

NATIVE VS. NON-NATIVE VARIATION IN ENGLISH FOOD BLOGS: A MULTIDIMENSIONAL ANALYSIS OF LANGUAGE AND STYLE

Muhammad Jawad Nasir (Kanzul Madaris Board International)

M Phil Scholar, Department of English, NUML University (Faisalabad Campus)

Email: kanz.english@gmail.com

Dr. Maimoona Abdulaziz

Prof. at National University of Modern Languages Faisalabad

Email: maimoona.aziz@numl.edu.pk

ABSTRACT

English food blogs constitute a rapidly growing form of digital discourse, yet limited research has examined how linguistic and stylistic patterns vary between native and non-native English bloggers. This study investigates linguistic and stylistic variations in English food blogs authored by native British and non-native Pakistani bloggers within Douglas Biber's (1988) Multidimensional Analysis (MDA) framework. Adopting an explanatory sequential mixed-methods design, the study compiled a specialized corpus of 100 food blog posts (50 native and 50 non-native) through purposive sampling. The corpus was analyzed using the Multidimensional Analysis Tagger (MAT) to identify lexicogrammatical patterns and multidimensional stylistic features. The findings revealed a shared genre core characterized by frequent use of nouns, determiners, prepositions, and adjectives, confirming the informational and descriptive baseline of food blogging. However, significant stylistic differences emerged between the two groups. Native bloggers demonstrated a highly informational, elaborated, and non-persuasive style consistent with Learned Exposition, and additionally relied more heavily on cardinal numbers to emphasize quantitative precision. In contrast, non-native Pakistani bloggers exhibited a more interactive style aligned with General Narrative Exposition and Involved Persuasion, marked by greater use of present-tense verbs and personal pronouns. Independent samples *t*-tests confirmed that these differences were systematic and statistically significant across all six functional dimensions, with five showing highly significant variance ($p < .001$). These findings suggest that language background shapes rhetorical identity in digital discourse. The study concludes that non-native bloggers adapt to the participatory nature of online communication by constructing community-oriented registers rather than deviating from native norms. These results challenge deficit-oriented perspectives in Second Language Acquisition and contribute to corpus linguistics and digital discourse studies by offering new insights into stylistic variation in online food blogging.

Keywords: Native vs. Non-native English, Food Blogs, Stylistic Variation, Corpus Linguistics, Multidimensional Analysis Tagger, Multidimensional Analysis.

1. INTRODUCTION

In the digital age, online platforms have significantly transformed the production, distribution, and consumption of language. Among these platforms, blogs have emerged as a prominent form of computer-mediated communication (CMC), offering spaces where personal expression, information sharing, and instructional discourse intersect (Herring, Stein, & Virtanen, 2013; Myers, 2010). Unlike traditional print genres, blogs are characterized by immediacy, interactivity, and a flexible blending of formal and informal linguistic styles. Within this broader category, food blogs represent a particularly rich hybrid genre, combining procedural instructions, narrative storytelling, and cultural commentary (Lomborg, 2014). In addition to providing recipes and cooking guidance, food blogs also function as sites of identity construction, where writers express personal voice, cultural affiliation, and audience engagement strategies.

From a linguistic perspective, food blogs are especially significant because they operate at the intersection of multiple discourse functions. They are simultaneously instructional (guiding readers through procedures), descriptive (detailing ingredients and processes), and interpersonal (engaging readers through evaluation and personal experience). This multifunctional nature makes food blogs an important domain for examining how language varies across users and communicative contexts (Swales, 1990).

English food blogs authored by native and non-native speakers often display noticeable differences in linguistic and rhetorical style. Second Language Acquisition (SLA) research has traditionally suggested that non-native speakers tend to use more explicit and formal language, whereas native speakers are expected to produce more natural and context-dependent discourse (Hinkel, 2002). However, in digital environments, these assumptions do not always hold, as online platforms allow for greater stylistic flexibility and hybridization of discourse norms. This raises an important question of whether native bloggers adhere more closely to traditional instructional conventions, while non-native bloggers adopt more interactive and engagement-oriented strategies shaped by digital communication practices (Nardi, Schiano, & Gumbrecht, 2004).

Despite the growing prominence of food blogs, there remains a lack of systematic empirical research investigating linguistic variation between native and non-native English bloggers within this genre. Existing Multidimensional Analysis (MDA) studies have primarily focused on academic writing, spoken registers, and general computer-mediated communication (Egbert, Biber, & Gray, 2022), while food blogs as a specialized digital genre remain underexplored. Consequently, there is limited understanding of how linguistic features pattern across native and non-native writers and how these patterns reflect broader communicative intentions in digital food discourse. This study addresses this gap by applying Biber's (1988) Multidimensional Analysis framework using the Multidimensional Analysis Tagger (MAT) (Nini, 2014) to provide a systematic, corpus-based investigation of stylistic variation.

The purpose of this study is to investigate linguistic and stylistic variation between native (British) and non-native (Pakistani) English food bloggers using the Multidimensional Analysis Tagger (MAT). Specifically, the study examines how linguistic features cluster across Biber's (1988) functional dimensions and how these patterns reflect differences in communicative style, informational density, and rhetorical orientation.

1.1 Statement of the Problem

Although previous research has examined linguistic variation in academic writing, spoken discourse, and general computer-mediated communication, limited attention has been given to food blogs as a distinct digital genre. Moreover, there is insufficient empirical evidence on how native and non-native English bloggers differ across Biber's multidimensional framework. In particular, the application of MAT to compare L1 and L2 rhetorical styles in specialized digital discourse remains underexplored. This gap highlights the need for a systematic, corpus-based analysis of food blogging discourse.

1.2 Significance of the Study

This study contributes both theoretically and practically. Theoretically, it extends Corpus Linguistics by applying MAT to a contemporary and underexplored digital genre. It also enhances Discourse Analysis by examining how linguistic features function in hybrid online food writing. Practically, the findings may benefit content creators, educators, and digital communicators by providing insights into effective stylistic strategies and audience engagement in online writing.

1.3 Objectives of the Study

The study aims to:

- To identify the most frequent linguistic features in native and non-native English food blogs using Multidimensional Analysis Tagger (MAT).
- To compare the multidimensional profiles of native and non-native English food blogs across Biber's six dimensions.

- To interpret the functional significance of any observed differences in terms of communicative style and discourse practices.
- To determine whether statistically significant differences exist between the two groups across the six dimensions.

1.4 Research Questions

1. What are the most frequent linguistic features identified by MAT in native and non-native English food blogs?
2. How do native and non-native English food blogs differ across Biber's six dimensions?
3. How can these differences be functionally interpreted in terms of language use and communicative style?
4. Which dimensions exhibit statistically significant differences between the two groups?

1.5 Limitations of the Study

This study has several limitations. The sample size is relatively limited (100 blog posts), which may restrict generalizability. The analysis relies on MAT's automated tagging system, which may occasionally misclassify ambiguous linguistic items. The study focuses exclusively on textual data, excluding multimodal elements such as images, layout, and video content. Finally, classification of bloggers' linguistic background is based on publicly available information, which may not always be fully verifiable.

