Making requests at work: An examination of phrase frames in workplace email communication

Detong Xia[1], **Matt Kessler**[2], **Yudi Chen**[3] **& Hye Pae**[4]

Xi'an Jiaotong-Liverpool University (China)[1], University of South Florida (United States)[2], Nanjing University of Aeronautics and Astronautics (China)[3] & University of Cincinnati (United States)[4]

detong.xia@xjtlu.edu.cn, mattjkess@gmail.com, ychen55@nuaa.edu.cn & paehe@ucmail.uc.edu

Abstract

In the realm of business communication, emails play a crucial role in fostering relationships and building support networks among colleagues, partners, and sponsors to achieve common business objectives. Research has suggested that nonnative speakers of English and individuals who are new to professional writing often encounter challenges when crafting formal email requests for business purposes. Previous studies have underscored the significance of phrase frames (p-frames), which are recurring multi-word sequences with a variable slot, in fulfilling rhetorical functions in academic discourse. However, little research has explored how p-frames are used for written communications within the business workplace context. This study investigated five-word p-frames used in rhetorical moves in business email requests. Using 1,148 authentic request emails from the Avocado Research Email Collection corpus, we analyzed the distribution and linguistic characteristics of p-frames across five rhetorical moves. Results showed an uneven distribution of p-frames across these rhetorical moves in business request emails. Notably, two moves, *making the inquiry* and *closing*, showed the highest degree of formulaicity. P-frames were found to be used in a variety of messages to soften demands and convey politeness. The findings have pedagogical implications for teaching English for business purposes, benefiting both English language learners and novice writers.

Keywords: workplace emails, requests, formulaic sequences, corpus-based genre analysis.

Resumen

Realizando solicitudes en el trabajo: Un análisis de los n-gramas discontinuos en la comunicación por correo electrónico en el entorno laboral

En el ámbito de la comunicación empresarial, los correos electrónicos desempeñan un papel crucial a la hora de fomentar relaciones y construir redes de apoyo entre compañeros, socios y patrocinadores con el fin de alcanzar objetivos comerciales comunes. La bibliografía ha sugerido que la redacción de solicitudes formales por correo electrónico con fines laborales es un reto tanto para los hablantes no nativos de inglés como para quienes carecen de experiencia en la escritura profesional. Algunos estudios previos han destacado la importancia de los marcos de frases o n-gramas discontinuos (*p-frames*), esto es, secuencias recurrentes de varias palabras con un espacio variable, para llevar a cabo ciertas funciones retóricas en el discurso académico. Sin embargo, se ha investigado poco sobre el modo en el que se utilizan los *p-frames* en las comunicaciones escritas en el contexto laboral empresarial. El presente estudio examina los *p-frames* de cinco palabras utilizados en movimientos retóricos en solicitudes de negocios por correo electrónico. Con base en 1.148 correos electrónicos auténticos de solicitudes del corpus *Avocado Research Email Collection*, analizamos la distribución y las características lingüísticas de los *p-frames* en cinco movimientos retóricos. Los resultados muestran una distribución desigual de los *p-frames* en los movimientos retóricos de los correos electrónicos de solicitud empresarial. Dos movimientos mostraron el mayor grado de formulaicidad: la *realización de la consulta* y el *cierre*. Asimismo, se ha detectado que los *p-frames* se utilizan en diferentes tipos de mensajes para mitigar las demandas y expresar cortesía. Estos hallazgos tienen implicaciones pedagógicas para la enseñanza de inglés con fines comerciales, lo cual resulta beneficioso tanto para los estudiantes de inglés como para los escritores noveles en este ámbito.

Palabras clave: correos electrónicos en el ámbito laboral, solicitudes, secuencias formulaicas, análisis de género basado en corpus.

1. Introduction

Email stands out as one of the most widely used forms of professional communication. It serves as the primary means for sharing information, coordinating tasks, and facilitating decision-making in professional settings. According to the survey by Radicati Group, the total daily volume of business and consumer emails surpassed 300 billion in 2021, and this figure is anticipated to continue its upward trajectory. In addition, email has been identified as a time-intensive activity for professionals, highlighting its central

role in day-to-day business operations. As a subset of the business correspondence genre, emails have attracted significant scholarly attention over the past two decades. Applied linguists have explored various facets of email communication, examining linguistic and stylistic conventions (Baron, 1998; Park et al., 2021; Qian & Pan, 2019), discourse strategies (Ho, 2018; Leopold, 2015), email structures (Van Herck et al., 2022; Warren, 2016), and differences in email composition between native and nonnative speakers of English (e.g., Giménez-Moreno & Skorczynska, 2013; Xia, Ai & Pae, 2022). Empirical research has identified the specific challenges faced by nonnative speakers in email writing, particularly related to word choice and the nuanced understanding of formality (e.g., Incelli, 2013; Xia, Sulzer & Pae, 2023). Crafting effective emails requires not only a command of language skills, but also the skillful application of linguistic functions tailored to specific situations (Carrió-Pastor & Muñiz-Calderón, 2013; Evans, 2014).

To address learners' difficulties in incorporating their linguistic knowledge into their writing in functionally appropriate ways, a growing body of research has integrated both genre analysis (of rhetorical functions) and corpus analysis (of linguistic features) in second language (L2) writing research (e.g., Casal & Kessler, 2020, 2024; Cortes, 2024; Lu, Casal & Liu, 2020; Yoon & Casal, 2020). Research in this line aims to tackle the concept of the "function-form gap" (Moreno & Swales, 2018, p. 41). This gap pertains to the relationship between the functional aspects of a text (its communicative functions or purposes) and its formal linguistic features (the specific language patterns and structures used to convey those functions). Lu, Casal & Liu (2021) dubbed such an approach a "corpus-based genre analysis" (p. 62), with a particular focus on the utilization of formulaic linguistic patterns to express the rhetorical functions of a particular genre. This involves the identification of the most prominent text patterns used in a particular rhetorical context that can effectively guide experienced readers to understand specific communicative purposes. A particular type of formulaic expression that is gaining increasing attention in corpus-based genre analysis research is phrase frames (or p-frames; Fletcher, 2012). A p-frame refers to a form of recurrent multi-word sequence with one variable slot (e.g., *as * as the [far, soon, well]*). P-frames have been identified as an ideal prefabricated formulaic pattern because their embedded lexico-grammatical characteristics are associated with the discourse function of a particular genre. P-frames also contain important characteristics of formulaic language, such as frequency, fixedness, and variability, such that they can

reveal the level of formulaicity of a text (Cortes, 2024; Forsyth & Grabowski, 2015; Golparvar & Barabadi, 2020).

