

## A Corpus Analysis of Non-Contiguous Phraseological Patterns in the Discussion Section of Microeconomics Research Articles

ANG LENG HONG  
School of Humanities  
Universiti Sains Malaysia, Malaysia  
[lenghong@usm.my](mailto:lenghong@usm.my)

### ABSTRACT

*This study explores the non-contiguous phraseological profile of Microeconomics research articles, specifically focusing on key lexical frames within the Discussion section. Key lexical frames are non-contiguous recurrent expressions with variable slots unique to specific contexts such as genres or disciplines. The analysis is based on a corpus of Discussion sections from research articles published in four leading Microeconomics journals between 2017 and 2022. Using the KfNgram program, four-word lexical frames were identified and refined according to exclusion criteria. The resulting list of lexical frames was compared with a general corpus, the Corpus of Contemporary American English, yielding 43 four-word key lexical frames unique to Microeconomics. These frames were analysed following the structural classification by Gray and Biber (2013) and the functional taxonomy by Simpson-Vlach and Ellis (2010). Structurally, most key lexical frames were found to be function word frames, while functionally, they primarily served referential functions. The findings reveal significant phraseological variation, encompassing both lexical and grammatical constructions unique to the field of Microeconomics. These insights contribute to a deeper understanding of discipline-specific academic writing and have important implications for English for Academic Purposes pedagogy. In particular, this research aims at improving the learners' part-genre awareness and competence. By identifying and analysing key lexical frames, the study provides valuable resources for enhancing the teaching of academic writing in Microeconomics, particularly for non-native English speakers.*

*Keywords: non-contiguous phraseology; research article; key lexical frame; discipline-specific writing; Microeconomics*

### INTRODUCTION

Research in phraseology indicates that language is predominantly made up of contiguous or non-contiguous multi-word expressions, which are key carriers of meaning (Römer, 2010; Wray, 2008). These expressions are known by various terms in the literature, including *phraseological expressions, formulaic sequences, n-grams, lexical bundles, collocations* and *lexical frames* (Biber et al., 1999; He et al., 2021; Lu et al., 2018; Wray, 2008). Despite the different labels and definitions, these multi-word expressions are characterised by a strong internal relationship that contributes to their meaning (Wray, 2008). In this study, the term *phraseological expressions* is used as an umbrella term to encompass both contiguous and non-contiguous multi-word expressions. Contiguous expressions include n-grams and lexical bundles, such as *the impact of globalisation*, and collocations like *verb + noun* (e.g., *conduct research*). Non-contiguous expressions, on the other hand, are exemplified by lexical frames, such as *taking a \* approach* (\* = *strategic/proactive/holistic*).

Phraseology, the study of contiguous and non-contiguous multi-word expressions, is essential for understanding socially-situated language practices (Biber et al., 1999, 2004). The use of these expressions indicates language proficiency and signals writing expertise and membership

within a discourse community (Wray, 2008). Academic discourse, in particular, relies heavily on phraseological expressions, with research traditionally focusing on contiguous forms like lexical bundles and collocations (Hyland, 2008a, 2008b). These expressions are practical for various contexts, offering utility in both written and spoken communication (Biber et al., 1999).

However, the focus on contiguous expressions has led to an oversight of non-contiguous phraseological expressions, or lexical frames, which contain variable slots that can be filled with different words (Römer, 2010; Wang, 2019). For example, the lexical frame *about the \* of* can include slot fillers like *importance*, *impact* or *role*. Lexical frames are particularly useful in pedagogical contexts due to their variability and adaptability (Lu et al., 2018). Despite their significance in English for Academic Purposes (EAP), lexical frames are under-researched, especially in different academic registers and specific sections of research articles (Lu et al., 2021). The Discussion section, in particular, has been identified as a crucial area needing more focus. This section plays a pivotal role in academic writing, where authors are expected to interpret findings and discuss their broader implications (Golparvar & Barabadi, 2020; Swales & Feak, 2012). The Discussion section is crucial as it reflects the author's ability to engage with discipline-specific concepts and contribute to ongoing scholarly debates. Focusing on the Discussion section allows for an in-depth examination of the language used in this critical part-genre, offering insights into how Microeconomics scholars construct and convey complex arguments. This study, therefore, addresses the need to better understand the phraseological patterns that are unique to this section, which has been under-researched in the existing literature.

The theory of lexical priming (Hoey, 2005) suggests that certain phrases are more likely to occur in specific sections of academic texts, making the identification of lexical frames in particular part-genres valuable. This study aims to fill the gap by identifying key lexical frames in the Discussion section of Microeconomics research articles, a field that has received relatively little attention in EAP research (Starfield, 2016). As Microeconomics is a foundational subject in economics education, understanding its phraseology can aid in developing teaching materials and strategies that improve comprehension for students, particularly those who are non-native speakers of English. Given the importance of English in global economic contexts, mastering these expressions can enhance the transferability of information and expertise. By developing a list of key lexical frames and relevant slot fillers, this study seeks to provide valuable reference materials for EAP students and instructors, which will contribute to more effective academic writing instruction in Microeconomics (Ang & Tan, 2019).

## LITERATURE REVIEW

### LANGUAGE AS PHRASEOLOGY

The study of phraseology, rooted in Firth's (1957) concept of *collocation*, emphasises the importance of understanding words through their associations with others. This notion has led to the contemporary view of language as phraseology (Hunston, 2002), where phraseological expressions are seen as key to facilitating communication and cognitive processing by making language predictable. While idioms like *Under the Weather* are idiomatic and invariant, they are less common compared to semantically transparent and regular phraseological expressions. Sinclair (1991) highlighted that language use predominantly involves common words in set patterns, forming a rich repertoire of multi-word patterns that convey meaning. Consequently,

scholars (Dahunsi & Ewata, 2022; Shirazizadeh & Amirfazlian, 2021; Wang, 2019) advocate for a pedagogic focus on recurrent lexical bundles, essential for effective language use and teaching.

