



A Corpus Based Study Of Common Lexical Chunks In Official Correspondence

¹Utkarsh Mishra

Scholar of MA Linguistics

¹Centre for Linguistics, School of Language, Literature and Culture Studies

¹Jawaharlal Nehru University, New Delhi, India

Abstract: This paper explores the critical role of lexical chunks in facilitating effective communication within official correspondence. Drawing upon a corpus of 40,626 words extracted from 250 official documents spanning seven sub-disciplines of office communication, the study analyzes recurring lexical patterns using AntConc software and identifies 100 common lexical chunks. Through detailed examination, representative lexical chunks are investigated for their usage and contextual nuances. The findings emphasize the centrality of lexical cohesion and coherence in promoting clarity, conciseness, and professionalism in formal communication settings. This research contributes to understanding the linguistic dynamics of official correspondence and offers practical insights for language educators, office correspondence trainers, government officials, and legal practitioners.

Index Terms - lexical chunks, corpus linguistics, official correspondence, AntConc, formulaic language.

Introduction

Language serves as a fundamental medium of communication, enabling individuals to express thoughts, ideas, feelings, and emotions to others. Jack C. Richards and Richard Schmidt define language as “the system of human communication, comprising the structured arrangement of sounds (or their written representation) into larger units, such as morphemes, words, sentences, and utterances.” Edward Sapir describes language as “a purely human and non-instinctive method of communicating ideas, emotions, and desires through voluntarily produced symbols.” Noam Chomsky articulates language as “a set of sentences, each finite in length and constructed from a finite set of elements.” These perspectives collectively emphasize the structural nature of language as it combines words and phrases as building blocks to express meaning in human interaction.

Linguists often perceive language learning as a process wherein the learner focuses on every linguistic element to acquire the language. While this method serves as one approach to learning, it is not the most effective one due to numerous challenges faced by new learners. Among the most prominent of these linguistic challenges are to memorise and arrange words in their proper usage and collocations following the correct norms of syntactic and sociopragmatic rules.

Grammar, with its multifaceted rules, presents a formidable hurdle for learners, necessitating the discernment and application of diverse syntactic structures (Alkhuli, 2006, p.61). Grammar rules can be complex and vary significantly between languages, leading learners to struggle with understanding and correctly applying them in sentences. Vocabulary acquisition poses its own set of challenges, particularly concerning the memorization and recall of words with multiple meanings (Celce- Murcia & Olshtain, 2000). While learning new words and phrases is essential for language acquisition, learners may encounter difficulties in memorizing and recalling vocabulary, and arranging them with proper collocation particularly when words have multiple meanings and usages. In addressing these challenges, the concept of lexical chunks emerges as a pivotal tool in language learning. Lexical chunks, comprising fixed expressions and common phrases, offer learners readily accessible linguistic units that occur frequently in everyday communication. Other terms for lexical chunks are formulaic sequences, lexicalized sentence stems, and multi-word items.

Learners easily learn those phrases or lexical chunks that occur frequently in their daily communication and are used together in fixed expressions. Examples of such phrases include “in my opinion,” “to cut a long story short,” “once upon a time,” and “how do you do,” among others. These types of lexical bundles are so repetitive in their day-to-day usage that the brain assimilates them as single entities without considering the syntactic structures. This enables the learner to use these lexical chunks seamlessly in their communication. Language chunks provide learners with ready-made phrases and expressions commonly used by native speakers. Instead of focusing solely on individual words, learners use lexical chunks as meaningful expressions in context. Learning language chunks helps learners write/speak more fluently and naturally, as they become accustomed to using language patterns that native speakers employ in everyday conversations.

Official correspondence refers to written communication exchanged between individuals, organizations, or government entities in a formal setting. This type of correspondence typically follows established protocols, formats, and traditions specific to the context in which it occurs. Official correspondence serves a variety of purposes, including conveying information, making requests, issuing instructions, and documenting agreements or decisions. The seven main criteria for effective office correspondence are clarity, conciseness, accuracy, completeness, courtesy, soundness, and thoughtfulness. By adhering to these seven 'C's, an individual can improve the effectiveness and professionalism of their official correspondence, thereby promoting clear communication and positive relationships in formal circumstances.

Lexical chunks play a crucial role in enhancing the clarity, conciseness, and consistency of official correspondence. They help to effectively convey meaning by expressing complex ideas. In official correspondence, where brevity is often important, the use of familiar phrases and expressions can streamline communication and ensure the clarity of the message. By using lexical chunks and common expressions, the writer can avoid unnecessary words and get straight to the point, allowing the recipient to better understand the intended meaning. The consistent use of common word groups contributes to the professionalism and overall coherence of official correspondence. It helps maintain a uniform tone and style across all documents.

