**Migration Letters**

Volume: 21, No: S9 (2024), pp. 1297-1326

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

# English Speech And Commentary: A Corpus Based Genre Analysis

*Jamal Shah[1] , Dr. Ayaz Ahmad[2] , Maryam Saba[3] , Sobia Siraj[4] , Shehzad[5]

**Abstract**

*The study explores the use of language in English commentary and speech using corpus-based genre analysis. The corpus includes English cricket commentaries and speeches of Barack Obama, categorized under various linguistic/lexical categories. Doglous Biber's (1988) Multidimensional Analysis has been used to quantitatively identify and compare the textual dimension and genre types of these speeches and commentaries. The comparative analysis revealed that commentary's textual dimension is positive on D2 and negative on D1, D3, D4, D5, and D6, while speech's textual dimension is positive on D3, D4, D6, and negative on D1, D2, D5. Biber's dimensions show the most similar and closest genre of commentary on dimension 1, 4 and 6 is broadcast, dimension 2 and 5 is prepared speeches and dimension 3 is general fiction, while speech's closest genre on dimension 1 is general fiction, dimension 2, 3 and 6 is academic prose, dimension 4 is personal letters and dimension 5 is press reportage.*

***Key Words:*** *Corpus-based Genre analysis, Commentary, Speech, Multidimensional Analysis.*

## 1. Introduction

This research aims to analyze English speech and commentary using corpus-based genre analysis. Genres are distinct types of texts with specific forms, styles, and content that authors adhere to. According to Bhatia (2004), genre differs across different disciplines and registers, demonstrating how arguments and evidence are accepted in one setting but rejected in anothe[1]r. Genre analysis is performed, to learn more about the text being used in a discourse community. Discourses within disciplines have specific conventions and guidelines in the world that are followed by the writers and speaker. According to Bhatia (1993), other criteria such as content, shape, channel or medium, and target audience, influence the genre's nature and creation, it is essentially defined by the communicative goal, designed to achieve. The internal structure of the genre is shaped by a shared set of communicative goals. As a result, the material is divided into many genres based on language being used. Obin, Dellwo, Lacheret, & Rodet, (2010) stated that depending on the setting in which it is delivered, speech can be divided into different discourse genres (e.g., "political speech, sport commentary speech, etc."). If speech and commentary are considered two different genres, then it is rightly said that each genre has a unique structure, layout, content, style, specific audience, and context in which it is used. Speech in the language is vocal communication between people, and each language uses

[1]*Lecturer, Department of English, Abdul Wali Khan University Mardan, Pakistan*
[2]*Professors and Head, Department of English, Abdul Wali Khan University Mardan, Pakistan.*
[3]*MPhil Scholar, Department of English,Abdul Wali Khan University Mardan, Pakistan*
[4]*Junior Teaching Mistress, Fazle Haq College Mardan, Pakistan.*
[5]*Southwest University Chongqing China ( School of International Studies, and English Literature)*
**Corresponding Email: jamalshahjs1996@gmail.com*

phonetic combinations of vowel and consonant sounds to create its words. In linguistics, speech is a system of communication that uses spoken words (or sound symbols). Commentary is an opinions expression or offering of explanations regarding a situation or events. According to Crystal and Davy (1969), "a spoken account of events which are taking place" is the common definition of the word commentary. According to Ferguson (1983), "It is the oral reporting of an ongoing activity, combined with the provision of background information and interpretation" (pp. 155–156). Speech and commentary constitute a significant portion of linguistic investigation from various perspectives.

The study focuses on the corpora of English discourses converted into text. These include different political English native speakers' speeches and cricket commentaries by commentators. Comparative corpus-based studies of speech and commentary revealed differences in textual dimensions and genre types, as per Biber's multidimensional analysis. The structure comprehension of speech and commentary indicates how to communicate in a comprehensible and effective manner. This study analyzes written discourses in English, including speeches by political English native speakers and cricket commentaries by commentators. Understanding these genres are essential for effective communication and understanding the structure of these documents.

### 1.1 Statement of the Research Problem

This research compares the corpus of commentary and speech in terms of textual dimensions and genre types. Communication follows specific channels, layouts, formats, linguistic patterns, and co-occurrences of features. The study identifies and compares the corpora of commentaries and speeches, focusing on their specific language features. Despite previous researches analyzing these genres from various perspectives, this study uses corpus-based genre analysis to investigate the use of language in English commentary and speech. The goal is to determine if these genres should be classified as different genres based on Biber's multidimensional analysis.

### 1.2 Research Questions

1. What are the textual dimensions in genre of commentary and speech?
2. What are the genre types of commentary and speech according to Biber’s multidimensional analysis?

## 2. Literature Review

In his study, Bhatia (2002) discussed genre analysis from two perspectives. He claims that it either represents the institutionalized communication of a discourse community or a tool for creating different teaching procedures and approaches that may be used in activities of the classroom. He described the genre as an inquiry into the conventions of a text in a specific field, as well as how participants of a specific disciplinary discourse community produce, evaluate and interpret distinct genres to bridge the communication gap. Therefore, speech and sports commentary are also considered different genres, having specific lexis, structures, and context.

Swales' creative work "Genre Analysis" (1990) explores the complexity that defines genre precisely, and he proposes the concept of a discourse community. He defined it, as a set of persons who cooperate to attain a common set of objectives. In order to attain their common

aims, these individuals are conversant with a variety of written and spoken communication formats (Swale, 1990). Swale (1990) argues, "a genre comprises a class of communicative events, the members of which share some set of communicative purposes" (p.58). Similarly, the speeches highlight one's opinion, ideas, views, etc. on different seniors and commentaries highlight the current seniors of different sports. The current study analyses these documents' written discourses in the English language.

Biber (1991) proposed a multidimensional technique to analyze registers. He looked at the differences between writing and speech and chose 67 linguistic variables to look at in both written and spoken data. These linguistic traits are thought to have a communication purpose in such texts. The purpose of the study was to better comprehend the distinctions between English's written and spoken registers. According to Swales and Feak (2000), one needs to consider elements like structure, purpose, presentation, and, most crucially, audience to categorize a genre. Swales put out the notion that moves can be utilized to identify recurring patterns in works belonging to a move's genre.

Yaakob (2014) conducted a corpus-based genre analysis of the lecture's introduction. In order to identify the generic features of the language used in the lectures, the study concentrated on the communication functions of the first section of the lecture and the disciplinary differences in lecture language. The BASE corpus was used to collect the data (British Academic Spoken English). 160 lectures and 39 seminars from various academic disciplines of the university were included in its corpus that were recorded on various occasions. These lectures came from four different areas of knowledge: "Arts and Humanities, Life Sciences, Physical Sciences, and Social Sciences." He compiled 89 lectures from lectures sub-genre introduction in the second portion of his investigation. In order to analyze the data, the corpus tool Wordsmith 5 was used. He looked at how often certain terms were used, as well as the employment of prepositions, infinitive markers, lexical bundles, and discourse markers. The findings of the study demonstrated that lectures as a genre had disciplinary variances, and the data corpus's analysis gave us a thorough description of the lexis used in these courses.

Villanueva, Dolom, and Belen (2018) investigated web-based genre analysis. They did their investigation on the university websites' "About Us" sections. They chose the websites of universities affiliated with the "Asian Association of Open Universities." The World Wide Web is used by these colleges to present themselves to the community. They generated a website corpus of 41 universities and analyzed the data using a corpus tool such as Antcone. The text was analyzed on three levels: keyword in context, concordance lines, and identification of moves. The study revealed that the keywords of the text portray the university image positively and that the rhetorical moves revealed that the language is informative and promotional in character used in the websites, with a communicative objective.

Amjad & Shakir (2014) view that this study observes the use of language in online university prospectuses using the first dimension of Biber's (1988) multidimensional analysis, which contrasts involved creation with informative production. The methodology utilized in this research is based on the corpus. The corpus was gathered from the Pakistan, UK, and India libraries' sections on "about us," "facilities," "departments," and vice chancellor messages. The data is tagged and examined using a multidimensional tagger (v. 1.0). By the first dimension of a multidimensional analysis, scores, and co-occurrences of linguistic features are computed statistically and qualitatively analyzed. The findings demonstrate the informative nature of online university prospectuses. According to their most negative score, India is informationally rich, while the UK is less informational because of the least negative scores from a cross-cultural perspective. Also, it has been shown that vice-chancellors tend to be less instructional

and more interactive than department sections. Very informative texts contain a wealth of information and are expertly written.

Ahmad, Naveed and Shakir (2014) examine the importance of non-narrative elements in Pakistan's online university prospectus. It was decided to do a corpus-based analysis. Multidimensional Analysis Tagger v 1.3 was used to tag and analyze the data during the analysis, with an emphasis on Dimension 2. According to the findings, factual/informational, argumentative, and non-narrative language is utilized in the online university prospectus.

Obin, Dellwo, Lacheret, & Rodet, (2010) stated that depending on the setting in which it is delivered, speech can be divided into different discourse genres (e.g., "political speech, sport commentary speech, etc."). This study sought to determine whether listeners could distinguish speech from various discourse genres only based on acoustic prosodic signals. 70 listeners with diverse levels of French expertise (no speakers, nonnative speakers, and native speakers) were asked to identify four different types of discourse genres in a perception experiment using delexicalized speech "church service, political, journal, and sports commentary." The results demonstrated a reasonable capacity to identify objects, with performance improving significantly with more French exposure. Discourse genres were grouped based on their perceptual similarities using identification confusion (pp. 3070-3073).

Meier-Vieracker (2021) argued that since the late 1990s, when football games online live text commentaries became popular, the genre has evolved significantly. Although linguistic research on live text commentary emphasizes its uniqueness, football has used this style of commentary for at least twenty years. Diachronic investigations, however, are still lacking. This study presents a corpus linguistic analysis of genre- and register-specific features of German live text commentary from 2003 to 2020. It emphasizes the distribution of linguistic features across time at various language levels (such as syntactical, lexical, graphemic, etc.). A more impersonal reporting style is observed as many signals that convey a colloquial tone and simulate orality in the written mode decline. Furthermore, individual perspective markers decline when one adopts a neutral attitude. As a result, the development of online text commentary might be defined as a standardized process (pp. 274-299).

This study attempts to define an emergent media genre, or an electronic register as the author prefers to refer to it: online sports commentary (OSC). The study is based on D. Biber and S. Conrad's (2009) methodological framework, which emphasizes the connection between situational factors and lexico-grammatical features of a language variety in register studies and focuses on English-language online live soccer match reports. In order to prove that the online commentary register combines spoken and written language, the study also plans to contrast it with other pertinent types, such as written sports commentary (WSC) and sports announcer talk (SAT). (Lewandowski, 2012, p. 65).