Lexical bundles, defined as contiguous sequences of words that function as single multi-word units, play a crucial role in various communicative contexts and are commonly identified through corpus-driven methods (Biber et al., 1999, 2004). These bundles, often seen as extended collocations, are widespread across different text types and fulfil multiple communicative functions, such as discourse organisation and precise information transfer, particularly in technical fields like aviation management and medicine. The identification process for lexical bundles typically involves semi-automatic extraction using corpus analysis tools, supplemented by manual verification to ensure accuracy. Researchers usually set frequency cut-offs between 10 and 40 occurrences per million words and a dispersion rate of at least 10% to avoid capturing idiosyncratic usages (Biber et al., 1999; Hyland, 2008b). Studies on lexical bundles have explored their structural and functional characteristics across genres and disciplines. For instance, Shirazizadeh and Amirfazlian (2021) found significant intra-disciplinary variation in the use of lexical bundles in applied linguistics, challenging the assumption of uniformity in EAP approaches. Reppen and Olson (2020) examined nine academic disciplines and found that most lexical bundles were unique to one or two disciplines, with only a few shared across all, highlighting their role as markers of discipline-specific language.

While research on lexical bundles has advanced our understanding of phraseological practices across different disciplines and genres, it has primarily focused on contiguous expressions. Less attention has been paid to non-contiguous phraseological expressions, initially termed *collocational frameworks* by Renouf and Sinclair (1991), *phrase frames* by Fletcher (2007), and more recently referred to as *lexical frames* by Gray and Biber (2013). These expressions, characterised by their conventional and variable elements, offer insights into the productive use of language across contexts and the balance between fixedness and variability in academic language (Römer, 2009, 2010). This perspective aligns with Sinclair's idiom and open-choice principles, which view language patterns as semi-preconstructed phrases that serve as single choices (Sinclair, 1991). Understanding lexical frames can thus enhance the teaching of EAP by providing a more comprehensive view of phraseological behaviour in academic discourse.

Lexical frames are defined as “sets of n-grams which are identical except for one word” (Römer, 2010, p. 98). Early research on lexical frames primarily focused on pre-determined grammatical structures, known as collocational frameworks, introduced by Renouf and Sinclair (1991). These frameworks consist of a non-contiguous sequence of two words separated by an intervening word, with their meaning depending on what fills the intervening slot (Renouf & Sinclair, p. 128). Examples include structures like *a \* of* and *many \* of*, where the asterisk (\*) represents a slot that can be filled by semantically related words. Before the development of advanced tools like *KfNgram* (Fletcher, 2007), these collocational frameworks were the main method for studying phraseological variation in language. *KfNgram* has since enabled a corpus-driven approach to analysing non-contiguous phraseological sequences, making it easier to identify and study lexical frames and their variations across different contexts. This tool has significantly advanced the study of phraseological variation, allowing researchers to explore a wider range of lexical frames automatically.

Lexical frames are valuable units of analysis because they provide insights into how non-contiguous phraseological expressions are used across different contexts (Lu et al., 2021; Renouf & Sinclair, 1991). For example, the frame *the \* of the*, often filled with words like *context*, *effect*, and *influence*, is common in academic writing for referential purposes. Understanding these frames

helps students grasp the structure and communicative functions of academic texts, which enhances their ability to identify common versus rare frames (Cunningham, 2017; Golparvar & Barabadi, 2020). Lexical frames have recently garnered attention in academic genres, with studies exploring their use and implications (Biber, 2009; Casal & Kessler, 2020; Cunningham, 2017; Golparvar & Barabadi, 2020; Grabowski, 2015; Gray & Biber, 2013; Lu et al., 2018, 2021; Römer, 2010; Win & Masada, 2015). Römer (2009) examined the use of lexical frames by non-native and native English speakers in academic writing, finding that both groups develop academic competence similarly, emphasising that expertise, rather than nativeness, is crucial for mastering academic conventions. She suggested that EAP courses should focus on the specific use of lexical frames by expert writers. Win and Masada (2015) proposed using technical lexical frames as query phrases in literature searches, highlighting their utility in refining research.

Despite increasing interest, research on the types, structures, functions, and variability of lexical frames across genres and registers remains limited. Biber (2009) compared academic prose and conversation, finding that academic discourse predominantly uses function word frames (e.g., *the \* of the*), with variable slots filled by content words like *nature* or *context*. In contrast, conversational discourse often features frames and slot fillers composed of function words (e.g., *do you have \**). Gray and Biber (2013) examined four-word lexical frames in academic prose and conversation, finding that academic writing typically includes more function word frames, while conversation favours verb-based frames. They also observed that academic lexical frames are more varied, linking them closely to grammatical constructions. Römer (2010) emphasised the importance of establishing a phraseological profile for text types, such as book reviews, to understand the “extent of the phraseological tendency of [a] language”, which provides insights into meaning creation in discourse (pp. 95–97). Grabowski (2015) highlighted significant intra-disciplinary variation in the use, structure, and functions of four-word lexical frames across different pharmaceutical documents, including textbooks, clinical trial protocols, product characteristics summaries, and patient information leaflets. A genre-based approach has been used to analyse the rhetorical functions of lexical frames in various academic contexts. For example, Cunningham (2017) explored Mathematics research articles, identifying key lexical frames associated with specific rhetorical functions. Casal and Kessler (2020) examined Fulbright grant application documents, finding strong relationships between lexical frames and rhetorical functions. Similarly, Yoon and Casal (2020) discovered that most lexical frames in applied linguistics conference abstracts are multi-functional, indicating that academic writers adapt their phraseological choices to their rhetorical aims.

Recent research has also focused on a part-genre approach, examining specific sections of research articles. Lu et al. (2018) classified five- and six-word lexical frames from Social Science research article introductions into structural and functional groups, finding that most five-word frames are non-verb content word frames serving referential functions, while six-word frames often act as discourse-organising markers. Golparvar and Barabadi (2020) identified that key lexical frames in Higher Education research article discussions are primarily non-verb content word frames serving referential functions. Expanding on this, Lu et al. (2021) matched lexical frames to rhetorical functions in the introduction sections of Social Science research articles, categorising them into specialised, semi-specialised, and non-specialised types.