For this study secondary data has been collected. From the website of KSE the monthly stock prices for the sample firms are obtained from Jan 2010 to Dec 2014. And from the website of SBP the data for the macroeconomic variables are collected for the period of five years. The time series monthly data is collected on stock prices for sample firms and relative macroeconomic variables for the period of 5 years. The data collection period is ranging from January 2010 to Dec 2014. Monthly prices of KSE -100 Index is taken from yahoo finance.

I. LITERATURE REVIEW

The concept of “prefabricated chunks” was first introduced by Becker (1975) and Bolinger (1975). Later, Pawley and Syder (1983) discussed “lexicalized sentence stems,” asserting that mastering such pre-formed sentence elements is essential for language proficiency. Moon (2002) elaborated on “multi-word items,” explaining that word groups of two or more words function as complete and inseparable units in both meaning and grammar. He further noted that multi-word items are closely related to collocations, often resembling intensified cases of collocational patterns.

Jespersen (1924) defined formulaic expressions as entire sentences or groups of words treated as fixed units, which cannot be analyzed or broken down like free combinations. These expressions can be regular or irregular and are retrieved directly from memory. He emphasized that “a language would be difficult to handle if its speakers had burden imposed on them of remembering every little item separately” (1924, p. 76).

Becker (1975) proposed a typology for lexical chunks, categorizing them into five sub-types: poly-words, phrasal constraints, meta-messages, sentence builders, and idioms. However, one notable shortcoming of this framework was its failure to clearly distinguish between form and function. Addressing this gap, Nattinger (1992, p. 25) classified lexical phrases into four categories based on form: poly-words, institutional expressions, phrasal constraints, and sentence builders. Lexical chunks found in official correspondence predominantly belong to the categories of poly-words and phrasal constraints. Common examples include fixed expressions such as in case of, with regard to and as far as, which reflect high-frequency usage in formal contexts.

Several studies have been conducted on the use of lexical chunks in foreign language teaching, translation and writing. Some of these important works have been reviewed here. Many scholars support the role of lexical chunks in improving foreign language learning. Lewis(1993), Wray(2002) and Thornbury(2017) agree that chunks are stored and used as complete units. These help learners in producing fluent and accurate outputs.

Shushma and Ramesh(2021) studied the essay writings of Chinese college students and introduced the idea of employing the use of lexical chunks in college-level English writing instruction. Their findings showed that students who practiced relevant chunks regularly showed noticeable improvements in writing. Similarly, Elmahadi and Bajri(2023) highlighted the significance of using formulaic expressions in first and second language learning. Their study concluded that students who used these expressions in language learning had a significant advantage over the others who did not use them. The research concluded that these expressions are important for achieving fluency and communicative competence.

Albaqami(2022) studied the role of lexical chunks in promoting English writing competence among EFL learners in Saudi Arabia. He concluded that increased exposure to lexical chunks significantly improves writing competence of Arabic speaking English learners. The study found a positive correlation between number of lexical chunks mastered and higher writing scores. The findings recommend incorporating exercises based on lexical chunks in EFL classes to enhance writing skills and communicative effectiveness. Further studies supported these findings. Al-Khazaali(2019) tested chunk based methods in classroom and found that learners improved their writings when exposed to ready made phrases. Zhao(2009) also confirmed a strong connection between usage of lexical chunks and better writing skills.

Overall, previous studies suggest that lexical chunks improve vocabulary use, reduce errors, increase fluency making them a useful tool for teaching EFL learners. However, little attention has been given to identifying the most commonly used lexical chunks in official written correspondence. Therefore, there is ample space for this study which seeks to focus on common lexical chunks in official correspondence in India.

This study aims to thoroughly explore the lexical landscape of official correspondence through a corpus study. Focusing on common lexical fragments in official communication, the study seeks to uncover the linguistic patterns inherent in formal communication in bureaucratic, governmental, and administrative contexts. The primary objectives are lexical profiling, identification of common lexical chunks, understanding linguistic features and assessment of contextual significance. The research aims to elucidate the contextual significance of common lexical chunks in official communication.

II. METHODOLOGY

This study is based on the analysis of a sizable corpus. Corpus is a large-scale collection of written and spoken natural language materials that serve as a warehouse for storing language materials. Because of the advancement and application of computers, the corpora we now use are essentially electronic corpora, which store a large number of language materials in the form of text via computers. The emergence of corpora provides more advanced technical means for linguistic research.

The corpus created for the study is a collection of 250 documents comprising official letters, memorandums, gazettes, notices, and guidelines published by various central authorities on their respective portals. The data collected for the corpus mainly originates from five central authorities, namely the Ministry of Home Affairs (MoHA), Ministry of Education (MoE), National Council of Educational Research and Training (NCERT), Delhi University (DU), and Jawaharlal Nehru University (JNU).