1.6 Delimitations of the Study

The study is delimited in several ways. It focuses exclusively on food blogs, excluding other forms of computer-mediated communication such as travel, fashion, and lifestyle blogs. The demographic comparison is limited to native British and non-native Pakistani English bloggers, excluding multilingual and code-switching practices. The study is grounded solely in Biber's (1988) Multidimensional Analysis framework using MAT 1.3, without incorporating alternative theoretical approaches such as Systemic Functional Linguistics or Critical Discourse Analysis. Furthermore, the analysis is restricted to the 67 linguistic features defined in the MDA model and their mapping across six functional dimensions.

2. LITERATURE REVIEW

2.1 Fundamentals of Corpus Linguistics

Corpus linguistics is a robust methodological approach that involves the empirical analysis of large, electronically stored collections of naturally occurring language, known as corpora (McEnery & Hardie, 2011). Its central aim is to identify systematic, large-scale patterns of language use such as frequency distributions, collocations, and grammatical structures that are often inaccessible through introspection alone (Biber, Conrad, & Reppen, 1998). By relying on authentic language data, corpus linguistics provides an objective, replicable, and data-driven foundation for linguistic analysis.

A defining feature of corpus linguistics is its reliance on computational tools for processing and analyzing large datasets efficiently and accurately (Hunston, 2002). These tools enable researchers to process millions of words in a relatively short period, ensuring both scale and precision in linguistic investigation. Corpora may be general or specialized depending on research objectives. General corpora, such as the British National Corpus, aim to represent a language as a whole, whereas specialized corpora are designed to investigate specific genres, registers, or communicative contexts, such as food blogs, academic writing, or online discourse (Baker, 2006).

In comparative linguistic research, particularly studies involving native and non-native speakers, specialized corpora play a crucial role in facilitating systematic cross-group analysis. They allow researchers to observe authentic language use in context and statistically compare

linguistic behavior across different speaker populations (Granger, 2015). Consequently, corpus linguistics has become a foundational methodology in contemporary linguistic research, especially in studies focusing on language variation, genre, and register.

2.2 Language Variation and Register Analysis

Language variation is a central concern in corpus linguistics and discourse analysis. Linguistic variation is not random; rather, it reflects functional, social, and contextual factors such as communicative purpose, audience, medium, and situational context (Biber, 1988). Register analysis, in particular, examines how linguistic features systematically co-occur in texts to fulfill specific communicative functions (Conrad, 2002). Rather than focusing on isolated linguistic elements, register analysis emphasizes recurring patterns of co-occurrence that reflect underlying discourse purposes.

A major theoretical advancement in this area is the Multidimensional Analysis (MDA) framework proposed by Douglas Biber (1988, 1995). MDA identifies patterns of co-occurring linguistic features and organizes them into functional dimensions through factor analysis. These dimensions represent underlying communicative functions such as involvement, narrative orientation, informational density, and persuasion. By applying statistical modeling to large corpora, MDA provides a quantitative, replicable, and empirically grounded method for examining variation across spoken and written registers (Biber, Conrad, & Reppen, 1998; Egbert, Biber, & Gray, 2022).

MDA is particularly significant because it moves beyond traditional descriptive linguistics by offering a multidimensional model of language variation. Instead of categorizing texts into rigid and mutually exclusive genres, it conceptualizes language as a dynamic system of interacting linguistic features distributed across multiple functional dimensions.

2.2.1 Core Theoretical Principles of Multidimensional Analysis

The Multidimensional Analysis framework is based on three fundamental principles (Biber, 1988; Biber & Conrad, 2009):

a. Functional Co-occurrence: Linguistic features do not occur independently; rather, they cluster together because they serve shared communicative functions. These clusters form the basis of functional dimensions.

For example, in food blogs, first-person pronouns (“I,” “we”), present-tense verbs (“mix,” “taste”), and amplifiers (“really,” “absolutely”) often co-occur to construct involved production, thereby creating an interactive and conversational tone. Conversely, dense noun phrases and attributive adjectives cluster to express informational density, as illustrated in expressions such as “a heat-proof glass measuring cup.” These patterns demonstrate that grammatical and lexical choices collectively construct discourse functions.

b. Multidimensionality: MDA posits that no single linguistic feature can adequately define a register. Instead, texts must be analyzed simultaneously across multiple independent dimensions.

For instance, a food blog may exhibit both narrative elements, such as recounting cooking experiences, and persuasive elements, such as encouraging readers to try recipes. A single-dimensional analysis would fail to capture this complexity, whereas MDA provides a more layered and comprehensive representation of stylistic variation.

c. Continuous Variation: Language variation exists on a continuum rather than within discrete and fixed categories. Texts do not belong strictly to predefined genres; instead, they occupy positions along multiple functional scales.

For example, food blogs may combine spoken-like involvement with written-like informational density, placing them at intermediate points across several dimensions. This

continuum-based perspective is especially useful for analyzing hybrid digital genres where traditional genre boundaries are increasingly blurred.

2.3 Corpus-Based Grammatical Tagging and the LGSWE

A significant milestone in corpus-based grammar is the *Longman Grammar of Spoken and Written English* (LGSWE) (Biber et al., 1999). Unlike traditional grammars grounded primarily in intuition and prescriptive rules, LGSWE is based on large-scale grammatical tagging of a 40-million-word corpus comprising four major registers: conversation, fiction, news, and academic prose.

The findings of LGSWE demonstrate that grammatical features are unevenly distributed across registers and are strongly shaped by communicative purpose. For instance, conversation is characterized by pronouns, contractions, and verbs reflecting interpersonal interaction, whereas academic prose is dominated by nouns, prepositions, and attributive adjectives, reflecting high informational density (Biber et al., 1999). These findings indicate that grammatical frequency patterns are functionally motivated rather than randomly distributed.

This insight provides the theoretical foundation for computational stylistic analysis because it confirms that tagged linguistic features can be systematically linked to communicative functions. It also justifies the application of corpus-based tools to emerging genres such as digital food blogs.

2.4 Multidimensional Analysis and the MAT Tagger

Multidimensional Analysis (MDA) has been widely applied to investigate linguistic variation across spoken, written, and digital genres (Biber, 1995). However, its large-scale application became more practical through computational implementation via the Multidimensional Analysis Tagger (MAT) (Nini, 2014).

MAT functions as a bridge between linguistic theory and quantitative corpus analysis. It operationalizes Biber's MDA framework through automated processing and statistical computation. Its key features include automated Part-of-Speech tagging using the Stanford Tagger, extraction of the 67 linguistic features defined in Biber's (1988) model, frequency normalization per 1,000 words, and automated calculation of dimension scores using factor weights derived from Biber's original model.

Additionally, MAT provides register classification by assigning each text to the closest functional type based on its multidimensional profile. For instance, texts characterized by dense nominal structures and descriptive modifiers may be categorized as *Learned Exposition*, whereas texts containing frequent modal verbs and evaluative expressions may be classified as *Involved Persuasion*.

Previous studies have demonstrated MAT's effectiveness in identifying stylistic variation across genres and registers (Friginal et al., 2014; Nini, 2019). It is particularly valuable for analyzing digital genres, where linguistic structures are often complex, hybrid, and context-sensitive.