This body of research underscores the need to explore how meaning is created in part genres, each with specialised grammar and vocabulary (Römer, 2010). The current study aims to expand this knowledge by generating a pedagogically useful list of key lexical frames specific to

the Discussion section of Microeconomics research articles, using the concept of keyness to identify context-specific lexical frames, following Cunningham (2017).

In light of the preceding explanations, the study intends to answer the following research questions:

- 1) What are the frequent four-word key lexical frames present in the Discussion section of Microeconomics research articles?
- 2) How are these key lexical frames distributed across the structural categories?
- 3) How are these key lexical frames distributed across the functional categories?

## METHODOLOGY

### THE CORPUS

The corpus for this study was compiled from the Discussion section of 239 research articles published in five indexed journals in the field of Microeconomics. These journals were selected based on their high citation indices and relevance to the discipline. This approach aligns with best practices in corpus compilation, which ensures that the study incorporates high-impact and widely recognised sources in the discipline. Articles published between 2019 and 2022 were chosen to ensure the most current data was analysed. This timeframe was chosen to capture recent developments and trends in Microeconomics research, which provides a contemporary perspective on the non-contiguous phraseological patterns prevalent in academic discourse. This selection process ensures that the corpus is both representative and relevant, allowing for the generalisation of the findings of this study to the broader field of Microeconomics. Only empirical research articles with a distinct Discussion section were included in the corpus. This criterion was set to focus the analysis on sections where authors interpret their findings, discuss implications, and provide context for their research, which are critical for examining phraseological patterns (Lu et al., 2018). Focusing on the Discussion section also aligns with the goals of analysing interpretive and contextual language use in academic writing, which is central to understanding phraseological patterns in Microeconomics. To ensure the accuracy of the extracted sections, a manual verification was conducted by a research assistant. This step is crucial for maintaining the integrity and validity of the corpus data. The corpus size is 1,200,020 words.

### THE PROCEDURE

The analysis of lexical frames in this study involved multiple stages. Initially, all four-word lexical frames were identified based on frequency and a set cut-off point. The keyness concept was then employed to filter key lexical frames from the broader inventory. These frames were subsequently examined for their structural and functional properties.

Lexical frames were extracted from the corpus using the *KfNgram* program (Fletcher, 2007), a software tool that can generate lists of n-grams of varying lengths from a corpus. Lexical frames were defined as n-grams that are identical except for one variable slot. For example, the n-grams, *the price effects of*, *the income effects of*, and *the substitution effects of* collectively generate the lexical frame *the \* effects of*. The focus on four-word lexical frames is consistent with prior

research, which predominantly uses this length for meaningful comparison (Golparvar & Barabadi, 2020; Grabowski, 2015; Römer, 2010).

The list of automatically generated lexical frames underwent a rigorous manual verification process to ensure the quality and relevance of the data. Initially, lexical frames with only one slot filler were excluded from further investigation, as these frames are essentially another form of lexical bundles. Lexical bundles are well-studied, and the focus of this research is on identifying more complex phraseological patterns. This step is critical in differentiating between simple recurrent phrases and more dynamic lexical frames that can vary meaningfully within their slots (Garner, 2016). Next, any lexical frames containing proper names or mathematical symbols were discarded. Proper names can skew frequency data and introduce bias. Similarly, mathematical symbols can disrupt the analysis of general phraseological patterns. Finally, lexical frames with variations or slot fillers at the beginning or end of the sequences were excluded from the analysis to maintain a focus on genuinely variable lexical frames. This exclusion is crucial as frames with such variations often resemble simple lexical bundles rather than dynamic phraseological patterns, thereby providing limited insights into the flexible and context-dependent use of language. This careful filtering process helps to focus the analysis on truly variable lexical frames that can reveal deeper insights into phraseological patterns (Garner, 2016; Römer, 2010).

After manual verification, the remaining four-word lexical frames underwent further scrutiny to meet minimum frequency and range criteria. According to Biber et al. (2004), for phrases to qualify as recurrent phraseological sequences, they must appear at least 20 times per million words. For this study's corpus, this equates to a minimum of 24 occurrences, ensuring that the frames are genuinely recurrent and not due to random or infrequent usage. Additionally, the frames must appear in at least 10% of the articles, which translates to 23 articles in this corpus. This range requirement prevents writer-specific idiosyncrasies and ensures the frames represent general usage within Microeconomics. Range identification was conducted using *AntConc* (version 3.5.9) (Anthony, 2021), a corpus analysis tool that enables precise identification and examination of lexical patterns across large datasets, ensuring a robust analysis of the frequency and distribution of the lexical frames.

Using the concept of keyness (Cunningham, 2017), key lexical frames specific to the Microeconomics field were identified by comparing their normalised frequencies against those in the academic section of the Corpus of Contemporary American English (COCA) (Davies, 2008). COCA (academic) was chosen as the reference corpus in the study for several reasons. COCA represents a broad and balanced collection of American English texts, which includes over 120 million words from various genres of general academic writing compiled between 1990 and 2019. This extensive compilation allows the researcher to effectively compare the frequency of specific lexical frames in Microeconomics research articles against a general corpus of academic English, thereby identifying which frames are distinctive to Microeconomics. Using COCA as the reference corpus helps ensure that the lexical frames identified as *key* are truly characteristic of the Microeconomics discipline and not just common across general academic writing. While COCA primarily represents American English, its comprehensive nature and academic focus make it suitable for studies aiming to understand the specificity and uniqueness of lexical frames within a specialised field like Microeconomics. This choice helps to control for general academic usage patterns, allowing for clearer identification of phraseological patterns that are distinctive to the discipline being studied. Additionally, the use of COCA provides a methodologically robust baseline for calculating *keyness*, or the degree to which a lexical frame is specific to a particular corpus compared to a general reference. In calculating keyness, lexical frames with higher

normalised frequencies in COCA were excluded. The symmetric Mean Absolute Percentage Error (sMAPE) was calculated for the remaining frames:  $[(\text{Microeconomics} - \text{COCA}) / (\text{Microeconomics} + \text{COCA}) / 2]$ . A threshold sMAPE score of 1.95 was set to include only those frames that occur 100 times more frequently in the Microeconomics corpus than in COCA. To ensure statistical significance, Fisher's exact test was applied ( $p < 0.0001$ ). The refined list of key lexical frames meeting these criteria was then subjected to detailed structural and functional analysis. The structures of key lexical frames in the study were studied following Gray and Biber's (2013) structural categorisation of lexical frames:

- 1) Lexical frames with content words (except verb) (e.g., *is \* likely to*).
- 2) Lexical frames with at least one verb (e.g., *is \* related to*).
- 3) Lexical frames are formed by function words, including conjunction, determiner, preposition and pronoun (e.g., *in \* to the*).