Previous studies have investigated p-frames in the realization of specific rhetorical functions in a variety of academic genres, such as applied linguistics conference abstracts (e.g., Yoon & Casal, 2020) and social science research article introductions (e.g., Lu, Yoon & Kisselev, 2021). These analyses have suggested that some linguistic features embedded in p-frames are strongly associated with the rhetorical goals of academic genres, which can provide important implications for both text analysis research and English for Academic/Specific Purposes (EAP/ESP) pedagogy. However, we are still less informed regarding how p-frames are used to fulfill rhetorical functions in professional workplace genres. As such, this study seeks to contribute to the growing body of corpus-based genre analysis research by investigating p-frames in terms of their aimed rhetorical functions in workplace request emails. This research can provide a better understanding of the phraseological features of workplace genres and help nonnative speakers of English or novice writers become familiar with typical discursive practices in workplace communities.

2. Literature review

2.1. Request emails

English serves as a common language that professionals use to collaborate and communicate with one another in the context of workplace discourse. As noted by Kankaanranta (2006), workplace emails facilitate a range of communicative functions. The primary purposes of emails are to provide information, issue directives, and make requests (Ghadessy & Webster, 1988). Among them, request emails may be the most difficult type, even for advanced L2 speakers of English. Making requests, which is the action of getting the addressee to do something in response to what is asked for by the addressor, can be a face-threatening act (Brown & Levinson, 1987). A face-threatening act can cause embarrassment or awkwardness, which may put professional relationships with members of the workplace at risk. Writers are expected to have substantial knowledge about the standards or norms of request emails and be able to articulate both linguistic (lexical or syntactic features) and pragmatic devices (e.g., politeness or directness) for social and situational appropriateness.

Request emails have drawn considerable attention from a number of EAP and ESP scholars since its emergence as the primary communication channel in the late 1970s (e.g., Baron, 2002; Ho, 2018; Nguyen & Miller, 2012; Park et al., 2021; Xia, Ai & Pae, 2022). English-speaking professionals typically follow certain norms of politeness (e.g., conventionalized opening and closing phrases) when making requests in email. Leopold (2015) investigated request strategies and linguistic devices used in the emails of English-speaking professionals in the United States. Her findings showed that working professionals preferred to use *let*-incorporated verb phrases to initiate requests (e.g., *let me know*, *let's make a time*). To tactfully soften demands, working professionals would often use the politeness marker *please* to make a command, such as *please share this with the team* and *please let me know if we agree* (Leopold, 2015, p. 12). Similarly, when examining English email exchanges between a British and an Italian company, Incelli (2013) reported that British working professionals used an extensive range of conventionalized phrases (e.g., *please find enclosed*) to maintain an objective tone.

Unlike native English speakers, nonnative speakers often take distinct approaches when composing email communication requests (Evans, 2014; Li & Volkov, 2017; Millot, 2017). For example, Millot (2017) found that, while workplace emails written by native English speakers relied on factual descriptions and documented interpretations with supporting evidence in detail, those written by nonnative English speakers from France relied more on opinion-based evaluations and statements. Additionally, Park et al. (2021) examined the rhetorical functions of English request emails written by naitve English professionals and Korean working professionals. The researchers identified nine rhetorical moves (i.e., functional units with specific communicative purposes) including *email subject*, *small talk*, *self-introduction*, *purpose of email*, *background*, *request*, *supportive move*, *asking about availability and requesting a reply*, and *closing*. Results indicated that Korean professionals used more direct expressions than did English professionals in addressing business matters. For example, while the English speakers used indirect language (e.g., *is it possible*, *it would be*), the Korean professionals composed their emails using a higher level of directness (e.g., *I want to ask*, *can you do*).

Notably, these studies underscore the significance of politeness and the use of conventionalized language in workplace emails, emphasizing the intricate nature of request emails. Given the significance and complexity of request

emails, there is a need to investigate request emails in authentic workplace communications both to understand the expectations of the discourse community and to inform ESP writing pedagogy.

2.2. Corpus-based genre analysis

ESP researchers have long been interested in understanding the language required for L2 writers to be successful in a variety of academic and professional genres. Swales (1990) defined genres as "a class of communicative events, the members of which share some set of communicative purposes" (p. 58). Within ESP, Swales (1990, 2004) proposed a genre analysis framework for analyzing and understanding different genres, and his gene analysis framework is characterized by segmenting texts into a series of rhetorical moves. A move is operationalized as a rhetorical unit of a text that performs a specific communicative purpose. Genre researchers have attempted to understand the typical communicative functions of a target genre by classifying texts into moves based on the prototypical communicative purposes, and then subsequently diving into the associated linguistic features of each move (e.g., Casal & Kessler, 2024; Flowerdew, 2005; Kessler, 2020).

Recent genre research has shown a growing interest in investigating formulaic sequences as a particular linguistic feature (e.g., Casal & Kessler, 2020; Kim & Kessler, 2022; Yoon & Casal, 2020). Formulaic sequences are highly frequent word-strings which have been found to account for a large proportion of the language production (Biber et al., 2004). In addition, a large number of empirical studies have shown that the use of formulaic sequences can facilitate fluency, accuracy, and formulaicity in L2 writing (e.g., Garner et al., 2019; Granger, 2018; Xia, Chen & Pae, 2023). In particular, one type of formulaic sequence, p-frames, has been used prominently in recent corpus-based genre research (e.g., Cortes, 2024; Casal & Kessler, 2020; Lu, Yoon & Kisselev, 2021; Yoon & Casal, 2020).

Research in this line has explored the relationships between p-frames and rhetorical moves in academic genres. For example, Lu, Yoon and Kisselev (2021) offered insights into how writers in the social sciences used p-frames to achieve different rhetorical functions in the introduction section of research articles (RAs). The researchers first manually annotated the rhetorical moves in a corpus of 600 RA introductions. Extracted five- and six-word p-frames were then carefully analyzed in terms of the functions

they identified in each rhetorical move. Results showed that the p-frames were unevenly distributed across rhetorical moves. Moreover, the p-frames displayed the different strengths of associations with rhetorical moves. For example, the p-frame *the * [purpose, goal, aim] of the paper* was only used for the rhetorical move of "announcing present research", while *due to the * [lack, deepening, passage] of* did not show an inherent semantic relationship to any rhetorical move in the RA introductions. The researchers also identified three types of p-frames based on their strength of association with the rhetorical moves, including specialized p-frames (occurring in a single rhetorical move), semi-specialized p-frames (at least two-thirds of the occurrences in one rhetorical move), and non-specialized p-frames (fewer than two-thirds of the occurrences in one rhetorical move). In another study, Yoon and Casal (2020) investigated the use of five- and six-word p-frames across rhetorical moves in a corpus of 625 applied linguistic conference abstracts. The researchers found that p-frames occurring in the moves of "announcing present research" and "describing future presentation" were mostly specialized p-frames, indicating high formulaicity of the two moves.