To conclude the discussion, speech and commentary are distinct genres due to their unique structure, content, style, audience, and context. While there is extensive research on these genres from various perspectives, there is a lack of focus on comparative corpus-based studies. This study aims to understand the textual dimensions and genre types of speech and commentary using Biber's multidimensional analysis, as well as the communicative purpose of these genres.

## 3. Research Methodology

### 3.1 Research Paradigm and Methods

The study uses a quantitative approach, categorizing data based on grammatical classes, frequency, and grammar annotations. It aims to identify the dimensions and genre types in Cricket's Commentaries and orators' speeches using this quantitative approach.

### 3.2. Data Collection and Sample

This study uses ESPNCRICINFO's cricket website for commentary data and Google's website for speech data. English speeches and commentaries from 2009 onwards are chosen for the data. A purposive non-random sampling technique is used to select the best fit population representation. The data consists of the last five T20 cricket world cup final matches commentaries and speeches of Barak Obama (2009-2017). Google webpages are used for the corpus, which is created in MS-word documents and converted into a text file using a corpus tool.

### 3.4 Data Analysis

The framework for this investigation is Biber's multidimensional analysis. A multi-dimensional tagger (MAT) is a Windows program that replicates by Biber (1988) for the analysis of English text. Six dimensions are used in this framework. To comprehend the nature of this dimension, it's critical to understand how these features of linguistics occur with each other frequently. Factors are also referred to in these dimensions. Six factors/dimensions recognized by Biber can be summed up as follows:

- **3.4.1.1 Dimension 1: Involved vs informational.**
- **3.4.1.2 Dimension 2: Narrative vs non-narrative**
- **3.4.1.3 Dimension 3: Context-Dependence vs Context-Independence**
- **3.4.1.4 Dimension 4: Measures Overt Expression of Persuasion**
- **3.4.1.5 Dimension 5: Abstract vs non-abstract**
- **3.4.1.6 Dimension 6: Informational Elaboration and Time Constraints**

### 3.5 Data Analysis Tools

#### 3.5.2 Multidimensional Analysis Tagger (V.1.3.3)

The Multidimensional Analysis Tagger (MAT) version 1.3.3 is a corpus tool used for investigating linguistic features. It produces a text corpus with grammatical annotations and plots the input text using Biber's six dimensions for genre development. The tool recognizes the frequency and percentages of linguistic features in the corpus data, allowing for a comprehensive understanding of the text.

## 4. Analysis and Discussion

This portion is based on data analysis and discussion on results. The multidimensional analysis proposed by Biber (1988) has been used to analyse the data. The model examined six dimensions in each genre corpus. Different textual dimensions and genre types are displayed by each genre. The genre of speech and commentary have been compared along each dimension. A variety of genre excerpts have been included to support the findings.

### 4.1.2 Dimension 1: 'Involved vs informational production.

The first-dimension graph divides the input data into two halves (informational versus involved), so the discourse is examined under mentioned the dimension. Biber (1988) asserts that a low score on dimension 1, denotes that the text is informative, whereas a high score on this dimension denotes an interactive and conversational text.

**4.1.2.1: Dimension 1: 'Involved vs informational production' for commentary.**

| Dimension 1 | Mean value |
|---|---|
| **CORPUS** | **-6.6** |

The data's graphical representation provides statistical information, with the data positioned at the bottom of the Dimension 1 graph. The corpus has low scores, indicating an emphasis on information rather than involvement. The mean score of -6.6 indicates accurate linguistic features, with nouns, adjectives, verbs, and long words being identified as linguistic qualities on Dimension 1.

**Text Type**

The negative D1 score indicates an average D1 score, which helps identify different text types. The closest genre is broadcast, with a mean value of -4.3. "Cricket Commentary" has average D1 scores similar to broadcast. Cricket commentary is informational, similar to broadcast. Broadcasts often have poor ratings due to lack of target audience and emphasis on unimportant or uninvolved external events. Broadcasts are often criticized for their lack of interactive and effective external events (Biber, 1988).

**4.1.2.2 Dimension 1: 'Involved vs informational production' for Speech.**

| Dimension 1 | Mean Value |
|---|---|
| **CORPUS** | **-0.4** |

The data's graphical representation provides statistical information, with the data positioned at the bottom of the Dimension 1 graph. The current corpus's mean score is -0.4. The corpus has low scores, indicating an emphasis on information rather than involvement. Dimension 1 linguistic features indicate that these features have been accurately tagged by the Biber tagger. Biber (1988) claims these linguistic features recognized on this dimension by the tagger are nouns, adjectives, verbs, and other long.

**Text Type**

The negative D1 score indicates an average D1 score, which helps identify different text types. The closest genre is general fiction having mean value of -0.8. Speech has average D1 scores that are most like those of general fiction. The comparison reveals that the speeches have characteristics that are clearly informational like those of the closest genre. Biber (1988) asserts that general fiction receives a comparatively low score because it lacks a clear addressee and concentrates on non-interactive, uninvolved external occurrences.

**4.1.2.3: Comparison of Commentary and Speech on Dimension 1:**

The text of commentary and speech comparison show that both the genres having negative mean scores of -6.6 and -0.4 respectively. Commentaries have lower mean scores, making them more informative, while the speeches corpus is less informative in nature because of low values of mean scores. The text type comparison reveals broadcast is the closest genre to the commentary's corpus, while general fiction is the closest genre to the speech's corpus.

**4.1.3: Dimension 2: Narrative vs non-Narrative**

The 2nd dimension of the graph divides data into two halves to examine the contrast between narrative and non-narrative texts. Biber (1988) asserts that low score indicates non-narrative text, while a high score indicates narrative text. Biber (1988) asserts that the highest values are found in fictional genres, which frequently use pronouns of the third person and past tense verbs, etc. While genres like broadcast, academic writing, and official papers have the lowest values due to the extremely low frequencies of pronouns of third person and past tense verbs, and public speeches, biographies, humor, press reports, and personal letters have intermediate levels on dimension 2 (p. 97).

**4.1.3.1: Dimension 2: Narrative vs non-Narrative for commentary**

| **Dimension 2** | **Mean Value** |
|---|---|
| **CORPUS** | **2.23** |

The Biber (1988) analysis of commentary genre reveals a narrative versus non-narrative dimension. The data's graphical depiction at the top of the graph shows that the commentaries' corpus has high score, indicating a strong emphasis on narrative rather than non-narrative features. The current corpus's mean score is 2.23. The Biber tagger accurately identifies linguistic features such as past and present tense verbs, third person pronouns, adjectives, and public verbs, among others.

**Text type**

The current corpus has a mean score is 2.23 on D2, making it very highly narrative rather than non-narrative. The text type comparison shows that the closest genre is prepared speeches to the supplied corpus, with a mean score of 0.7, "commentaries" get average D2 ratings that are most comparable to prepared speeches while comparing the different genres. The closest genre comparison shows that the commentary shares qualities with prepared speeches.

**4.1.3.2: Dimension 2: Narrative versus Non-Narrative for speech**

| Dimension 2 | Mean Value |
|---|---|
| **CORPUS** | **-2.12** |

The text analyzes the narrative versus non-narrative component of Biber's data, indicating which text type corresponds to the speech genre (1988). The data is presented graphically, with a low score indicating non-narrative features. The present corpus has a mean score of -2.12, indicating a greater focus on non-narrative elements than narrative ones. The Biber tagger correctly identified linguistic features such as past and present tense verbs, third-person pronouns, adjectives, and public verbs.

**Text type**

The current corpus has the mean score is -2.12 on D2, making it very non-narrative. The text type comparison shows that academic prose is the closest genre to the supplied corpus, with a mean score of -2.6, "Speech" gets average D2 ratings that are most comparable to academic prose while comparing the different genres. The closest genre comparison shows that the speech shares qualities with academic prose that are obviously non-narrative like the closest genre.

**4.1.3.3: Comparison of Commentary and Speech on Dimension 2**

Commentary, which has 2.23 as the mean score, lends narrative value to the text while speech has the mean score of -2.12, which lend non-narrative value to the text. The text type comparison shows that the prepared speeches are the closest genre to the commentary's corpus, while the closest genre is academic prose to the supplied corpus of speeches. The former corpus has positive scores because of having high frequencies of positive loading linguistic features, while the latter one has high frequencies of negative loading linguistic features on dimension 5.

**4.1.4 Dimension 3: Context Dependent vs Context Independent**

Biber (1988) defines dimension 3 as a measure of a text's context-dependent and context-independent features. Higher scores indicate context independence, while lower scores indicate context dependence. High scores indicate numerous nouns, while low scores indicate numerous adverbs. Additionally, according to Biber (1988), the presence of numerous WH-relative clauses in subject and object positions indicates a definite reference, while lack indicates context independence.

**4.1.4.1 Dimension 3: Context Dependent versus Context Independent for commentary**

| D3 | Mean |
|---|---|
| **CORPUS** | **-3.06** |

The corpus text is graphically represented that the data has low scores, making it context-dependent. The graph reveals that the mean D3 scores are -3.06, indicating that cricket commentator language is context-sensitive. Below are listed the prevalent linguistic traits on dimension 3.

**Text Type**

The current corpus has the mean score is -3.06 on D3, making it very context-dependent. The text type comparison shows that the supplied corpus's closest genre is general fiction, with a mean score of -3.1, "Commentary" gets average D3 ratings that are most comparable to general fiction while comparing the different genres. The closest genre comparison shows that the commentary shares qualities with general fiction that are obviously situational dependent on the context like the closest genre.

#### 4.1.4.2 Dimension 3: Context Dependent versus Context Independent for Speech

| Dimension 3 | Mean Value |
|---|---|
| **CORPUS** | **4.35** |

The text of the speech's corpus is represented graphically, and it is clear that the data is independent on the context because it is located above the graph and has high scores. The graph demonstrates that the mean D3 scores are 4.35, indicating that the language used in speeches by orators independent on the context. Here are listed the prevalent linguistic traits on dimension 3.

**Text Type**

The current corpus having the mean score is 4.35 on D3, making it very independent on context. The text type comparison shows that the supplied corpus's closest genre is academic prose, with a mean score of 4.2, "Speech" gets average D3 ratings that are most comparable to academic prose while comparing the different genres. The closest genre comparison shows that the commentary shares qualities with academic prose that are obviously situational and independent on the context like the closest genre.

#### 4.1.4.3 Comparison of Commentary and Speech on Dimension 3:

Commentary, which has the mean score -3.06 on dimension 3 makes the text dependent on context while speech has the mean score of 4.35, which lends context-independent value to the text. The text type comparison shows that the closest genre to the commentary's corpus is general fiction, while the closest genre is academic prose to the supplied corpus of speeches.