The final stage of the key lexical frames analysis involved a detailed examination of their discourse functions to determine how these frames are used in various contexts. This analysis utilised Simpson-Vlach and Ellis's (2010) functional taxonomy, which classifies lexical frames into three main categories: referential expressions, stance expressions, and discourse organising expressions. This taxonomy was chosen because it provides a comprehensive framework for understanding the functional roles that lexical frames play in academic writing. Referential expressions are used to provide specific information, describe entities, or define concepts within the text. Stance expressions convey the author's attitudes, judgments, or evaluations about the subject matter, while discourse organising expressions help to structure the text and guide the reader through the logical progression of the argument.

To classify the key lexical frames functionally, this study followed the methodology of Lu et al. (2018), which involves analysing the semantics of frame variants and their contextual usage. Each lexical frame was examined within its specific textual environment to accurately determine its functional category, ensuring that the classification reflects both the inherent meaning of the frame and its role in the discourse. This contextual analysis provides deeper insights into how lexical frames contribute to meaning-making in academic writing. Understanding these functional categories is crucial for grasping the role of phraseological patterns in different sections of research articles, especially in the Discussion section, where the interpretation and contextualisation of findings are essential for effective academic communication.

## RESULTS

### THE MOST FREQUENT KEY LEXICAL FRAMES

To qualify for inclusion, key lexical frames need to occur at least 20 times per million words in the corpus, exhibit a sMAPE value of 1.95, and achieve a significance level of  $p < 0.0001$  on Fisher's exact test. In total, 43 four-word lexical frames met these stringent criteria. Table 1 provides a detailed overview of the 20 most frequently occurring four-word lexical frames. This includes information on the normalised frequencies of these key lexical frames as well as the most common slot fillers identified within them. Appendix A contains the complete list of the 43 key lexical frames, along with their frequent slot fillers. This list serves as a valuable resource for further

research and pedagogical applications, offering a deeper exploration of the language used in the Discussion section of Microeconomics research articles.

TABLE 1. The 20 most frequent 4-word lexical frames and the fillers

Lexical frame	Frequency (per million words=pmw)	Most frequent fillers
<b>the * of the</b>	189	impact, role, effects, analysis, elasticity
<b>to the * of</b>	187	dynamics, distribution, allocation, efficiency, optimisation
<b>the * effects of</b>	183	price, income, substitution, cross, cumulative
<b>in the * of</b>	172	context, case, framework, domain, scope
<b>for the * of</b>	150	purpose, objective, goal, function, task
<b>the * in the</b>	132	role, importance, significance, relevance, prominence
<b>on the * of</b>	123	basis, premise, assumption, foundation, consideration
<b>under the * of</b>	110	conditions, circumstances, policies, principles, guidelines
<b>within the * of</b>	102	context, framework, scope, setting, domain
<b>due to * of</b>	93	impacts, effects, limitations, consequences, challenges
<b>as a * for</b>	92	basis, substitute, method, rationale, strategy
<b>the * of these</b>	86	results, findings, implications, outcomes, conclusions
<b>likely to * in</b>	72	result, lead, culminate, develop, emerge
<b>role of * in</b>	61	technology, investment, competition, innovation, government
<b>to be * in</b>	60	effective, costly, projected, presumed, anticipated
<b>effects of * on</b>	52	subsidies, tariffs, taxes, policies, regulations
<b>the * to the</b>	43	impact, relation, contribution, effect, significance
<b>the * by the</b>	42	decision, action, study, analysis, report
<b>the * of such</b>	40	effect, impact, importance, relevance, necessity
<b>from the * of</b>	40	perspective, viewpoint, analysis, angle, approach

As shown in Table 1, *the \* of the*, with 189 occurrences pmw, is the most frequent four-word lexical frame in the Microeconomics corpus. The most frequent slot fillers of this lexical frame include *impact*, *role*, and *effects*, indicating that sequences like *the impact of the*, *the role of the* and *the effects of the* are common realisations. This highlights the emphasis on discussing the impacts, roles, and effects of various factors in Microeconomics research. Another frequently occurring frame is *to the \* of*, appearing 187 times pmw, typically filled by words such as *dynamics* (to the *dynamics* of), *distribution* (to the *distribution* of), and *allocation* (to the *allocation* of). This frame often introduces topics related to the dynamics, distribution, and allocation efficiency within economic systems.

Similarly, *the \* effects of*, with 183 occurrences pmw, frequently include slot fillers such as *price* (the *price* effects of), *income* (the *income* effects of), and *substitution* (the *substitution* effects of). This frame is used to discuss the effects of pricing, income changes, and substitution on economic outcomes. The lexical frame *in the \* of*, occurring 172 times pmw, is often filled by *context* (in the *context* of), *case* (in the *case* of), and *framework* (in the *framework* of), situating discussions within specific contexts, case studies, or theoretical frameworks. The lexical frame *for the \* of*, with a frequency of 150 pmw, is commonly filled by *purpose* (for the *purpose* of), *objective* (for the *objective* of), and *goal* (for the *goal* of), indicating a focus on the aims and objectives of economic policies or studies. The lexical frame, *the \* in the*, appearing 132 times



pmw, frequently includes *role* (the *role* in the), *importance* (the *importance* in the), and *significance* (the *significance* in the), highlighting crucial aspects and roles within the economic context. The lexical frame *on the \* of*, with 123 occurrences pmw, typically includes fillers such as *basis* (on the *basis* of), *premise* (on the *premise* of), and *assumption* (on the *assumption* of), which suggests discussions about foundational assumptions or premises of economic theories. The lexical frame *under the \* of*, appearing 110 times pmw, often includes *conditions* (under the *conditions* of), *circumstances* (under the *circumstances* of), and *policies* (under the *policies* of), indicating various conditions or policies affecting economic outcomes.

