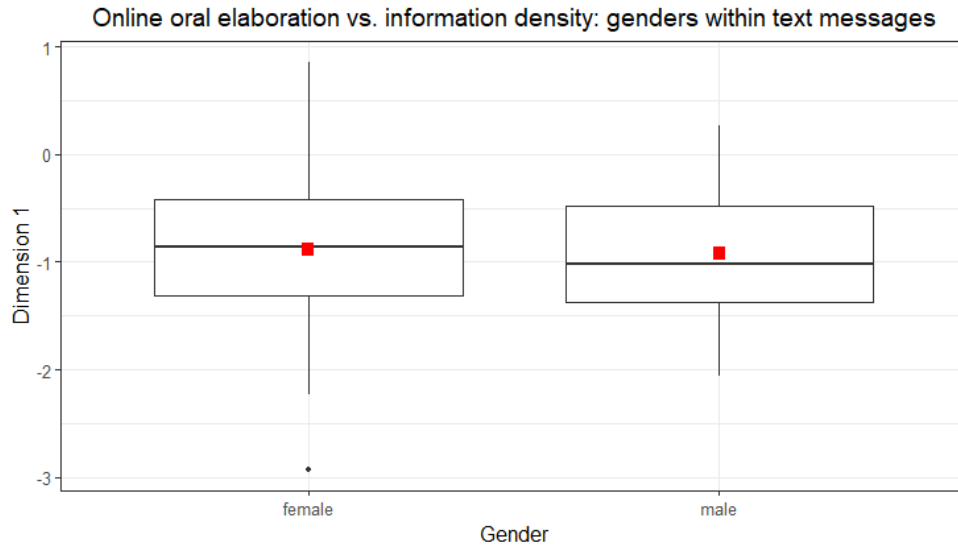
Gender groups within evaluations. Dimension 1



Gender groups within interviews. Dimension 1



Gender groups within text messages. Dimension 1



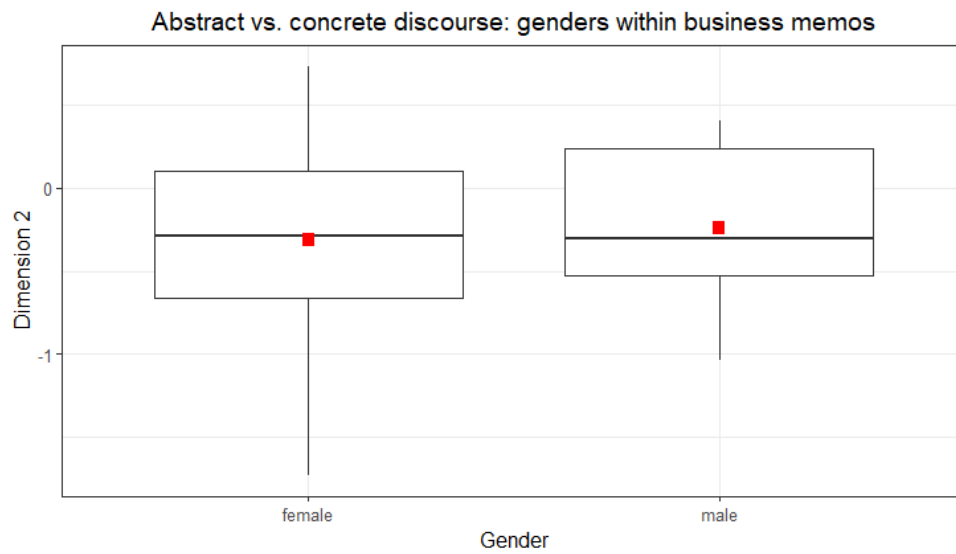
Gender groups within registers. Dimension 2

Gender	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Men	-	0.47	-1.18	0.45	0.6	0.38	1.3	0.26	0.58	0.29	-	0.12	-	0.87
Women	0.24										0.05		0.46	
Women	-	0.55	-1.34	0.37	0.72	0.5	1.3	0.3	0.59	0.32	-	0.17	-	0.65
men	0.31										0.06		0.72	

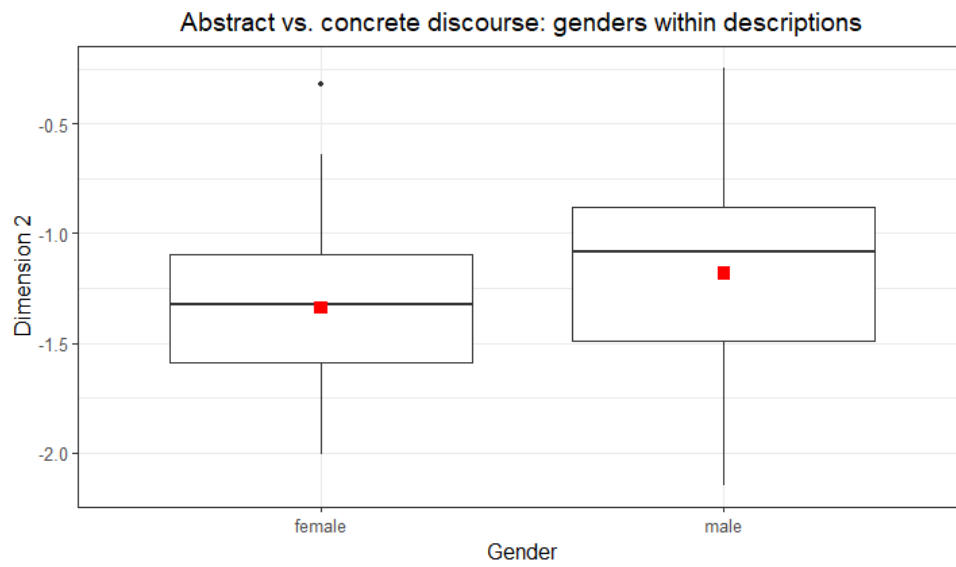
Variance explained by gender within registers. Dimension 2

Register	<i>df</i>	<i>F</i>	<i>p</i>	<i>Cohen's d</i>
Business memos	1	0.28	>0.001	0.1
Descriptions	1	2.73	>0.001	0.4
Emails	1	1.06	>0.001	-0.3
Essays	1	0.001	>0.001	0
Evaluations	1	0.02	>0.001	-0.03
Interviews	1	0.01	>0.001	0.07
Text Messages	1	2.28	>0.001	0.3

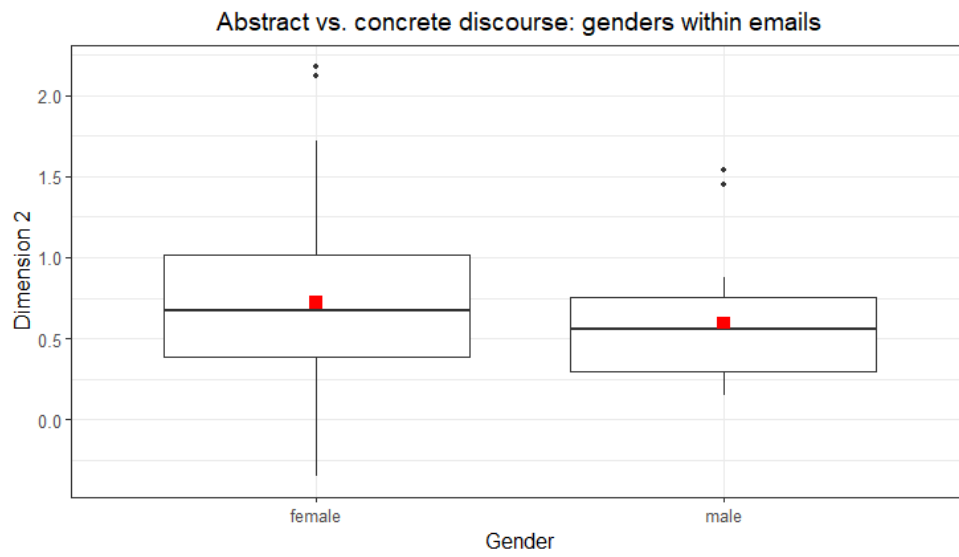
Gender groups within business memos. Dimension 2



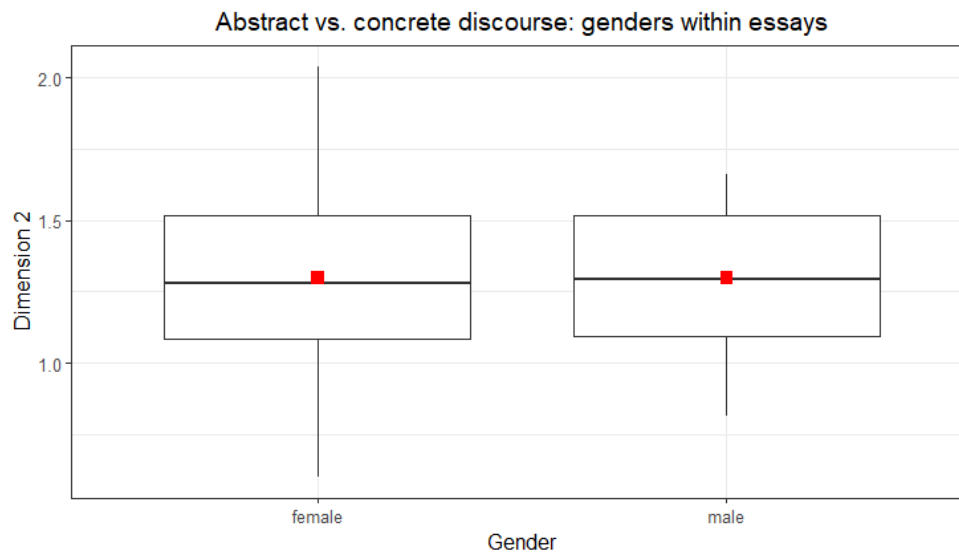
Gender groups within descriptions. Dimension 2



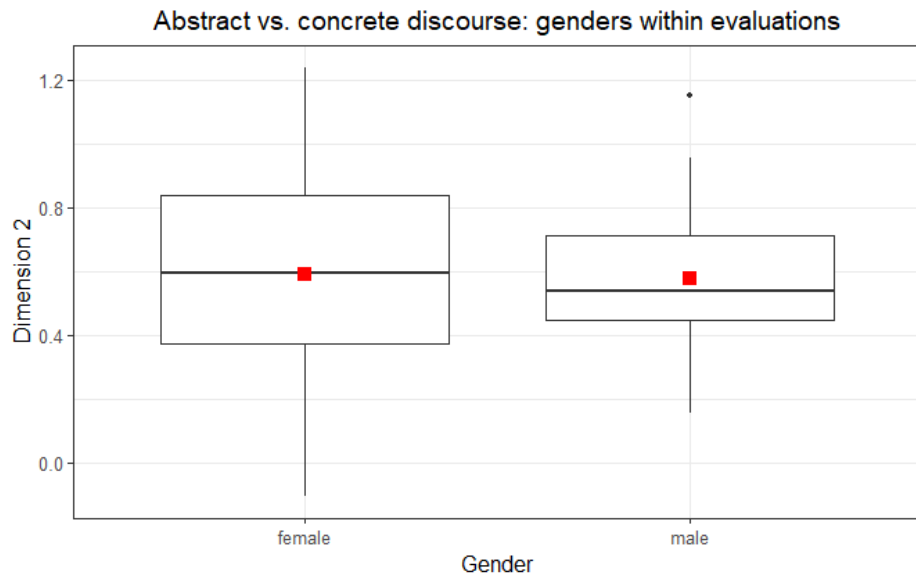
Gender groups within emails. Dimension 2