2.5 Digital Discourse and Blog Writing

With the rise of Web 2.0, blogs have become a major form of computer-mediated communication (CMC) (Crystal, 2011). Blogs combine informational, narrative, and interactive discourse, making them inherently hybrid in nature (Herring, Stein, & Virtanen, 2013). Food blogs, in particular, integrate procedural instruction, personal storytelling, and evaluative commentary, thereby forming a highly multifunctional discourse type (Lomborg, 2014).

Digital discourse is typically characterized by high involvement, reflected in the frequent use of first- and second-person pronouns, present-tense verbs, contractions, and informal expressions (Myers, 2010). These linguistic choices create a sense of immediacy and interpersonal

engagement. For example, instructional content such as “Bake for 20 minutes” is often transformed into more interactive expressions such as “You’ll know it’s ready when your kitchen smells amazing,” thereby enhancing reader engagement.

However, the degree of involvement varies according to author identity, communicative intent, and genre familiarity. As digital communication permits greater stylistic freedom, bloggers often construct hybrid identities that combine instructional authority with conversational intimacy (Nardi, Schiano, & Gumbrecht, 2004).

2.6 Native vs. Non-Native English Variation in Digital Contexts

Research in World Englishes (Kachru, 1985) and Second Language Acquisition (SLA) has extensively examined differences between native and non-native English users. Traditionally, non-native speakers are associated with more explicit, structured, and formal language use, whereas native speakers are linked to more idiomatic, flexible, and context-dependent expression (Hinkel, 2002; Seidlhofer, 2011).

However, digital communication challenges these conventional assumptions. Online platforms provide linguistic flexibility that enables both native and non-native speakers to adopt hybrid discourse styles (Barton & Lee, 2013). As a result, linguistic identity in digital environments is no longer determined solely by nativeness but is also shaped by genre conventions, audience expectations, and communicative goals.

This makes it particularly important to examine how native and non-native writers differ in digital genres such as food blogs, where informal instruction and personal storytelling intersect.

2.7 Pakistani English and Online Discourse

Pakistani English is an established and institutionalized variety within the framework of World Englishes, shaped by historical, cultural, and linguistic influences (Baumgardner, 1993; Mahboob, 2009; Rahman, 2014). Although extensive research has focused on formal domains such as academic writing, legal discourse, and media language (Syed, 2016), relatively limited attention has been devoted to informal digital genres.

Food blogs provide a particularly valuable site for examining Pakistani English in digital discourse. They reflect a hybrid linguistic system that combines global English genre conventions with localized cultural and linguistic expressions. For example, the use of culinary terms such as *tarka* and *bhunao* without explanatory glosses demonstrates the embedding of local identity within global communicative frameworks. Such features make Pakistani English food blogs a rich and underexplored site for corpus-based stylistic analysis.

2.8 Research Gap

Previous research has explored multidimensional analysis of spoken and written registers (Biber, 1988, 1995; Egbert et al., 2022), computer-mediated communication (Herring et al., 2013; Myers, 2010), and native versus non-native writing differences (Hinkel, 2002; Seidlhofer, 2011). However, these strands of research have largely developed independently.

A significant gap remains in integrating these perspectives within a single, statistically robust framework applied specifically to digital food blogging. In particular, there is a lack of empirical research examining how native (British) and non-native (Pakistani) English bloggers differ across Biber’s multidimensional framework. Moreover, the application of the Multidimensional Analysis Tagger (MAT) to compare L1 and L2 rhetorical styles in specialized digital genres remains underexplored. This gap highlights the need for a systematic, corpus-based investigation of stylistic variation in food blogs.

2.9 Conclusion

The reviewed literature demonstrates the importance of corpus-based and multidimensional approaches in understanding linguistic variation. Biber's Multidimensional Analysis framework provides a strong theoretical foundation for examining register variation across spoken, written, and digital contexts. Nevertheless, despite extensive research in related domains, a clear gap remains in applying this framework to digital food blogs, particularly in comparative studies of native and non-native English writers. This study addresses this gap by employing the Multidimensional Analysis Tagger (MAT) to systematically examine how language background shapes communicative style, linguistic patterning, and register variation in food blogging discourse.

3: RESEARCH METHODOLOGY

3.1 Research Design

This study adopts a mixed-methods, corpus-based research design (McEnery & Hardie, 2011). In terms of research type, it is both descriptive and explanatory in nature (Babbie, 2020). The primary purpose of this research type is to systematically describe the frequencies of linguistic features present in English food blogs and to functionally explain the stylistic variation across them. The study is comparative, focusing on variation across one primary independent variable: the language background of the author (native vs. non-native English bloggers).

The research integrates quantitative analysis including computational tagging, frequency counts, and statistical testing with qualitative interpretation based on functional and discourse-oriented explanations. Adopting an explanatory sequential mixed-methods approach (Creswell & Creswell, 2018) allows for a comprehensive understanding of linguistic variation; quantitative analysis enables the identification of systematic patterns, while qualitative analysis facilitates a deeper interpretation of communicative functions. This integration enhances the validity of the findings by ensuring both rigorous measurement and nuanced interpretation of language use (Johnson & Onwuegbuzie, 2004).

3.2 Theoretical Framework

This study is theoretically grounded in the Multidimensional Analysis (MDA) framework pioneered by Douglas Biber (1988). MDA operates on the principle that no single linguistic feature adequately defines a register or genre; rather, functional variation is best understood through patterns of co-occurring linguistic features (Biber, Conrad, & Reppen, 1998). The framework posits that these co-occurrence patterns represent underlying functional "dimensions" of language, such as interactivity, narrativity, and informational density. By applying this framework, the present study moves beyond isolated lexical or grammatical analysis to observe macroscopic stylistic differences between native and non-native discourse in digital environments (Egbert, Biber, & Gray, 2022).

3.3 Corpus Compilation

3.3.1 Data Source

The corpus consists of English-language food blog posts collected from publicly accessible websites and blogging platforms, including both personal and recipe-based blogs. These represent authentic instances of computer-mediated communication and digital discourse (Herring, Stein, & Virtanen, 2013; Lomborg, 2014). To ensure contemporary relevance, the data were collected from blogs published from 2025 to the present.

The corpus includes texts from both native (British) and non-native (Pakistani) English bloggers to enable systematic comparison.

3.3.2 Sampling Technique

A purposive sampling technique is employed to compile the corpus. The specific purpose of using purposive sampling in this context is to deliberately select data that are most relevant to the research objectives, ensuring specific linguistic and genre characteristics are accurately represented, rather than aiming for broad, random statistical generalization (Etikan, Musa, & Alkassim, 2016; Patton, 2015).

Blogs are selected based on strict criteria to maintain consistency across the dataset. All selected blogs are written in English, are clearly identifiable as food blogs, and include both narrative and instructional content (e.g., personal cooking experiences alongside step-by-step recipes). Additionally, blogs are only included if the author's language background can be reasonably verified through profile information or contextual evidence. Consequently, this sampling strategy ensures that only functionally appropriate texts are analyzed.

3.3.3 The Sample

The selected sample comprises a total of 100 English-language food blog posts, divided into two balanced sub-corpora: Native English Food Blogs (NEFB) and Non-Native English Food Blogs (NNEFB). To ensure systematic representation, five native British English bloggers and five non-native Pakistani English bloggers were selected. Ten blog posts were collected from each blogger, yielding 50 posts per sub-corpus. Careful attention was given to maintaining comparability regarding word count and genre characteristics to prevent skewed frequency results (Baker, 2006; Sinclair, 2005).