The lexical frame *within the \* of*, with 102 occurrences pmw, is frequently filled by *context* (within the *context* of), *framework* (within the *framework* of), and *scope* (within the *scope* of), situating discussions within specific contexts or frameworks. *Due to \* of*, appearing 93 times pmw, includes fillers such as *impacts* (due to *impacts* of), *effects* (due to *effects* of), and *limitations* (due to *limitations* of), explaining the reasons behind certain economic phenomena. The lexical frame *as a \* for*, with a frequency of 92 pmw, is typically filled by *basis* (as a *basis* for), *substitute* (as a *substitute* for), and *method* (as a *method* for), indicating discussions about the foundational basis or alternative methods in economic analysis. The frame *the \* of these*, appearing 86 times pmw, often includes *results* (the *results* of these), *findings* (the *findings* of these), and *implications* (the *implications* of these), summarising outcomes or implications of studies. The lexical frame *likely to \* in*, with 72 occurrences pmw, includes fillers such as *a result* (likely to *result* in), *lead* (likely to *lead* in), and *culminate* (likely to *culminate* in), predicting likely outcomes or developments in economic contexts. The lexical frame *role of \* in*, appearing 61 times pmw, is frequently filled by *technology* (role of *technology* in), *investment* (role of *investment* in), and *competition* (role of *competition* in), discussing the roles of various factors in economic development.

The lexical frame *to be \* in*, with a frequency of 60 pmw, commonly includes *effective* (to be *effective* in), *costly* (to be *costly* in), and *projected* (to be *projected* in), often discussing anticipated outcomes or costs. The frame *effects of \* on*, appearing 52 times pmw, is typically filled by *subsidies* (effects of *subsidies* on), *tariffs* (effects of *tariffs* on), and *taxes* (effects of *taxes* on), analysing the impacts of various economic policies or interventions.

The lexical frame *the \* to the*, with 43 occurrences pmw, often includes *impact* (the *impact* to the), *relation* (the *relation* to the), and *contribution* (the *contribution* to the), discussing relationships or contributions in economic contexts. The frame *the \* by the*, appearing 42 times pmw, is commonly filled by *decision* (the *decision* by the), *action* (the *action* by the), and *study* (the *study* by the), indicating discussions about actions or studies conducted. *The \* of such*, with a frequency of 40 pmw, frequently includes *effect* (the *effect* of such), *impact* (the *impact* of such), and *importance* (the *importance* of such), emphasising the significance or impact of certain factors. The frame *from the \* of*, appearing 40 times pmw, often includes *perspective* (from the *perspective* of), *viewpoint* (from the *viewpoint* of), and *analysis* (from the *analysis* of), presenting discussions from various perspectives or analytical angles. These key lexical frames and their frequent slot fillers provide valuable insights into the common phraseological patterns in Microeconomics research articles, reflecting the key themes and focus within the discipline.

#### THE STRUCTURAL CLASSIFICATION

Using the structural classification of lexical frames as proposed by Gray and Biber (2013), the key lexical frames identified in the Discussion section of Microeconomics research articles were categorised according to their structural correlates. This classification involved grouping the lexical frames based on the types of words that appear within them. The frames were divided into

three distinct groups: non-verb content word frames, verb-based frames, and function word frames. Table 2 provides a comprehensive overview of these categories, presenting both the types (different lexical frames) and tokens (total number of lexical frames) of four-word key lexical frames by their structural characteristics.

TABLE 2. The distribution of four-word key lexical frame types and tokens across the structural categories

Structure	Type	%	Token	%
Non-verb content word frame	9	21	801	27
Verb-based frame	4	9	205	7
Function word frame	30	70	2010	66

In the Discussion section of Microeconomics research articles, the majority of key lexical frames are function word frames, comprising 70% of the total types. Non-verb content word frames account for 21%, and verb-based frames represent 9%. This distribution shows a strong prevalence of function word frames in structuring discussions within this field. A similar trend is seen in the frame token proportions: function word frames dominate with 66% of total tokens, underscoring their significant presence in these sections. Non-verb content word frames constitute 27% of the tokens, while verb-based frames make up just 7%. The high percentage of function word frame tokens relative to their type proportion highlights their frequent and repetitive use in academic discourse.

The prevalence of function word frames can be attributed to their essential role in constructing complex sentence structures and linking ideas coherently. Function words, being part of closed word classes with limited membership, form the backbone of many grammatical constructions, which explains their significant presence despite the limited variety of function word frame types. On the other hand, non-verb content word frames, although fewer in type, are relatively frequent in token count. This suggests that specific non-verb content word frames are repeatedly employed to introduce and elaborate on key concepts and findings within the discussions.

Verb-based frames, while the least prevalent, still play a crucial role in conveying actions and processes related to the research findings. The lower proportion of verb-based frames may reflect the specific nature of discussions in Microeconomics, which often focus more on describing states, conditions, and relationships rather than actions.

Overall, the distribution of key lexical frame types and tokens highlights the importance of function word frames in organising and presenting information in the Discussion section of Microeconomics research articles. The frequent use of these frames aids in achieving clarity and coherence, which is essential for effective academic communication.

- a. Non-verb content word frames:  
*likely to [culminate/ emerge] in  
 impact of [policy/investment] on*
- b. Verb-based frames:  
*To be [effective/costly] of  
 this [result/observation] suggests that*
- c. Function word frames:  
*the [evaluation/assessment] of their  
 to the [dynamics/distribution] of*

### THE FUNCTIONAL CLASSIFICATION

In examining the functions of key lexical frames within the Discussion section of Microeconomics research articles, the present study utilised Simpson-Vlach and Ellis's (2010) functional taxonomy to categorise the discourse functions these frames serve. Table 3 illustrates the distribution of the key lexical frames across three functional categories, segmented by types and tokens.