Casal and Kessler's (2020) study is one of the few studies that have explored the linguistic realization of rhetorical functions in a non-academic genre. They analyzed the distribution of five-word p-frames across rhetorical moves in 148 purpose statements written for grant applications for the US Fulbright program. Their findings showed that most p-frames were strongly associated with a particular rhetorical move. Their research also confirmed the usefulness and pedagogical value of p-frames for teaching and learning genre-specific phraseological and functional features based on a survey of both L2 writing instructors and learners regarding their opinions of the utility and usefulness of the extracted p-frames.

Taken together, the reviewed literature uncovers the prevalence of formulaic sequences in texts, the strong association between p-frames and particular rhetorical functions, and writers' varying linguistic choices based on different rhetorical goals. Cortes (2013, 2024) has suggested that an analysis of associations between rhetorical functions and p-frames can inform the move-schema theory, as these linguistic features contribute to a better description and explanation of communicative functions in different rhetorical moves. However, as noted, most studies to date have focused on academic genres, leaving workplace genres underexplored, particularly from a corpus-based genre analysis approach.

2.3. The current study

With the increased focus on formulaic sequences in writing research yet a limited number of studies delving into the use of formulaic language in workplace discourse, this study occupies this gap by investigating how p-frames are employed to fulfill rhetorical functions in workplace request emails. Such an investigation could uncover how working professionals integrate linguistic features into their email writing and ensure their functional appropriateness. Additionally, we sought to provide insights that can inform corpus- and genre-based pedagogies. Two questions were addressed in this study:

1. How are p-frames distributed across rhetorical moves in a corpus of workplace request emails?

2. What are the linguistic characteristics of p-frames in workplace request emails?

3. Method

3.1. Corpus description

A workplace corpus of authentic email messages from the Avocado Research Email Collection[1] was used in this study. Avocado (pseudonym) was a US Information Technology software and services company that developed products for the mobile Internet market, operating from the late 1990s to the middle of the first decade of the 21st century. The Avocado collection comprises 937,958 email messages derived from 279 employee accounts. We chose the Avocado dataset for this study because it represented the most up-to-date real-world email data that were publicly accessible. To make the data suitable for linguistic analyses, we deleted all annotated information, such as dates, subject lines, and email addresses of senders/recipients. Only the main body of each email message was kept for further analysis.

It is important to note additional information about the Avocado dataset before proceeding. Because the first language status of Avocado's employees was not available during the study, there is a possibility that the Avocado dataset may include emails composed by nonnative speakers of English. Past research suggests that in business communication, professional competence,

skill set, and work experience be considered more important than English language proficiency (see Kankaanranta et al., 2018; Millot, 2017). Given that the dataset mainly contains senior employees' emails, we contend that the Avocado dataset could sufficiently function as a representative example of authentic workplace email correspondence within the information technology sector.

The Avocado dataset contains emails for different communicative purposes such as requests, information sharing, and directions or guidelines. As we were only interested in emails written for request purposes, we manually identified such emails. An email was considered as a request if it contained an act of request and expressed expectations from the sender to the recipient about prospective action required in the workplace. We randomly sampled 20 emails from the top 200 email writers in the Avocado dataset (4,000 emails in total). After that, the first author and one experienced business English teacher independently determined whether the sampled 4,000 emails belonged to the category of request emails. Only those that were marked as request emails by both coders were included for analysis. As a result, we had 1,148 emails coded as request emails (total number of words = 125471, M length = 109.5, SD = 54.8) in our workplace request email (WRE) corpus. Email communications in the WRE corpus primarily focused on software development, client relationships, and information technology infrastructure with various business situations covered, such as scheduling meetings, exchanging ideas, discussing contracts, coordinating events, and seeking assistance for business-related matters or technology support.

3.2. Rhetorical move framework and annotation

Park et al.'s (2021) rhetorical move framework for request emails was selected for the move analysis in this study. However, there were differences in writing contexts, tasks, and topics between their study and the current one. While Park et al. had working professionals write request emails in a single discourse completion writing task for their study, we used an existing corpus containing authentic email messages. The use of natural language production allowed us to minimize task effects in this study and to focus on the nature of requests more broadly. To address the differences and to accommodate the particulars found in the corpus for this study, we modified Park et al.'s categorizations. In the pilot coding, the first author and one experienced business English teacher used Park et al.'s framework and marked chunks for further discussion if they did not directly correspond to

this framework. The two coders had multiple rounds of discussions to modify the framework as necessary. Table 1 shows our modified framework with five rhetorical moves including Move 1 (*background*), Move 2 (*purpose of email*), Move 3 (*making the inquiry*), Move 4 (*supportive move*), and Move 5 (*closing*).

Moves	Brief explanation of the move	Examples from the WRE corpus
Move 1 Background	Writer chats about previous contact, introduces the attachment(s), and/or talks about current situations (e.g., difficulties faced).	*Hi XXX, I hope you are doing well.*
Move 2 Purpose of email	Writer specifies the general purpose of the email.	*I just wanted to follow up with you on my previous email.*
Move 3 Making the inquiry	Writer asks for something such as information, action, availability, opinions, permissions, and/or funding.	*Did you get any feedback from XXX company yet?*
Move 4 Supportive move	Writer provides additional information about their previous inquiry (e.g., further explanations or offering help), justifies their inquiry, and/or expresses their apology.	*I will be out of the office from 9/24 to 9/28 but feel free to contact Dr. XXX if you need any additional information during my absence. His number is (XXX) XXX-XXXX.*
Move 5 Closing	Writer expresses their appreciation, asks for a reply from the recipient, and/or opens up for further communication.	*I appreciate your help.*

Table 1. Moves identified in the request emails in the WRE corpus

The two coders independently annotated all request emails based on the modified framework. Both coders followed a simple tag set for annotation that required move tags to be placed within brackets at move boundaries. Each move had two tags, one signaling the start of each move (e.g., [M2_start]) and the other one indicating the end of the move (e.g., [M2_end]). The tags helped the following procedure to match p-frames to the rhetorical moves (to be discussed in section 3.4). The inter-coder agreement for the move coding was high (0.95). Despite the high level of inter-coder agreement, the two coders diligently addressed all discrepancies through thorough discussions until their consensus was reached 100%.

3.3. P-frame extraction

Methodological decisions were made for the identification and extraction of p-frames. In a pilot study, we extracted p-frames containing four-, five-, and six-word sequences from our corpus. We finally decided to focus on five-word frames because they entailed greater information than four-word frames and displayed greater variability than six-word frames. The frequency threshold for p-frames extraction was set at 6 (roughly corresponding to 43 times per million words). In addition, each p-frame needed to have more

than two variants and appear in at least 5 emails. Such frequency and range decisions were made to ensure that only highly frequent multi-word sequences were included in this study.