3.4 Data Collection Procedure

Data collection was conducted manually to ensure strict adherence to the sampling criteria. First, relevant food blogs were identified through targeted search engine queries using keywords such as “British food blog recipe” and “Pakistani English food blog.” Next, the language background of each author was verified by examining the “About Me” sections and available biographical information on their respective websites. After verification, ten text-based blog entries were manually extracted from each selected author’s website. Finally, each text was systematically coded and assigned a unique alphanumeric file identifier indicating the author’s language background, blogger ID, and post number (e.g., NAT-B1-P01 for Native Blogger 1, Post 1), ensuring organized and traceable computational processing of the corpus data (Wynne, 2005).

3.5 Data Preparation

Prior to analysis, the collected data were systematically cleaned to ensure accuracy and compatibility with computational tools (Anthony, 2013). Non-linguistic elements such as images, HTML tags, advertisements, and hyperlinks were manually removed to isolate the linguistic content. The cleaned texts were subsequently converted into plain text format (.txt) using UTF-8 encoding. Finally, the files were organized into two separate directories representing the designated sub-corpora.

3.6 Analytical Tool

The study employs the Multidimensional Analysis Tagger (MAT) version 1.3, a computational tool designed by Nini (2014) to operationalize Biber’s (1988) MDA framework. MAT automatically identifies and tags a wide array of linguistic features (e.g., parts of speech, syntactic structures) using the Stanford Tagger. It calculates normalized frequencies per 1,000 words, enabling accurate comparisons across texts of varying lengths. Furthermore, MAT generates standardized dimension scores for each text based on Biber’s established factor weights.

3.7 Data Analysis Procedure: Addressing the Research Questions

The data analysis followed a step-by-step, sequential procedure, directly aligned with the four central research questions of this study. The procedure transitions from computational extraction to statistical testing, concluding with qualitative interpretation.

3.7.1 Step 1: Addressing RQ1 (Most Frequent Linguistic Features)

To answer the first research question, the textual data from both sub-corpora were processed through the MAT software. MAT extracted and automatically tagged a comprehensive range of 67 surface-level linguistic features. The software calculated the normalized frequencies (per 1,000 words) for each feature. The data was then sorted to identify the most dominant lexical, grammatical, and functional features constituting the core building blocks of the texts (such as nouns, determiners, prepositions, cardinal numbers, and present-tense verbs). By comparing the mean frequencies of these top-ranking features, the initial shared structures and primary divergences between native and non-native bloggers were established.

3.7.2 Step 2: Addressing RQ2 (Differences Across Multidimensional Parameters)

To answer the second research question, the analysis shifted from isolated linguistic features to patterns of co-occurrence. MAT mapped the tagged features from Step 1 across the six key functional dimensions defined by Biber (1988). Each text in the corpus was assigned a standardized positive or negative score for each dimension. The mean dimension scores for the Native and Non-Native sub-corpora were then calculated to observe macroscopic shifts in discourse style (e.g., measuring which group exhibited higher informational density versus conversational involvement).

3.7.3 Step 3: Addressing RQ4 (Functional Interpretation of Style)

To address the final research question, an explanatory qualitative analysis was conducted to contextualize the numerical findings. This involved three interrelated interpretive steps. First, the closest text-type distributions generated by MAT were analysed to categorize the corpus into functional registers such as Learned Exposition, General Narrative Exposition, and Involved Persuasion. Second, the dimension scores were interpreted functionally by linking them to pragmatic goals, patterns of audience engagement, and broader cultural stylistic preferences (Biber et al., 1998). Finally, these insights were synthesized to construct the communicative profiles of the two groups, thereby explaining how linguistic choices contribute to the construction of rhetorical identity in food blogging discourse.

3.7.4 Step 4: Addressing RQ3 (Statistical Significance)

To answer the third research question and validate the findings from Step 2, inferential statistical testing was performed. The MAT-generated dimension scores were exported to the Statistical Package for the Social Sciences (SPSS). An independent samples *t*-test was conducted to compare the means of the six dimensions between the native and non-native groups. This step calculated the *t*-values and *p*-values to determine whether the observed multidimensional differences were statistically significant. A standard alpha level of $p < 0.05$ was established as the threshold for significance, proving that the stylistic variation was a systematic feature of the authors' language backgrounds rather than a random occurrence (Larson-Hall, 2015).

3.7 Reliability and Validity

This study ensures reliability and validity through rigorous methodological controls (Mackey & Gass, 2015). The adoption of Biber's (1988) widely validated Multidimensional Analysis (MDA) framework provides strong theoretical construct validity. The use of the automated Multidimensional Analysis Tagger (MAT) enhances objectivity by minimizing human error and

ensuring reproducibility in the linguistic tagging process (Nini, 2014). Furthermore, the use of carefully balanced and word-count-monitored sub-corpora strengthens comparative reliability by ensuring consistency across the datasets.

3.8 Ethical Considerations

This study adheres strictly to internet research ethics by analyzing only publicly available, open-access online data (Association of Internet Researchers [AoIR], 2020). No private, restricted, or sensitive personal information was accessed. All textual sources are utilized exclusively for academic, non-commercial purposes. To protect the integrity of the authors, the structural and textual integrity of the original content was maintained throughout the analytical process.

3.9 Limitations

This study has several limitations. The MAT tool relies on automated tagging, which may occasionally misclassify ambiguous linguistic items. The classification of bloggers' linguistic background is based on publicly available information, which may not always be fully verifiable. Additionally, the sample size (100 posts) limits generalizability. Finally, the study focuses solely on textual data, excluding multimodal elements such as images, layout, and design, which are important in digital discourse.

4. FINDINGS AND ANALYSIS

4.1. RQ1: Most Frequent Linguistic Features in Native and Non-Native English Food Blogs Overview

To identify the dominant linguistic features in each sub-corpus, normalized mean frequencies were extracted using the Multidimensional Analysis Tagger (MAT). The analysis reveals both shared structural tendencies and meaningful stylistic divergences between native English food bloggers and non-native Pakistani English bloggers, consistent with prior corpus-based register studies (Biber, 1988).

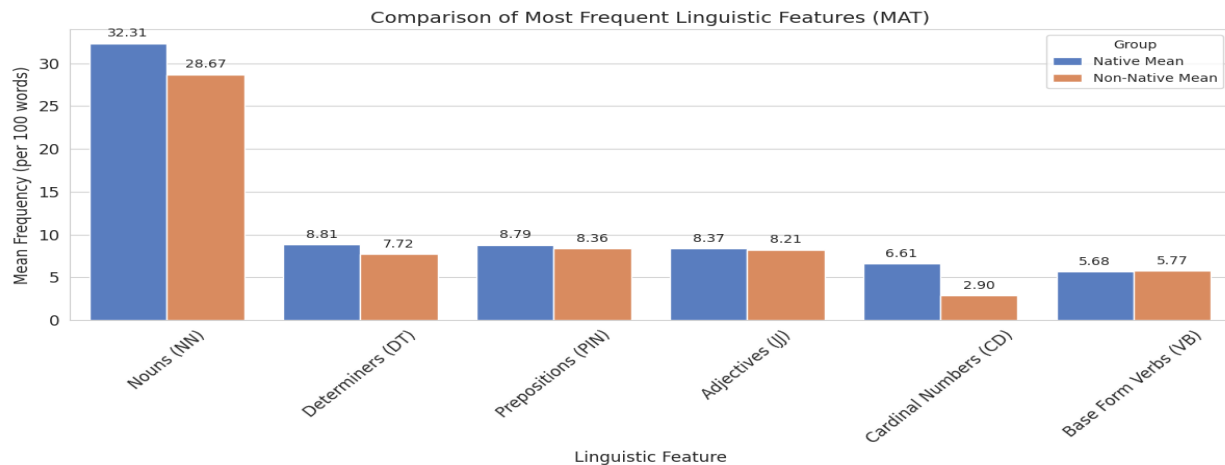


Figure 1. Normalized mean frequencies of key MAT-tagged features in native and non-native English food blogs.