TABLE 3. The distribution of 4-word key lexical frame types and tokens across the functional categories

Structure	Type	%	Token	%
Referential	32	75	2302	76
Discourse	7	16	484	16
Stance	4	9	230	8

Referential key lexical frames constitute the largest category, representing 75% of the types and 76% of the tokens. Examples of these referential frames include *the \* of the* with frequent fillers like *impact*, *role*, and *effects* and *the \* of* with common fillers such as *context*, *case*, and *framework*. These frames are essential for providing detailed descriptions and contextualising discussions, making them predominant in the corpus. By frequently employing these frames, authors can efficiently introduce and elaborate on crucial economic concepts, ensuring that the discourse remains focused and informative. For instance, a phrase like *the impact of the policy on market dynamics* succinctly conveys a significant analysis point, directly linking an economic policy to its observed effects.

Discourse-organising expressions, making up 16% of the types and tokens, play a crucial role in structuring the argument and guiding the reader through the text. Examples include *as a \* for* with fillers like *basis*, *substitute* and *method*, and *due to \* of* with fillers such as *impacts*, *effects*, and *limitations*. These frames are pivotal in organising information and linking different parts of the discourse, ensuring coherence and clarity. The use of discourse-organising frames allows authors to connect various sections of their arguments seamlessly, thereby enhancing the overall readability and logical flow of the article. For example, a phrase like *due to the impacts of regulatory changes* helps to clarify causal relationships and integrate different discussion points, leading to a more coherent narrative.

Stance expressions, which account for 9% of the types and 8% of the tokens, are the least frequent category. These frames convey the author's perspective or evaluation, such as *likely to \* in* with fillers like *result*, *lead* and *culminate*, and *to be \* in* with fillers like *effective*, *costly* and *projected*. These stance frames help articulate predictions, assessments, and evaluations within the discussions. Although less frequent, stance frames are critical for expressing the author's interpretive and evaluative judgments. They provide a means to speculate on future developments, assess the implications of findings, and convey subjective interpretations. For example, using a phrase like *likely to result in significant economic shifts* allows the author to project future trends and their potential impact based on the current analysis.

The distribution of functional categories in this study highlights the predominance of referential frames in conveying detailed content, the significance of stance frames in expressing evaluations and perspectives, and the crucial role of discourse-organising frames in maintaining textual coherence. These findings demonstrate the complex interplay of different types of lexical frames in constructing effective academic discourse in Microeconomics. The use of these frames not only aids in the clear communication of complex ideas but also reflects the methodological rigour and analytical depth typical of academic writing in this field. The dominance of referential

frames points to a focus on providing comprehensive analyses, which are essential for discussing and interpreting economic phenomena. The notable presence of discourse-organising frames underscores the importance of clear, logical argument structuring, which is vital for reader comprehension. Although less frequent, stance frames highlight the value of expert interpretation and the anticipation of future implications, adding depth and foresight to academic discussions.

## DISCUSSION AND CONCLUSION

The current study's findings show both similarities and differences compared to previous research. A notable similarity is the prominence of function word frames, which accounted for 70% of the types and 66% of the tokens, highlighting their crucial role in structuring academic discourse. This aligns with findings by Gray and Biber (2013) and He et al. (2021), where function word frames were also predominant. However, studies by Lu et al. (2018) and Golparvar and Barabadi (2020) found non-verb content word frames to be more dominant. This discrepancy can be attributed to differences in the corpora: Gray and Biber (2013) used general academic books and articles, He et al. (2021) focused on Business Management articles, while Lu et al. (2018) and Golparvar and Barabadi (2020) examined Social Science and Higher Education articles. These differences reflect the discipline-specific use of phraseological patterns and underscore the need for tailored approaches in teaching academic writing.

In terms of functional categories, referential key lexical frames were the most prevalent, making up 75% of the types and 76% of the tokens. These frames are crucial for providing detailed descriptions and contextualising discussions, highlighting their importance in academic writing. This trend is consistent with Lu et al. (2018) and Golparvar and Barabadi (2020), indicating that referential expressions are fundamental in both the introduction and Discussion sections of research articles. The prevalence of referential frames in this study underscores their role in elucidating and framing key concepts, which is essential for clear communication of complex economic phenomena. This facilitates a thorough understanding of the presented research, ensuring that readers can grasp the implications and nuances of the findings.

Discourse-organising expressions made up 16% of the types and tokens, playing a crucial role in structuring arguments and guiding readers through the text. While these frames were the second most common in this study, they were the least common in previous studies by Lu et al. (2018) and Golparvar and Barabadi (2020). The prominence of discourse-organising frames in this corpus underscores their importance in creating a coherent narrative flow, linking different sections of the text, clarifying causal relationships, and emphasising logical progression. This ensures that the discussion remains well-structured and that complex arguments are accessible. The discrepancy in the frequency of discourse-organising frames between this study and previous ones may be due to the specific nature of Microeconomics research articles, where clear and logical presentation of arguments is essential. This suggests that different academic disciplines may prioritise different types of lexical frames according to their unique communicative needs.

Stance expressions, which accounted for 9% of the types and 8% of the tokens, were the least frequent in this study. In contrast, they were the second most frequent frames in the studies by Golparvar and Barabadi (2020) and Lu et al. (2018), where they were used to convey the author's perspective or evaluation. The difference in frequency of stance expressions across studies likely stems from variations in the focus and nature of the corpora. For example, Golparvar and Barabadi (2020) analysed higher education research articles, where authors frequently engage in

evaluative commentary on pedagogical practices and educational outcomes. Similarly, Lu et al. (2018) focused on social science research articles, where authors often express stances on social phenomena and theoretical interpretations. In contrast, Microeconomics research articles tend to emphasise empirical findings and theoretical analyses over evaluative commentary, which may account for the lower frequency of stance expressions in this study. However, the presence of stance frames, even in smaller numbers, is significant for providing critical assessments and projecting future implications. Frames like *likely to \* in* and *to be \* in* enable authors to speculate on potential outcomes and offer insights into their research findings. The use of stance expressions is essential for engaging with the broader academic community, positioning the authors' findings within ongoing scholarly debates, and contributing to the development of the field by inviting further inquiry and discussion.