### 4.1.5 Dimension 4: Overt Expression of Persuasion

Dimension 4 measures the overt expression of persuasion in language, with positive and negative values in two equal halves. High scores indicate the author's point of view, while low ratings indicate objectivity. Biber (1988), multidimensional analysis of text provides details on persuasion expressions, indicating the author's perspective or objectivity.

#### 4.1.5.1 Dimension 4: Overt Expression of persuasion for commentary

| Dimension 4 | Mean Value |
|---|---|
| **CORPUS** | **-3.07** |

According to the graphical expression of D4, the corpus of commentary is in the area with the negative scores. The mean score of the present corpus's data is -3.07 which demonstrates that the language of commentary lacks arguments and has a more factual and assertive style on dimension 4. When speakers address the audience, there are some components of persuasion

present to make their point of view appear true in a highly factual and assertive manner. The linguistic characteristics on dimension 4 are listed below.

**Text Type**
The text is general and less persuasive, with a mean score of -3.07 on D4. The closest genre to the provided corpus is broadcast, with a mean score of -4.4. The comparison reveals that "commentary" receives average D4 ratings and is most likely to be broadcast. The commentary shares characteristics with broadcasts that are unpersuasive in tone. Biber (1988) claim that broadcasts are basic reports of occurrences without argumentation or opinion (p. 151).

**4.1.5.2 Dimension 4: Overt Expression of persuasion for Speech**

| Dimension 4 | Mean value |
|---|---|
| **CORPUS** | **2.05** |

According to the graphical expression of D4, the corpus of speech has the highest positive mean scores was that of professional letters of the original study (Biber, 1988). The mean score of the present corpus's data is 2.05 which demonstrates that the language of speech is argumentative in nature on dimension 4. When speakers address the audience, there are components of persuasion present in the text which makes their point of view more explicit and their assessment of likelihood and certain. The employment of modal verbs is one of the linguistic characteristics that is indicated by the high scores. The linguistic characteristics on dimension 4 are listed below.

**Text Type**
The text is specific and persuasive in expression with a mean score of 2.05 on D4. The closest genre to the provided corpus is "personal letters," with a mean score of 1.5, are the closest genre that matches the provided corpus. Comparing the various genres, "speech" receives average D4 ratings and is most likely to be personal letters. The closest genre comparison reveals that the speech has characteristics in common with personal letters that are more persuasive in tone.

**4.1.5.3 Comparison of Commentary and Speech on Dimension 4:**
Commentary, which has the mean score of -3.07 on dimension 4 made that the corpus lacks persuasiveness, while speech has the mean score of 2.05, which shows that the text of the speech is persuasive in nature. The text type's comparison shows that the commentary's corpus has the closest genre to the broadcast, while the closest genre is personal letters to the supplied corpus of speeches.

**4.1.6 Dimension 5: Abstract Versus Non-Abstract Information**
Dimension 5 categorizes non-technical and technical discourses, or abstract and non-abstract information, as first proposed by Biber (1988, p. 152). It includes personal and non-personal styles. The text has abstract qualities, agentless passive, past participial clauses, by passive, past participle whiz deletion, adverbial subordination, and others if it has a high score on this dimension, (Biber, 1988). These formal, abstract, and technical conversations include formal characteristics. It presents information using straightforward phrase structure and syntax.

#### 4.1.6.1 Dimension 5: Abstract Versus Non-Abstract Information for Commentary

| Dimension 5 | Mean value |
|---|---|
| CORPUS | -2.18 |

According to the graphic for D5, the corpus's mean score is -2.18, for commentaries. The corpus graph position demonstrates that the corpus is non-abstract because the score is negative. Biber (1998) asserts that texts with low scores on this dimension are subjective and non-impersonal in nature. The commentators' use of language in their commentaries is informal and non-technical. The following table displays dimension 5's linguistic characteristics.

**Text Type**

The text is non-abstract and informative, with a mean score of -2.18 on D5. The closest genre to the provided corpus is "prepared speeches," with a mean score of -1.9. Commentaries receive average D5 ratings and are most likely to be prepared speeches. The closest genre comparison reveals that commentaries share characteristics with prepared speeches, as they are planned but not written. Written genres typically have greater low scores (abstract information), while spoken genres typically have lower scores on dimension 5 which is non-abstract information (Biber, 1988, p. 161).

#### 4.1.6.2 Dimension 5: Abstract Versus Non-Abstract Information for Speeches

| Dimension 5 | Mean value |
|---|---|
| CORPUS | -0.32 |

According to the graphic for D5, the corpus's mean score is -0.32, for speeches. The corpus's graph position demonstrates that the text of the corpus is non-abstract because of the negative score. Biber (1998) asserts that texts with low scores on this dimension are subjective and non-impersonal in nature. The speakers' use of language in their speeches is less informal and non-technical. The following table displays dimension 5's linguistic characteristics.

**Text Type**

The text is non-abstract and informative, with a mean score of -0.32 on D5. The closest genre to the provided corpus is "press reportage," with a mean score of 0.6. "Speeches" receives average D5 ratings and is most likely to be press reportage. The closest genre comparison reveals that speeches share characteristics with press reportage. Due to the dual aims of press reportage, press review, and editorial genres, reporting on events involving concrete, frequently human, referent, and abstract discussion of the implications of the events in conceptual terms, these genres display intermediate values on this dimension (Biber, 1988, p. 154).

#### 4.1.6.3 Comparison of Commentary and Speech on Dimension 5:

Commentary, which has the mean score of -2.18 on dimension 5 made the corpus non-abstract information, while speech has the mean score of -0.32, which also shows that the non-abstract information of speeches. The comparison shows that the text of commentaries has more non-abstract information qualities as compared to speeches. The text type comparison shows that the genre which is closest to the corpus of commentary is prepared speeches, while the closest genre is 'press reportage' to the supplied corpus of speeches. Both the corpus (commentaries and speeches) has negative scores because of having linguistic features and low frequencies on dimension 5.

### 4.1.7 Dimension-6: Measures on Online Informational Elaboration

Biber (1988) claims that dimension 6 evaluates the elaboration of online information. The sixth dimension distinguishes between informational texts and those produced in real-time. This dimension divides the data into halves of the plot. The discourse is information-rich, as shown by the high ratings on the graph, but it is created under time constraints. Because of this dimension, poor scores show that text is informational, but not just by time limited.

#### 4.1.7.1 Dimension-6: Measures on Online Informational Elaboration for commentary

| Dimension 6 | Mean value |
|---|---|
| **Corpus** | **-1.17** |

The location of the data is depicted graphically in the current corpus, and the corpus's mean scores are -1.17. The low scores present corpus show that the language is informative but not a time constraint. The following table of dimension 6 shows the list of linguistic features which occur thoroughly:

**Text Type**

The supplied corpus's mean score is -1.17 on D6. The given corpus's mean score on D6 indicates that the text is informative but not limited to time constrained. The graphical representation displays the genres that are most like the provided corpus. The resulting graph shows that "broadcast" having mean score of -1.3 is the genre that most closely matches the supplied corpus. The closest genre comparison reveals that the commentaries have characteristics in common with the broadcast.

#### 4.1.7.2 Dimension 6: Measures on Online Informational Elaboration for Speech

| D6 | Mean |
|---|---|
| **CORPUS** | **1.22** |

The location of the data is depicted graphically in the current corpus, and the corpus's mean scores are 1.22. The high scores show that the present corpus's language is informative, but the corpus is a time constraint. The following table of dimension 6 shows the list of linguistic features which occur thoroughly:

**Text Type**

The supplied corpus's mean score is 1.22 on D6. The given corpus's mean score on D6 indicates that the text is informative but limited to time constrained. The graphical representation displays the genres that are most like the provided corpus. The resulting graph shows that "academic prose" having mean score of 0.5 is the genre that most closely matches the supplied corpus. The other closest genre is "press editorial" having mean score of 1.5. The closest genre comparison reveals that the speeches have characteristics in common with academic prose and press editorial.

#### 4.1.7.3 Comparison of Commentary and Speech on Dimension 6:

Commentary, which has the mean score of -1.17 on dimension 6 shows that the text is informative but not limited to time-constrained, while speech has the mean score of 1.22, which also indicate that the text is informative but limited to time constrained. The comparison shows

that the commentaries' text is informative in nature, but not time specific and limited, while the speeches corpus is also informative in nature but time specific and limited. The text type comparison shows that the closest genre is broadcast to the commentary's corpus, while the closest genre is academic prose and press editorial to the supplied corpus of speeches.

**Table 4.13: Textual Dimension of Commentary and Speech**

| Textual Dimension | Commentary | Speech |
|---|---|---|
| Dimension 1: Involved vs informational production | -6.6 (Negative) | -0.4 (Negative) |
| Dimension 2: Narrative versus Non-Narrative | 2.23 (Positive) | -2.12 (Negative) |
| Dimension 3: Context dependent vs Context Independent | -3.06 (Negative) | 4.35 (Positive) |
| Dimension 4: Overt Expression of Persuasion | -3.07 (Negative) | 2.05 (Positive) |
| Dimension 5: Abstract Versus Non-Abstract Information | -2.18 (Negative) | -0.32 (Negative) |
| Dimension 6: Measures on Online Informational Elaboration | -1.17 (Negative) | 1.22 (Positive) |

The commentary textual dimension is positive on D2 and negative on D1, D3, D4, D5, D6, while speech textual dimension is positive on D3, D4, D6 and negative on D1, D2, D5.

**Table 4.14: Commentary and Speech Closest Genre Type**

| Genre Type | | | |
|---|---|---|---|
| **Commentary** | | **Speech** | |
| Dimension 1 | Broadcast | Dimension 1 | General fiction |
| Dimension 2 | Prepared speeches | Dimension 2 | Academic prose |
| Dimension 3 | General fictions | Dimension 3 | Academic prose |
| Dimension 4 | Broadcast | Dimension 4 | Personal letters |
| Dimension 5 | Prepared speeches | Dimension 5 | Press reportage |
| Dimension 6 | Broadcast | Dimension 6 | Academic prose |

The commentary closest genre type on dimension 1, 4 and 6 is broadcast, dimension 2 and 5 is prepared speeches and dimension 3 is general fiction, while speech closest genre on dimension 1 is general fiction, dimension 2, 3 and 6 is academic prose, dimension 4 is personal letters and dimension 5 is press reportage.

