Jason Miin-Hwa Lim*  

Centre for the Promotion of Knowledge and Language Learning  
University of Malaysia Sabah, Malaysia  
jasonlim@ums.edu.my

Xianqing Luo  

School of Foreign Languages  
Jiangsu University of Technology, People’s Republic of China  
lxq@jsut.edu.cn

WRITING RESEARCH QUESTIONS AND HYPOTHESES: A GENRE-BASED INVESTIGATION INTO WRITERS’ LINGUISTIC RESOURCES IN SOCIAL SCIENCES

Abstract

Novice writers and university students often encounter problems with regard to (i) the extent to which they should include research questions and hypotheses in their research reports, and (ii) how they can effectively present these information elements to enhance clarity. To resolve these problems, we used a genre-based analytical framework and inputs provided by specialist informants to ascertain the prevalence of these information elements and the range of linguistic resources employed by writers in the research article introductions (RAIs) in two social sciences, namely Ethnic Studies (ES) and Industrial Relations (IR). We have found that most RAIs in IR incorporate research questions and hypotheses, while less than half of those in ES include them. The difference is attributable to writers’ greater tendency to employ epistemological methodologies in IR, and ontological methodologies in ES. Polar questions are rarely used in both disciplines, while factual, evaluative and circumstantial wh-questions are far more prevalent. Active anticipatory verbs and that-clauses are used to postulate hypotheses in both disciplines, whereas first-person pronouns are far more common in ES than in IR. Our findings can be flexibly adapted to design instructional materials aimed at guiding novice writers in the formulation of research questions and hypotheses.

Key words

English for academic purposes, genre-based analysis, research questions, hypotheses, linguistic resources.

* Corresponding address: Jason Miin-Hwa Lim, Centre for the Promotion of Knowledge and Language Learning, University of Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia.
INTRODUCTION

Academic scholars and postgraduate students across the world are often expected to write and publish research articles in a quest to achieve professional advancement. In fact, it is widely acknowledged that “the importance of the research article to the creation of academic knowledge cannot be overestimated” (Hyland, 2005: 89), especially as we consider how scholars are expected to use the research genre to advance knowledge in their respective fields of study. In particular, in the writing of research articles in social sciences, writers are generally expected to formulate research questions and/or hypotheses, which have been considered as “directional determinants” (Feak & Swales, 2011: 112; Lim, 2018: 2) that guide their entire research. Notwithstanding the importance of such directional determinants, novice writers and university students often encounter difficulties in formulating proper research questions (RQs) and hypotheses which can enlighten readers about the direction in which their reports will develop.

‘Presenting research questions or hypotheses’ constitutes a newly added rhetorical step in Swales’s (2004) Create a Research Space (CaRS) model, and his inclusion of this step in Move 3 (i.e. ‘presenting the present work’) has resulted in a greater emphasis on these information elements in recent research (Lim, 2014a). These two sub-steps have also formed the focus of numerous recent studies (e.g. Lim, Loi, & Hashim, 2014; Loi & Evans, 2010; Martín & Pérez, 2014; Ozturk, 2007; Sheldon, 2011; Soler-Monreal, Carbonell-Olivares, & Gil-Salom, 2011). Despite the emphasis, research questions and hypotheses do not seem to be equally important in research article introductions (RAIs) written in different languages. For example, research questions are used in more than a third (35%) of the English RAIs in some fields, such as Educational Psychology, but only 10% of Chinese RAIs (Loi & Evans, 2010). Likewise, it was reported that hypotheses are found in 75% of the English RAIs in Educational Psychology, but they are completely absent in Chinese introductions in the same discipline (Loi, 2010). Such findings seem to indicate that research questions and hypotheses play a relatively more important role in English research introductions.

Considering the use of research questions alone, we notice that the prevalence of research questions varies considerably across disciplines. For instance, only 7% of the RAIs in Computer Science include research questions (Shehzad, 2010), but over a third (35%) of the Educational Psychology research articles incorporate such questions (Loi & Evans, 2010). Likewise, the prevalence of research hypotheses may also differ markedly across different fields of study. Research hypotheses appear in 15% of the Political Philosophy RAIs, but they are incorporated in 30% of the Political Science RAIs, 45% of the Dermatological RAIs, and even 65% of the Clinical and Health Psychology RAIs (Martín & Pérez, 2014). Although these previous studies have indicated that the percentages of texts containing research questions and hypotheses do differ across various disciplines,
we are uncertain about the degrees of prevalence of these elements in some social science disciplines such as Ethnic Studies (ES) and Industrial Relations (IR). Neither are we definite about the rationale leading to the different degrees of prevalence of research questions and hypotheses. In this regard, this study looks into the frequencies and percentages of texts containing these two information elements and explores the reasons relating to the higher and lower degrees of prevalence of these information elements using a genre-based move analysis and inputs provided by specialist informants.