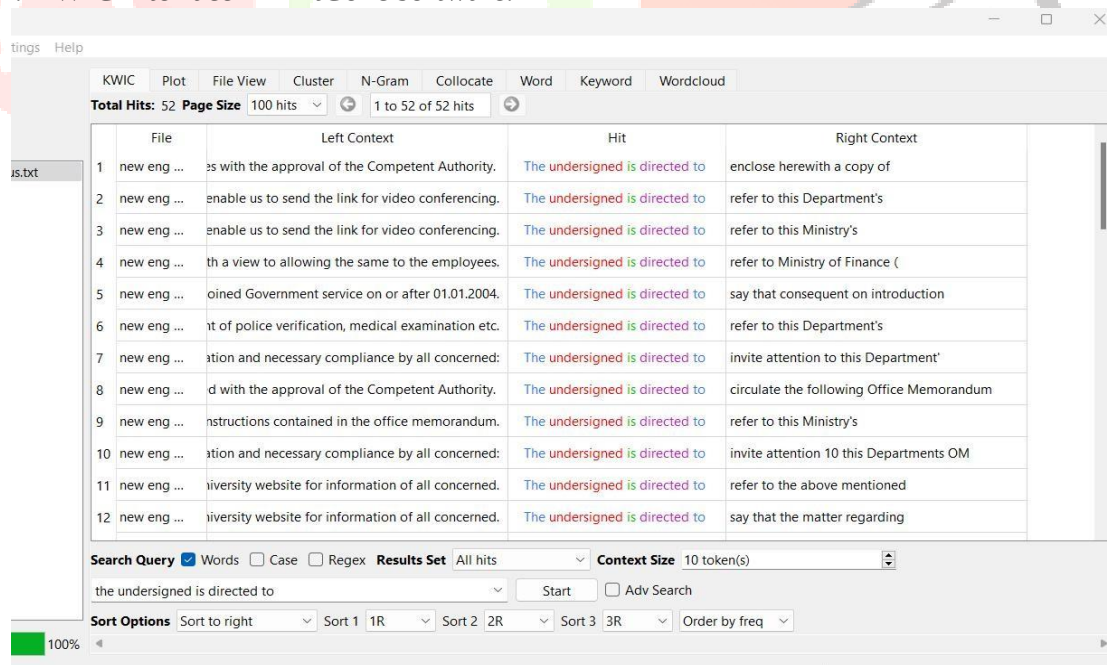
From the government portals of the Ministry of Home Affairs (<https://www.mha.gov.in/>) and Ministry of Education (<https://www.education.gov.in/>), I selectively downloaded gazettes, guidelines, notices, and circulars. I excluded some forms of official communication such as advertisements, tenders, and press releases as they did not contain the most common lexical chunks found in office-level literature, and cleaning of data was not possible in such texts. From the websites of central universities, namely NCERT(<https://ncert.nic.in/>), DU(<https://www.du.ac.in/>), and JNU(<https://www.jnu.ac.in/>), I downloaded notices, transfer letters, promotion letters, letters of appointments, circulars, orders, and office memorandums.

After downloading these files, I converted them to text format. Some of the PDFs I downloaded were in image format, so I used an online Optical Character Reader (OCR) tool to convert them to text format. After converting all the downloaded files into text (.txt format), I manually compiled them into a single corpus. I meticulously checked the converted files to correct spelling errors that occurred during the conversion phase. The initial corpus I assembled lacked cleanliness. I cleaned some of it manually and some using a Python script, which deleted the heading, dates, and signature part of official texts. I did this to allow the software to process the raw data and discard factors irrelevant to the lexical composition of the corpus. I was left only with the main body of official texts. After cleaning, the total word count of the corpus was 40,626.

In the later stage, the analysis of the corpus was done using a corpus analyzing tool called AntConc, developed by Laurence Anthony. AntConc is widely used in corpus linguistics and text analysis for exploring and analyzing large amounts of text data. It provides various functionalities for concordance, clustering, n-grams, keyword analysis, and text visualization, making it a valuable tool for researchers and linguists. I opened the corpus in AntConc and used its n-gram functionality to identify frequently occurring three, four, and five-word lexical chunks. The Concgram procedure takes a whole corpus of text and finds all sorts of combinations like these, whether consecutive or not, and is able to group the resulting patterns that share the same items but vary in their ordering. Cheng et al. (2006: 414) define a concgram as “all of the permutations of constituency variation and positional variation generated by the association of two or more words.” This means that the associated words comprising a particular concgram may be the source of a number of ‘collocational patterns’ (Sinclair et al. 2004: xxvii). I found often occurring lexical chunks for each category. I then created tables for each of these categories and input these lexical bundles one by one. Some of the three and four-word chunks were subsets of five-word lexical bundles. In this case, I dropped the three and four-word lexical chunks and only used the five-word lexical bundles.

After preparing the tables, I needed to determine the context in which these lexical chunks occurred. For this, I used the Key Word in Context (KWIC) feature of AntConc. For each lexical chunk input, I obtained the context in which it occurred throughout the corpus. Recent corpus studies have provided evidence that certain lexical items have a tendency to occur at particular points in a text, i.e., the beginning or end of texts, paragraphs or sentences (Scott and Tribble 2006; Hoey and O'Donnell 2007b, 2008a, 2008b; Mahlberg 2009; Csomay and Cortes 2010; Römer 2010; O'Donnell and Römer).