## 5. CONCLUSION

The research examines the textual dimensions of speeches and commentary. Both corpora had negative mean scores (-0.4 for commentary and -6.6 for speeches) on dimension 1, showing informational value to the text. Dimension 2 shows that speeches have a negative mean score of -2.12, indicating non-narrative value, while commentary has a positive mean score of 2.23, suggesting narrative value. Dimension 3 indicates that speeches are considered context-independent with a positive mean score of 4.35, while commentary has a negative mean score of -3.06 are considered context-dependent. Dimension 4 highlights commentary lacking persuasiveness having negative mean score of -3.07, while speeches are persuasive with positive mean score of 2.05. Both corpora are non-abstract on dimension 5, with negative mean scores i.e. -2.18 for commentary and -0.32 for speeches. On dimension 6, commentaries mean score is -1.17, shows that the text is informative but not limited to time-constrained, while speeches mean scores is 1.22, which also indicates that the text is informative but limited to time constrained. **The second** study question analyzed the closest genres that commentary corpus having the closest genre on dimensions 1, 4, and 6 is broadcast, on dimensions 2 and 5, is prepared speeches and on dimension 3, is closer to general fiction. Thus, the Cricket commentary's corpus analysis discovered that the text type is broadcast because of transmission method. The text is written since it is found on websites in written form and has been informed of and released like a press report. The source of this commentary genre is broadcasting which is available to the general public. On contrast, speech's corpus has the closest genre on dimension 2, 3, and 6 is academic prose because that the speech corpus is included in academic prose, on dimension 1, 4 and 5 it is close to general fiction, personal letters and press reportage respectively. It is like a press report because the orator sought to offer information about a certain matter to the public. These messages are released on websites like press reports. The mode of the text is written because it is found in written form on the websites and has been notified and released like a press report. The academic prose having a low score on dimension 1 and 5 while having a high score on dimension 3 make the corpus informational exposition which is formal and concentrates on delivering information. This study concludes the genre of commentary shows remarkable variations in textual dimensions and genre type compared to the genre of speech, so according to Biber's multidimensional analysis, the genres of commentary and speech are considered different genres.

## REFERENCE

Ahmad, N. N., & Shakir, A. (2014). Significance of Non-Narrative Features in Online University Prospectus of Pakistan: A Corpus-Based Investigation. Educational Research International.3 (6), 1-10.

Amjad, T., & Shakir, A. (2014). Study of information generating linguistic features in online university prospectuses. Research on Humanities and Social Sciences, 4(25), 122-128.

Bhatia, V. (1993). Analyzing genre: Language use in professional settings. London and New York: Longman.

Bhatia, V. (2002). Applied genre analysis: a multi-perspective model. Ibérica: Revista de la Asociación Europea de Lenguas para fines específicos (AELFE), 4, 3-19.

Bhatia, V. (2004). Worlds of written discourse: A genre-based view. Bloomsbury Publishing.

Biber, D. (1988). Variation across speech and writing. Cambridge: Cambridge University Press.

Biber, D. (1991). Variation across speech and writing. Cambridge University Press.

Biber, D. (1993). The multidimensional approach to linguistic analyses of genre variation: An overview of methodology and findings. Computers and the Humanities 26(1): 331-345.

Biber, D. (1993). Representativeness in corpus design. Literary and linguistic computing, 8(4), 243-257.

Biber, D., & Finegan, E. (1994). Multidimensional analyses of authors' styles: Some case studies from the eighteenth century. Research in Humanities Computing 3(1): 3-17.

Conrad. S, (2009). Genre analysis: Structural and linguistic evolution of the English-medium medical research article. English for Specific Purposes, 28(2), 93-104.

Crystal. D., & Davy. D. (1969) Prosodic systems and intonation in English. Cambridge: University Press.

Ferguson, Charles A. (1983) “Sports announcer talk: Syntactic aspects of register variation.” In: Huebner, Thom (ed.) Sociolinguistic perspectives: papers on language in society, 1959 – 1994. New York, NY: Oxford University Press, 148-166.

Lewandowski, M. (2012). The language of online sports commentary in a comparative perspective. Lingua posnaniensis, 54(1), 65.

Meier-Vieracker, S. (2021). The evolution of football live text commentaries: A corpus linguistic case study on genre change. AILA Review, 34(2), 274-299.

Obin, N., Dellwo, V., Lacheret, A., & Rodet, X. (2010). Expectations for discourse genre identification: a prosodic study. In Interspeech (pp. 3070-3073).

Swales, J.M. (1990). Genre analysis: English in academic and research settings. Cambridge: Cambridge University Press.

Swales, J. M., & Feak, C. A. (2000). English in today's research world: A writing guide. University of Michigan Press.

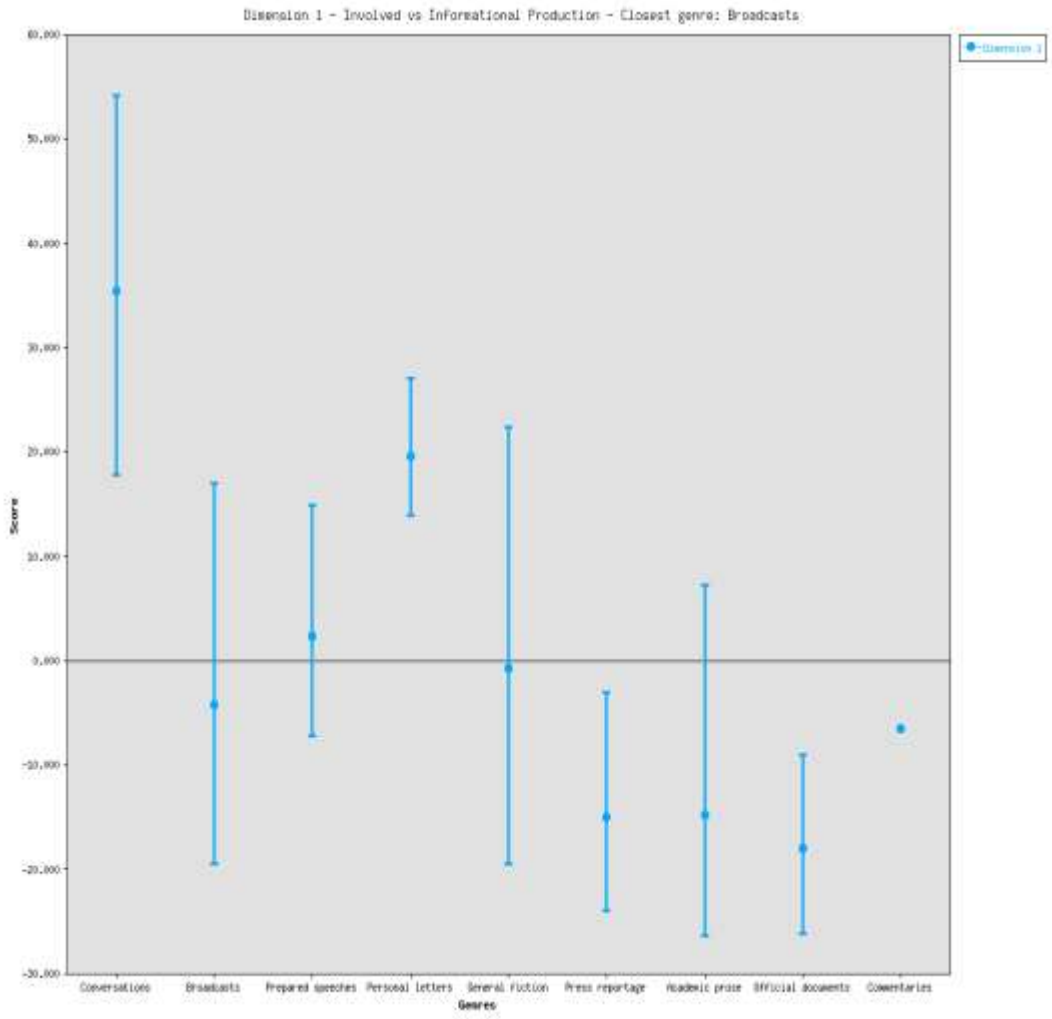
Villanueva, L. S., Dolom, M. A. C., & Belen, J. S. (2018). Genre analysis of the "About Us" sections of Asian Association of Open University websites. Asian Association of Open Universities Journal, 13(1), 37-59.

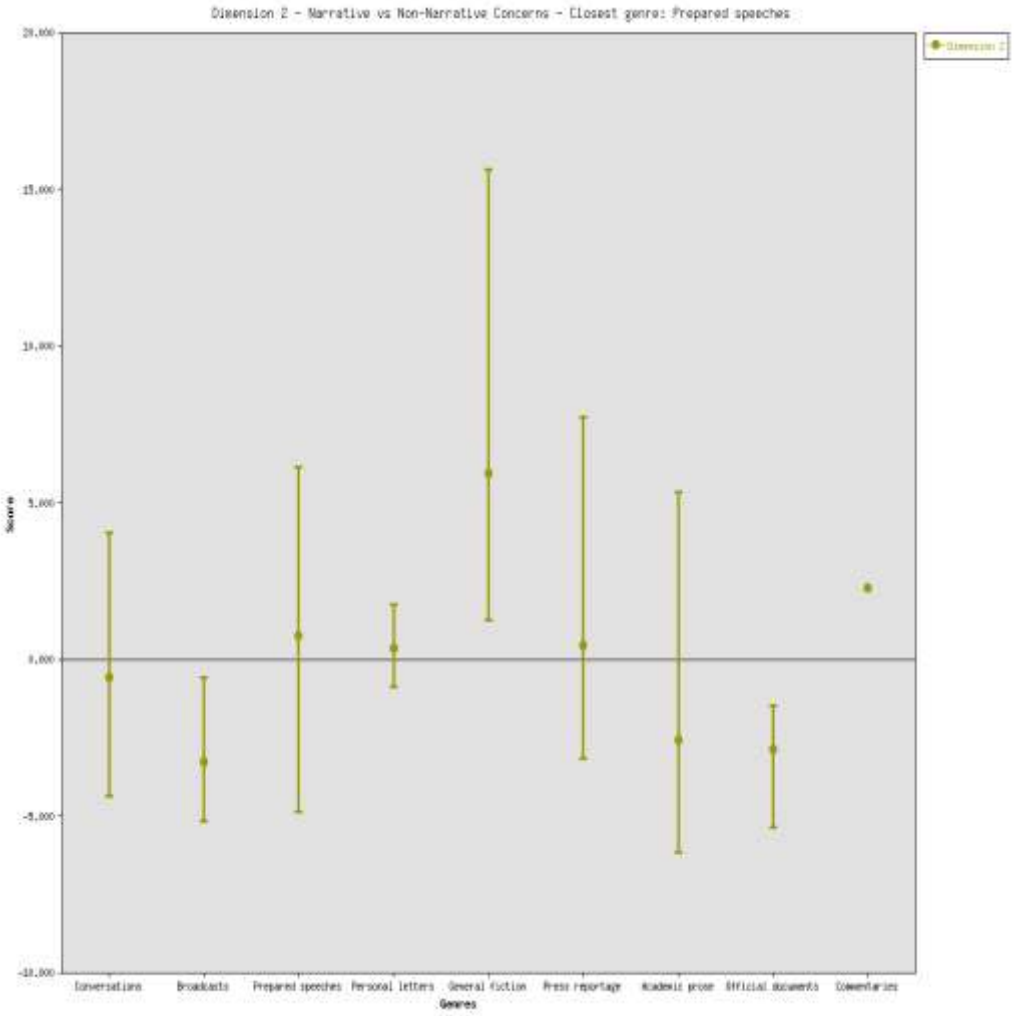
Yaakob, S. B. (2014). A genre analysis and corpus-based study of university lecture introductions (Doctoral dissertation, University of Birmingham).