We used *kfNgram* (Fletcher, 2012) to extract all five-word p-frames in the WRE corpus. The extraction included two steps: In the first step, the *kfNgram* software generated a list of five-grams (i.e., sequences of five words) from the WRE corpus (e.g., *look forward to working with*). In the second step, *kfNgram* created five-word p-frames by combining five-grams that were identical with only one variable slot. For example, the five-grams *look forward to working with* (6 occurrences), *look forward to speaking with* (6 occurrences), *look forward to talking with* (3 occurrences) would generate a five-word p-frame *look forward to * with* (15 occurrences). The output of the *kfNgram* contained all types of p-frames, the corresponding token counts, and the variants of each p-frame together with the token counts for each variant. The type refers to the number of distinct p-frames observed, while the token refers to the total number of occurrences. Take the aforementioned p-frame *look forward to * with* as an example; it is one type of p-frame with 3 variants and 15 tokens.

Some researchers have eliminated p-frames with a variable slot in the initial and final positions since these p-frames always crossed phrasal or clausal boundaries without conveying complete meaning (e.g., Lu, Yoon & Kisselev, 2018) or might be subsumed into larger frames (e.g., Römer, 2010). However, Yoon and Casal (2020) identified a subset of productive and meaningful p-frames with an initial or final slot that could not be subsumed into larger frames, and thus deserved further investigation. As such, we decided to include p-frames with a variable slot in all positions in our study.

The two coders manually checked all extracted p-frames and their variants in context using the concordance program *AntConc* (Anthony, 2020). If a p-frame crossed the clausal or phrasal boundaries (e.g., *have any questions please * [do, feel contact]*) or lacked a complete linguistic meaning (e.g., *me know when you * [think, need, get]*) or semantic coherence (e.g., *to hearing from you * [regards, we]*), they were deleted. The final list consisted of 167 p-frames.

3.4. Matching p-frames to rhetorical moves

We wrote a script in Python 2.7.16[2] to automatically map a p-frame to one of the rhetorical moves in which they occurred. The output generated the following information: p-frames with frequency, their variants with

frequency, the occurrences of p-frames in each rhetorical move, and the text chunks in which the frame occurred. The association strength of a p-frame was calculated by dividing the number of times a frame occurred in its primary function by the total occurrences of the p-frame (e.g., Yoon & Casal, 2020). For example, *feel free to * me* had 9 occurrences in Move 5 (*closing*) and 1 occurrence in Move 4 (*supportive move*) in the WRE corpus. Therefore, the primary function of *feel free to * me* was for accomplishing Move 5 (*closing*), and its association strength was 0.90 (9 divided by 10). When a p-frame equally occurred in two or more moves (out of the 10 total occurrences of *want to make sure **, four times in Move 2, and four times in Move 3), the association strength was counted once. The measure of association strength was used to examine the distribution of p-frames across rhetorical moves in addition to frequency (Lu, Yoon & Kisselev, 2021; Yoon & Casal, 2020).

Following Lu, Yoon & Kisselev (2021), we divided p-frames based on their association strength with their primary rhetorical functions, including the specialized, semi-specialized, and non-specialized p-frames. As the specialized p-frames had all their occurrences in one specific rhetorical move in the corpus, they were unique to one rhetorical function. The semi-specialized p-frames occurred in two or more rhetorical moves but had the majority (at least two-thirds) of the occurrences found in a specific move. The non-specialized p-frames had fewer than two-thirds of their occurrences in any particular move, and thus displayed no strong association with any single move.

4. Results

4.1. The distribution of p-frames across rhetorical moves in the request emails

To determine the frequency distribution of rhetorical moves and p-frames in the request emails, we calculated the number of emails containing each rhetorical move as well as the type and token of p-frames occurring in each rhetorical move. Table 2 displays the top 5 most frequent p-frames identified in each rhetorical move along with their raw frequency in each move. Table 3 shows the number and percentage of emails containing each rhetorical move and corresponding p-frames. Considering that all texts in the WRE corpus were request emails, it was not surprising that Move 3 (*making the inquiry*) was the most frequently used move. Move 3 also contained the

largest number of p-frames both by type and token. Move 2 (*purpose of email*) was not only the least frequently used move by writers, but also contained the lowest number of p-frames both by type (8%) and token (3%). While Move 3 and Move 2 showed a connection between rhetorical move frequency and p-frame occurrences, others displayed a somewhat different picture. Take Move 5 (*closing*) as an example, as it was ranked second in terms of move frequency (91%), yet the p-frames occurring in this move were placed second-to-last (23.1%). P-frames were only closely associated with certain rhetorical moves in the request emails.

Moves	P-frames (raw frequency)
Move 1 - Background	*in the process of * (12), we are in the * (10), are in the * of (9), by the * of the (8), the * of the year (8)*
Move 2 - Purpose of email	*wanted to follow up * (7), I just wanted to * (5), would like to * the (4), to follow up with * (4)*
Move 3 - Making the inquiry	*please let me know * (88), let me know what * (36), so that we can * (18), do you have any * (18), let * know what you (17)*
Move 4 - Supportive move	*we * be able to (12), will be able to * (9), to be able to * (6), it would be * to (6), we need to * this (6)*
Move 5 - Closing	*please let me know * (130), if you have any * (50), I look forward to * (27), know if you * any (20), look forward to * with (15)*

Table 2. Top 5 most frequent p-frames in each rhetorical move

Moves	Texts (%) N = 1,148	P-frame types (%) N = 376	P-frame tokens (%) N = 1,784
Move 1. Background	802 (70%)	102 (27.1%)	366 (20.5%)
Move 2. Purpose of email	106 (9%)	30 (8.0%)	53 (3.0%)
Move 3. Making the inquiry	1,148 (100%)	123 (32.7%)	806 (45.2%)
Move 4. Supportive move	539 (47%)	87 (23.1%)	241 (13.5%)
Move 5. Closing	1,041 (91%)	34 (9.0%)	318 (17.8%)

Table 3. The number and percentage of texts containing each move and p-frames

To understand how p-frames were used to fulfill rhetorical functions in the request emails, we computed the association strength between each p-frame and the move. Following Lu, Yoon & Kisselev (2021), we further classified p-frames into three categories based the association strength measure, including the specialized, semi-specialized, and non-specialized p-frames. As shown in Table 4, most p-frames were the non-specialized (46.1%), followed by the specialized (27.4%). Out of the 74 p-frames in Move 3, 30 p-frames fell into the specialized category. In Move 5, 13 out of 21 were the specialized p-frames. Move 3 and Move 5 displayed high formulaicity regarding the use of p-frames since almost a half of the total number of p-frames in the two moves fell into the specialized category. However, all p-

frames in Move 2 and Move 4 fell into either semi-specialized or non-specialized categories, suggesting that no p-frames were exclusively associated with both moves. Overall, our results indicated that some moves showed stronger associations with p-frames than others in the request emails.