Figure 1 : KWIC interface in AntConc software.



Using this feature saved a lot of time, reducing the need for manual effort. I examined the collocations and compiled sentences in which each lexical bundle occurred. Later, I analyzed each lexical bundle and its context in terms of formality, linguistic mannerism, tone, and politeness.

The research employs both quantitative and qualitative methodologies; quantitative analysis is conducted using Laurence Anthony's AntConc, while qualitative analysis of chunks is performed manually. The data under scrutiny is extracted from a comprehensive corpus consisting of official memos, notifications, gazettes, guidelines, circulars, and orders sourced from five central government repositories. The data is categorized based on the number of words each lexical chunk contains, namely 3, 4, and 5-word lexical chunks. Each

category is represented in a separate table containing the forty most frequently encountered lexical bundles. Subsequently, representatives of these bundles are subjected to manual analysis to gain deeper insights.

III. ANALYSIS

The corpus was analysed with AntConc to identify common lexical chunks used in official correspondence. A detailed analysis has been presented below.

Analysis of 3 Word Lexical Chunks

From the in-depth analysis of the corpus using the AntConc software, approximately 500 recurring three-word lexical chunks were identified. This analysis employed the n-gram functionality of the software, with the parameter 'n' set to 3 (tri-gram). Out of these roughly 500 lexical chunks, only 31 are provided below as they emerged as the most relevant within their respective contexts. These chunks exhibit a level of formality, linguistic mannerism, official tone, and politeness that align with the characteristics of office correspondence.

Table 1: Thirty-One Frequently Occurring Three Word Lexical Chunks

With immediate effect	With/in regard to	Of the said
The provisions of	Deemed to be	The under mentioned
Powers conferred by	Be entitled to	In pursuance of
Requested to furnish	Is hereby posted	For regulation of
In exercise of	In addition to	As per annexure/notification/rule
In terms of	To refer to	Is/are hereby notified
In the prescribed	In/with reference to	For inclusion in
Is hereby transferred	Is directed to	Pleased to decide
In continuation of	With effect from	Taking into consideration
Until/till further orders	Is/are hereby conveyed/informed	In view of
The prescribed tenure		

These phrases are like little packages of language that appear frequently in day to day office communication. These phrases might show urgency, authority, or the way things are done in that place. Five of these are discussed below :

1. with immediate effect : This phrase is commonly used in official correspondence to denote urgency and prompt action. It is typically formal and carries an official tone. Its occurrence suggests a sense of efficiency and decisiveness in administrative actions. This lexical chunk is straightforward and direct in its meaning, leaving little room for ambiguity or indirect interpretation. The phrase is grammatically correct and commonly used in various formal contexts, such as legal documents, corporate announcements, or official notices. It functions as an adverbial phrase modifying the action of the verb being taken. As can be seen in the examples below:

A) Shri Navin Soi, CSS officer of SGSL is hereby appointed as Joint Secretary with immediate effect in his current place of posting.

B) The services of Shri Bani Brata Roy, Director are placed at the disposal of Central Vigilance Commission on transfer from Department of Legal Affairs with immediate effect and until further orders.

In both cases, this expression is used to convey a sense of urgency and importance. It adds a formal and official tone to the conversation, indicating that action must be taken immediately based on the given instructions or decisions. Using this expression aligns with the expected language in formal correspondence, where clarity, accuracy, and timeliness are valued. Moreover, while it conveys urgency, this expression also maintains a degree of politeness by providing clear and prompt instructions.

2. taking into consideration : This is a formal phrase used to express that a decision or action is influenced by certain reasons or instructions. This framework is organized in a formal manner, suitable for formal communication or official announcements. It conveys a sentiment of carefully considering and respecting the established instructions or recommendations. As can be seen in the examples below:

A) After taking into consideration the guidelines issued by the Commission, expenditure has been updated and released.

B) Taking into consideration the advice of UGC, the Central Government hereby withdraws its Notification.

The phrase 'taking into consideration' is utilized in both contexts to denote a formal, deliberate approach to decision-making. It signals that various factors, guidelines, or advice have been carefully weighed before arriving at a conclusion or taking action. This usage contributes to the official tone, formality, and politeness of the communication.

3. in reference to: It is a formal phrase commonly used in official or professional correspondence. It serves as a formal way to indicate the relation and basis of the information being discussed. The phrase lends an official tone to the sentences, indicating a formal acknowledgment of a specific document, letter, or notification. It establishes a clear connection between the current communication and the referenced source. The phrase contributes to clarity and transparency in communication by specifying the reference point.

A) In reference to Ministry of Education letter No. 13-59/2022, Prof. Amarjiva Lochan has been appointed as Compliance Officer at the University of Delhi.