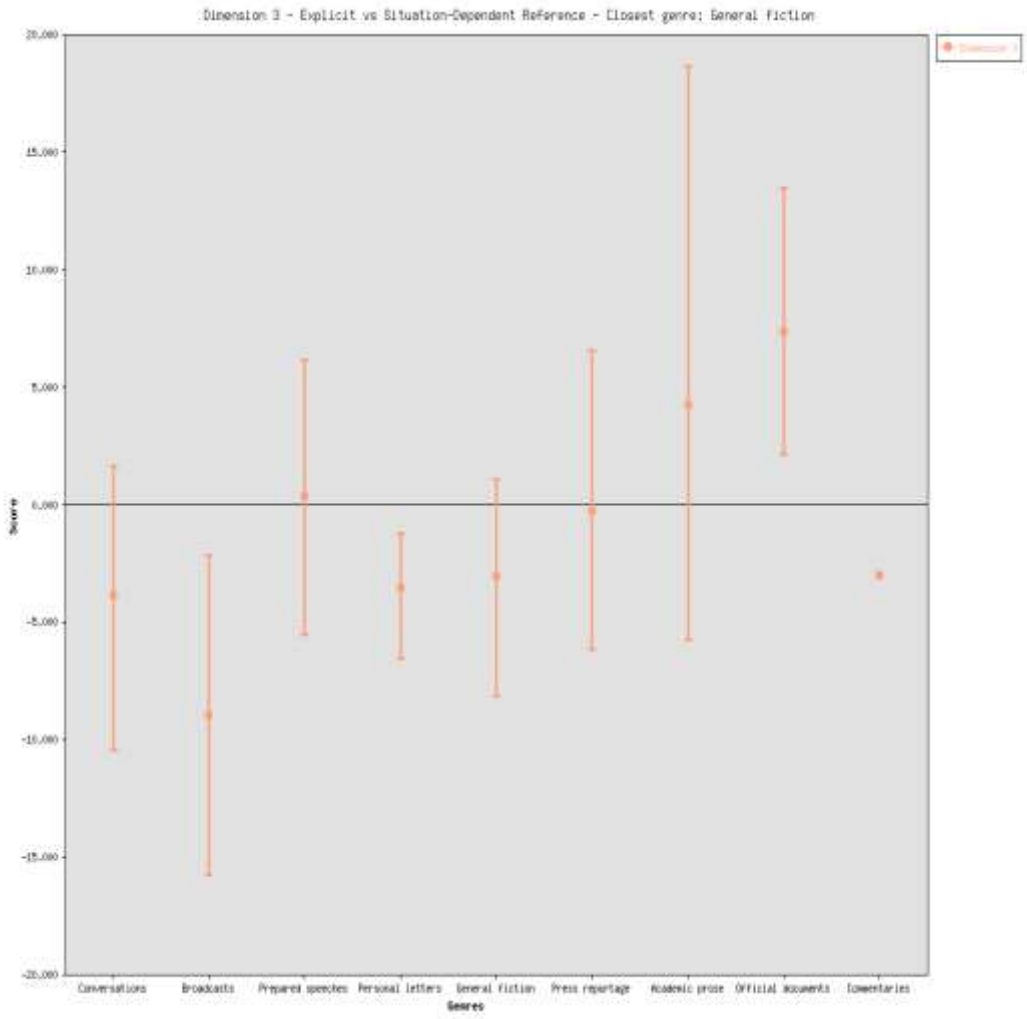
## Appendix

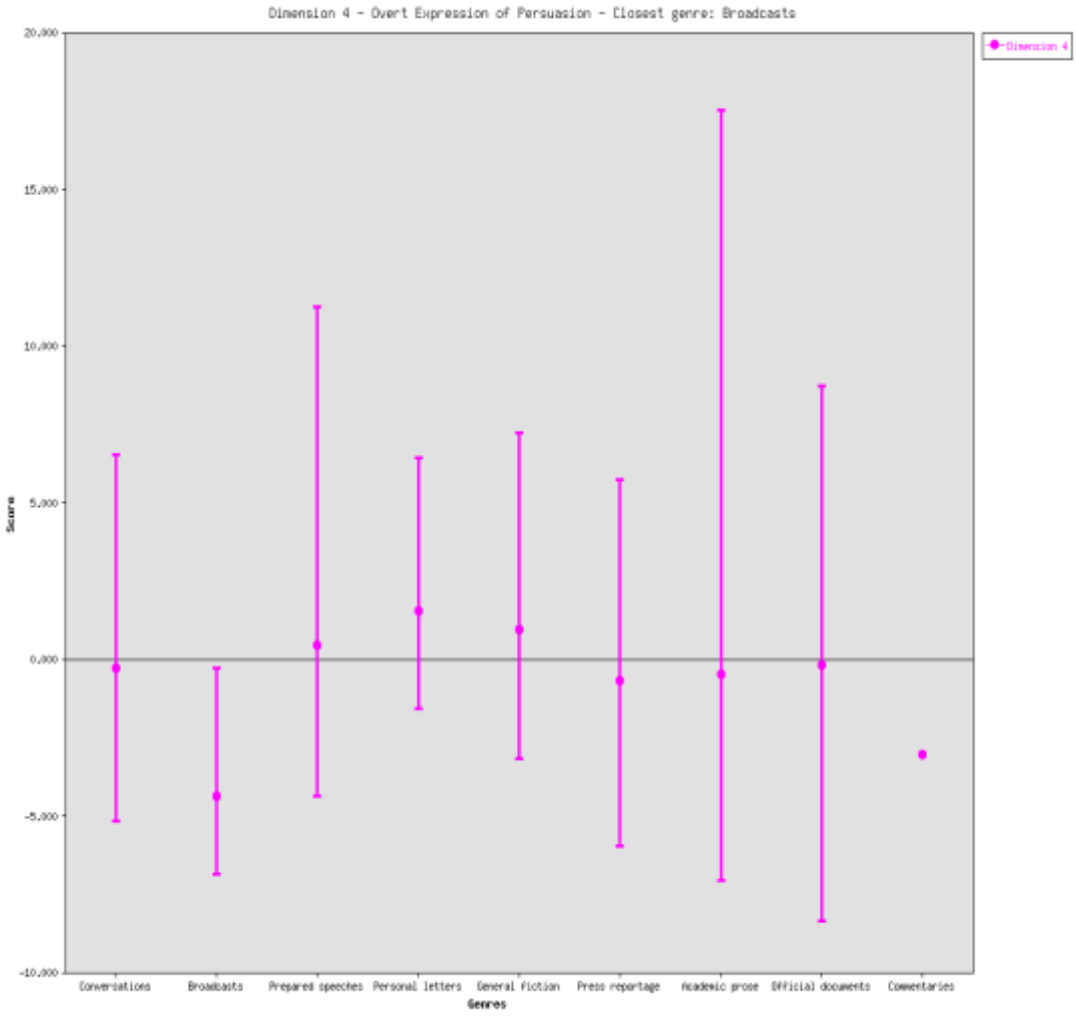
## Graphs

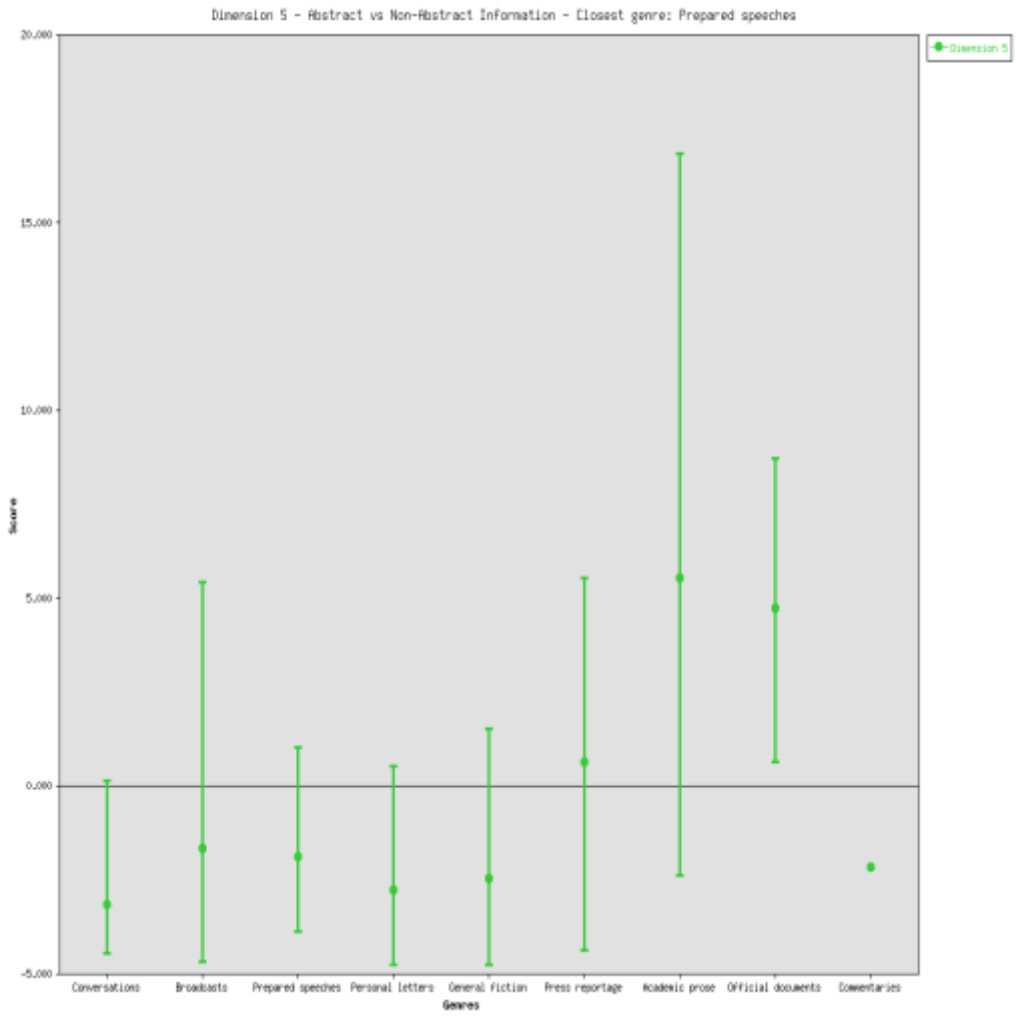
### Graphs of Commentary Dimension

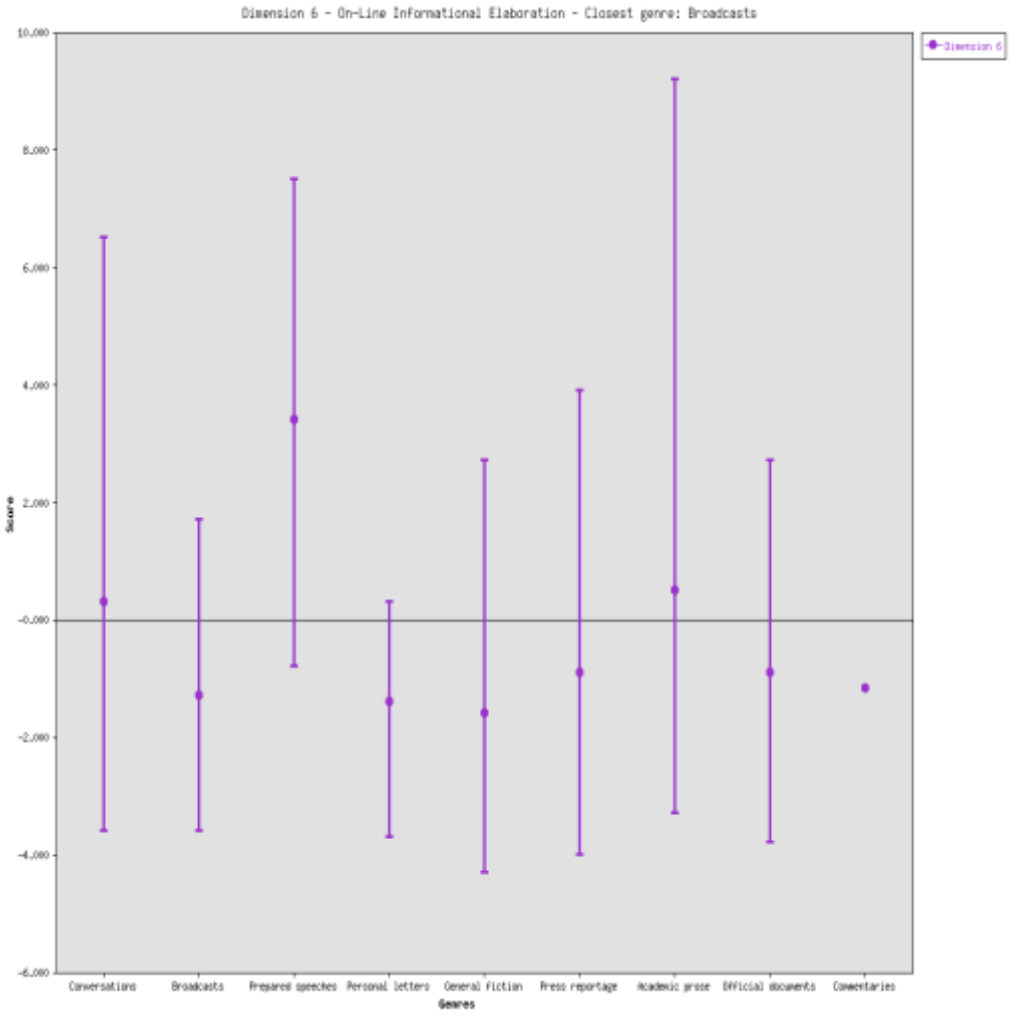

Dimension 1 - Involved vs Informational Production - Closest genre: Broadcasts
Score
Genres
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Commentaries

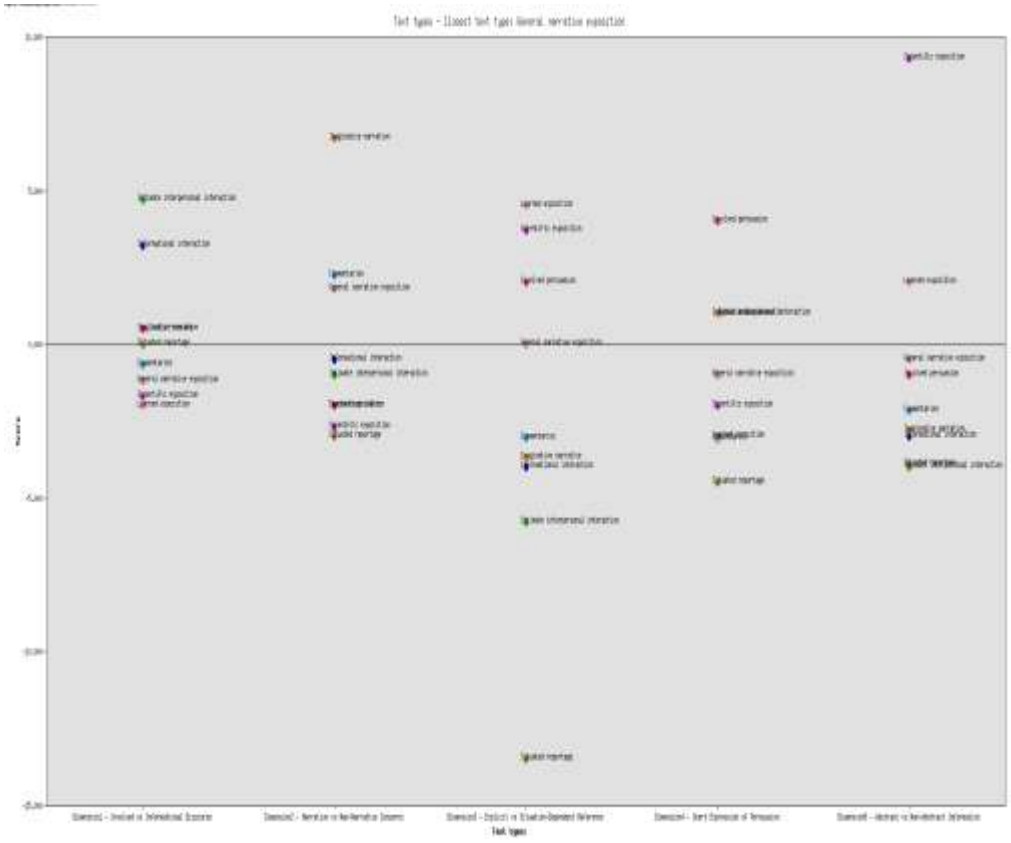

Dimension 2 - Narrative vs Non-Narrative Concerns - Closest genre: Prepared speeches
Dimension 2
20,000