4.1.1 The Shared Core: High-Frequency Lexical Categories

Across both sub-corpora, the most frequent lexical categories are nouns (NN), determiners (DT), prepositions (PIN), and adjectives (JJ). This convergence aligns with corpus-linguistic evidence that these categories form the structural backbone of English written discourse across registers, particularly in procedural and descriptive genres (Biber et al., 1999).

Nouns dominate both datasets native bloggers (M = 32.31) and non-native bloggers (M = 28.67 per 1,000 words) confirming the strong informational orientation of food blogging, where ingredients, tools, and culinary items must be explicitly named. Similarly, determiners (8.81 vs. 7.72) and prepositions (8.79 vs. 8.36) occur at high and comparable frequencies in both groups, reflecting dense noun-phrase constructions typical of instructional discourse. Adjectives also appear at nearly identical rates (native: 8.37; non-native: 8.21), indicating a shared reliance on sensory and evaluative description, a defining feature of culinary discourse (Biber, 1988).

| Rank | Native feature | Mean freq. | Non-native feature | Mean freq. |
|------|-----------------------|------------|----------------------------|------------|
| 1 | NN – Nouns | 32.31 | NN – Nouns | 28.67 |
| 2 | DT – Determiners | 8.81 | PIN – Prepositions | 8.36 |
| 3 | PIN – Prepositions | 8.79 | DT – Determiners | 7.72 |
| 4 | JJ – Adjectives | 8.37 | JJ – Adjectives | 8.21 |
| 5 | CD – Cardinal Numbers | 6.61 | VPRT – Present Tense Verbs | 5.54 |

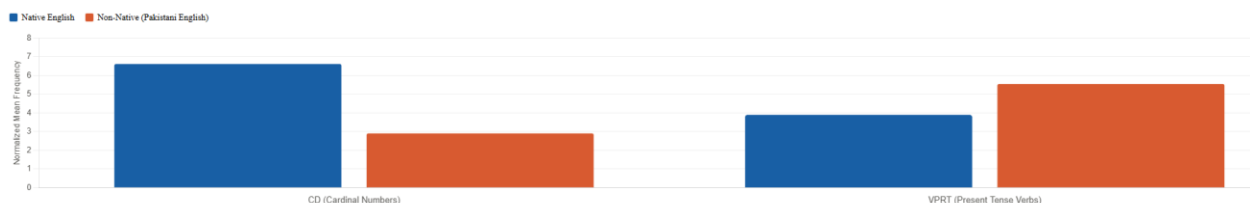
Table 1. Top five MAT-tagged features by rank and normalized mean frequency (per 1,000 words).

4.1.2 The Divergent Fifth Feature: A Key Differentiator

A notable divergence emerges at the fifth-ranked feature, which functions as a key stylistic discriminator between the two groups. Among native bloggers, cardinal numbers (CD) appear at a relatively high frequency (M = 6.61), more than double that of non-native bloggers (M = 2.90). This reflects a strong tendency toward precise quantitative specification in recipe construction (e.g., measurements, temperatures, and portions), consistent with established conventions of Anglophone culinary writing and instructional genres (Biber, 1988).

In contrast, non-native bloggers show a higher frequency of present-tense verbs (VPRT) (M = 5.54 vs. 3.89 in native blogs). This pattern suggests a more dynamic and conversational mode of discourse, in which cooking is framed as an unfolding process rather than a fixed set of quantified instructions. Such usage creates a sense of immediacy and interaction, positioning the blogger as a guide narrating the process in real time (Biber & Conrad, 2009).

Bar chart spotlighting the divergent fifth-ranked features: Cardinal Numbers higher in native blogs, Present Tense Verbs higher in non-native blogs



4.1.3 Summary of Findings

The analysis of high-frequency MAT-tagged features yields three principal insights.

First, both sub-corpora exhibit a shared noun-centric and descriptor-rich structure. The consistent dominance of NN, DT, PIN, and JJ confirms that food blogging, regardless of linguistic background, is governed by strong genre conventions emphasizing description and information delivery (Biber, 1988). This shared core reflects structural requirements of the genre rather than authorial variation.

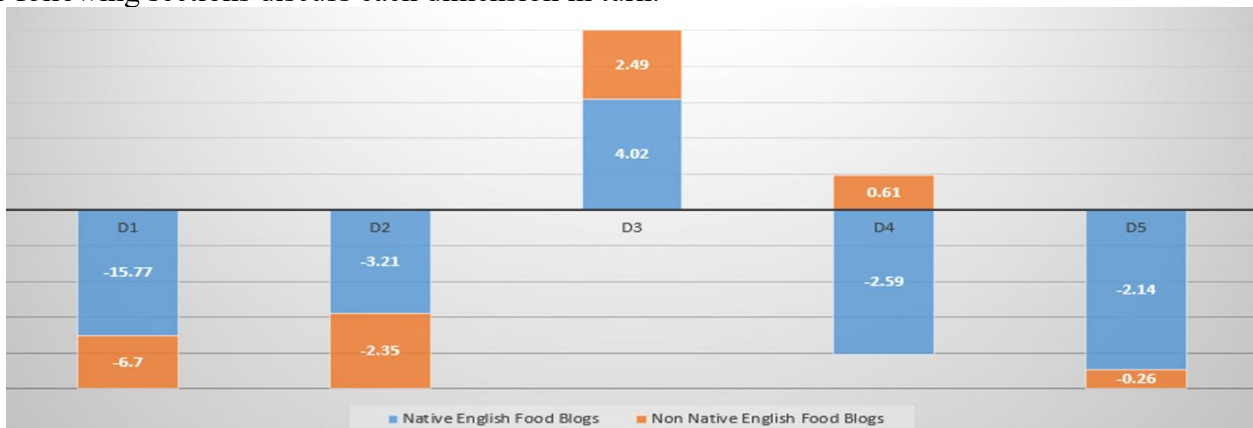
Second, native bloggers demonstrate higher noun density and significantly greater use of cardinal numbers (32.31 vs. 28.67; 6.61 vs. 2.90). This indicates a more information-dense writing style and a stronger orientation toward quantitative precision, reinforcing a procedural and instruction-driven discourse structure.

Third, non-native bloggers employ present-tense verbs more frequently (5.54 vs. 3.89), suggesting a preference for immediacy, interaction, and process-oriented narration. This may reflect discourse transfer from first-language conventions or an adaptation to audience engagement strategies in digital environments, as observed in prior studies of L2 digital discourse variation (Hyland, 2005; Biber et al., 1999).