Moves	Specialized	Semi-specialized	Non-specialized	Total
Move 1 - Background	5 (3.0%)	13 (7.8%)	24 (14.4%)	**42 (25.1%)**
Move 2 - Purpose of email	0 (0.0%)	2 (1.2%)	2 (1.2%)	**4 (2.4%)**
Move 3 - Making the inquiry	30 (18.0%)	18 (10.8%)	26 (15.6%)	**74 (44.3%)**
Move 4 - Supportive move	0 (0.0%)	4 (2.4%)	14 (8.4%)	**18 (10.8%)**
Move 5 - Closing	13 (7.8%)	5 (3.0%)	3 (1.8%)	**21 (12.6%)**
Multiple	0 (0.0%)	0 (0.0%)	8 (4.8%)	**8 (4.8%)**
Total	**48 (28.7%)**	**42 (25.1%)**	**77 (46.1%)**	**167 (100%)**

Table 4. The distribution of the specialized, semi-specialized, and non-specialized p-frames across rhetorical moves

4.2. The characteristics of p-frames in the request emails

The specialized p-frames were the most prominent among all three categories in the request emails. Since the specialized p-frames only occurred in a particular move, they could reveal the communicative function specific to a rhetorical move. The majority of p-frames in Move 5 (*closing*) were the specialized p-frames representing conventionalized expressions to conclude an email. Many writers expressed their expectations of continued communication (e.g., *look forward to * with*), delivered an appreciation to the recipient (e.g., *thanks in advance for **), or indicated their willingness to offer assistance as necessary (e.g., *do not hesitate to **). In (1a), *I look forward to ** was frequently used by writers to show their expectations for future replies or face-to-face communications with a variety of fillers such as *speaking, talking, hearing, working*, and *meeting*. The p-frame *please feel free to ** in (1b) was another common expression to conclude an email with an offer of potential assistance via a phone call (e.g., *please feel free to [call]*) or extra information (e.g., *please feel free to [check]*).

(1a) **I look forward to *** [hearing $_{(9)}$, speaking $_{(5)}$, talking $_{(3)}$, working $_{(2)}$, reviewing $_{(2)}$, meeting $_{(1)}$, seeing $_{(1)}$]

I look forward to [hearing] back from the appropriate individual(s) within your company to discuss this matter further.

I look forward to [speaking] with you.

(1b) **please feel free to *** [call (2), give (2), contact (2), check (1)]

> If there is any further information that I can provide, <u>please feel free to [call]</u> me on my cell phone at XXX-XXX-XXXX.

> If you need directions to any of the offices, <u>please feel free to [give]</u> me a call.

> In the meantime, <u>please feel free to [check]</u> out our website <u>at</u> <u>www.abcdefg.com</u>.

In Move 3 (*making the inquiry*), the specialized p-frames were frequently used to ask for information, opinion, or action. Many of the specialized p-frames contained consultative devices to mitigate potential risks for inappropriateness by using the politeness maker *please*, downtoners (e.g., *possible*), or hedges (e.g., *any*) to strengthen points and add to the action of formalizing the request (e.g., *is it possible to * [set, specify, add, bring, get]*, *do you have any * [idea, information, suggestions, tips, data]*, *can you please * the [update, give, provide, document, call, check, confirm]*). The utility of these mitigation devices in the p-frames was largely to decrease the degree of imposition possibly implied in the request emails, and thus to increase the likelihood of cooperation from the receiver. In addition, the specialized p-frames frequently appeared in the initial position of a sentence to directly initiate a new move. As shown in (2a-b), the two specialized p-frames *can you please * the* and *thanks in advance for ** were used at the beginning of a sentence to signal Move 3 (*making the inquiry*) and Move 5 (*closing*), respectively. These p-frames for signaling move shifts were similar to the so-called "triggers" (e.g., the use of "*the purpose of this study is to*" for the function of "announcing present research descriptively and/or purposefully") in Cortes (2013, p. 39).

(2a) **can you please * the** [call, check, confirm, document, give, provide, update] in Move 3

> <u>Can you please [call] the</u> number in the message and see if we can coordinate a 30 minute phone call for some time next week?

> <u>Can you please [confirm] the</u> time you expect to start this load testing & the expected duration?

> <u>Can you please [update] the</u> teams and aliases with the following information?

(2b) **thanks in advance for *** [your, all, working] in in Move 5

> Thanks in advance for [your] help with this.
>
> Thanks in advance for [all] of your help and please let me know if you have any questions.
>
> Thanks in advance for [working] to prioritize this date.

Another prominent feature for the specialized p-frames in Move 3 (*making the inquiry*) was the frequent use of the *let* imperative when making requests (e.g., *let me know * any*, *let me know your **). When examining the concordance lines of the *let*-embedded specialized p-frames, we found that they were always proceeded by the politeness marker *please* or consultative modal devices (e.g., *could* or *would*). For example, *let * know if this* (12 total occurrences) was headed by *please* (6 times), *could you* (2 times), and *can you* (2 times). There were only two occurrences when *let * know if this* was used in the initial position of a sentence to directly initiate the request.

Working professionals also used linguistic devices to minimize the degree of impositions or demands through p-frames in Move 3 (*making the inquiry*). One of the most prominent devices was the politeness marker *please*. Nine out of 30 specialized p-frames included *please* (e.g., *please let * know how*, *can you please send **, *could you please * me*). Another noticeable feature of the specialized p-frames was embedding, particularly with the embedded *if* clause (e.g., *let * know if this*, *if you could * me*). The use of a hypothetical *if*-clause provides the addressee with the possibility to question the proposition followed or the chance to decide whether to comply with the request. Notably, when *if*-embedded p-frames occurred in Move 3, they were frequently used along with lexical modifiers. For example, when the email writers asked about information, they used the p-frame *if you could * me* together with gratitude expressions (e.g., *I would appreciate* in (3a)) or subjectivizers (e.g., *I was wondering* shown in (3b)). The combined use of *if*-clauses and lexical modifiers enhanced the mitigation effect on the imposition of requests facilitating perceived politeness.

if you could * me [give, let]

(3a) I would appreciate <u>if you could [let] me know</u> of the results of this test.

(3b) Have you had a chance to test it yet? I was wondering <u>if you could [let] me</u> know how it is going.