15,000
10,000
5,000
0,000
-5,000
-10,000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Commentaries
Genres


Dimension 3 - Explicit vs Situation-Dependent Reference - Closest genre: General Fiction
Score
Genres
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Commentaries

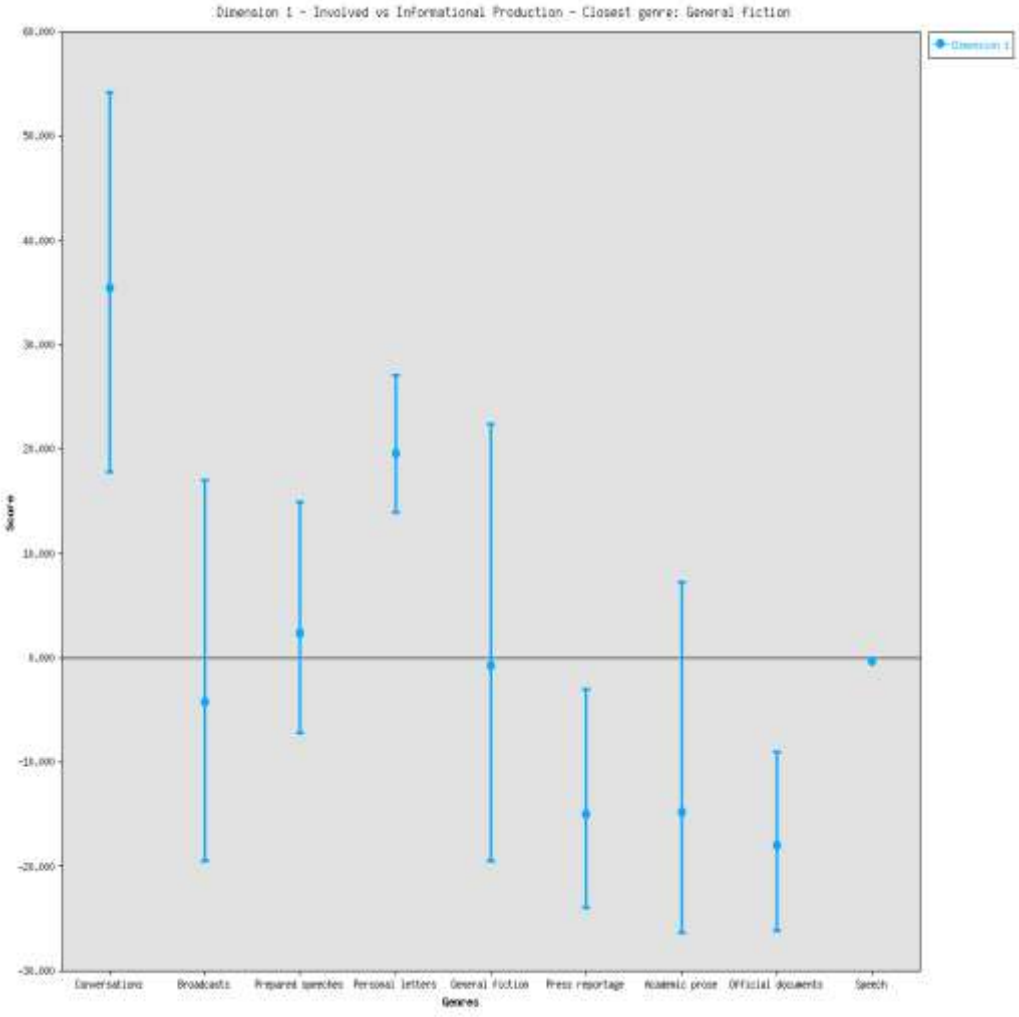

Dimension 4 - Overt Expression of Persuasion - Closest genre: Broadcasts
Dimension 4
20,000
15,000
10,000
5,000
0,000
-5,000
-10,000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Commentaries
Genres