B) Candidates who have applied earlier in reference to Notification Ref. No. 1V/046 need not apply again.

This chunk is used in both contexts to establish a formal connection between the current communication and a specific document, letter, or notification. It contributes to the formality, clarity, and transparency of the message by indicating the source or basis of the information being conveyed.

4. in addition to: This is a formal expression used to present additional information or requirements. It's often used in formal or professional settings to specify additional elements or tasks. It clearly indicates that, in addition to the primary requirement, additional elements must be considered or met. This expression helps organize instructions or information in a clear and organized manner. It ensures that the recipient understands all necessary elements or responsibilities associated with a specific task.

A) In addition to the documents required to be furnished by the candidates, two passport size photos must also be carried by the applicant.

B) Dr. Pawan Sudhir, Professor in Department of Education will continue to look after the work of Head in addition to his existing duties till further orders.

This lexical chunk is utilized in both contexts to specify additional requirements or responsibilities beyond the main task or role. It adds formality and clarity to the communication ensuring that recipients understand all relevant components or duties associated with a specific directive.

5. is hereby conveyed: This expression is a formal and official way of conveying a decision or information. It is very formal. It is often used in official documents, announcements, or correspondence where formality is required. Its use suggests a precise and formal communication style. It demonstrates the authority and clarity of the message being conveyed. The tone of this expression is formal and authoritative. It suggests that the information being communicated is important and should be taken seriously. It uses the passive voice to emphasize the action of conveying the information. As can be seen in the examples below:

A) Approval of the President is hereby conveyed for inclusion of Shri Rakesh Kumar in the Senior Selection Grade Select List of Central Secretariat Services.

B) The decision to extend the deadline for submission of proposals until the end of the month is hereby conveyed to all participating research teams.

The phrase 'is hereby conveyed' is a formal, authoritative, and grammatically correct way to communicate information or decisions in official correspondence.

Analysis of 4 Word Lexical Chunks

After analysis of the corpus approximately 240 recurring four-word lexical chunks were identified. This examination utilized the n-gram feature of the software, with the parameter 'n' set to 4 (quad-gram). Among these 240 lexical chunks, only 33 are provided in the table below, as they emerged as the most relevant within their respective contexts, having level of formality, linguistic style, official tone, and politeness.

Table 2: Thirty-Three Frequently Occurring Four Word Lexical Chunks

It is proposed to	Conditions of service of	Consequent upon inclusion of
In the event of	Applicable in case of	Applicants are requested to
On the advice of	In the prescribed proforma	During the term of
For a period of	On the subject mentioned	For information of
In the light of	I accept the said	In continuation to the
Date of assumption of	Applicable in case of	As per their requirement
Advised to refer to	Consequent upon his appointment	In order to facilitate
Terms and conditions of	In the context of	Consequent upon transfer of
It has been decided	Regarding the appointment of	In partial modification of
Guidelines for regulation of	On the request of	Is notified for compliance
In order to observe	As per the notification	Received in response of

Five of the above-mentioned lexical chunks are discussed below:

1. date of assumption of: The phrase exhibits a formal and procedural linguistic mannerism typical of official documents and correspondence. It conveys a sense of official procedure and protocol. The tone conveyed by the phrase is official and authoritative. It implies adherence to established procedures and protocols in official matters.

Usage in different contexts:

A) The appointment of officers will be effective from the date of assumption of charge of the post.

B) The undermentioned officers of the Central Secretariat Service are hereby promoted as Directors on a regular basis, effective from the date of assumption of charge.

The lexical chunk is formal, exhibits a procedural linguistic mannerism, maintains an official tone, adheres to politeness norms within administrative contexts.

2. consequent upon transfer of: The phrase is formal and commonly used in official or bureaucratic contexts to denote actions or decisions that follow a transfer. The phrase indicates a sequence of events in relation to transfers. The tone conveyed by the phrase is official and authoritative, signaling that decisions or actions are taken in response to a transfer. It effectively communicates the cause-effect relationship between a transfer and succeeding actions.

A) Consequent upon transfer of Smt. Usha, Administrative Officer to Shillong, she is hereby posted to CIET as Junior Secretary.

B) Consequent upon transfer of Dr. Rakesh to the regional office, Prof. Saloni Dubey has been appointed as the new HOD for the department.

Its occurrences in the provided sentences highlights its role in signaling actions or decisions following transfers in organizational settings.

3. of the competent authority: It carries a sense of authority and establishment, suitable for formal documents and official correspondence. its usage in official correspondence suggests adherence to formalities and procedures, which can be seen as a form of respect and politeness within bureaucratic contexts. It acts as a prepositional phrase modifying the noun it precedes. It specifies the source of authorization or decision-making.

A) This is issued with the approval of the competent authority.

B) It has been decided with the approval of the competent authority that the following officers are hereby posted as Directors in the Ministry.

In both provided contexts, it serves to denote the source of authority behind the decisions being communicated.