Overall, the findings from this study offer valuable insights into the use and functions of lexical frames in the Discussion section of Microeconomics research articles. When compared to lexical frames identified in other disciplines, such as Social Sciences or Applied Linguistics, it is evident that Microeconomics relies heavily on function word frames to structure complex arguments and convey precise meanings. For instance, the prevalence of frames like *the \* of the* and *effects of \* on* in Microeconomics highlights the field's focus on detailed, systematic analysis and the explanation of relationships between economic variables. This contrasts with fields like Social Sciences, where more stance-based lexical frames may be used to convey evaluative judgments or theoretical perspectives. These differences suggest that Microeconomics, as a discipline, prioritises clarity, precision, and the logical progression of ideas, reflecting its foundational role in economic theory and policy-making. Understanding these distinct phraseological patterns can, therefore, provide deeper insights into the communicative norms and expectations within the field of Microeconomics, which are crucial for both scholars and students engaged in academic writing in this area. By understanding the structural and functional roles these frames play, researchers and educators can better appreciate the intricacies of academic writing and improve the clarity and impact of their own scholarly work. The strategic use of these lexical frames facilitates the construction of well-rounded, coherent, and insightful academic articles that contribute meaningfully to the field of Microeconomics. The interplay of these frames supports the articulation of complex economic analyses, the logical structuring of discourse, and the expression of evaluative and predictive insights, thereby enhancing the overall quality and impact of the research presented.

Key lexical frames can significantly enhance the teaching and learning of academic writing in Microeconomics by providing insights into the specific linguistic patterns used in the discipline (Cunningham, 2017). To incorporate the findings into EAP teaching materials, educators can develop targeted exercises that help students identify and practice using these frames in context. For example, instructors could create fill-in-the-blank activities where students select appropriate slot fillers for frames like *the \* effects of* or *due to \* of*, helping them understand how these frames are employed to convey specific economic concepts. Additionally, writing assignments could encourage students to integrate these lexical frames into their own academic writing, reinforcing their ability to produce clear and coherent Microeconomics discourse. By focusing on these key lexical frames, EAP courses can better prepare students, particularly non-native speakers, to meet the discipline-specific writing standards expected in Microeconomics. Understanding the specific uses of lexical frames helps learners produce language that aligns with native speakers' norms (Gray & Biber, 2013).

Focusing on the Discussion section of research articles offers both advantages and limitations. While it narrows the application of key lexical frames to this specific part-genre, it also opens up opportunities for future research to explore how these frames function in other sections of Microeconomics research articles, such as the Introduction, Methodology, and Conclusion. Future studies could also investigate the presence and role of similar lexical frames in related disciplines, like Macroeconomics, Finance, or Business Studies, to determine whether these patterns are consistent across the broader field of Economics or unique to Microeconomics. Additionally, extending the analysis to include longer lexical frames (e.g., five-word or six-word frames) could provide deeper insights into the complexity of academic discourse. These avenues for future research would not only enhance our understanding of phraseological patterns in academic writing but also support the development of more effective teaching materials for EAP learners across disciplines. This study is particularly innovative in highlighting the importance of phraseological variation across different part genres, which is essential for improving EAP learners' awareness and competence. Lastly, evaluating the pedagogical utility of these key lexical frames and researching their effectiveness in EAP classrooms are crucial steps for future educational practices.

#### ACKNOWLEDGEMENTS

*This work was supported by Universiti Sains Malaysia, Bridging with Project No: R501-LR-RND003-0000000736-0000.*

#### REFERENCES

- Ang, L. H., & Tan, K. H. (2019). From lexical bundles to lexical frames: Uncovering the extent of phraseological variation in academic writing. *3L: The Southeast Asian Journal of English Language Studies*, 25(2), 99–112. <https://doi.org/10.17576/3L-2019-2502-08>
- Anthony, L. (2021). *AntConc (Version 3.5.9)* [Computer Software]. Tokyo: Waseda University. <https://www.laurenceanthony.net/software>
- Biber, D. (2009). A corpus-driven approach to formulaic language in English. *International Journal of Corpus Linguistics*, 14, 275–311. <https://doi.org/10.1075/ijcl.14.3.08bib>
- 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>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Grammar of spoken and written English*. Essex: Longman. <https://doi.org/10.1075/z.232>
- Casal, J. E., & Kessler, M. (2020). Form and rhetorical function of phrase-frames in promotional writing: A corpus- and genre-based analysis. *System*, 95, 102370. <https://doi.org/10.1016/j.system.2020.102370>
- Cunningham, K. J. (2017). A phraseological exploration of recent mathematics research articles through key lexical frames. *Journal of English for Academic Purposes*, 25, 71–83. <https://doi.org/10.1016/j.jeap.2016.11.005>
- Dahunsi, T. N., & Ewata, T. O. (2022). An exploration of the structural and colligational characteristics of lexical bundles in L1–L2 corpora for English language teaching. *Language Teaching Research*. <https://doi.org/10.1177/13621688211066572>
- Davies, M. (2008). *The corpus of contemporary American English: 450 million words, 1990-present*. <https://www.english-corpora.org/coca/>
- Firth, J. R. (1957). Modes of meaning. In F. R. Palmer (Ed.), *Papers in linguistics 1934–1951* (pp. 190–215). Oxford: Oxford University Press.
- Fletcher, W. H. (2007). *KfNgram (Version 1.3.1)* [Computer Software]. Annapolis, MD: USNA. <http://kwicfinder.com/kfNgram/kfNgramHelp.html>
- Garner, J. R. (2016). A phrase-frame approach to investigating phraseology in learner writing across proficiency levels. *International Journal of Learner Corpus Research*, 2, 31–67.
- Golparvar, S. E., & Barabadi, E. (2020). Key lexical frames in the discussion section of research articles of higher education. *Lingua*, 236, 1–15. <https://doi.org/10.1016/j.lingua.2020.102804>