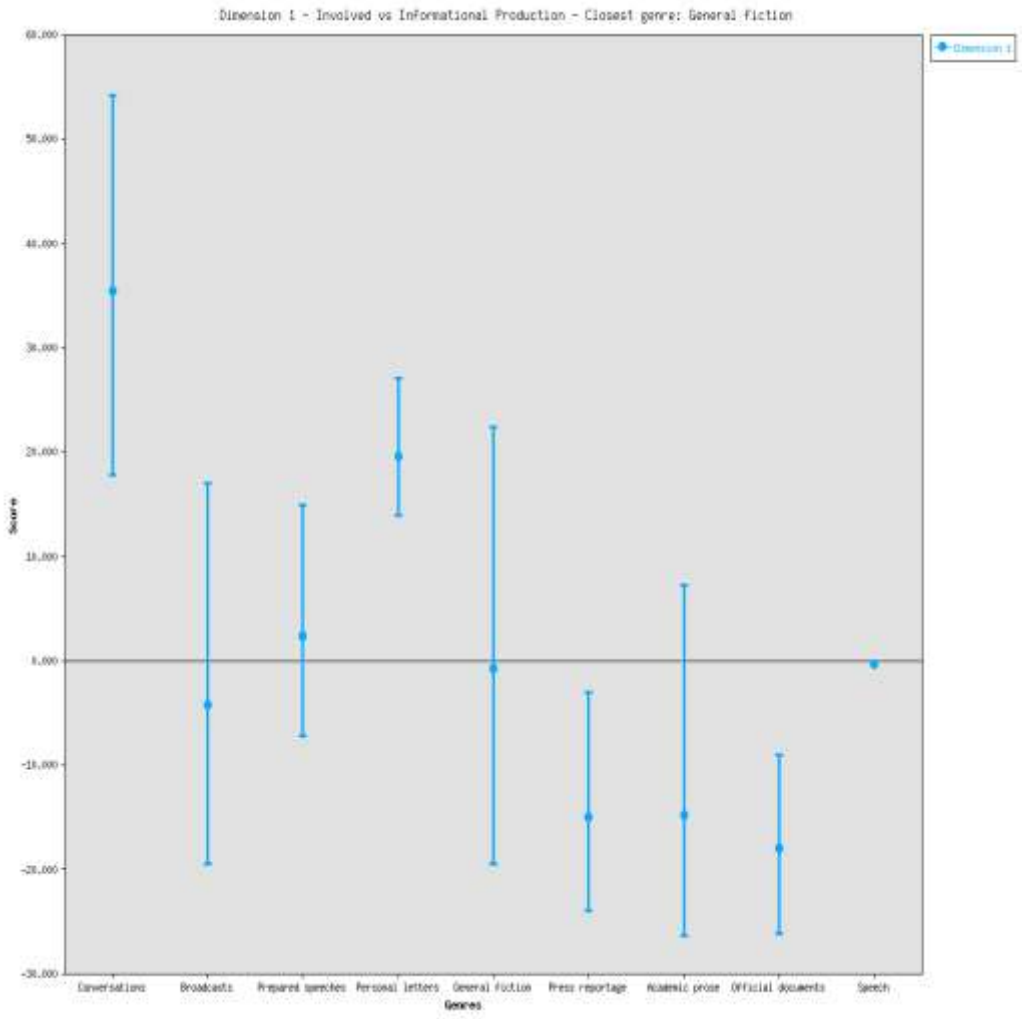

Dimension 5 - Abstract vs Non-Abstract Information - Closest genre: Prepared speeches
Dimension 5
20,000
15,000
10,000
5,000
0,000
-5,000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Commentaries
Genres

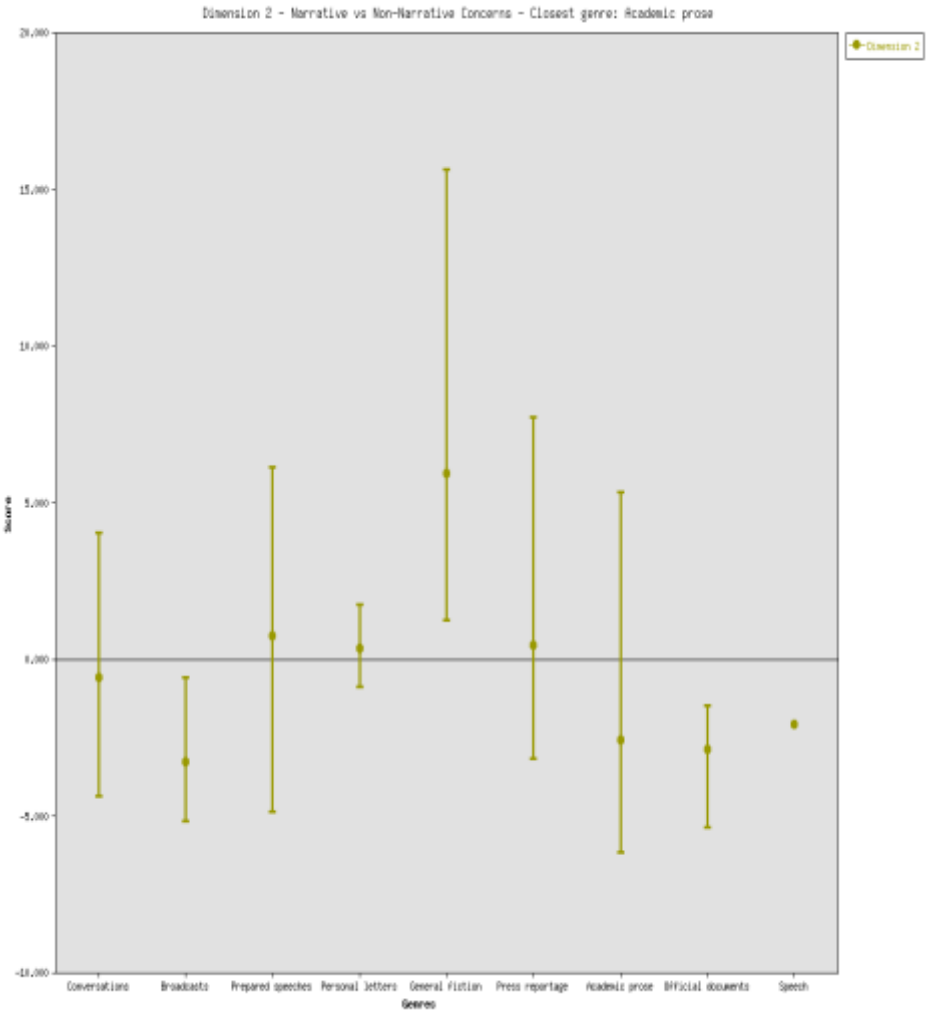

Dimension 6 - On-Line Informational Elaboration - Closest genre: Broadcasts
Dimension 6
10.000
8.000
6.000
4.000
2.000
-0.000
-2.000
-4.000
-6.000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Commentaries
Genres