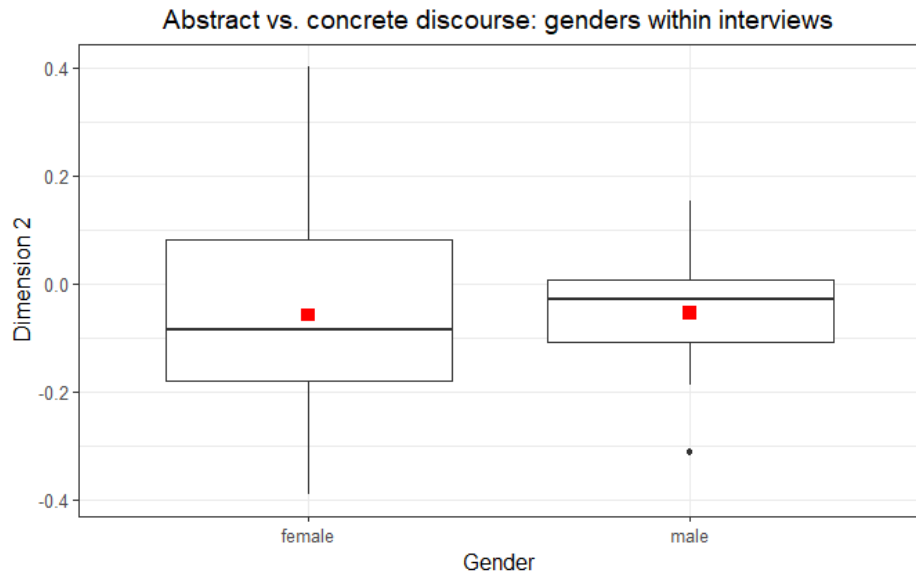
Gender groups within essays. Dimension 2



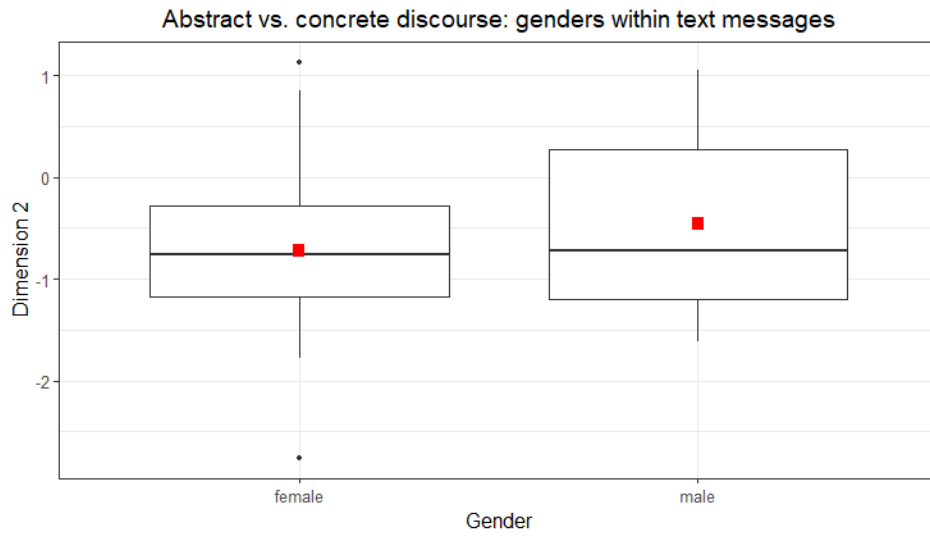
Gender groups within evaluations. Dimension 2



Gender groups within interviews. Dimension 2



Gender groups within text messages. Dimension 2



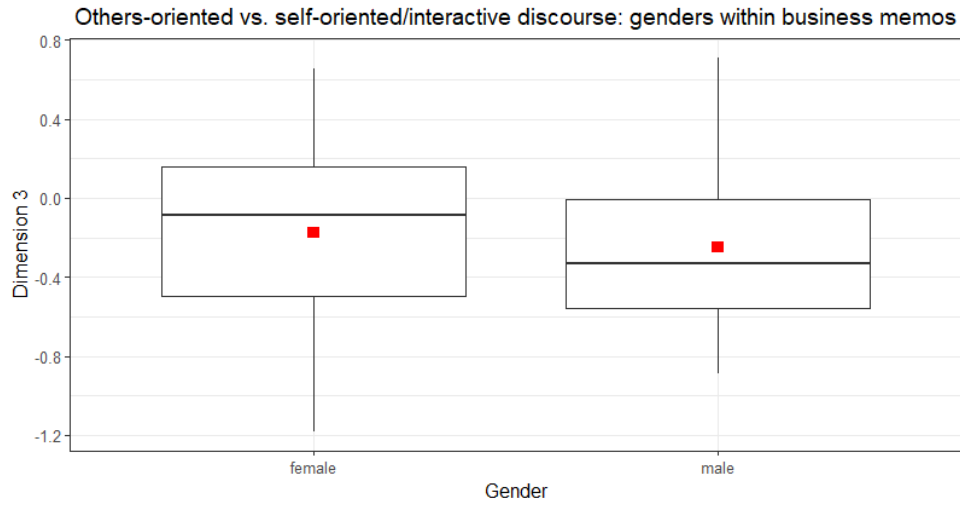
Gender groups within registers. Dimension 3

Gender	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Men	-	0.41	1.59	0.37	-0.4	0.56	0.62	0.27	0.21	0.2	-	0.22	-1.11	0.94
Women	-	0.45	1.57	0.43	-	0.42	0.69	0.25	0.09	0.26	-	0.18	-1.02	0.91
	0.25				0.44						0.61			
	0.17										0.68			

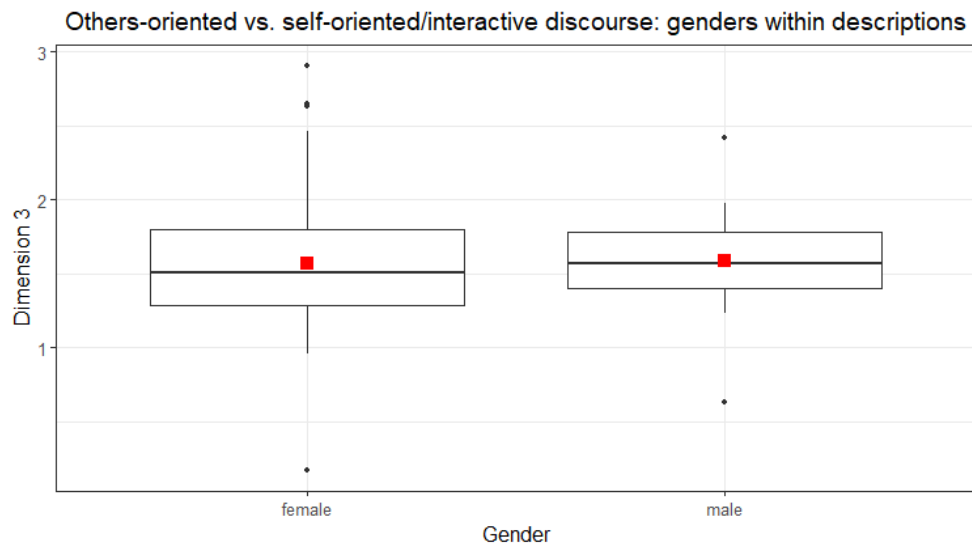
Variance explained by gender within registers. Dimension 3

Register	<i>df</i>	<i>F</i>	<i>p</i>	<i>Cohen's d</i>
Business memos	1	0.45	>0.001	-0.2
Descriptions	1	0.02	>0.001	0.05
Emails	1	0.14	>0.001	0.1
Essays	1	1.26	>0.001	-0.2
Evaluations	1	2.13	>0.001	0.5
Interviews	1	2.17	>0.001	0.3
Text Messages	1	0.15	>0.001	-0.1

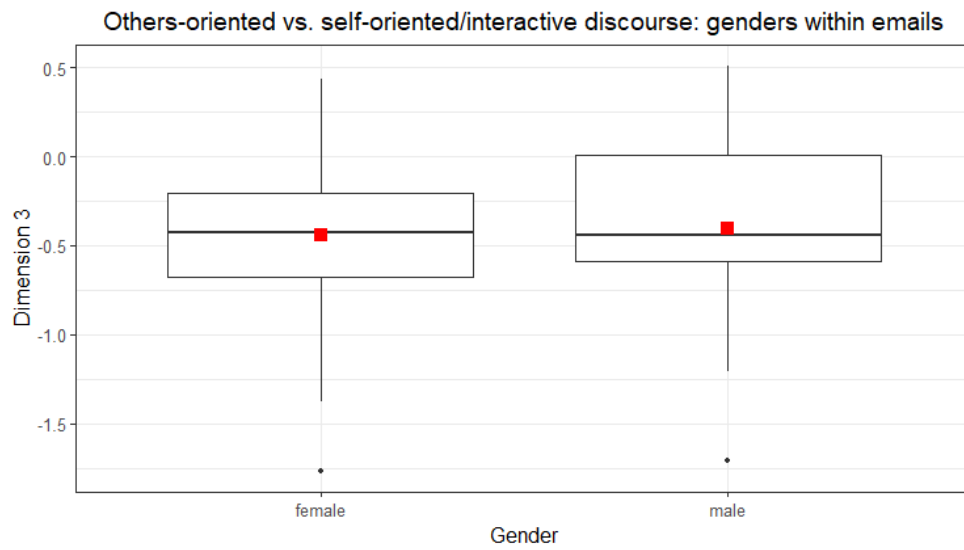
Gender groups within business memos. Dimension 3



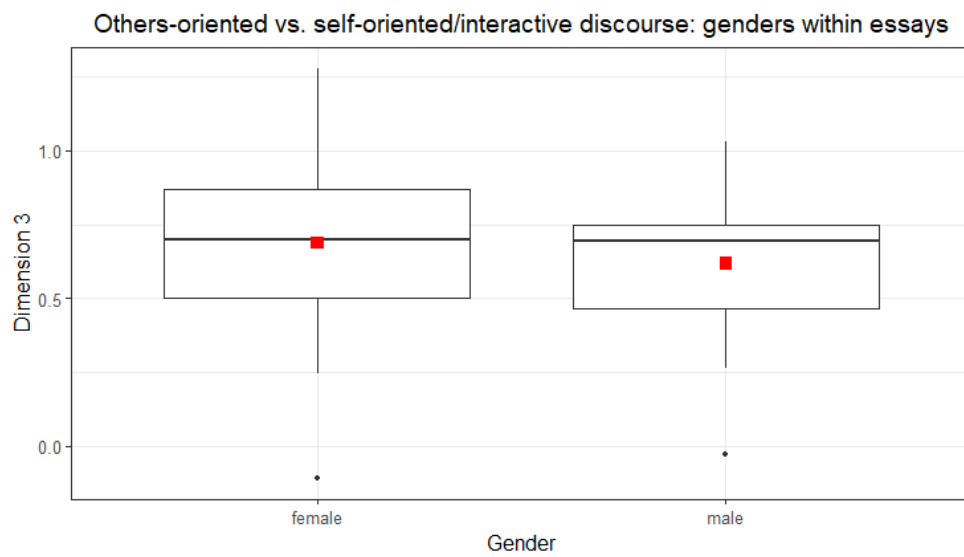
Gender groups within descriptions. Dimension 3