The semi-specialized frames were the least frequently used p-frames in the WRE corpus, accounting for 25.1%. Although they appeared in more than

one move, they displayed strong association with one specific move. For instance, *we need to * the* occurred in both Move 1 (*background*) and Move 3 (*making the inquiry*). Primarily, it was associated with Move 3 given that 12 of the 15 total occurrences were found in this move. When the p-frame *we need to * the* was used in Move 1, it was to provide a reason for the subsequent request as shown in (4a). In Move 3, writers used *we need to * the* to direct the recipient to carry out desired actions as shown in (4b).

> **we need to * the** [add, capture, compile, discuss, do, extend, finalize, have, know, move, provide, reschedule, run]
>
> (4a) We are getting this because <u>we need to [provide] the</u> lang setting in the connection string. [Move 1 - Background] Please let us know how to set the language and send the parameter in the connection string. [Move 3 - Making the inquiry]
>
> (4b) Please let me know if <u>we need to [extend] the</u> same for other devices also. [Move 3 - Making the inquiry]

In Move 2 (*purpose of email*), we identified only four p-frames in the corpus with two falling into the semi-specialized category, (i.e., *I just wanted to ** and *wanted to follow up **). The use of past tense in both frames and the downtoner *just* implied that Move 2 was often expressed for mitigation purposes by writers. As all Move 2 appeared prior to Move 3 in our corpus, the mitigation effect expressed in Move 2 might be motivated by the need to draw attention to a given subject and entice cooperation (5):

> (5) **I just wanted to** * [drop, bring, follow, take]
>
> I <u>just wanted to [drop]</u> you a quick note to say that we are still interested in signing an agreement with you. [Move 2 - Purpose of email] Can you clear up the question I asked in my last email regarding section 1? [Move 3 - Making the inquiry]

The non-specialized p-frames accounted for the largest proportion (46.1%) of all frames. For instance, *would like to * the* appeared 16 times across the four moves. Although it occurred more frequently in Move 1 (*background*) than any other move (7 out of 16), it did not directly lead to the interpretation of Move 1. As shown in (6a-c), the p-frame *would like to * the* could be used to introduce a situational background in Move 1 (6a), to express a request in Move 3 (6b), or to provide additional information for the request in Move 4 (6c).

> **would like to * the** [arrange, ask, attend, confirm, do, establish, get, have, know, link, meet, move, purpose, restrict, schedule, see]

> (6a) Sanjay at XXX Trade is the one who will be manipulating our report data. He <u>would like to [link] the</u> report subjects below with what he should be looking for in the database. [Move 1 - Background]

> (6b) If you are interested in, we <u>would like to [arrange] the</u> meeting with you next week. [Move 3 - Making the inquiry]

> (6c) Pat has graciously offered to take the lead on this while I'm out of the country. I <u>would like to [get] the</u> business terms nailed down before I go. [Move 4 - Supportive move]

Finally, Move 1 (*background*) and Move 4 (*supportive move*) had more than half of the p-frames belonging to the non-specialized category (24 out of 42, and 14 out of 18, respectively). A careful analysis of the non-specialized p-frames in Move 1 and Move 4 suggested that there was a large proportion of shared p-frames in the two moves. This was not surprising given that the two moves were mainly information-oriented. When looking at the overlapping frames in both moves, we found that they mainly served the function of introducing a focus or a topic. Most of these p-frames were used to express a desire (e.g., *I would be * to*, *we need to * this*) or intention (e.g., *in order to * the*, *so that we can **) with various modal verbs incorporated, such as *can*, *could*, *might*, *would*, *will* and *need to*. Regarding the reference to the requestee in these p-frames for desire or intention, the email writers almost unanimously chose the speaker-oriented perspective through the first-person pronoun *I* (e.g., *<u>I</u> think we * to*, *<u>I</u> need to * a*), the speaker- and hearer-oriented perspective through the first-person pronoun *we* (e.g., *<u>we</u> need to get **, *<u>we</u> need to make **), or the impersonal perspective (e.g., *<u>it</u> would be * to*).

5. Discussion

Given that previous studies have mostly focused on the connection between rhetorical moves and formulaic expressions within academic genres (e.g., conference abstracts, research articles, grant proposals), the current study is one of the few endeavors to probe such connections in a workplace genre. We investigated the use of p-frames in relation to the rhetorical moves in workplace request emails. To summarize, our results revealed that Move 3 (*making the inquiry*) showed the largest number of p-frames, which was also

the most frequently used rhetorical move in our corpus. There were cases in which Move 3 occurred more than once within a single email. The writers, at times, made several requests that required separate tasks in one email message. The move that occurred the least frequently was Move 2 (*purpose of email*), which also had the smallest number of p-frames. Overall, the results showed that the rhetorical moves in request emails relied on p-frames, yet the degree of such reliance differed across the moves. Generally, the more frequent a move was, the greater the number of p-frames it contained. The seldom use of the move *purpose of email* is also consistent with the findings of Park et al. (2021), showing that working professionals oftentimes do not prefer to specify the general purpose at the beginning of an email. Notably, emails in the business workplace (as in our corpus) are often connected with previous emails or face-to-face communications. Given the intertextual nature of workplace emails, specifically mentioning the purpose at the beginning of an email might be regarded as redundant or unnecessary. In addition, the fast-paced work culture and shared interests within an organization might prevent some professionals from mentioning the purpose of an email.

Our results are also in line with previous findings on formulaic sequences (e.g., Casal & Kessler, 2020; Lu, Yoon & Kisselev, 2021; Yoon & Casal, 2020), showing that p-frames displayed different degrees of associations across rhetorical functions in email writing. No p-frame was found to be tied solely to the move of *purpose of email* or *supportive move*. In contrast, the moves of *making the inquiry* and *closing* displayed relatively strong associations with p-frames, as the two moves had a sizable proportion of p-frames solely occurring in them.

Consistent with prior studies on business genres (Evans, 2014; Handford, 2010; Xia, Ai & Pae, 2022), this study also reaffirms the prevalence of high formality in the language used in business correspondence. The identified p-frames showing strong associations with the rhetorical moves have the capacity to contribute to our understanding of the communicative needs and the style of specific rhetorical functions in workplace request emails. The specialized p-frames were prominent in the moves of *making the inquiry* and *closing*, which indicated that writers used highly formulaic language to pose their inquiries or to signal the ending of an email. Such preference could be partially explained by their busy schedules and conventionalized email writing in a real-life workplace. As a popular workplace genre, emails are used not only to deliver quick messages, but also to preserve a record of

communication. The ultimate aim of a request email is to have the requestee perform the request adequately. Thus, using clear and formulaic expressions to make a request can save time and effort for both parties involved in email correspondence.