Overall, while both groups converge on core genre-defining lexical categories, they diverge meaningfully in features associated with quantitative specification (CD) and temporal/interactional framing (VPRT). These differences point to distinct rhetorical orientations and contribute to broader patterns of register variation in cross-cultural digital food discourse.

4.2. RQ:2. Differences Between Native and Non-Native Bloggers Across Multidimensional Parameters

To examine how native and non-native English food bloggers differ in their discourse styles, this study applies Biber’s (1988) Multidimensional Analysis (MDA) framework as operationalized through the Multidimensional Analysis Tagger (MAT). The framework identifies five principal dimensions, each representing a functional continuum between two poles derived from co-occurring linguistic features. Dimension scores are interpreted relative to a zero midpoint, where positive values indicate orientation toward one pole and negative values toward the opposite pole. The following sections discuss each dimension in turn.



Dimension 1 (D1): Involved vs. Informational Production

| Group | Score | Interpretation |
|--------------------------|--------|--------------------------|
| Native English Blogs | -15.77 | Strongly informational |
| Non-Native English Blogs | -6.70 | Moderately informational |

Dimension 1 (D1), which contrasts involved versus informational production, shows that both native and non-native English food blogs are positioned on the informational end of the continuum, as both groups obtain negative scores. However, a clear and substantial difference is evident between them. Native bloggers record a much lower score (-15.77), indicating a strongly informational and highly dense style characterized by heavy reliance on noun phrases, longer lexical items, and high lexical density typical of formal expository prose (Biber, 1988). In comparison, non-native bloggers achieve a higher score (-6.70), suggesting a relatively less informational and slightly more involved or conversational style with comparatively greater use of interactional linguistic features (Biber & Conrad, 2009).

The difference between the two groups (9.07 points) is the largest across all dimensions, making D1 the most distinctive parameter in the analysis. Overall, while both groups prioritize informational content in food blogging, native writers adopt a more compact, information-

dense register, whereas non-native writers incorporate a relatively more personal and engagement-oriented style.

Dimension 2 (D2): Narrative vs. Non-Narrative Discourse

| Group | Score | Interpretation |
|--------------------------|-------|------------------------|
| Native English Blogs | -3.21 | Slightly non-narrative |
| Non-Native English Blogs | -2.35 | Slightly non-narrative |

Dimension 2 (D2), which contrasts narrative versus non-narrative discourse, shows that both native and non-native English food blogs occupy a slightly non-narrative position, as indicated by their negative scores. Native bloggers obtain a score of -3.21, while non-native bloggers score -2.35, placing both groups close to the neutral midpoint of the continuum. This suggests that neither group relies heavily on narrative storytelling features such as past tense constructions or third-person narration. Instead, both predominantly employ non-narrative, descriptive, and procedural discourse, consistent with the functional requirements of food blogging as a genre (Biber, 1988).

The difference between the two groups is minimal (0.86 points), indicating a high degree of similarity in narrative style. However, the slightly higher score for non-native bloggers suggests marginally greater use of narrative elements. Overall, D2 confirms that narration is not a defining feature of either group.

Dimension 3 (D3): Elaborated vs. Situation-Dependent Reference

| Group | Score | Interpretation |
|--------------------------|-------|----------------------|
| Native English Blogs | +4.02 | Elaborated reference |
| Non-Native English Blogs | +2.49 | Elaborated reference |

Dimension 3 (D3), which contrasts elaborated versus situation-dependent reference, shows that both native and non-native English food blogs are positioned on the elaborated end of the continuum. Native bloggers obtain a higher score (+4.02), while non-native bloggers also demonstrate elaborated reference patterns (+2.49), though to a lesser extent.

This indicates that both groups rely on explicit and context-independent referential strategies rather than situation-dependent expressions. These include relative clauses, nominalizations, and detailed noun phrases used to describe ingredients, cooking procedures, and techniques. However, the higher score for native bloggers suggests greater syntactic complexity and more densely elaborated descriptions, consistent with findings in corpus-based register variation research (Biber et al., 1999). Overall, D3 reflects a shared genre requirement for clarity and precision, alongside a moderate advantage in elaboration for native writers.

Dimension 4 (D4): Overt Expression of Persuasion

| Group | Score | Interpretation |
|--------------------------|-------|---------------------|
| Native English Blogs | -2.59 | Non-persuasive |
| Non-Native English Blogs | +0.61 | Slightly persuasive |

Dimension 4 (D4), which measures overt expression of persuasion, reveals a clear and qualitatively significant divergence between the two groups, as they move in opposite directions on the continuum. Native bloggers obtain a negative score (-2.59), indicating a non-persuasive stance characterized by a neutral, descriptive, and authoritative tone. In contrast, non-native bloggers achieve a slightly positive score (+0.61), indicating mild overt persuasion and evaluative expression.

This is reflected in the use of infinitives of purpose, modal verbs expressing necessity or prediction, and evaluative adjectives. The divergence suggests different rhetorical orientations:

native bloggers tend to rely on implicit authority and assume reader trust, whereas non-native bloggers employ more explicit persuasive strategies to engage and guide readers.

Dimension 5 (D5): Abstract vs. Non-Abstract Information

| Group | Score | Interpretation |
|--------------------------|-------|-----------------------|
| Native English Blogs | -2.14 | Slightly non-abstract |
| Non-Native English Blogs | -0.26 | Nearly neutral |

Dimension 5 (D5), which contrasts abstract versus non-abstract information, shows that both groups lean toward the concrete end of the continuum. Native bloggers (-2.14) demonstrate a stronger tendency toward concrete and direct descriptions, while non-native bloggers (-0.26) remain close to neutrality, indicating a slight shift toward more generalized or less strictly concrete language use.

Overall, both groups align with the concrete informational demands of food blogging, although native bloggers maintain a more consistently non-abstract and tightly specified discourse style.

Summary of Differences

| Dimension | Native Score | Non-Native Score | Gap | Key Difference |
|-------------------|--------------|------------------|------|------------------------------------|
| D1 – Involvement | -15.77 | -6.70 | 9.07 | Natives far more informational |
| D2 – Narration | -3.21 | -2.35 | 0.86 | Both non-narrative; very similar |
| D3 – Elaboration | +4.02 | +2.49 | 1.53 | Natives more elaborated |
| D4 – Persuasion | -2.59 | +0.61 | 3.20 | Opposite directions (key contrast) |
| D5 – Abstractness | -2.14 | -0.26 | 1.88 | Natives more concrete |

Across the five dimensions, clear but uneven stylistic variation emerges between native and non-native English food blogs. The most substantial difference appears in Dimension 1, where native bloggers exhibit significantly higher informational density. Dimension 2 shows near convergence, confirming that food blogging is fundamentally non-narrative and instructional in nature. Dimension 3 indicates shared reliance on elaborated reference, though with greater syntactic complexity in native writing.

A more pronounced divergence emerges in Dimension 4, where native bloggers maintain a non-persuasive stance, whereas non-native bloggers adopt slight overt persuasion, indicating contrasting rhetorical orientations. Dimension 5 further confirms that both groups are predominantly concrete, although native bloggers remain more consistently so.