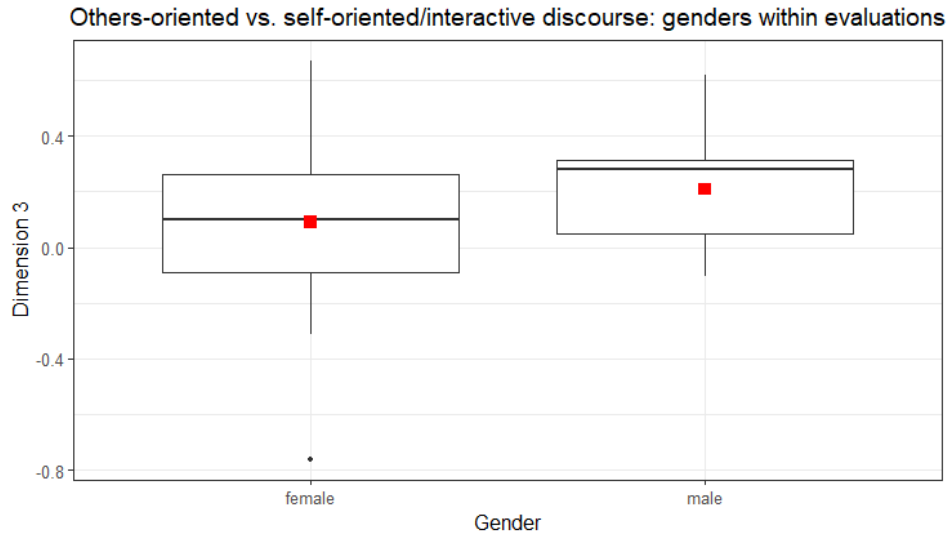
Gender groups within emails. Dimension 3



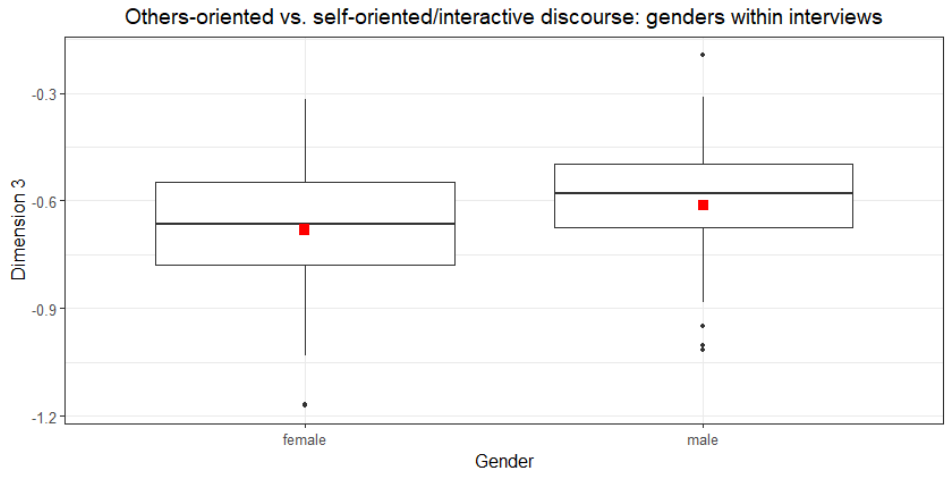
Gender groups within essays. Dimension 3



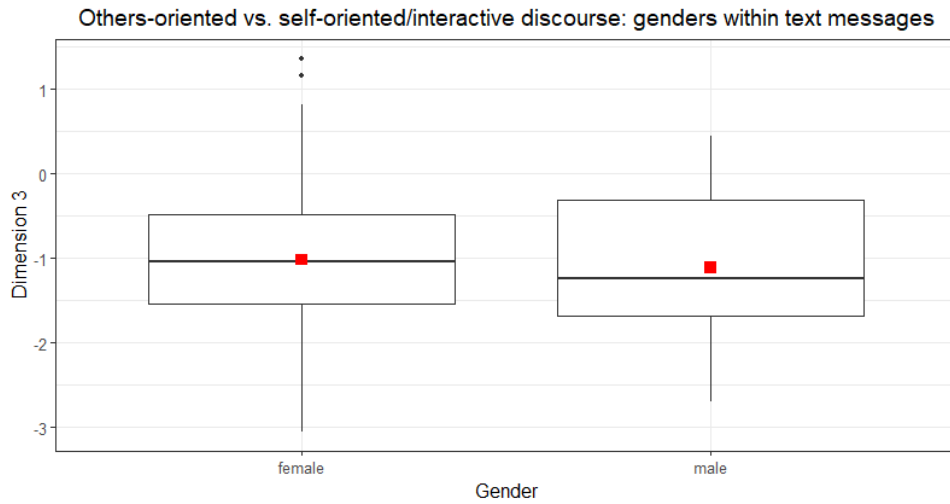
Gender groups within evaluations. Dimension 3



Gender groups within interviews. Dimension 3



Gender groups within text messages. Dimension 3



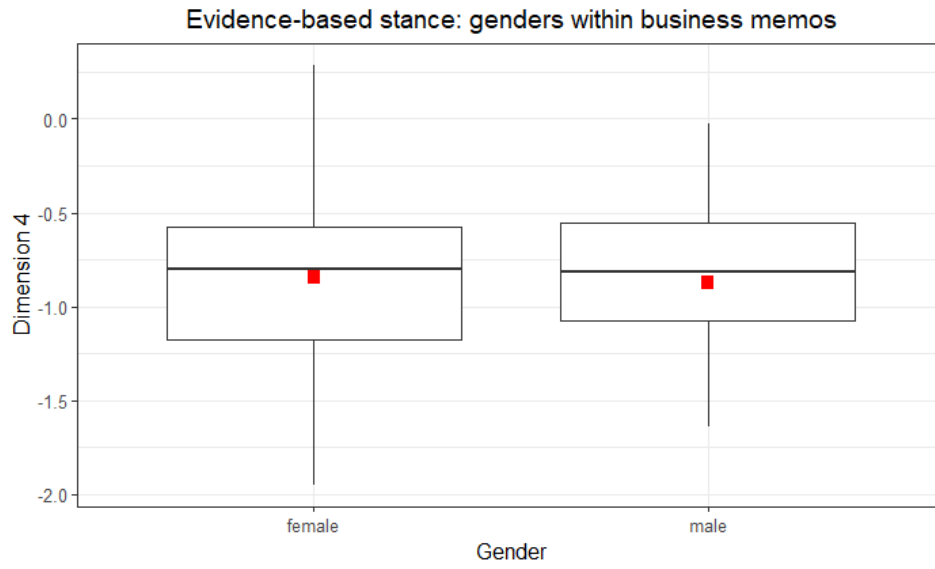
Gender groups within registers. Dimension 4

Gender	Business memos		Descriptions		Emails		Essays		Evaluations		Interviews		Text Messages	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Men	–	0.41	0.41	0.6	0.27	0.62	–0.2	0.36	–	0.44	–	0.24	1.09	1.62
	0.87								0.34		0.64			
Women	–	0.46	0.26	0.57	0.42	0.67	0.02	0.44	–	0.36	–	0.26	0.93	1.66
	0.84								0.25		0.58			

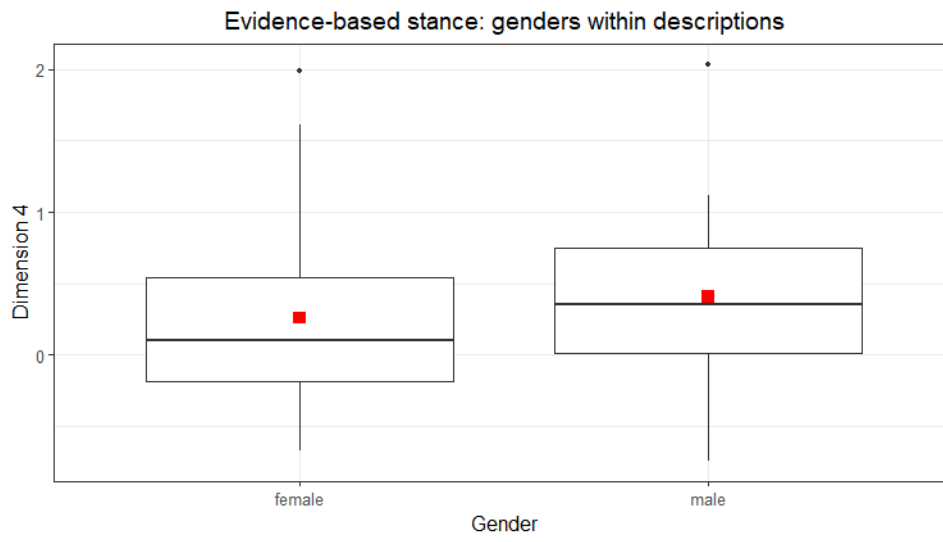
Variance explained by gender within registers. Dimension 4

Register	<i>df</i>	<i>F</i>	<i>p</i>	<i>Cohen's d</i>
Business memos	1	0.05	>0.001	–0.1
Descriptions	1	1.04	>0.001	0.2
Emails	1	0.88	>0.001	–0.2
Essays	1	4.49	>0.001	–0.5
Evaluations	1	0.55	>0.001	–0.2
Interviews	1	0.94	>0.001	–0.2
Text Messages	1	0.15	>0.001	0.1

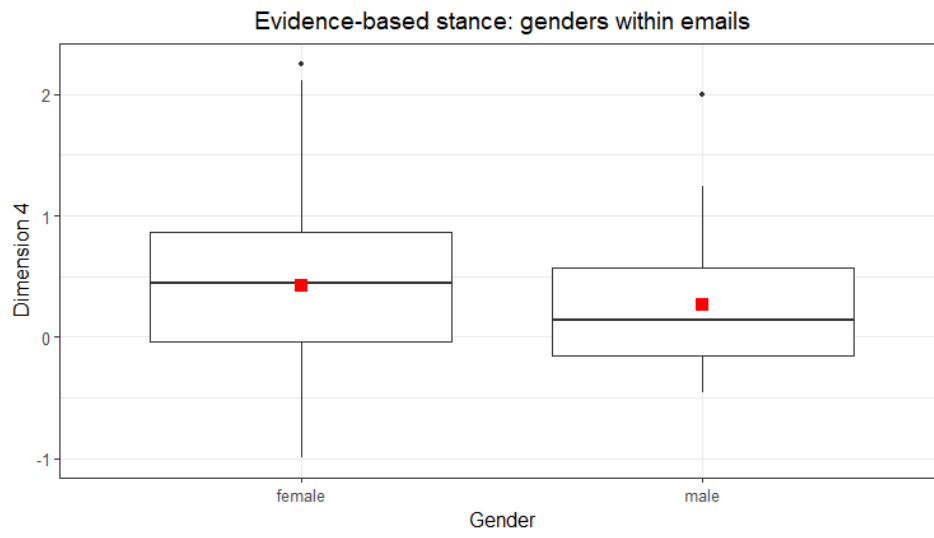
Gender groups within business memos. Dimension 4