4. in order to facilitate: It is a formal phrase commonly used in official and professional contexts. It denotes a deliberate action taken to make something easier or smoother. Its use implies a consideration for making processes easier or more efficient. It functions as an adverbial phrase modifying the verb or action that follows. It indicates the purpose or intention behind the action described in the sentence.

A) In order to facilitate quicker decisions by this Department, all applications should be supported with the required information, as indicated in the policy circular.

B) The new regulations have been put in place in order to facilitate compliance with international standards.

In both provided contexts, the phrase clarifies the intention behind specific actions or regulations, emphasizing the aim of simplifying procedures or achieving compliance with standards.

5. in the light of: It is a formal and professional phrase used to introduce considerations, decisions, or actions based on certain circumstances or information. It implies a careful review or assessment of the given situation. Its usage in formal contexts indicates a respectful acknowledgment of the circumstances or developments being considered. It functions as a prepositional phrase introducing the context or basis for a decision or action. It sets the stage for the subsequent explanation or announcement.

A) In the light of the above notification, it has been decided that while fixing remuneration in the case of Central Government retirees, an amount equivalent to 30% of the last basic pay shall be deducted from the new basic pay.

B) In the light of recent developments, the department has decided to revise its procurement policies.

'In the light of' is a formal phrase used to introduce the basis or context for a decision, action, or revision. In both provided contexts, the phrase precedes announcements or decisions that are informed by specific notifications or recent developments, highlighting the importance of context in official decision-making processes.

Analysis of 5 Word Lexical Chunks

After analyzing the corpus, approximately 120 recurring five-word lexical chunks were identified. This examination utilized the n-gram feature of the software, with the parameter 'n' set to 5 (pent-gram). Among these 120 lexical chunks, only 36 are provided in the table below, as they emerged as the most relevant within their respective contexts. These chunks exhibit a level of formality, linguistic style, official tone, and politeness suitable for office correspondence.

Table 3: Thirty-Six Frequently Occurring Five Word Lexical Chunks

Based on the recommendations of	Be effective from the date
I would like to inform	Regulated in accordance with
Issued with the approval of	consequent upon completion of
On the subject mentioned below	With reference to your inquiry
It may please be certified	The undersigned is directed to
To furnish the requisite information	in accordance with the regulations,
To the best of my knowledge	Of the powers conferred in
Is enclosed for necessary action	It is imperative to adhere to
As per details given below	with the approval of the competent authority
It is for information to	Required to take the appropriate measures
Regarding filling up posts of	In case of any further clarification
On the subject cited above	To contact the concerned authorities
Under posting in this department	As per the provisions of
A true and fair view	Your prompt attention to this matter,
From the date of assumption	We regret to inform you that
Shall come into force on	For information of all concerned
To Bring to the notice of	With regard to the aforementioned matter
on the subject mentioned above	It is to inform you that

Five of the most occurring lexical chunks are discussed below:

1. based on the recommendations of: This lexical chunk is structured in a formal and official manner, indicating reliance on recommendations for decision-making. The tone is authoritative and procedural, reflecting adherence to established recommendations. It maintains a neutral and respectful tone, acknowledging the authority of the recommendations.

A) Based on the recommendations of the said committee, requisite order for promotion of CSSS officials is issued.

B) The Government of India, based on the recommendations of the Reserve Bank of India, hereby declares that the provisions of sub-clauses of Section 10 of the said Act shall not apply to the Indian Overseas Bank.

In both the sentences provided, the phrase signifies the reliance on expert opinions or regulatory guidelines in making official decisions or declarations within governmental or bureaucratic frameworks. This chunk is used in a formal setting, likely within a bureaucratic or governmental framework where decisions are made based on expert recommendations or committee findings.

2. the undersigned is directed to: It adopts a formal and structured manner of communication, indicating the delegation of a specific task or instruction to the person identified as 'the undersigned'. The tone is authoritative and directive, suggesting that the action being conveyed is mandated or instructed by a higher authority. The lexical chunk consists of a subject 'the undersigned', a form of the verb 'is', and a passive voice construction with the verb 'directed to' followed by the action or instruction.

A) The undersigned is directed to circulate the following Office Memorandum received from the Ministry of Personnel, Government of India on the subject as given below.

B) The undersigned is directed to enclose herewith the vacancy circular for filling up the post of Deputy Director General.

This chunk refers to a structured, hierarchical style of communication within a bureaucratic or organizational environment, where tasks, information dissemination, or document sharing are carried out according to established protocols. Its use refers to formal communication channels within a bureaucratic structure, used to issue directives to ensure the dissemination of information and compliance with official protocols. In both sentences, the expression refers to the transmission of official directives or instructions through formal communication channels to ensure compliance with organizational procedures.

3. of the powers conferred in: The phrase is formal and legalistic, often used in official documents, regulations, and legal contexts. The tone is authoritative, indicating the delegation or authorization of specific powers or responsibilities. The phrase consists of a prepositional phrase 'of the powers' followed by a passive construction 'conferred in', indicating the source or origin of the conferred powers.