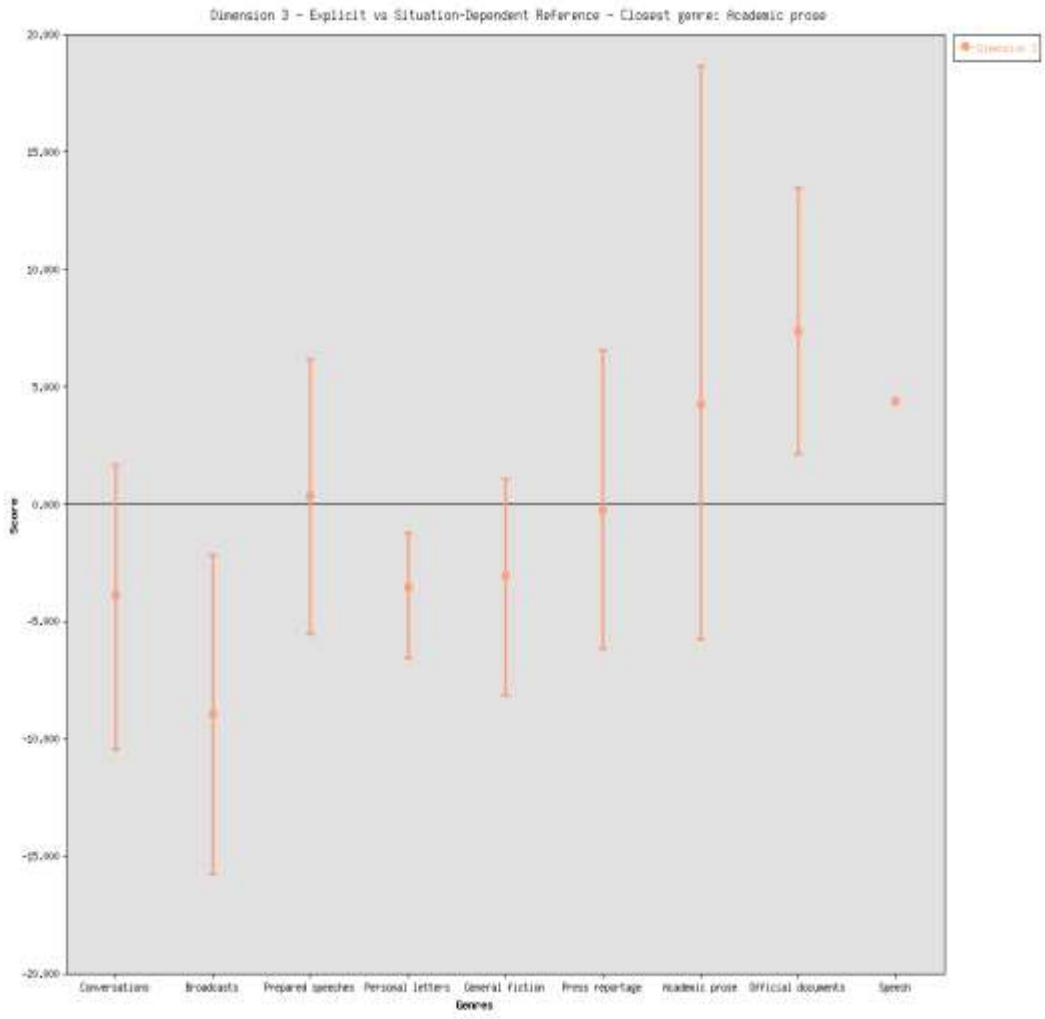


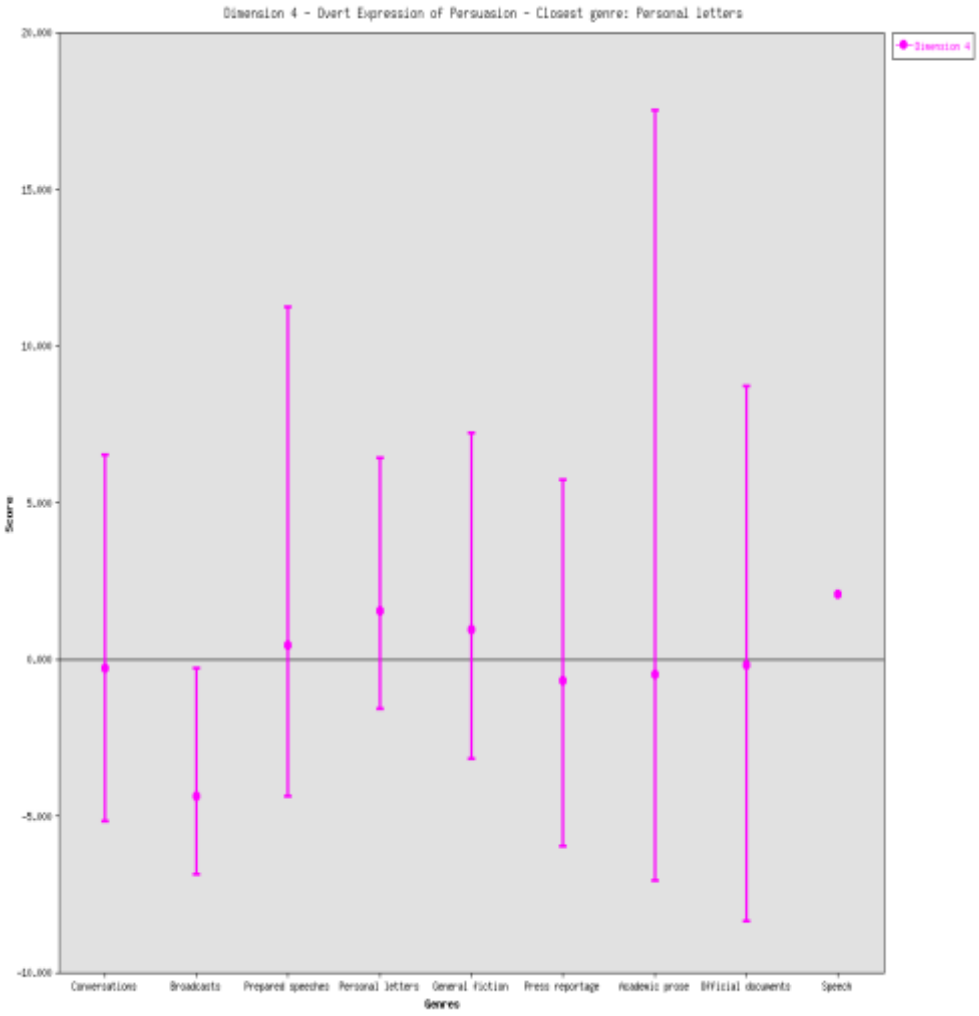
## Graphs of Speech's Dimensions

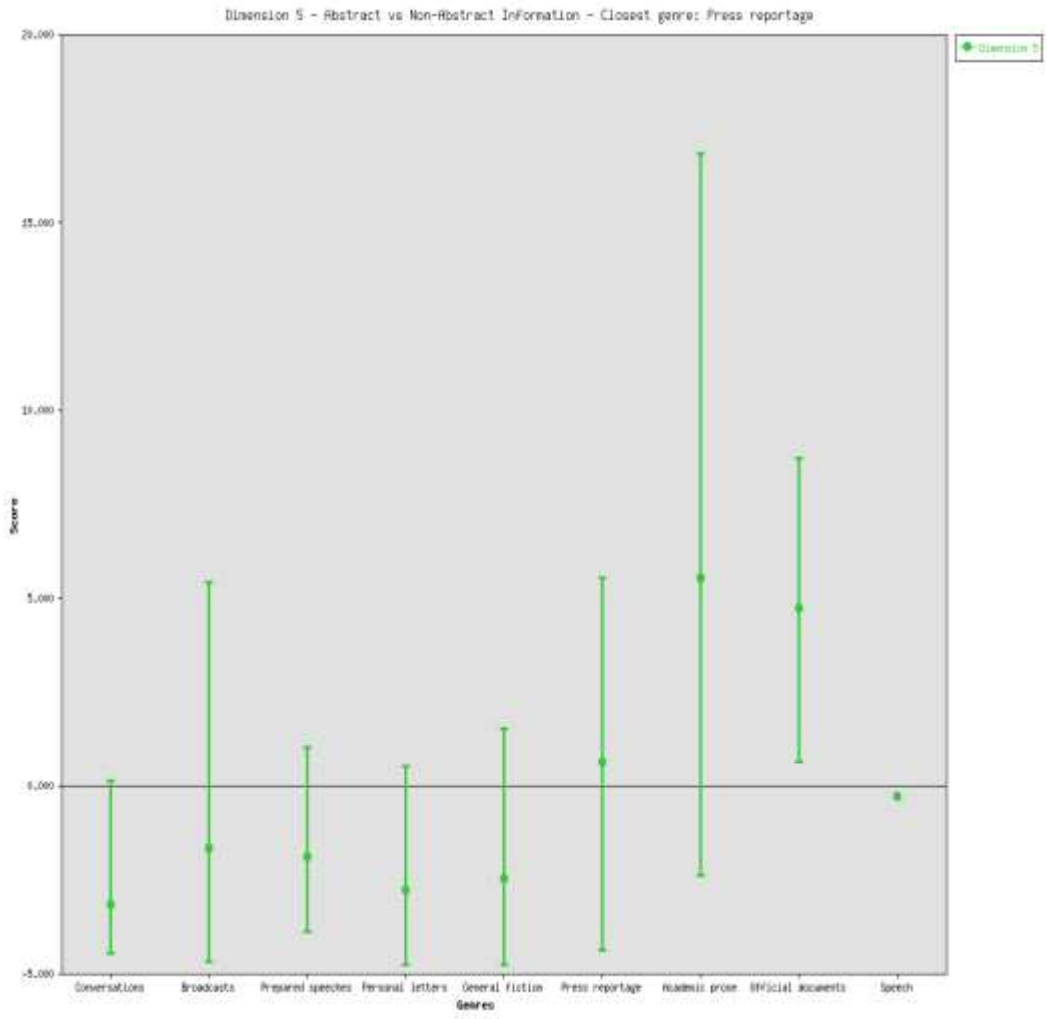


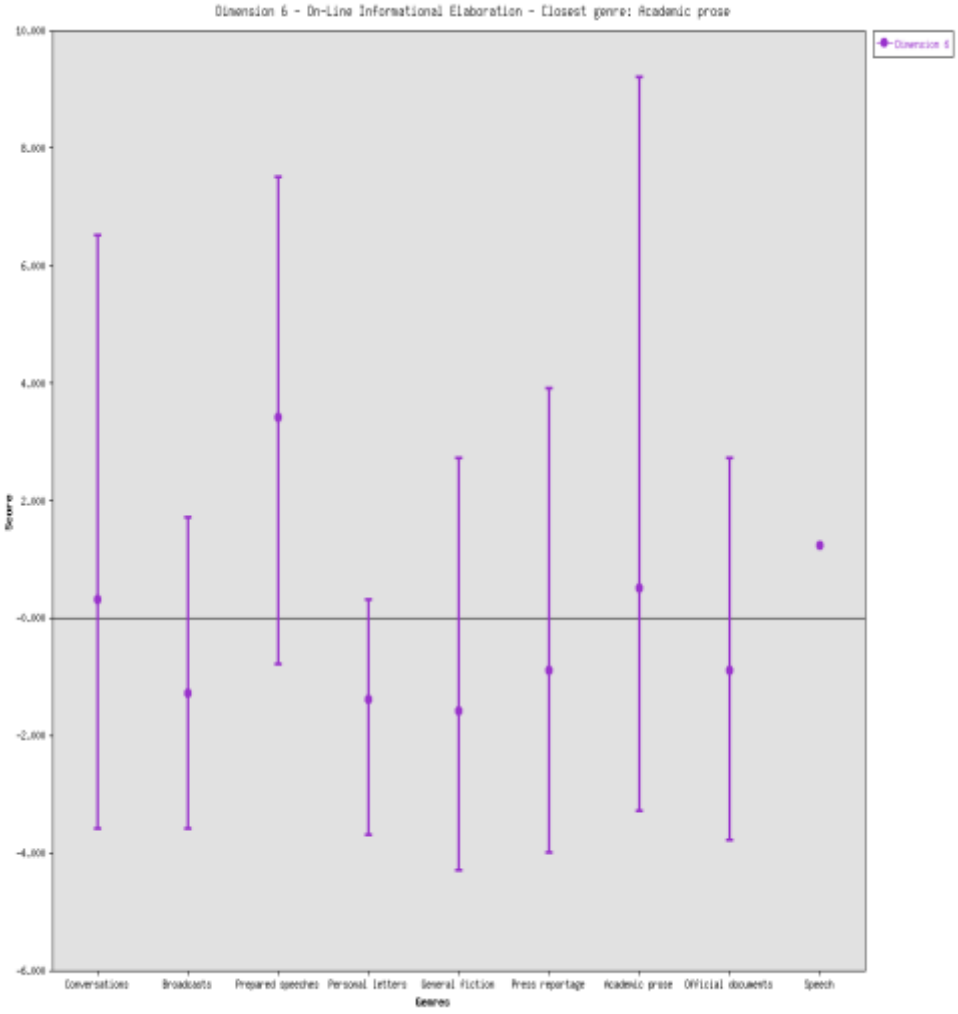

Dimension 1 - Involved vs Informational Production - Closest genre: General Fiction
Dimension 1
60.000
50.000
40.000
30.000
20.000
10.000
0.000
-10.000
-20.000
-30.000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General Fiction
Press reportage
Academic prose
Official documents
Speech
Genres


Dimension 2 - Narrative vs Non-Narrative Concerns - Closest genre: Academic prose
Dimension 2
20.000
15.000
10.000
5.000
0.000
-5.000
-10.000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General Fiction
Press reportage
Academic prose
Official documents
Speech
Genres


Dimension 3 – Explicit vs Situation-Dependent Reference – Closest genre: Academic prose
Score
Genres
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Speech


Dimension 4 - Overt Expression of Persuasion - Closest genre: Personal letters
Dimension 4
20.000
15.000
10.000
5.000
0.000
-5.000
-10.000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Speech
Genres


Dimension 5 - Abstract vs Non-Abstract Information - Closest genre: Press reportage
Dimension 5
20.000
15.000
10.000
5.000
0.000
-5.000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Speech
Genres


Dimension 6 - On-Line Informational Elaboration - Closest genre: Academic prose
Dimension 6
10.000
8.000
6.000
4.000
2.000
-0.000
-2.000
-4.000
-6.000
Score
Conversations
Broadcasts
Prepared speeches
Personal letters
General fiction
Press reportage
Academic prose
Official documents
Speech
Genres