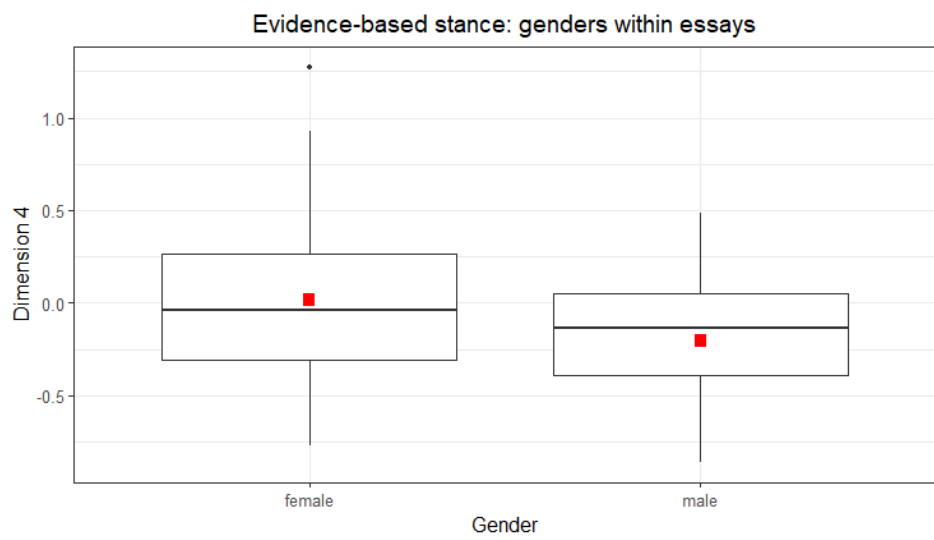
Gender groups within descriptions. Dimension 4



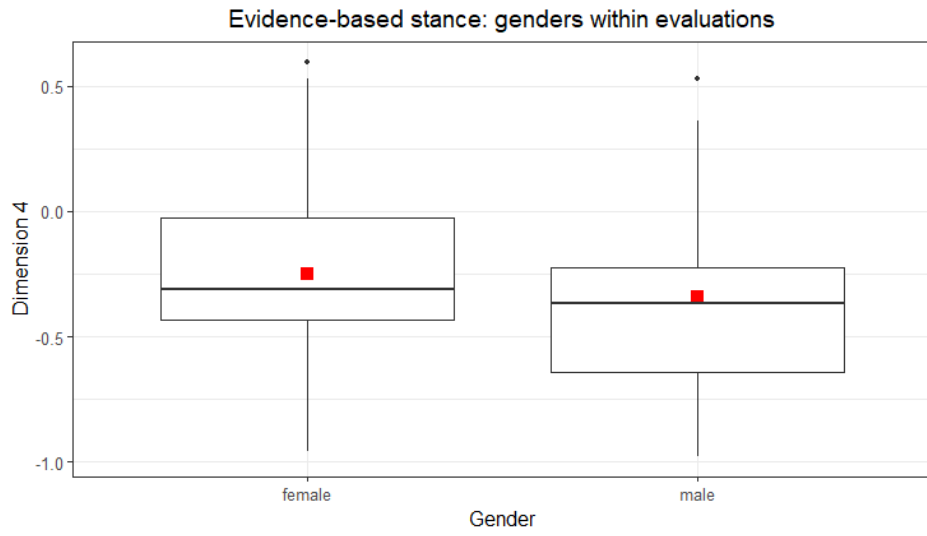
Gender groups within emails. Dimension 4



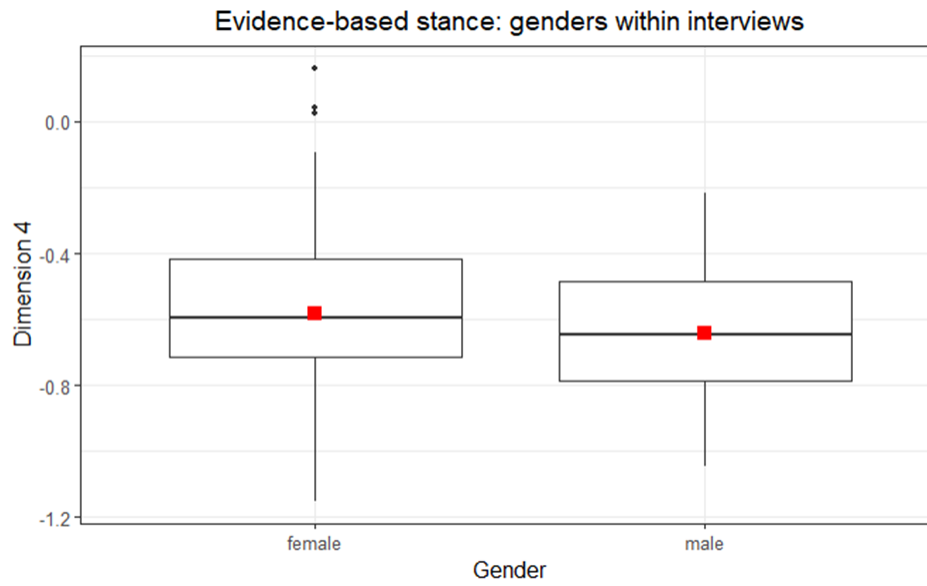
Gender groups within essays. Dimension 4



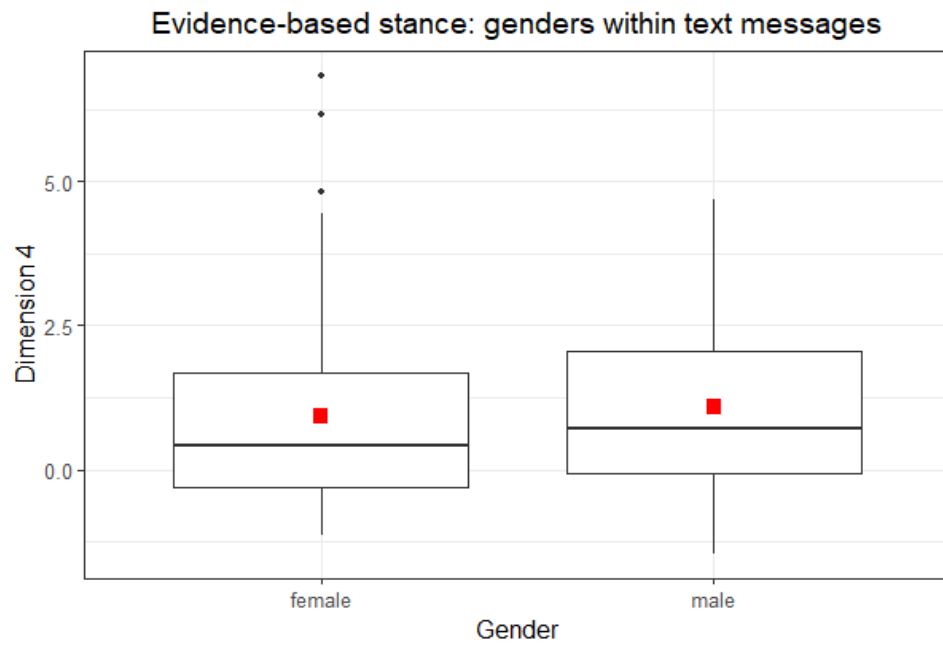
Gender groups within evaluations. Dimension 4



Gender groups within interviews. Dimension 4



Gender groups within interviews. Dimension 4



Appendix 5. Text scores, register means, measures of distance. Dimension 1

Author	Business memos			Descriptions			Emails			Essays			Evaluations			Interviews			Text messages		
	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean
1	-1.84	-1.01	0.83	0.98	0.44	0.54	-0.25	0	0.25	-0.37	-0.24	0.13	na	0.15	na	1.72	1.62	0.1	-1.71	-0.9	0.81
2	-0.77	-1.01	0.24	0.36	0.44	0.08	0.70	0	0.7	-0.36	-0.24	0.12	0.17	0.15	0.02	1.58	1.62	0.04	-0.42	-0.9	0.48
3	-1.42	-1.01	0.41	0.24	0.44	0.2	0.38	0	0.38	-0.25	-0.24	0.01	0.26	0.15	0.11	1.37	1.62	0.25	-0.42	-0.9	0.48
4	-1.40	-1.01	0.39	0.35	0.44	0.09	-0.22	0	0.22	-0.18	-0.24	0.06	na	0.15	na	2.05	1.62	0.43	-0.87	-0.9	0.03
5	-0.76	-1.01	0.25	0.41	0.44	0.03	-0.02	0	0.02	-0.23	-0.24	0.01	0.72	0.15	0.57	1.53	1.62	0.09	-1.01	-0.9	0.11
6	-1.21	-1.01	0.2	0.36	0.44	0.08	0.22	0	0.22	-0.49	-0.24	0.25	na	0.15	na	1.75	1.62	0.13	-0.21	-0.9	0.69
7	-0.40	-1.01	0.61	0.24	0.44	0.2	-0.05	0	0.05	-0.16	-0.24	0.08	na	0.15	na	1.54	1.62	0.08	-1.24	-0.9	0.34
8	-2.14	-1.01	1.13	-0.42	0.44	0.86	-0.46	0	0.46	-0.15	-0.24	0.09	na	0.15	na	1.42	1.62	0.2	-0.29	-0.9	0.61
9	-2.47	-1.01	1.46	1.21	0.44	0.77	0.07	0	0.07	-0.28	-0.24	0.04	0.18	0.15	0.03	1.54	1.62	0.08	-1.35	-0.9	0.45
10	-0.30	-1.01	0.71	0.36	0.44	0.08	0.17	0	0.17	-0.13	-0.24	0.11	0.43	0.15	0.28	1.53	1.62	0.09	-0.56	-0.9	0.34
11	-0.47	-1.01	0.54	1.05	0.44	0.61	-0.21	0	0.21	-0.03	-0.24	0.21	0.40	0.15	0.25	1.94	1.62	0.32	-0.10	-0.9	0.8
12	-0.63	-1.01	0.38	0.50	0.44	0.06	-0.25	0	0.25	-0.50	-0.24	0.26	0.20	0.15	0.05	1.79	1.62	0.17	-2.23	-0.9	1.33
13	-0.04	-1.01	0.97	0.78	0.44	0.34	0.12	0	0.12	-0.10	-0.24	0.14	0.30	0.15	0.15	1.84	1.62	0.22	-1.46	-0.9	0.56
14	-0.08	-1.01	0.93	0.31	0.44	0.13	0.27	0	0.27	0.01	-0.24	0.25	0.12	0.15	0.03	1.20	1.62	0.42	-1.72	-0.9	0.82
15	-0.95	-1.01	0.06	0.32	0.44	0.12	-0.10	0	0.1	-0.27	-0.24	0.03	0.27	0.15	0.12	1.76	1.62	0.14	-1.11	-0.9	0.21
16	-0.49	-1.01	0.52	0.70	0.44	0.26	0.45	0	0.45	-0.47	-0.24	0.23	na	0.15	na	1.87	1.62	0.25	-0.76	-0.9	0.14
17	-0.48	-1.01	0.53	0.95	0.44	0.51	0.09	0	0.09	0.11	-0.24	0.35	na	0.15	na	1.62	1.62	0	-1.32	-0.9	0.42
18	-1.54	-1.01	0.53	0.85	0.44	0.41	0.32	0	0.32	-0.66	-0.24	0.42	0.18	0.15	0.03	1.59	1.62	0.03	-0.93	-0.9	0.03
19	-0.62	-1.01	0.39	-0.02	0.44	0.46	0.36	0	0.36	-0.41	-0.24	0.17	0.11	0.15	0.04	1.63	1.62	0.01	-1.37	-0.9	0.47
20	-2.34	-1.01	1.33	0.94	0.44	0.5	-0.28	0	0.28	-0.18	-0.24	0.06	-0.14	0.15	0.29	1.53	1.62	0.09	-0.35	-0.9	0.55
21	-0.39	-1.01	0.62	0.05	0.44	0.39	-0.52	0	0.52	-0.12	-0.24	0.12	na	0.15	na	1.49	1.62	0.13	-0.77	-0.9	0.13
22	-1.91	-1.01	0.9	-0.20	0.44	0.64	0.34	0	0.34	-0.40	-0.24	0.16	na	0.15	na	1.29	1.62	0.33	-0.05	-0.9	0.85
23	-2.00	-1.01	0.99	0.34	0.44	0.1	0.07	0	0.07	-0.40	-0.24	0.16	na	0.15	na	1.35	1.62	0.27	-0.40	-0.9	0.5
24	-0.33	-1.01	0.68	0.02	0.44	0.42	-0.32	0	0.32	-0.13	-0.24	0.11	0.07	0.15	0.08	1.30	1.62	0.32	-0.50	-0.9	0.4
25	-2.47	-1.01	1.46	0.27	0.44	0.17	-0.21	0	0.21	0.05	-0.24	0.29	0.63	0.15	0.48	1.65	1.62	0.03	-1.34	-0.9	0.44
26	-1.04	-1.01	0.03	0.67	0.44	0.23	0.14	0	0.14	0.25	-0.24	0.49	0.31	0.15	0.16	1.65	1.62	0.03	-0.97	-0.9	0.07
27	-2.07	-1.01	1.06	0.57	0.44	0.13	-0.35	0	0.35	-0.85	-0.24	0.61	0.19	0.15	0.04	1.38	1.62	0.24	-0.23	-0.9	0.67