A) In exercise of the powers conferred in terms of O.M. No. 21, the services of Shri Bani Brata Roy, Director, are placed at the disposal of the Central Vigilance Commission on transfer from the Department of Legal Affairs.

B) In light of the powers conferred in the relevant regulations, the committee has decided to proceed with the proposed amendments to the policy.

'Of the powers conferred in' is a formal and legalistic lexical chunk used in official documents and communications to indicate the authority or jurisdiction under which specific powers or responsibilities are delegated or exercised.

4. as per the provisions of: The phrase is formal and commonly used in legal and regulatory contexts to indicate compliance with established rules or guidelines. It adopts a formal and structured manner of expression, typical of legal and bureaucratic language. The phrase consists of a prepositional phrase "as per the provisions of" followed by the specific regulation, statute, or guideline being referenced.

A) As per the provisions of Statute 38, the Vice-Chancellor has appointed Prof. A. Venkat Raman, Professor in the Department of Business Management, as the Head of the Department of Business Management for a term of 3 years.

B) A person shall impart coaching or establish, run, manage, or maintain a coaching center only with prior registration of such coaching center as per the provisions of these guidelines.

It conveys an authoritative tone and emphasizes adherence to specific provisions or regulations governing a particular activity or decision. In both contexts provided, the phrase indicates alignment with

institutional or regulatory requirements, ensuring that actions or decisions are conducted in accordance with prescribed guidelines.

5. to furnish the requisite information: This phrase is formal and often used in official communications and instructions. It uses a formal and directive tone, indicating the need to act in order to provide specific information. Its tone is official and directive, indicating the obligation to provide the requested information. This phrase is composed of an infinitive verb phrase “to furnish” followed by noun phrase “the requisite information” which indicates both the action required and the specific information.

A) All Cadre Units are requested to furnish the requisite information properly in the prescribed Annexures to this Department.

B) The Departments concerned are, therefore, requested to furnish the requisite information in the prescribed proforma to this Department without any further delay.

In the two sentences given, this phrase refers to the formal submission of a request for information within an organizational or bureaucratic structure to ensure compliance with procedural requirements and improve the efficiency of administrative processes. This phrase reflects the formal and directive nature of requesting information from the relevant institution, thus ensuring the smooth flow of administrative processes.

IV. DISCUSSION

The study conducted talks about the significance of common lexical chunks in official correspondence, exploring their frequency, usage patterns, and implications for effective communication. Through a corpus-based analysis, I uncover the important lexical bundles in official documents, shedding light on their role in enhancing clarity, conciseness, and coherence in written communication.

The findings of the study align with previous research in corpus linguistics, which emphasizes the importance of lexical cohesion and collocation in text processing and comprehension (Sinclair, 1991, 2004; Stubbs, 1996; Evert, 2005). Words exhibit strong preferences for collocating with specific lexical items, contributing to the cohesive and coherent structure of texts (Biber, 1988; Biber et al., 2007). My study of lexical chunks in official correspondence underlines the role of these linguistic units in facilitating efficient processing and understanding of written communication. Furthermore, this study contributes to the understanding of lexical distribution in texts and reveals trends in the presence of certain lexical items in specific locations within documents. The identification of lexical bundles and their location within texts provides insight into textual organization and discourse structure, thus deepening our understanding of the linguistic patterns in official correspondence.

This study employed a systematic corpus analysis approach, using AntConc to identify frequent lexical segments in a large corpus of official documents. By examining lexical bundles of three, four, and five words, this paper provides a comprehensive overview of common linguistic patterns in official correspondence, deepening our understanding of language use in formal contexts.

In addition, my research highlights the importance of lexical chunks in various fields, including language teaching, administrative correspondence training, government and public administration. The identification of common lexical clauses can guide the development of language curricula, professional training programs, and administrative practices, thereby improving the effectiveness and efficiency of communication in various contexts.

This corpus-based study provides valuable insights into the role of lexical chunks in official correspondence, and highlights their importance in promoting clear, concise, and coherent communication. By analyzing the frequency, distribution, and usage patterns of lexical clauses, I have gained a deeper understanding of language use in formal contexts and laid the foundation for future research in corpus linguistics and applied linguistics.

V. CONCLUSION

Effective academic writing requires a variety of language skills, including efficient use of grammar, pragmatics, and vocabulary. However, this article argues that effective language use in official correspondence can depend largely on the author's vocabulary. Based on this approach, I analyzed a corpus of 40,626 words from 250 official written documents and ultimately identified 100 lexical segments. The results highlight the importance of lexical cohesion and coherence in promoting clarity, conciseness, and professionalism in official correspondence. By elucidating the usage patterns and contextual significance of lexical segments, this study provides valuable insights into the linguistic dynamics of formal communication. Identifying recurring lexical segments provides practical tools to improve written communication skills,

promote transparency, and foster effective communication in formal situations. Furthermore, this study highlights the interdisciplinary relevance of corpus-based research in linguistic studies, applied linguistics, and administrative communication. By combining theoretical insights with practical applications, this study contributes to the advancement of knowledge in this field and informs instructional practices, communication strategies, and policy initiatives aimed at improving language proficiency and communication effectiveness in formal settings.