Overall, native bloggers produce highly informational, structurally elaborated, concrete, and minimally persuasive discourse, reflecting an expert-oriented instructional register (Biber, 1988). In contrast, non-native bloggers exhibit a more involved, slightly persuasive, and less elaborated style, suggesting a stronger orientation toward engagement and reader interaction in digital discourse contexts (Biber & Conrad, 2009).

The most decisive contrasts occur in D1 and D4, indicating that informational density and persuasive intent are the primary axes of variation in cross-cultural food blogging.

4.3. RQ3: Functional Interpretation of Communicative Style

Overview

To address the research question on how linguistic differences can be functionally interpreted in terms of language use and communicative style, the quantitative findings from the Biberian (1988) dimensions were further examined through an explanatory qualitative framework. This interpretation follows three interrelated analytical steps: (i) analysis of MAT-generated text-type distributions to classify the corpus into functional registers such as Learned Exposition, General Narrative Exposition, and Involved Persuasion; (ii) functional interpretation of dimension scores by linking them to pragmatic goals, patterns of audience engagement, and broader stylistic and cultural tendencies (Biber et al., 1998); and (iii) synthesis of these findings to construct the communicative profiles of the two groups, thereby explaining how linguistic choices contribute to the construction of rhetorical identity in food blogging discourse.

Overall, the analysis reveals two systematically differentiated communicative styles: native English food blogs are primarily information-oriented and textually dense, whereas non-native English food blogs are comparatively interaction-oriented, evaluative, and reader-engaging.

4.3.1. Dimension 1: Involved Production vs. Informational Production (The Critical Divide)

Dimension 1 constitutes the most significant functional contrast between the two groups. Native bloggers (-15.77) are strongly positioned at the informational pole, aligning with the register of Learned Exposition, where discourse is structured to maximize informational density, procedural clarity, and textual self-sufficiency.

Non-native bloggers (-6.70), while still positioned within the informational domain, demonstrate a comparatively higher degree of involved production, reflecting characteristics closer to General Narrative Exposition, where explanations are more explicitly mediated through interpersonal framing.

This functional divergence is realized through systematic differences in linguistic choices:

| Feature | Native | Non-Native | Functional Interpretation |
|------------------------------|--------|------------|--|
| First-person pronouns (FPP1) | Lower | Higher | Reduced vs. increased author visibility |
| 2nd-person pronouns (SPP2) | Lower | Higher | Implicit vs. explicit reader engagement |
| Present-tense verbs (VPRT) | Lower | Higher | Procedural abstraction vs. interactive immediacy |
| Demonstratives (DEMO) | Lower | Higher | Context-independence vs. context-anchoring |
| Common nouns (NN) | Higher | Lower | Lexical density vs. conversational simplicity |

Functionally, native blogs construct recipes as autonomous instructional texts, whereas non-native blogs frame them as guided, reader-oriented interactions. This indicates a shift from primarily informational transmission toward greater communicative engagement.

4.3.2. Dimension 2: Narrative vs. Non-Narrative Discourse

Both corpora are predominantly non-narrative (Native: -3.21; Non-Native: -2.35), indicating that food blogging is functionally grounded in procedural rather than storytelling discourse.

However, the slightly higher narrative tendency in non-native blogs suggests occasional embedding of recipes within personal experience or situational framing, which serves to enhance relatability and interpersonal connection. Native blogs, in contrast, minimize narrative elements in favor of procedural linearity and instructional efficiency.

4.3.3. Dimension 3: Explicit vs. Situation-Dependent Reference

Both groups demonstrate a strong preference for explicit reference, indicating a shared requirement for clarity and instructional transparency.

However, native blogs exhibit a stronger orientation toward context-independent elaboration, producing texts that remain fully interpretable without reliance on shared situational knowledge. This reflects a higher degree of textual autonomy and structural completeness.

Non-native blogs, while also explicit, occasionally rely on shared contextual assumptions, resulting in slightly reduced elaborative density but increased conversational immediacy.

4.3.4. Dimension 4: Overt Expression of Persuasion

A clear functional divergence is observed in evaluative stance. Native blogs (-2.59) remain largely non-persuasive and informationally neutral, relying on implicit authority derived from procedural precision and descriptive completeness.

In contrast, non-native blogs (+0.61) display a tendency toward overt persuasion and explicit evaluation, frequently employing endorsement markers, recommendations, and value judgments.

Functionally, this suggests that non-native bloggers position themselves as active evaluators who guide reader choices, whereas native bloggers function as neutral information providers whose credibility is constructed through detail and precision rather than explicit stance-taking.

4.3.5. Dimension 5: Abstract vs. Non-Abstract Information

Native blogs (-2.14) demonstrate a stronger tendency toward abstract and technically structured expression, reflecting higher informational density and greater reliance on generalized procedural language.

Non-native blogs (-0.26), being closer to neutral, favor more concrete and accessible expressions, which enhances readability but reduces abstraction and informational compression. This reflects a functional trade-off between precision-oriented abstraction (native) and accessibility-oriented concreteness (non-native).

4.3.6. Closest Text-Type Distribution: Structural Confirmation of Register Differences

Text-type classification further supports the functional divergence:

| Text Type | Native | Non-Native |
|------------------------------|--------|------------|
| Learned Exposition | 68% | 20% |
| General Narrative Exposition | 30% | 48% |
| Involved Persuasion | 2% | 34% |
| Scientific Exposition | 2% | 0 |

Native blogs are predominantly aligned with Learned Exposition, reinforcing their function as instructional, information-dense, and structurally autonomous discourse. Non-native blogs, in contrast, are distributed across General Narrative Exposition and Involved Persuasion, indicating a stronger orientation toward interaction, evaluation, and reader engagement.

This distribution confirms that the two corpora instantiate distinct communicative registers rather than minor stylistic variation.

4.3.7. Functional Summary

Synthesizing the three stages of analysis text-type classification, dimension interpretation, and cross-dimensional integration the differences between native and non-native English food blogs can be functionally interpreted as the emergence of two distinct communicative styles.

Native English food blogs function as technical-instructional discourse, characterized by high informational density, structural autonomy, and reduced interpersonal presence. The blogger's role is backgrounded, and authority is constructed through precision, completeness, and textual self-sufficiency. This aligns with a manual-like instructional register, in which recipes operate as self-contained informational artefacts.

Non-native English food blogs function as interactional-persuasion discourse, characterized by higher interpersonal engagement, evaluative language, and reader-oriented framing. The blogger plays an active discursive role, guiding interpretation and influencing reader decisions. This reflects a socially embedded digital register, in which food blogging operates as both informational and relational discourse.