28	-0.11	-1.01	0.9	0.01	0.44	0.43	-0.79	0	0.79	0.18	-0.24	0.42	na	0.15	na	1.30	1.62	0.32	-0.85	-0.9	0.05
29	-1.95	-1.01	0.94	0.36	0.44	0.08	-0.27	0	0.27	-0.20	-0.24	0.04	na	0.15	na	1.82	1.62	0.2	-2.06	-0.9	1.16
30	-1.66	-1.01	0.65	-0.24	0.44	0.68	0.18	0	0.18	0.10	-0.24	0.34	0.50	0.15	0.35	1.45	1.62	0.17	-0.81	-0.9	0.09
31	-2.12	-1.01	1.11	0.40	0.44	0.04	0.32	0	0.32	-0.14	-0.24	0.1	0.07	0.15	0.08	1.75	1.62	0.13	-1.14	-0.9	0.24
32	-0.47	-1.01	0.54	1.48	0.44	1.04	0.34	0	0.34	-0.69	-0.24	0.45	-0.03	0.15	0.18	1.91	1.62	0.29	-0.27	-0.9	0.63
33	-0.32	-1.01	0.69	0.91	0.44	0.47	-0.13	0	0.13	0.28	-0.24	0.52	0.32	0.15	0.17	1.73	1.62	0.11	-0.40	-0.9	0.5
34	-0.05	-1.01	0.96	0.21	0.44	0.23	-0.09	0	0.09	-0.01	-0.24	0.23	na	0.15	na	1.54	1.62	0.08	0.86	-0.9	1.76
35	-0.62	-1.01	0.39	1.22	0.44	0.78	0.44	0	0.44	-0.03	-0.24	0.21	na	0.15	na	1.77	1.62	0.15	-1.09	-0.9	0.19
36	-0.01	-1.01	1	0.13	0.44	0.31	0.34	0	0.34	-0.21	-0.24	0.03	-0.07	0.15	0.22	1.42	1.62	0.2	-0.77	-0.9	0.13
37	-0.26	-1.01	0.75	0.03	0.44	0.41	0.18	0	0.18	-0.26	-0.24	0.02	na	0.15	na	1.52	1.62	0.1	0.04	-0.9	0.94
38	-0.40	-1.01	0.61	1.15	0.44	0.71	-0.42	0	0.42	-0.30	-0.24	0.06	na	0.15	na	1.48	1.62	0.14	-0.37	-0.9	0.53
39	-0.79	-1.01	0.22	-0.45	0.44	0.89	-0.01	0	0.01	0.06	-0.24	0.3	na	0.15	na	1.44	1.62	0.18	-1.60	-0.9	0.7
40	-1.56	-1.01	0.55	0.82	0.44	0.38	-0.58	0	0.58	-0.06	-0.24	0.18	0.10	0.15	0.05	1.55	1.62	0.07	-1.39	-0.9	0.49
41	-0.39	-1.01	0.62	-0.06	0.44	0.5	-0.59	0	0.59	-0.29	-0.24	0.05	-0.05	0.15	0.2	1.70	1.62	0.08	-0.83	-0.9	0.07
42	-1.41	-1.01	0.4	-0.30	0.44	0.74	-0.31	0	0.31	0.09	-0.24	0.33	na	0.15	na	1.33	1.62	0.29	-1.42	-0.9	0.52
43	-0.14	-1.01	0.87	0.13	0.44	0.31	0.12	0	0.12	-0.39	-0.24	0.15	0.06	0.15	0.09	1.30	1.62	0.32	-1.37	-0.9	0.47
44	-1.16	-1.01	0.15	0.13	0.44	0.31	0.11	0	0.11	-0.63	-0.24	0.39	0.10	0.15	0.05	1.86	1.62	0.24	-1.91	-0.9	1.01
45	-2.42	-1.01	1.41	0.84	0.44	0.4	0.51	0	0.51	0.19	-0.24	0.43	na	0.15	na	2.13	1.62	0.51	-0.83	-0.9	0.07
46	-0.82	-1.01	0.19	0.42	0.44	0.02	0.17	0	0.17	-1.10	-0.24	0.86	-0.01	0.15	0.16	1.53	1.62	0.09	-0.62	-0.9	0.28
47	-2.03	-1.01	1.02	0.80	0.44	0.36	0.11	0	0.11	-0.55	-0.24	0.31	0.47	0.15	0.32	1.67	1.62	0.05	0.58	-0.9	1.48
48	-0.21	-1.01	0.8	0.19	0.44	0.25	-0.07	0	0.07	0.04	-0.24	0.28	na	0.15	na	1.65	1.62	0.03	-1.57	-0.9	0.67
49	-2.30	-1.01	1.29	0.59	0.44	0.15	0.19	0	0.19	-0.53	-0.24	0.29	na	0.15	na	1.84	1.62	0.22	-2.93	-0.9	2.03
50	-1.76	-1.01	0.75	0.62	0.44	0.18	-0.32	0	0.32	-0.06	-0.24	0.18	0.42	0.15	0.27	1.75	1.62	0.13	-0.49	-0.9	0.41
51	-0.63	-1.01	0.38	-0.08	0.44	0.52	-0.11	0	0.11	0.09	-0.24	0.33	na	0.15	na	1.48	1.62	0.14	-1.25	-0.9	0.35
52	-0.31	-1.01	0.7	-0.44	0.44	0.88	0.35	0	0.35	0.03	-0.24	0.27	0.09	0.15	0.06	1.80	1.62	0.18	-1.12	-0.9	0.22
53	-0.65	-1.01	0.36	-0.05	0.44	0.49	0.48	0	0.48	-0.26	-0.24	0.02	0.19	0.15	0.04	1.47	1.62	0.15	-0.76	-0.9	0.14
54	-0.29	-1.01	0.72	1.08	0.44	0.64	0.28	0	0.28	-0.03	-0.24	0.21	0.16	0.15	0.01	1.42	1.62	0.2	-0.38	-0.9	0.52
55	-0.62	-1.01	0.39	0.62	0.44	0.18	0.39	0	0.39	-0.26	-0.24	0.02	-0.34	0.15	0.49	1.60	1.62	0.02	0.10	-0.9	1
56	-2.08	-1.01	1.07	0.61	0.44	0.17	-0.01	0	0.01	-0.91	-0.24	0.67	-0.01	0.15	0.16	2.04	1.62	0.42	-1.71	-0.9	0.81
57	-1.58	-1.01	0.57	0.36	0.44	0.08	0.51	0	0.51	-0.51	-0.24	0.27	na	0.15	na	1.42	1.62	0.2	-0.60	-0.9	0.3
58	-0.71	-1.01	0.3	-0.23	0.44	0.67	-1.25	0	1.25	-0.19	-0.24	0.05	na	0.15	na	1.87	1.62	0.25	-1.52	-0.9	0.62
59	-1.30	-1.01	0.29	0.78	0.44	0.34	-0.32	0	0.32	-0.15	-0.24	0.09	-0.08	0.15	0.23	1.62	1.62	0	-1.15	-0.9	0.25
60	-0.52	-1.01	0.49	0.63	0.44	0.19	-0.73	0	0.73	0.07	-0.24	0.31	0.17	0.15	0.02	1.25	1.62	0.37	-1.71	-0.9	0.81