Future research could explore the dynamic evolution of lexical chunks in response to communication trends, technological advances, and sociocultural changes. Additionally, comparative studies across diverse language groups and cultural contexts can provide valuable insights into the universality and variability of lexical patterns in formal communication.

All in all, this study opens the door for more research into how language, communication, and administrative discourse connect and influence each other. By highlighting the importance of lexical chunks in formal correspondence, this study underlines the importance of language proficiency in understanding complex communication environments and fostering meaningful interactions in formal settings.

REFERENCES

- [1]Albaqami, R. (2022). The role of lexical chunks in promoting English writing competence among foreign language learners in Saudi Arabia. *Arab World English Journal*, 13(2), 1–15.
- [2]Al-Khazaali, H. (2019). The impact of lexical chunks instruction in developing Iraqi students' writing fluency. *Advances in Language and Literary Studies*, 10(5), 70–78.
<https://doi.org/10.7575/aiac.all.v.10n.5p.70> Amsterdam: John Benjamins
- [3]Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Bloomsbury Publishing
- [4]Becker, J. (1975). The phrasal lexicon. In R. Schank & B. L. Nash-Webber (Eds.), *Theoretical issues in natural language processing*. Cambridge, MA: Bolt Beranek & Newman.
<https://doi.org/10.3115/980190.980212>
- [5]Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honor of Stig Johansson* (pp. 181–189). Amsterdam: Rodopi
- [6]Biber, D., Conrad, S., & Cortes, V. (2004). If You Look at ...: Lexical Bundles in University Teaching and Textbooks. *Applied Linguistics*, 25, 371–405.
<https://doi.org/10.1093/applin/25.3.371>
- [7]Church, Kenneth and Patrick Hanks. 1990. Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics* 16 (1), 22–29.
- [8]Elmahdi, O. E. H., & Bajri, I. A. (2023). The crucial role of formulaic expressions in fluent communication and language acquisition. *Journal of Namibian Studies*, 34, 33–50.
- [9]Evert, Stefan. 2005. *The statistics of word co-occurrences: Word pairs and collocations*. University of Stuttgart, Stuttgart.
- [10]Hoey, Michael & Matthew Brook O'Donnell. 2009. The Chunking of Newspaper Text. In M. Shiro, P. Bentivoglio & F. Erlich (eds.), *Haciendo discurso. Homenaje a Adriana Bolívar*, 433–452. Comisi n de Estudios de Postgrado de la Facultad de Humanidades y Educaci n de la Universidad Central de Venezuela.
- [11]Hoey, Michael & Matthew Brook O'Donnell. 2008. Lexicography, grammar, and textual position. *International Journal of Lexicography* 21(3). 293–309
- [12]Hoey, Michael. 1991. *Patterns of Lexis in Text*. Oxford: Oxford UP.
- [13]Jespersen, O. (1924). *The philosophy of grammar*. London: Allen and Unwin.
- [14]Lewis, M. (1993). *The lexical approach*. England: Language Teaching Publications.
- [15]Mahlberg, Michaela. 2005. *English general nouns: A corpus theoretical approach*.
- [16]Moon, R. (2002). Vocabulary connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 31–33). Shanghai: Foreign Language Education Press.
- [17]Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- [18]O'Donnell, Matthew Brook & Ute R mer. Positional variation of n-grams and phrase-frames in a new corpus of student writing.
- [19]Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Native-like selection and native-like fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 101–121). New York: Longman.

- [20]Sinclair, John. 1991. Corpus, concordance, collocation. Oxford: Oxford University Press
- [21]Stubbs, Michael. 1996. Text and corpus analysis. Oxford: Blackwell.
- [22]Sushma, G., & Ramesh, N. (2021). A study based on lexical chunks present among college students. International Journal of Creative Research Thoughts (IJCRT), 9(7).
<https://doi.org/10.13140/RG.2.2.35342.41285>
- [23]Thornbury, S. (2017). Scott Thornbury's 30 language teaching methods. Cambridge: Cambridge University Press.
- [24]Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. Applied Linguistics, 21(4), 463–489. <https://doi.org/10.1093/applin/21.4.463>
- [25]Zhao, Y. (2009). An empirical study of L2 learners' use of lexical chunks and language production (Unpublished doctoral dissertation). Kristianstad University College, Sweden.

CORPUS AVAILABILITY

The corpus used in this study is publicly available at:

https://drive.google.com/file/d/1pAyhlj_mwE5R42BT2jODtp4XmryHR7aw/view?usp=drive_link