- Grabowski, Ł. (2015). Lexical frames in English pharmaceutical discourse: A corpus-driven study of interdisciplinary register variation. *Research in Language*, 13, 266–291. <https://doi.org/10.1515/rela-2015-0025>
- Gray, B., & Biber, D. (2013). Lexical frames in academic prose and conversation. *International Journal of Corpus Linguistics*, 18, 109–136. <https://doi.org/10.1075/ijcl.18.1.08gra>
- He, M., Ang, L. H., & Tan, K. H. (2021). A corpus-driven analysis of lexical frames in research articles on business management. *Southern African Linguistics and Applied Language Studies*, 39, 139–151. <https://doi.org/10.2989/16073614.2021.1920438>
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge. <https://doi.org/10.4324/9780203327630>
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524773>
- Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18, 41–62. <https://doi.org/10.1111/j.1473-4192.2008.00178.x>
- Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4–21. <https://doi.org/10.1016/j.esp.2007.06.001>
- Lu, X., Yoon, J., & Kisselev, O. (2018). A phrase-frame list for social science research article introductions. *Journal of English for Academic Purposes*, 36, 76–85. <https://doi.org/10.1016/j.jeap.2018.09.004>
- Lu, X., Yoon, J., & Kisselev, O. (2021). Matching phrase-frames to rhetorical moves in social science research article introductions. *English for Specific Purposes*, 61, 63–83. <https://doi.org/10.1016/j.esp.2020.10.001>
- Renouf, A., & Sinclair, J. (1991). Collocational frameworks in English. In K. Aijmer, & B. Altenberg (Eds.), *English corpus linguistics* (pp. 128–143). New York: Longman.
- Reppen, R., & Olson, S. (2020). Lexical bundles across disciplines: A look at consistency and variability. In U. Römer, V. Cortes, E. Friginal (Eds.), *Advances in corpus-based research on academic writing: Effects of discipline, register and writer expertise* (pp. 169–182). Amsterdam: John Benjamins Publishing. <https://doi.org/10.1075/scl.95.07rep>
- Römer, U. (2009). English in academia: Does nativeness matter? *Anglistik: International Journal of English Studies*, 20, 89–100.
- Römer, U. (2010). Establishing the phraseological profile of a text type: The construction of meaning in academic book reviews. *English Text Construction*, 3, 95–119. <https://doi.org/10.1075/etc.3.1.06rom>
- Shirazizadeh, M., & Amirfazlian, R. (2021). Lexical bundles in theses, articles and textbooks of applied linguistics: Investigating intradisciplinary uniformity and variation. *Journal of English for Academic Purposes*, 49, 100946. <https://doi.org/10.1016/j.jeap.2020.100946>
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31, 487–512. <https://doi.org/10.1093/applin/amp058>
- Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press: Oxford.
- Starfield, S. (2016). English for specific purposes. In G. Hall (Ed.), *The Routledge handbook of English language teaching* (pp. 150–163). London: Routledge.
- Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: Essential tasks and skills*. Ann Arbor: University of Michigan Press.
- Wang, Y. (2019). A functional analysis of text-oriented formulaic expressions in written academic discourse: Multi-word sequences vs. single words. *English for Specific Purposes*, 54, 50–61. <https://doi.org/10.1016/j.esp.2018.12.002>
- Win, Y., & Masada, T. (2015, March 24–27). Exploring technical lexical frames from research paper titles. In *Proceedings – IEEE 29th International Conference on Advanced Information Networking and Applications Workshops (WAINA)* (pp. 558–563). Gwangju, South Korea.
- Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.
- Yoon, J., & Casal, J. E. (2020). P-frames and rhetorical moves in applied linguistics conference abstracts. In U. Römer, V. Cortes, & E. Friginal (Eds.), *Advances in corpus-based research on academic writing: Effects of discipline, register and writer expertise* (pp. 282–305). Amsterdam: John Benjamins Publishing. <https://doi.org/10.1075/scl.95.12yoo>

APPENDIX A

List of key four-word lexical frames in the Discussion section of Microeconomics research articles

Lexical frame	Frequent filler
the * of the	impact, role, effects, analysis, elasticity
to the * of	dynamics, distribution, allocation, efficiency, optimisation
the * effects of	price, income, substitution, cross, cumulative
in the * of	context, case, framework, domain, scope
for the * of	purpose, objective, goal, function, task
the * in the	role, importance, significance, relevance, prominence
on the * of	basis, premise, assumption, foundation, consideration
under the * of	conditions, circumstances, policies, principles, guidelines
within the * of	context, framework, scope, setting, domain
due to * of	impacts, effects, limitations, consequences, challenges
as a * for	basis, substitute, method, rationale, strategy
the * of these	results, findings, implications, outcomes, conclusions
likely to * in	result, lead, culminate, develop, emerge
role of * in	technology, investment, competition, innovation, government
to be * in	effective, costly, projected, presumed, anticipated
effects of * on	subsidies, tariffs, taxes, policies, regulations
the * to the	impact, relation, contribution, effect, significance
the * by the	decision, action, study, analysis, report
the * of such	effect, impact, importance, relevance, necessity
from the * of	perspective, viewpoint, analysis, angle, approach
necessary to * the	explore, analyse, investigate, expect
to * in the	contribute, engage, invest, spend
a significant * in	increase, decrease, impact, change, shift
impact of * on	changes, policy, investment, regulation, innovation
the * among the	differences, similarities, correlations, patterns, trends
to the * and	understanding, interpretation, explanation, clarification, analysis
the * between the	relationship, correlation, connection, association, link
the * for the	basis, justification, rationale, explanation, reasoning
the * is likely	outcome, result, effect, conclusion, implication
this * suggests that	finding, result, analysis, study, observation
the * to be	approach, methodology, framework, strategy, model
the * of this	implication, consequence, result, significance, relevance
the * with the	relationship, interaction, correlation, association, link
the * of their	understanding, interpretation, evaluation, assessment, insight
to * in order	analyse, evaluate, understand, investigate, assess
the * as a	use, role, function, basis, justification
the * that the	assumption, hypothesis, theory, assertion, proposition
the * as an	indicator, measure, example, instance, representation
the * by this	analysis, study, investigation, research, exploration



---

<b>the * of which</b>	extent, nature, impact, scope, breadth
<b>the * as the</b>	basis, justification, foundation, explanation, rationale
<b>the * through the</b>	implementation, application, use, adoption, integration
<b>the * from the</b>	outcome, result, effect, consequence, conclusion

---