This study also demonstrated that highly frequent multi-word sequences were useful linguistic devices to display genre-specific characteristics. The results showed that writers used a wide range of mitigation strategies to minimize any potential impact of their emails through careful linguistic choices. This was shown in the syntactic and lexical mitigation devices embedded in p-frames, especially in the move of *making the inquiry*, including politeness markers (e.g., *could you please * me*), embedded *if*-clauses (e.g., *let * know if this*), progressive aspect (e.g., *I was wondering if *), downtoners (e.g., *is it possible to *), and hedges (e.g., *do you have any *). Since p-frames are frequently occurring multi-word sequences in the target genre, the prevalence of mitigation devices in these frames serve to soften the tone and thus assure compliance through greater politeness.

The findings of this study also have important pedagogical implications. In particular, since the p-frames identified in this study were extracted from real-world business workplace emails, they can thus be regarded as representing authentic and practical examples that can be leveraged for business English courses. Besides, to successfully participate in workplace communication, business English learners need to develop a strong command of rhetorical conventions and linguistic knowledge of the target genre. Thus, we believe that the p-frame list categorized by rhetorical moves in this article can be used as a pedagogical tool[3].

6. Conclusion

This study adopted a corpus-based genre analysis approach to analyze the recurrent phraseological patterns in the genre of request emails, and specifically, the rhetorical goals that working professionals employ to realize those requests. Following a systematic investigation of the use of p-frames across rhetorical moves, we found that p-frames were embedded within a variety of mitigation devices to express politeness in business workplace communications. The current study offers empirical support for the utilization of formulaic language in email communication and posits that p-frames serve as a highly effective framework for characterizing the

prevalence of formulaic expressions in email. As email communications constitute a subset of a broader genre of business correspondence, our contention is that formulaic sequences, exemplified by p-frames, construct potent lexico-grammatical building blocks for a business genre. We also suggest that greater emphasis and scrutiny be devoted to these formulaic expressions in future research concerning business correspondence analyses.

Although the current study provided meaningful resources that are useful for researchers and practitioners, it is not without limitations. Firstly, it is possible that individual cultural characteristics or variances are displayed in email communications, particularly in terms of individuals' formality, promptness, and interpersonal relationships (Holtbrügge et al., 2013). Considering that our emails were collected exclusively from a single country, they may be reflective of a homogeneous cultural aspect. An investigation of the presence or absence of cultural differences in email or cultural considerations in business email in terms of language and tone, opening and closing, and directness and indirectness will allow us to better identify cross-cultural implications for business email writing and teaching.

Secondly, given that the length of p-frames might impact the phraseological patterns captured, our focus on five-word frames in this study might only partially reveal the phraseological characteristics of request emails. Thus, we suggest that future researchers consider comparing the use of p-frames of varying lengths (e.g., three-word, four-word, and six-word) to reveal a fuller description of the phraseological units in business discourse.

Lastly, p-frames were only matched to the move-level in this study. This decision was made primarily because we chose to work from an existing move framework, which was developed by Park et al. (2021). However, there also exists a lower-level analytical unit of a move (i.e., a step), which could be used to describe more fine-grained segments that writers use to construct a move (for more, see Casal & Kessler, 2024). For example, for Move 3 (*making the inquiry*) in this study, we identified lower-level steps such as *asking for information*, *asking for action*, and *asking for opinions*. Thus, future studies can adopt a bottom-up approach to understand the rhetorical moves and steps of workplace request emails, and then analyze how p-frames are used at both the move and step levels. Such studies will contribute to a more comprehensive understanding of p-frames in request emails, which can result in additional implications that are beneficial for both EAP and ESP pedagogy.

Acknowledgements

References

Anthony, L. (2020). *AntConc (Version 3.5.9)* [computer software]. Waseda University. http://www.laurenceanthony.net/software

Baron, N. S. (1998). Letters by phone or speech by other means: The linguistics of email. *Language & Communication*, *18*(2), 133-170.

Baron, N. S. (2002). Who sets e-mail style? Prescriptivism, coping strategies, and democratizing communication access. *Information Society*, *18*(5), 403-413. https://doi.org/10.1080/01972240290108203

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371-405. https://doi.org/10.1093/applin/25.3.371

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

Carrió-Pastor, M. L., & Muñiz-Calderón, R. (2013). Variation of English business e-mails in Asian countries. *Ibérica, Journal of the European Association of Languages for Specific Purposes*, *26*, 55-76. https://www.revistaiberica.org/index.php/iberica/article/view/273

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

Casal, J. E., & Kessler, M. (2024). Rhetorical move-step analysis. In M. Kessler & C. Polio (Eds.), *Conducting genre-based research in applied linguistics: A methodological guide* (pp. 82-104). Routledge.

Cortes, V. (2013). *The purpose of this study is to*: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, *12*(1), 33-43. https://doi.org/10.1016/j.jeap.2012.11.002

Cortes, V. (2024). Lexical bundles and phrase frames. In M. Kessler & C. Polio (Eds.), *Conducting genre-based research in applied linguistics: A methodological guide* (pp. 105-126). Routledge.

Evans, S. (2014). Teaching business correspondence: Lessons from the globalised workplace. *The Asian Journal of Applied Linguistics*, *1*(2), 102-120. https://caes.hku.hk/ajal/index.php/ajal/article/view/47

Fletcher, W. H. (2012). *KfNgram*. USNA.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes*, *24*(3), 321-332. https://doi.org/10.1016/j.esp.2004.09.002

Forsyth, R. S., & Grabowski, L. (2015). Is there a formula for formulaic language? *Poznan Studies in*

Contemporary Linguistics, *51*(4), 511-549. https://doi.org/10.1515/psicl-2015-0019

Garner, J., Crossley, S., & Kyle, K. (2019). N-gram measures and L2 writing proficiency. *System*, *80*(2), 176-187. https://doi.org/10.1016/j.system.2018.12.001

Ghadessy, M., & Webster, J. (1988). Form and function in English business letters: Implications for computer-based learning. In M. Ghadessy (Ed.), *Registers of written English: Situational factors and linguistic features* (pp. 110-127). UNKNO.

Giménez-Moreno, R., & Skorczynska, H. (2013). Business communication across three European cultures: A contrastive analysis of British, Spanish and Polish email writing. *Ibérica, Journal of the European Association of Languages for Specific Purposes*, *26*, 77-98. https://revistaiberica.org/index.php/iberica/article/view/274

Golparvar, S. E., & Barabadi, E. (2020). Key phrase frames in the discussion section of research articles of higher education. *Lingua, 236*, 102804. https://doi.org/10.1016/j.lingua.2020.102804

Granger, S. (2018). Formulaic sequences in learner corpora: Collocations and lexical bundles. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 228-247). Routledge.

Handford, M. (2010). *The language of business meetings*. Cambridge University Press.