61	-2.56	-1.01	1.55	0.78	0.44	0.34	-0.04	0	0.04	-0.39	-0.24	0.15	-0.10	0.15	0.25	1.69	1.62	0.07	-1.08	-0.9	0.18
62	-0.36	-1.01	0.65	1.58	0.44	1.14	0.28	0	0.28	-0.02	-0.24	0.22	na	0.15	na	1.67	1.62	0.05	-0.38	-0.9	0.52
63	-0.81	-1.01	0.2	-0.04	0.44	0.48	-0.07	0	0.07	-0.69	-0.24	0.45	na	0.15	na	1.50	1.62	0.12	-0.83	-0.9	0.07
64	-0.90	-1.01	0.11	1.51	0.44	1.07	-0.01	0	0.01	-0.27	-0.24	0.03	0.05	0.15	0.1	1.66	1.62	0.04	-1.12	-0.9	0.22
65	0.09	-1.01	1.1	0.51	0.44	0.07	-0.29	0	0.29	-0.35	-0.24	0.11	0.51	0.15	0.36	1.33	1.62	0.29	-0.98	-0.9	0.08
66	-2.14	-1.01	1.13	-0.28	0.44	0.72	-0.77	0	0.77	-0.24	-0.24	0	na	0.15	na	1.41	1.62	0.21	-1.26	-0.9	0.36
67	-1.57	-1.01	0.56	0.11	0.44	0.33	-0.74	0	0.74	-0.08	-0.24	0.16	na	0.15	na	1.64	1.62	0.02	-0.59	-0.9	0.31
68	-1.21	-1.01	0.2	1.16	0.44	0.72	0.52	0	0.52	0.04	-0.24	0.28	0.02	0.15	0.13	1.68	1.62	0.06	-0.96	-0.9	0.06
69	-1.77	-1.01	0.76	1.10	0.44	0.66	-0.59	0	0.59	-0.27	-0.24	0.03	na	0.15	na	2.01	1.62	0.39	-1.26	-0.9	0.36
70	-0.50	-1.01	0.51	-0.04	0.44	0.48	-0.35	0	0.35	-0.13	-0.24	0.11	-0.33	0.15	0.48	1.34	1.62	0.28	-1.24	-0.9	0.34
71	-0.31	-1.01	0.7	0.02	0.44	0.42	0.42	0	0.42	-0.02	-0.24	0.22	-0.16	0.15	0.31	1.50	1.62	0.12	-2.05	-0.9	1.15
72	-1.56	-1.01	0.55	1.06	0.44	0.62	-0.52	0	0.52	0.28	-0.24	0.52	0.22	0.15	0.07	1.52	1.62	0.1	-2.32	-0.9	1.42
73	-0.36	-1.01	0.65	-0.19	0.44	0.63	-0.01	0	0.01	0.17	-0.24	0.41	0.33	0.15	0.18	1.32	1.62	0.3	-1.10	-0.9	0.2
74	-0.53	-1.01	0.48	1.22	0.44	0.78	0.14	0	0.14	-0.34	-0.24	0.1	na	0.15	na	1.59	1.62	0.03	-0.92	-0.9	0.02
75	-0.35	-1.01	0.66	0.94	0.44	0.5	-0.41	0	0.41	-0.13	-0.24	0.11	na	0.15	na	1.75	1.62	0.13	-1.52	-0.9	0.62
76	-0.30	-1.01	0.71	0.90	0.44	0.46	-0.08	0	0.08	0.04	-0.24	0.28	0.41	0.15	0.26	1.45	1.62	0.17	-0.98	-0.9	0.08
77	-2.51	-1.01	1.5	0.62	0.44	0.18	0.31	0	0.31	-0.40	-0.24	0.16	na	0.15	na	1.29	1.62	0.33	-2.01	-0.9	1.11
78	-1.02	-1.01	0.01	-0.28	0.44	0.72	0.02	0	0.02	-0.32	-0.24	0.08	0.07	0.15	0.08	1.45	1.62	0.17	-1.34	-0.9	0.44
79	-1.15	-1.01	0.14	1.00	0.44	0.56	0.20	0	0.2	-0.87	-0.24	0.63	na	0.15	na	1.62	1.62	0	-0.82	-0.9	0.08
80	0.09	-1.01	1.1	1.04	0.44	0.6	0.38	0	0.38	-0.51	-0.24	0.27	0.22	0.15	0.07	2.00	1.62	0.38	-0.68	-0.9	0.22
81	-2.82	-1.01	1.81	0.43	0.44	0.01	-0.59	0	0.59	-0.66	-0.24	0.42	na	0.15	na	1.69	1.62	0.07	-0.56	-0.9	0.34
82	-1.06	-1.01	0.05	0.70	0.44	0.26	-0.09	0	0.09	-0.29	-0.24	0.05	0.56	0.15	0.41	2.17	1.62	0.55	-1.20	-0.9	0.3
83	-0.44	-1.01	0.57	1.01	0.44	0.57	0.12	0	0.12	-0.96	-0.24	0.72	na	0.15	na	1.55	1.62	0.07	-0.87	-0.9	0.03
84	-0.88	-1.01	0.13	0.91	0.44	0.47	0.09	0	0.09	-0.02	-0.24	0.22	na	0.15	na	1.68	1.62	0.06	-0.63	-0.9	0.27
85	-0.70	-1.01	0.31	-0.42	0.44	0.86	0.04	0	0.04	0.19	-0.24	0.43	0.37	0.15	0.22	1.65	1.62	0.03	-1.33	-0.9	0.43
86	-0.08	-1.01	0.93	1.40	0.44	0.96	-0.72	0	0.72	-0.44	-0.24	0.2	0.51	0.15	0.36	1.91	1.62	0.29	-0.69	-0.9	0.21
87	-1.32	-1.01	0.31	0.42	0.44	0.02	0.58	0	0.58	0.05	-0.24	0.29	na	0.15	na	1.96	1.62	0.34	0.17	-0.9	1.07
88	-0.56	-1.01	0.45	-0.13	0.44	0.57	0.07	0	0.07	-0.63	-0.24	0.39	0.05	0.15	0.1	1.67	1.62	0.05	-1.71	-0.9	0.81
89	-0.60	-1.01	0.41	1.06	0.44	0.62	0.01	0	0.01	-0.25	-0.24	0.01	na	0.15	na	1.80	1.62	0.18	-0.92	-0.9	0.02
90	-2.10	-1.01	1.09	0.16	0.44	0.28	0.54	0	0.54	-0.35	-0.24	0.11	-0.11	0.15	0.26	1.58	1.62	0.04	-1.11	-0.9	0.21
91	-0.50	-1.01	0.51	0.76	0.44	0.32	0.33	0	0.33	-0.15	-0.24	0.09	0.35	0.15	0.2	1.77	1.62	0.15	0.06	-0.9	0.96
92	-0.89	-1.01	0.12	0.65	0.44	0.21	-0.46	0	0.46	-0.16	-0.24	0.08	na	0.15	na	2.08	1.62	0.46	-0.83	-0.9	0.07
93	-1.86	-1.01	0.85	0.58	0.44	0.14	0.38	0	0.38	-0.24	-0.24	0	-0.30	0.15	0.45	1.95	1.62	0.33	0.41	-0.9	1.31

94		-1.01	0.96		0.44	0.33	-0.0001	0	0.0001	-0.30	-0.24	0.06	na	0.15	na	1.71	1.62	0.09	-0.33	-0.9	0.57
95	-1.97	-1.01	0.23	0.11	0.44	0.42	0.32	0	0.32	-0.08	-0.24	0.16	0.21	0.15	0.06	1.60	1.62	0.02	-0.79	-0.9	0.11
96	-0.78	-1.01	0.78	0.79	0.44	0.35	0.32	0	0.32	-0.36	-0.24	0.12	0.34	0.15	0.19	1.64	1.62	0.02	-0.23	-0.9	0.67
97	-2.13	-1.01	1.12	-0.32	0.44	0.76	0.40	0	0.4	-0.48	-0.24	0.24	0.13	0.15	0.02	1.32	1.62	0.3	-1.18	-0.9	0.28
98	-1.98	-1.01	0.97	0.38	0.44	0.06	-0.14	0	0.14	-0.46	-0.24	0.22	na	0.15	na	1.43	1.62	0.19	-1.53	-0.9	0.63
99	-0.49	-1.01	0.52	1.42	0.44	0.98	0.11	0	0.11	-1.06	-0.24	0.82	0.35	0.15	0.2	1.49	1.62	0.13	-1.06	-0.9	0.16
100	0.58	-1.01	1.59	-0.30	0.44	0.74	0.21	0	0.21	-0.07	-0.24	0.17	-0.13	0.15	0.28	1.50	1.62	0.12	-0.07	-0.9	0.83
101	-0.98	-1.01	0.03	0.23	0.44	0.21	-0.61	0	0.61	-0.88	-0.24	0.64	0.03	0.15	0.12	1.56	1.62	0.06	-0.88	-0.9	0.02
102	-0.53	-1.01	0.48	0.52	0.44	0.08	-0.03	0	0.03	0.05	-0.24	0.29	-0.43	0.15	0.58	1.51	1.62	0.11	0.05	-0.9	0.95
103	-1.91	-1.01	0.9	-0.24	0.44	0.68	0.87	0	0.87	-0.32	-0.24	0.08	na	0.15	na	1.73	1.62	0.11	-0.32	-0.9	0.58
104	-0.53	-1.01	0.48	-0.05	0.44	0.49	0.20	0	0.2	0.26	-0.24	0.5	na	0.15	na	1.48	1.62	0.14	0.26	-0.9	1.16
105	-0.70	-1.01	0.31	-0.08	0.44	0.52	-0.28	0	0.28	0.72	-0.24	0.96	na	0.15	na	1.67	1.62	0.05	0.72	-0.9	1.62
106	0.14	-1.01	1.15	-0.52	0.44	0.96	0.48	0	0.48	-2.18	-0.24	1.94	na	0.15	na	1.91	1.62	0.29	-2.18	-0.9	1.28
107	-2.20	-1.01	1.19	0.68	0.44	0.24	-0.51	0	0.51	-1.44	-0.24	1.2	0.07	0.15	0.08	1.60	1.62	0.02	-1.44	-0.9	0.54
108	-0.30	-1.01	0.71	0.48	0.44	0.04	0.25	0	0.25	-0.79	-0.24	0.55	-0.003	0.15	0.153	1.24	1.62	0.38	-0.79	-0.9	0.11
109	-0.46	-1.01	0.55	0.17	0.44	0.27	0.05	0	0.05	-0.66	-0.24	0.42	na	0.15	na	1.76	1.62	0.14	-0.66	-0.9	0.24
110	-0.33	-1.01	0.68	-0.18	0.44	0.62	0.32	0	0.32	-0.43	-0.24	0.19	na	0.15	na	1.55	1.62	0.07	-0.43	-0.9	0.47
111	-0.26	-1.01	0.75	0.49	0.44	0.05	0.30	0	0.3	-0.95	-0.24	0.71	-0.11	0.15	0.26	2.10	1.62	0.48	-0.95	-0.9	0.05
112	-2.37	-1.01	1.36	0.45	0.44	0.01	0.02	0	0.02	-1.45	-0.24	1.21	na	0.15	na	1.64	1.62	0.02	-1.45	-0.9	0.55

Text scores, register means, measures of distance. Dimension 4

Author	Business memos			Descriptions			Emails			Essays			Evaluations			Interviews			Text messages		
	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean
1	-0.95	-0.85	-0.10	-0.46	0.28	-0.74	0.00	0.39	-0.39	-0.22	-0.02	-0.20	na	-0.28	na	-0.54	-0.59	0.05	2.94	0.95	1.99
2	-0.63	-0.85	0.22	0.24	0.28	-0.04	-0.05	0.39	-0.44	-0.32	-0.02	-0.30	-0.01	-0.28	0.27	-0.60	-0.59	-0.01	-0.87	0.95	-1.82
3	-1.18	-0.85	-0.33	0.19	0.28	-0.09	0.87	0.39	0.48	-0.74	-0.02	-0.72	-0.06	-0.28	0.22	-0.20	-0.59	0.39	3.37	0.95	2.42
4	-0.55	-0.85	0.30	0.17	0.28	-0.11	0.13	0.39	-0.26	-0.48	-0.02	-0.46	na	-0.28	na	-0.82	-0.59	-0.23	1.00	0.95	0.05
5	-0.77	-0.85	0.08	0.67	0.28	0.39	1.44	0.39	1.05	-0.06	-0.02	-0.04	0.40	-0.28	0.68	-0.52	-0.59	0.07	2.22	0.95	1.27
6	-0.95	-0.85	-0.10	-0.64	0.28	-0.92	-0.23	0.39	-0.62	-0.37	-0.02	-0.35	na	-0.28	na	-1.13	-0.59	-0.54	1.79	0.95	0.84
7	-0.68	-0.85	0.17	-0.08	0.28	-0.36	0.61	0.39	0.22	0.25	-0.02	0.27	na	-0.28	na	-0.54	-0.59	0.05	-0.28	0.95	-1.23
8	-1.44	-0.85	-0.59	-0.06	0.28	-0.34	0.21	0.39	-0.18	-0.08	-0.02	-0.06	na	-0.28	na	-0.26	-0.59	0.33	3.38	0.95	2.43
9	-1.25	-0.85	-0.40	0.22	0.28	-0.06	-0.28	0.39	-0.67	0.05	-0.02	0.07	-0.76	-0.28	-0.48	-0.72	-0.59	-0.13	4.69	0.95	3.74
10	-0.68	-0.85	0.17	0.79	0.28	0.51	0.08	0.39	-0.31	-0.25	-0.02	-0.23	-0.60	-0.28	-0.32	-0.54	-0.59	0.05	1.01	0.95	0.06
11	-0.35	-0.85	0.50	0.39	0.28	0.11	0.03	0.39	-0.36	0.65	-0.02	0.67	-0.41	-0.28	-0.13	0.03	-0.59	0.62	-0.31	0.95	-1.26
12	-0.45	-0.85	0.40	1.14	0.28	0.86	0.35	0.39	-0.04	-0.31	-0.02	-0.29	-0.62	-0.28	-0.34	-0.66	-0.59	-0.07	4.82	0.95	3.87
13	-0.60	-0.85	0.25	-0.13	0.28	-0.41	0.28	0.39	-0.11	-0.40	-0.02	-0.38	-0.76	-0.28	-0.48	-0.36	-0.59	0.23	0.03	0.95	-0.92
14	-0.64	-0.85	0.21	0.86	0.28	0.58	0.56	0.39	0.17	-0.50	-0.02	-0.48	-0.23	-0.28	0.05	-0.41	-0.59	0.18	1.59	0.95	0.64
15	-1.25	-0.85	-0.40	-0.26	0.28	-0.54	0.98	0.39	0.59	0.01	-0.02	0.03	-0.42	-0.28	-0.14	-0.43	-0.59	0.16	-0.41	0.95	-1.36
16	-0.49	-0.85	0.36	-0.26	0.28	-0.54	-0.25	0.39	-0.64	-0.21	-0.02	-0.19	na	-0.28	na	-0.74	-0.59	-0.15	0.41	0.95	-0.54
17	-0.63	-0.85	0.22	-0.26	0.28	-0.54	1.02	0.39	0.63	0.20	-0.02	0.22	na	-0.28	na	-0.82	-0.59	-0.23	-0.90	0.95	-1.85
18	-0.79	-0.85	0.06	0.42	0.28	0.14	0.14	0.39	-0.25	-0.34	-0.02	-0.32	-0.59	-0.28	-0.31	-0.89	-0.59	-0.30	-0.54	0.95	-1.49
19	-0.30	-0.85	0.55	0.04	0.28	-0.24	0.51	0.39	0.12	-0.20	-0.02	-0.18	0.17	-0.28	0.45	-0.54	-0.59	0.05	-0.33	0.95	-1.28
20	-1.49	-0.85	-0.64	-0.25	0.28	-0.53	-0.10	0.39	-0.49	0.26	-0.02	0.28	-0.31	-0.28	-0.03	-1.03	-0.59	-0.44	-0.60	0.95	-1.55
21	-0.78	-0.85	0.07	0.05	0.28	-0.23	0.04	0.39	-0.35	0.11	-0.02	0.13	na	-0.28	na	-0.60	-0.59	-0.01	2.91	0.95	1.96
22	-1.25	-0.85	-0.40	0.02	0.28	-0.26	0.07	0.39	-0.32	0.20	-0.02	0.22	na	-0.28	na	-0.59	-0.59	0.00	3.51	0.95	2.56