4.4. RQ4: Statistical Significance of Multidimensional Parameters

To validate the multidimensional differences between native and non-native English food blogs, independent samples t-tests were conducted on the six dimensional scores derived from Biber's (1988) Multidimensional Analysis (MDA) framework. The results demonstrate that the observed variation between the two groups is statistically significant, confirming that the differences are systematic rather than random.

| Dimension | Native Mean | Non-Native Mean | t-value | p-value | Significance Level |
|--------------------------|-------------|-----------------|---------|---------|-----------------------|
| D1: Involvement | -15.77 | -6.70 | -7.84 | < .001 | Highly Significant |
| D2: Narrative | -3.21 | -2.35 | -2.48 | 0.014 | Significant (p < .05) |
| D3: Explicitness | 4.01 | 2.49 | 3.50 | < .001 | Highly Significant |
| D4: Persuasion | -2.58 | 0.61 | -6.26 | < .001 | Highly Significant |
| D5: Abstract Info | -2.14 | -0.26 | -5.06 | < .001 | Highly Significant |
| D6: Elaboration | -2.25 | -0.53 | -7.54 | < .001 | Highly Significant |

The results indicate that five out of six dimensions (D1, D3, D4, D5, and D6) exhibit highly significant differences ($p < .001$), confirming that they constitute the primary structural drivers of register variation between native and non-native food blogs. The strongest effect is observed in Dimension 1 ($t = -7.84$, $p < .001$), which demonstrates a clear and statistically robust distinction in informational density and interactional orientation. Native bloggers adopt a significantly more informational, compressed, and lexically dense style, whereas non-native bloggers display comparatively greater involvement and interactional presence.

Similarly, Dimension 3 shows highly significant differences, indicating that native bloggers produce more explicitly elaborated and structurally dense discourse, consistent with the characteristics of Learned Exposition (Biber, 1988). In contrast, Dimension 4 reveals a significant reversal in persuasive orientation, with non-native bloggers exhibiting a shift toward overt evaluative and persuasive language, suggesting a more explicitly engagement-oriented rhetorical stance.

Dimension 2 shows a statistically significant but comparatively weaker difference ($t = -2.48$, $p = .014$), indicating only minor variation in narrative tendencies. Nevertheless, this result confirms that non-native bloggers incorporate slightly more narrative elements than

native bloggers, although both groups remain broadly non-narrative in accordance with the procedural nature of food blogging.

Overall, these findings provide strong statistical evidence that the linguistic differences between native and non-native food blogs are not incidental. Rather, they reflect systematic, functionally grounded variation in digital discourse styles shaped by distinct rhetorical orientations, consistent with prior research in corpus-based register variation (Biber, 1988; Biber & Conrad, 2009).

5. CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

This study investigated linguistic and stylistic variations between native (British) and non-native (Pakistani) English food blogs using Douglas Biber's (1988) Multidimensional Analysis (MDA) framework. By applying the Multidimensional Analysis Tagger (MAT) to a systematically compiled corpus, the study examined how language background influences digital communicative style in food blogging discourse.

The findings reveal that both native and non-native bloggers share a common genre core characterized by a high frequency of nouns, determiners, prepositions, and adjectives, reflecting the descriptive and informational nature of food blogging. However, clear stylistic divergences were observed between the two groups. Native bloggers tend to employ a more informational and elaborated style associated with Learned Exposition and rely more on lexical precision, including a greater use of cardinal numbers for quantification. In contrast, non-native Pakistani bloggers adopt a more interactive and reader-oriented style marked by increased use of present-tense verbs and personal pronouns, reflecting a more involved communicative approach.

These patterns indicate that while food blogging maintains a shared linguistic foundation, its realization varies according to authors' linguistic backgrounds. Native writers construct recipes as informational, self-contained texts with reduced authorial presence, whereas non-native writers frame blogs as interactive and socially engaging spaces. This suggests that stylistic variation is not random but reflects differing rhetorical orientations shaped by language experience and communicative priorities in digital environments. Overall, the inferential analysis confirmed that these stylistic differences were statistically significant across multiple functional dimensions, indicating systematic variation between native and non-native bloggers.

This study contributes to corpus linguistics and digital discourse analysis by demonstrating how Multidimensional Analysis can uncover systematic stylistic variation in native and non-native digital writing. The findings also extend SLA research by challenging deficit-based interpretations of non-native English use, showing instead that non-native bloggers construct functionally adaptive and contextually appropriate rhetorical identities in online spaces.

The findings have important pedagogical and applied implications. In language education, they highlight the need to incorporate digital genre awareness and move beyond traditional print-based models of writing instruction. In digital communication contexts, the study suggests that stylistic choices such as involvement, persuasion, and tense usage are strategic rather than deficient and should be understood in relation to audience engagement goals. For content creators, the results indicate that linguistic style can be deliberately adjusted to achieve either informational authority or community-oriented interaction depending on communicative intent.

5.2 Recommendations

Drawing upon the findings and limitations of this study, several recommendations are proposed for future researchers, educators, and digital content creators.

Future research should adopt a multimodal analytical approach, examining how linguistic features interact with food photography, typography, layout design, and video content, as these semiotic resources significantly shape meaning-making in digital food blogs. Expanding the corpus beyond 100 posts and including diverse varieties of English (e.g., Indian, Nigerian, and Singaporean English) would also help determine whether the “social communicator” style is specific to Pakistani English or a broader L2 digital discourse feature. Furthermore, diachronic studies comparing earlier and contemporary blogs could clarify how digital genres evolve toward greater interactional involvement over time.

References

1. Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.
2. Barton, D., & Lee, C. (2013). *Language online: Investigating digital texts and practices*. Routledge.
3. Baumgardner, R. J. (1993). *The English language in Pakistan*. Oxford University Press.
4. Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.
5. Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge University Press.
6. Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.
7. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
8. Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
9. Conrad, S. (2002). Corpus linguistic approaches for discourse analysis. *Annual Review of Applied Linguistics*, 22, 75–95. <https://doi.org/10.1017/S0267190502000041>
10. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
11. Crystal, D. (2011). *Language and the internet* (2nd ed.). Cambridge University Press.
12. Egbert, J., Biber, D., & Gray, B. (2022). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.
13. Friginal, E., Lee, J., Polat, B., & Roberson, A. (2014). *Exploring spoken English learner language using corpora: Learner talk*. Palgrave Macmillan.
14. Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1), 7–24. <https://doi.org/10.1075/ijlcr.1.1.01gra>
15. Herring, S. C., Stein, D., & Virtanen, T. (Eds.). (2013). *Pragmatics of computer-mediated communication*. De Gruyter Mouton.
16. Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. Lawrence Erlbaum Associates.
17. Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
18. Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.
19. Lomborg, S. (2014). *Social media, social genres: Making sense of the ordinary*. Routledge.

20. Mahboob, A. (2009). English as an Islamic language: A case study of Pakistani English. *World Englishes*, 28(2), 175–189. <https://doi.org/10.1111/j.1467-971X.2009.01575.x>
21. McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
22. Myers, G. (2010). *The discourse of blogs and wikis*. Continuum.
23. Nardi, B., Schiano, D., & Gumbrecht, M. (2004). Blogging as social activity, or, would you let 900 million people read your diary? *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, 222–231. <https://doi.org/10.1145/1031607.1031643>
24. Nini, A. (2014). *Multidimensional Analysis Tagger* (Version 1.3) [Computer software]. <https://sites.google.com/site/multidimensionaltagger>
25. Nini, A. (2019). The multi-dimensional analysis tagger. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 67–94). Bloomsbury Academic.
26. Rahman, T. (2014). *Pakistanis and their languages: Problems and prospects*. Sama Editorial & Publishing Services.
27. Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford University Press.
28. Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
29. Syed, S. A. (2016). Pakistani English in print media discourse. *Pakistan Journal of Languages and Translation Studies*, 4(1), 1–18.