Ho, V. (2018). Using metadiscourse in making persuasive attempts through workplace request emails. *Journal of Pragmatics*, *134*, 70-81. https://doi.org/10.1016/j.pragma.2018.06.015

Holtbrügge, D., Weldon, A., & Rogers, H. (2013). Cultural determinants of email communication styles. *International Journal of Cross Cultural Management*, *13*(1), 89-110. https://doi.org/10.1177/1470595812452638

Incelli, E. (2013). Managing discourse in intercultural business email interactions: A case study of a British and Italian business transaction. *Journal of Multilingual and Multicultural Development*, *34*(6), 515-532. https://doi.org/10.1080/01434632.2013.807270

Kankaanranta, A. (2006). "Hej Seppo, could you pls comment on this!"—internal email communication in lingua franca English in a multinational company. *Business Communication Quarterly*, *69*(2), 216-225. https://doi.org/10.1515/multi-2017-0077

Kankaanranta, A., Karhunen, P., & Louhiala-

Salminen, L. (2018). "English As Corporate Language" in the multilingual reality of multinational companies. *Multilingua*, *37*(4), 331-351. https://doi.org/10.1515/multi-2017-0077

Kessler, M. (2020). A text analysis and gatekeepers' perspectives of a promotional genre: Understanding the rhetoric of Fulbright grant statements. *English for Specific Purposes*, *60*, 182-192. https://doi.org/10.1016/j.esp.2020.07.003

Kim, S., & Kessler, M. (2022). Examining L2 English University students' uses of lexical bundles and their relationship to writing quality. *Assessing Writing*, *51*, 100589. https://doi.org/10.1016/j.asw.2021.100589

Leopold, L. (2015). Request strategies in professional e-mail correspondence: insights from the United States workplace. *TESL Canada Journal*, *32*(2), 1-29. https://doi.org/10.18806/tesl.v32i2.1205

Li, Z., & Volkov, A. (2017). "To whom it may concern": A study on the use of lexical bundles in email writing tasks in an English proficiency test. *TESL Canada Journal*, *34*(3), 54-75. https://doi.org/10.18806/tesl.v34i3.1273

Lu, X., Casal, J. E., & Liu, Y. (2020). The rhetorical functions of syntactically complex sentences in social science research article introductions. *Journal of English for Academic Purposes*, *44*, 100832. https://doi.org/10.1016/j.jeap.2019.100832

Lu, X., Casal, J. E., & Liu, Y. (2021). Towards the synergy of genre- and corpus-based approaches to academic writing research and pedagogy. *International Journal of Computer-Assisted Language Learning and Teaching*, *11*(1), 59-71. https://doi.org/10.4018/IJCALLT.2021010104

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

Millot, P. (2017). Inclusivity and exclusivity in English as a Business Lingua Franca: The expression of a professional voice in email communication. *English for Specific Purposes*, *46*, 59-71. https://doi.org/10.1016/j.esp.2016.12.001

Moreno, A. I., & Swales, J. M. (2018). Strengthening move analysis methodology

towards bridging the function-form gap. *English for Specific Purposes*, *50*, 40-63. https://doi.org/10.1016/j.esp.2017.11.006

Nguyen, H., & Miller, J. (2012). Exploring business request genres. *Business Communication Quarterly*, *75*(1), 5-28. https://doi.org/10.1177/1080569911430379

Park, S., Jeon, J., & Shim, E. (2021). Exploring request emails in English for business purposes: A move analysis. *English for Specific Purposes*, *63*, 137-150. https://doi.org/10.1016/j.esp.2021.03.006

Qian, D. D., & Pan, M. (2019). Politeness in business communication: Investigating English modal sequences in Chinese learners' letter writing. *RELC Journal*, *50*(1), 20-36. https://doi.org/10.1177/0033688217730142

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

Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge University Press.

Van Herck, R., Decock, S., & Fastrich, B. (2022). A unique blend of interpersonal and transactional strategies in English email responses to customer complaints in a B2C setting: A move analysis. *English for Specific Purposes*, *65*, 30-48.

https://doi.org/10.1016/j.esp.2021.08.001

Warren, M. (2016). Signalling intertextuality in business emails. *English for Specific Purposes*, *42*, 26-37. https://doi.org/10.1016/j.esp.2015.11.001

Xia, D., Ai, H., & Pae, H. K. (2022). "Please let me know" Lexical bundles in business emails by business English learners and working professionals. *International Journal of Learner Corpus Research*, *8*(1), 1-30. https://doi.org/10.1075/ijlcr.20019.xia

Xia, D., Chen, Y., & Pae, H. (2023). Lexical and grammatical collocations in beginning and intermediate L2 argumentative essays: a bigram study. *International Review of Applied Linguistics in Language Teaching*, *61*(4), 1421-1453. https://doi.org/10.1515/iral-2021-0188

Xia, D., Sulzer, M., & Pae, H. (2023). Phrase-frames in business emails: a contrast between learners of business English and working professionals. *Text & Talk*. https://doi.org/10.1515/text-2022-0027

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, registers and writer expertise* (pp. 281-306). John Benjamins.

Detong Xia is an Assistant Professor in the Department of Applied Linguistics at Xi'an Jiaotong-Liverpool University. Her research focuses on corpus linguistics, L2 writing and English for Specific Purposes. Her work appears in research journals such as *International Journal of Learner Corpus Research*, *International Review of Applied Linguistics in Language Teaching* and *Text & Talk*.

Matt Kessler is an Assistant Professor of Applied Linguistics at the University of South Florida. His research explores issues pertaining to L2 writing, genre-based teaching and learning, and computer-assisted language learning. He is the co-editor of the volume, *Conducting genre-based research in applied linguistics: A methodological guide* (Routledge, 2024), and the author of the book *Digital multimodal composing: Connecting theory, research and practice in second language acquisition* (Multilingual Matters, 2024).

Yudi Chen is an Associate Professor in the College of Economics and

Management at Nanjing University of Aeronautics and Astronautics. He specializes in machine learning and social media mining.

Hye K. Pae, Ph.D., is a Professor in the Literacy and Second Language Studies program at the University of Cincinnati, USA. Her research centers on psycholinguistics, second language acquisition, and assessment. She has published two authored books with Springer, one edited book with John Benjamins, and about 100 publications including journal articles and book chapters.

NOTES

[1] Detailed information of the Avocado Research Email Collection is available at: https://catalog.ldc.upenn.edu/LDC2015T03.

[2] The python script is available at: https://github.com/xdt1991/Data-for-ESP-paper/blob/main/Python%20code%20for%20the%20ESP%20paper.

[3] The full list of p-frames sorted by rhetorical function is available at: https://github.com/xdt1991/Data-for-ESP-paper/blob/main/full%20list_phrase%20frames.csv.