432

Author	Business memos			Descriptions			Emails			Essays			Evaluations			Interviews			Text messages		
	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean
23	-1.41	-0.85	-0.56	-0.29	0.28	-0.57	0.59	0.39	0.20	-0.29	-0.02	-0.27	na	-0.28	na	-0.64	-0.59	-0.05	-0.28	0.95	-1.23
24	-0.71	-0.85	0.14	0.46	0.28	0.18	-0.12	0.39	-0.51	-0.53	-0.02	-0.51	-0.45	-0.28	-0.17	-0.78	-0.59	-0.19	-0.04	0.95	-0.99
25	-1.59	-0.85	-0.74	0.05	0.28	-0.23	0.41	0.39	0.02	0.08	-0.02	0.10	0.09	-0.28	0.37	-0.74	-0.59	-0.15	0.10	0.95	-0.85
26	-0.65	-0.85	0.20	1.46	0.28	1.18	1.15	0.39	0.76	0.00	-0.02	0.02	0.27	-0.28	0.55	-0.99	-0.59	-0.40	0.51	0.95	-0.44
27	-0.76	-0.85	0.09	-0.23	0.28	-0.51	1.40	0.39	1.01	-0.66	-0.02	-0.64	-0.47	-0.28	-0.19	-0.68	-0.59	-0.09	1.75	0.95	0.80
28	-0.86	-0.85	-0.01	1.46	0.28	1.18	-0.42	0.39	-0.81	-0.27	-0.02	-0.25	na	-0.28	na	-0.57	-0.59	0.02	-0.01	0.95	-0.96
29	-1.55	-0.85	-0.70	0.50	0.28	0.22	0.80	0.39	0.41	-0.86	-0.02	-0.84	na	-0.28	na	-0.74	-0.59	-0.15	3.83	0.95	2.88
30	-1.05	-0.85	-0.20	0.50	0.28	0.22	0.51	0.39	0.12	0.23	-0.02	0.25	0.09	-0.28	0.37	-0.35	-0.59	0.24	1.34	0.95	0.39
31	-1.23	-0.85	-0.38	0.55	0.28	0.27	0.36	0.39	-0.03	0.76	-0.02	0.78	-0.34	-0.28	-0.06	-0.55	-0.59	0.04	0.86	0.95	-0.09
32	-0.64	-0.85	0.21	-0.01	0.28	-0.29	-0.25	0.39	-0.64	0.27	-0.02	0.29	-0.58	-0.28	-0.30	-0.63	-0.59	-0.04	-0.81	0.95	-1.76
33	-0.75	-0.85	0.10	-0.05	0.28	-0.33	2.00	0.39	1.61	0.06	-0.02	0.08	0.53	-0.28	0.81	-0.60	-0.59	-0.01	0.68	0.95	-0.27
34	0.27	-0.85	1.12	0.49	0.28	0.21	0.98	0.39	0.59	0.06	-0.02	0.08	na	-0.28	na	-0.41	-0.59	0.18	1.53	0.95	0.58
35	-0.80	-0.85	0.05	-0.55	0.28	-0.83	0.62	0.39	0.23	0.03	-0.02	0.05	na	-0.28	na	-0.94	-0.59	-0.35	-0.40	0.95	-1.35
36	-0.51	-0.85	0.34	0.87	0.28	0.59	0.44	0.39	0.05	-0.08	-0.02	-0.06	0.08	-0.28	0.36	-0.15	-0.59	0.44	2.01	0.95	1.06
37	-0.44	-0.85	0.41	0.04	0.28	-0.24	0.53	0.39	0.14	-0.30	-0.02	-0.28	na	-0.28	na	-0.65	-0.59	-0.06	-0.49	0.95	-1.44
38	-1.16	-0.85	-0.31	0.04	0.28	-0.24	-0.07	0.39	-0.46	-0.21	-0.02	-0.19	na	-0.28	na	-0.81	-0.59	-0.22	1.47	0.95	0.52
39	-1.06	-0.85	-0.21	1.99	0.28	1.71	0.20	0.39	-0.19	0.62	-0.02	0.64	na	-0.28	na	-0.61	-0.59	-0.02	3.72	0.95	2.77
40	-1.09	-0.85	-0.24	0.56	0.28	0.28	1.03	0.39	0.64	0.25	-0.02	0.27	-0.24	-0.28	0.04	-0.31	-0.59	0.28	-0.27	0.95	-1.22
41	-0.71	-0.85	0.14	1.02	0.28	0.74	-0.70	0.39	-1.09	-0.58	-0.02	-0.56	-0.75	-0.28	-0.47	-0.69	-0.59	-0.10	-0.64	0.95	-1.59
42	-1.07	-0.85	-0.22	0.80	0.28	0.52	0.10	0.39	-0.29	0.05	-0.02	0.07	na	-0.28	na	-0.37	-0.59	0.22	1.69	0.95	0.74
43	-0.37	-0.85	0.48	-0.17	0.28	-0.45	0.09	0.39	-0.30	-0.77	-0.02	-0.75	-0.42	-0.28	-0.14	-0.51	-0.59	0.08	-1.01	0.95	-1.96
44	-1.13	-0.85	-0.28	-0.05	0.28	-0.33	0.67	0.39	0.28	-0.13	-0.02	-0.11	-0.68	-0.28	-0.40	-0.80	-0.59	-0.21	3.74	0.95	2.79
45	-1.64	-0.85	-0.79	-0.27	0.28	-0.55	1.52	0.39	1.13	0.86	-0.02	0.88	na	-0.28	na	-0.70	-0.59	-0.11	0.09	0.95	-0.86
46	-1.05	-0.85	-0.20	-0.14	0.28	-0.42	0.06	0.39	-0.33	-0.57	-0.02	-0.55	-0.31	-0.28	-0.03	-0.23	-0.59	0.36	-0.07	0.95	-1.02
47	-1.10	-0.85	-0.25	-0.01	0.28	-0.29	-0.08	0.39	-0.47	-0.35	-0.02	-0.33	-0.43	-0.28	-0.15	-0.69	-0.59	-0.10	-0.99	0.95	-1.94
48	-0.15	-0.85	0.70	1.20	0.28	0.92	0.47	0.39	0.08	0.02	-0.02	0.04	na	-0.28	na	-0.40	-0.59	0.19	1.70	0.95	0.75
49	-1.51	-0.85	-0.66	-0.52	0.28	-0.80	0.49	0.39	0.10	1.28	-0.02	1.30	na	-0.28	na	-0.60	-0.59	-0.01	6.84	0.95	5.89
50	-0.92	-0.85	-0.07	0.31	0.28	0.03	0.18	0.39	-0.21	-0.10	-0.02	-0.08	-0.35	-0.28	-0.07	-0.66	-0.59	-0.07	-0.12	0.95	-1.07
51	-1.11	-0.85	-0.26	1.24	0.28	0.96	0.85	0.39	0.46	0.11	-0.02	0.13	na	-0.28	na	-0.09	-0.59	0.50	-0.85	0.95	-1.80
52	-0.25	-0.85	0.60	1.08	0.28	0.80	-0.76	0.39	-1.15	-0.39	-0.02	-0.37	-0.39	-0.28	-0.11	-0.97	-0.59	-0.38	1.52	0.95	0.57
53	-0.25	-0.85	0.60	0.03	0.28	-0.25	-0.13	0.39	-0.52	0.54	-0.02	0.56	-0.52	-0.28	-0.24	-0.57	-0.59	0.02	-0.03	0.95	-0.98
54	0.18	-0.85	1.03	-0.32	0.28	-0.60	1.16	0.39	0.77	0.16	-0.02	0.18	-0.25	-0.28	0.03	-0.22	-0.59	0.37	1.52	0.95	0.57

Author	Business memos			Descriptions			Emails			Essays			Evaluations			Interviews			Text messages		
	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean
55	-1.15	-0.85	-0.30	-0.19	0.28	-0.47	0.34	0.39	-0.05	0.91	-0.02	0.93	-0.32	-0.28	-0.04	-0.72	-0.59	-0.13	-0.49	0.95	-1.44
56	-1.12	-0.85	-0.27	0.74	0.28	0.46	0.79	0.39	0.40	-0.52	-0.02	-0.50	-0.21	-0.28	0.07	-1.01	-0.59	-0.42	1.82	0.95	0.87
57	-1.25	-0.85	-0.40	0.89	0.28	0.61	0.50	0.39	0.11	-0.14	-0.02	-0.12	na	-0.28	na	-0.59	-0.59	0.00	0.37	0.95	-0.58
58	-0.52	-0.85	0.33	0.67	0.28	0.39	-0.34	0.39	-0.73	-0.33	-0.02	-0.31	na	-0.28	na	-1.01	-0.59	-0.42	-0.38	0.95	-1.33
59	-1.51	-0.85	-0.66	0.88	0.28	0.60	-0.46	0.39	-0.85	-0.17	-0.02	-0.15	-0.71	-0.28	-0.43	-0.90	-0.59	-0.31	1.86	0.95	0.91
60	-0.85	-0.85	0.00	0.21	0.28	-0.07	2.11	0.39	1.72	0.33	-0.02	0.35	-0.11	-0.28	0.17	-0.26	-0.59	0.33	0.95	0.95	0.00
61	-1.69	-0.85	-0.84	0.57	0.28	0.29	-0.58	0.39	-0.97	0.28	-0.02	0.30	-0.38	-0.28	-0.10	-0.72	-0.59	-0.13	2.74	0.95	1.79
62	-0.60	-0.85	0.25	0.00	0.28	-0.28	0.59	0.39	0.20	0.12	-0.02	0.14	na	-0.28	na	-0.64	-0.59	-0.05	1.14	0.95	0.19
63	-0.17	-0.85	0.68	-0.07	0.28	-0.35	1.13	0.39	0.74	0.15	-0.02	0.17	na	-0.28	na	-0.49	-0.59	0.10	1.00	0.95	0.05
64	-1.08	-0.85	-0.23	-0.74	0.28	-1.02	0.17	0.39	-0.22	-0.27	-0.02	-0.25	-0.33	-0.28	-0.05	-0.29	-0.59	0.30	-0.19	0.95	-1.14
65	-0.63	-0.85	0.22	-0.25	0.28	-0.53	0.83	0.39	0.44	0.93	-0.02	0.95	-0.83	-0.28	-0.55	-0.67	-0.59	-0.08	0.68	0.95	-0.27
66	-1.33	-0.85	-0.48	-0.67	0.28	-0.95	1.02	0.39	0.63	-0.54	-0.02	-0.52	na	-0.28	na	-0.84	-0.59	-0.25	4.44	0.95	3.49
67	-1.44	-0.85	-0.59	-0.21	0.28	-0.49	-0.55	0.39	-0.94	-0.21	-0.02	-0.19	na	-0.28	na	-0.58	-0.59	0.01	-0.77	0.95	-1.72
68	-1.23	-0.85	-0.38	0.01	0.28	-0.27	0.23	0.39	-0.16	-0.39	-0.02	-0.37	-0.14	-0.28	0.14	-0.46	-0.59	0.13	0.34	0.95	-0.61
69	-1.11	-0.85	-0.26	0.14	0.28	-0.14	2.01	0.39	1.62	-0.51	-0.02	-0.49	na	-0.28	na	-0.65	-0.59	-0.06	-0.72	0.95	-1.67
70	-0.49	-0.85	0.36	1.35	0.28	1.07	-0.54	0.39	-0.93	-0.33	-0.02	-0.31	-0.32	-0.28	-0.04	-0.18	-0.59	0.41	0.67	0.95	-0.28
71	-0.27	-0.85	0.58	-0.10	0.28	-0.38	0.01	0.39	-0.38	0.72	-0.02	0.74	-0.15	-0.28	0.13	-0.62	-0.59	-0.03	3.07	0.95	2.12
72	-1.46	-0.85	-0.61	-0.53	0.28	-0.81	0.32	0.39	-0.07	0.12	-0.02	0.14	-0.74	-0.28	-0.46	-0.20	-0.59	0.39	-0.16	0.95	-1.11
73	-0.97	-0.85	-0.12	1.61	0.28	1.33	-0.15	0.39	-0.54	0.68	-0.02	0.70	0.60	-0.28	0.88	-0.41	-0.59	0.18	1.23	0.95	0.28
74	-0.50	-0.85	0.35	-0.33	0.28	-0.61	-0.99	0.39	-1.38	-0.42	-0.02	-0.40	na	-0.28	na	-0.43	-0.59	0.16	0.97	0.95	0.02
75	-0.69	-0.85	0.16	0.73	0.28	0.45	1.25	0.39	0.86	0.01	-0.02	0.03	na	-0.28	na	-0.88	-0.59	-0.29	-0.04	0.95	-0.99
76	-0.02	-0.85	0.83	-0.21	0.28	-0.49	0.56	0.39	0.17	0.15	-0.02	0.17	-0.62	-0.28	-0.34	-0.46	-0.59	0.13	3.08	0.95	2.13
77	-1.33	-0.85	-0.48	0.28	0.28	0.00	0.80	0.39	0.41	-0.02	-0.02	0.00	na	-0.28	na	-1.15	-0.59	-0.56	1.37	0.95	0.42
78	-0.78	-0.85	0.07	-0.26	0.28	-0.54	-0.34	0.39	-0.73	0.77	-0.02	0.79	-0.32	-0.28	-0.04	-0.50	-0.59	0.09	0.41	0.95	-0.54
79	-0.88	-0.85	-0.03	-0.07	0.28	-0.35	0.74	0.39	0.35	-0.20	-0.02	-0.18	na	-0.28	na	-0.54	-0.59	0.05	-0.44	0.95	-1.39
80	-0.07	-0.85	0.78	-0.36	0.28	-0.64	0.23	0.39	-0.16	-0.46	-0.02	-0.44	0.03	-0.28	0.31	-1.05	-0.59	-0.46	1.23	0.95	0.28
81	-1.18	-0.85	-0.33	0.22	0.28	-0.06	0.99	0.39	0.60	-0.16	-0.02	-0.14	na	-0.28	na	-0.86	-0.59	-0.27	0.29	0.95	-0.66
82	-1.40	-0.85	-0.55	-0.06	0.28	-0.34	0.16	0.39	-0.23	-0.11	-0.02	-0.09	-0.96	-0.28	-0.68	-0.50	-0.59	0.09	1.08	0.95	0.13
83	-0.70	-0.85	0.15	0.25	0.28	-0.03	0.19	0.39	-0.20	-0.58	-0.02	-0.56	na	-0.28	na	-0.34	-0.59	0.25	0.46	0.95	-0.49
84	-1.27	-0.85	-0.42	0.43	0.28	0.15	0.54	0.39	0.15	0.18	-0.02	0.20	na	-0.28	na	-0.39	-0.59	0.20	-0.84	0.95	-1.79
85	-0.87	-0.85	-0.02	0.32	0.28	0.04	1.79	0.39	1.40	0.57	-0.02	0.59	-0.27	-0.28	0.01	-0.87	-0.59	-0.28	-0.03	0.95	-0.98
86	-0.39	-0.85	0.46	0.50	0.28	0.22	0.52	0.39	0.13	0.39	-0.02	0.41	0.31	-0.28	0.59	-0.64	-0.59	-0.05	-0.05	0.95	-1.00

Author	Business memos			Descriptions			Emails			Essays			Evaluations			Interviews			Text messages		
	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean	Text score	Register mean	Difference from register mean
87	-1.07	-0.85	-0.22	0.17	0.28	-0.11	0.15	0.39	-0.24	0.49	-0.02	0.51	na	-0.28	na	-0.26	-0.59	0.33	0.08	0.95	-0.87
88	-0.52	-0.85	0.33	2.04	0.28	1.76	1.03	0.39	0.64	0.05	-0.02	0.07	-0.26	-0.28	0.02	-0.63	-0.59	-0.04	-0.60	0.95	-1.55
89	-0.77	-0.85	0.08	0.41	0.28	0.13	0.44	0.39	0.05	-0.01	-0.02	0.01	na	-0.28	na	-0.54	-0.59	0.05	-1.13	0.95	-2.08
90	-1.64	-0.85	-0.79	-0.17	0.28	-0.45	-0.14	0.39	-0.53	0.03	-0.02	0.05	-0.98	-0.28	-0.70	-0.76	-0.59	-0.17	0.51	0.95	-0.44
91	-0.81	-0.85	0.04	0.33	0.28	0.05	-0.81	0.39	-1.20	0.39	-0.02	0.41	0.53	-0.28	0.81	-0.50	-0.59	0.09	0.03	0.95	-0.92
92	0.29	-0.85	1.14	-0.25	0.28	-0.53	1.24	0.39	0.85	0.36	-0.02	0.38	na	-0.28	na	-0.34	-0.59	0.25	3.14	0.95	2.19
93	-1.06	-0.85	-0.21	-0.21	0.28	-0.49	1.09	0.39	0.70	-0.11	-0.02	-0.09	-0.62	-0.28	-0.34	-0.57	-0.59	0.02	-0.27	0.95	-1.22
94	-0.74	-0.85	0.11	0.36	0.28	0.08	0.17	0.39	-0.22	0.23	-0.02	0.25	na	-0.28	na	-1.03	-0.59	-0.44	-0.58	0.95	-1.53
95	-0.63	-0.85	0.22	0.80	0.28	0.52	0.66	0.39	0.27	0.60	-0.02	0.62	-0.26	-0.28	0.02	-0.70	-0.59	-0.11	-0.36	0.95	-1.31
96	-0.70	-0.85	0.15	0.52	0.28	0.24	0.95	0.39	0.56	-0.21	-0.02	-0.19	-0.43	-0.28	-0.15	-0.32	-0.59	0.27	0.47	0.95	-0.48
97	-1.29	-0.85	-0.44	1.06	0.28	0.78	-0.27	0.39	-0.66	-0.26	-0.02	-0.24	0.33	-0.28	0.61	0.16	-0.59	0.75	-0.52	0.95	-1.47
98	-1.42	-0.85	-0.57	1.03	0.28	0.75	0.87	0.39	0.48	-0.23	-0.02	-0.21	na	-0.28	na	-1.06	-0.59	-0.47	-0.07	0.95	-1.02
99	-0.80	-0.85	0.05	-0.16	0.28	-0.44	0.22	0.39	-0.17	0.08	-0.02	0.10	-0.12	-0.28	0.16	-0.22	-0.59	0.37	0.77	0.95	-0.18
100	-0.82	-0.85	0.03	0.04	0.28	-0.24	-0.16	0.39	-0.55	-0.36	-0.02	-0.34	-0.38	-0.28	-0.10	-0.59	-0.59	0.00	1.02	0.95	0.07
101	-0.56	-0.85	0.29	0.48	0.28	0.20	0.67	0.39	0.28	0.06	-0.02	0.08	-0.23	-0.28	0.05	-0.29	-0.59	0.30	0.90	0.95	-0.05
102	-0.15	-0.85	0.70	-0.32	0.28	-0.60	0.94	0.39	0.55	-0.31	-0.02	-0.29	-0.79	-0.28	-0.51	-0.65	-0.59	-0.06	0.21	0.95	-0.74
103	-1.60	-0.85	-0.75	1.18	0.28	0.90	-0.74	0.39	-1.13	0.57	-0.02	0.59	na	-0.28	na	-0.66	-0.59	-0.07	0.16	0.95	-0.79
104	-1.02	-0.85	-0.17	0.94	0.28	0.66	-0.33	0.39	-0.72	-0.32	-0.02	-0.30	na	-0.28	na	-0.63	-0.59	-0.04	-1.46	0.95	-2.41
105	-0.09	-0.85	0.76	1.24	0.28	0.96	0.23	0.39	-0.16	0.51	-0.02	0.53	na	-0.28	na	-0.83	-0.59	-0.24	0.16	0.95	-0.79
106	-0.74	-0.85	0.11	0.22	0.28	-0.06	-0.51	0.39	-0.90	0.75	-0.02	0.77	na	-0.28	na	0.04	-0.59	0.63	6.15	0.95	5.20
107	-1.31	-0.85	-0.46	0.17	0.28	-0.11	1.22	0.39	0.83	-0.28	-0.02	-0.26	0.19	-0.28	0.47	-0.81	-0.59	-0.22	3.75	0.95	2.80
108	-0.24	-0.85	0.61	-0.51	0.28	-0.79	-0.17	0.39	-0.56	-0.22	-0.02	-0.20	-0.42	-0.28	-0.14	-0.54	-0.59	0.05	1.99	0.95	1.04
109	-0.87	-0.85	-0.02	-0.01	0.28	-0.29	-0.39	0.39	-0.78	-0.58	-0.02	-0.56	na	-0.28	na	-0.53	-0.59	0.06	0.25	0.95	-0.70
110	-0.35	-0.85	0.50	1.12	0.28	0.84	-0.01	0.39	-0.40	-0.82	-0.02	-0.80	na	-0.28	na	-0.50	-0.59	0.09	2.61	0.95	1.66
111	-0.56	-0.85	0.29	0.40	0.28	0.12	0.60	0.39	0.21	-0.80	-0.02	-0.78	0.36	-0.28	0.64	-1.05	-0.59	-0.46	2.86	0.95	1.91
112	-1.95	-0.85	-1.10	0.28	0.28	0.00	2.25	0.39	1.86	0.32	-0.02	0.34	na	-0.28	na	-0.68	-0.59	-0.09	3.18	0.95	2.23