While some previous researchers have reported interesting data on the proportions of texts containing the two directional determinants mentioned above, only some research (i.e. Lim, 2014a; Lim et al., 2014), to our knowledge, has looked into the salient language resources used in research questions and hypotheses. It was reported that a wide spectrum of language resources are adopted in formulating research questions although the average length of a research question is 24.2 words (Lim, 2014a). Polar questions, which begin with the present-tense operators, including ‘do/does’ and ‘is/are’, occur in 69.2% of the experimental introductory chapters, while 42.3% of the experimental introductory chapters include wh-questions, which may be factual, evaluative, and circumstantial questions that begin with ‘what’, ‘to what extent’, and ‘how’, respectively (Lim, 2014a). In postulating hypotheses in experimental doctoral dissertations on Applied Linguistics, it was found that 62.5% of doctoral dissertations use the simple past and simple present to present research hypotheses, while 50% of them employ the simple future in postulating hypotheses (Lim et al., 2014). The aforementioned studies have indicated that disciplinary variations do occur with regard to the proportions of research articles which incorporate research questions and hypotheses, thus suggesting that more research needs to be conducted to ascertain if the salient language resources reported in the few past studies so far actually occur in other social science disciplines.

In response to the lacuna mentioned above, we have designed an investigation that (i) ascertains the extent to which research questions and hypotheses are incorporated in journal articles in ES and IR, and (ii) explores the salient linguistic resources used to formulate the research questions and hypotheses concerned. Based on Swales’s (1990, 2004) genre-based analytical framework and Lim’s (2011, 2014b) descriptions of recurring linguistic resources in rhetorical steps, this study is designed to answer the following research questions:

(1) To what extent are research questions and hypotheses incorporated in research article introductions on ES and IR?

(2) How do writers in ES and IR use linguistic resources to formulate research questions and hypotheses in their research article introductions?

A combination of both quantitative and qualitative research approaches was used in this study. Quantitative data were collected to answer the first research question about the frequencies and degrees of prevalence of the two information
elements. The second research question sought largely qualitative data to reveal the salient linguistic resources employed to formulate research questions and hypotheses that meet the conventions accepted in their discourse communities. Apart from this, additional spoken data were elicited from specialist informants to explain the circumstances under which the two directional determinants are used in each discipline.

2. RESEARCH METHODS

The following sections report the procedures for collecting articles, analyzing data, and seeking information from specialist informants.

2.1. Data collection procedure

A total of 60 research articles, consisting of 30 in Ethnic Studies and 30 in Industrial Relations, were selected purposively from six Quartile 1 or Quartile 2 journals indexed in the Web of Science (WoS). Four aspects were considered when selecting journals for this study: (i) the prestige of the journals in each discipline, (ii) the numbers of journals, (iii) topic coverage of each field, and (iv) the recency of the journals, given that the “latest journal(s)” need to be chosen (Swales, 1990: 161). Having considered the basic principles in data collection, three WoS-indexed journals were selected from the discipline of ES, and they included Ethnicities, Ethnic and Racial Studies, and Ethnicity and Health. As for IR, the three selected WoS-indexed journals were Industrial Relations, Journal of Industrial Relations, and European Journal of Industrial Relations. In terms of representativity, having three established journals for each discipline was considered as an acceptable practice in genre-based research (Wang & Yang, 2015).

After the six representative journals were selected from the two disciplines, the research articles were subsequently chosen from each of the reputed journals. In the first round, 35 most recently published research articles (RAs) were selected from each journal in the two disciplines, thus resulting in a total of 105 RAs in each field. This means that a pool of 210 RAs (105 ES RAs and 105 IR RAs) was compiled in the first round. The next stage was to calculate the percentage of empirical data-based RAs among these 105 RAs in each discipline to check whether the empirical data-based RAs exhibited common practices in disseminating the academic knowledge in each discipline. The empirical data-based RAs in both disciplines were found to constitute over 70% of the all the RAs, thus implying that empirical data-based RAs were worthy of being studied. The third stage was carried out to select 10 most recently published empirical data-based RAs from each journal at the commencement of the study. A total of 30 RAs
were eventually selected in each field to build a corpus of 60 RAs covering both disciplines.

2.2. Data analysis procedure

Swales’s (1990, 2004) Create a Research Space (CaRS) model was used for analyzing rhetorical steps of presenting research questions and hypotheses in RAs in ES and IR. In addition, Wong and Lim’s (2014) procedure for describing recurring linguistic resources was used to analyse ‘presenting research questions and hypotheses’ in the RAs. Initially guided by Swales’s (2004), Create a Research Space (CaRS) model, we came up with three moves. While Move 1 deals with reviews of previous research with a focus on centrality claims, Move 2 concerns indications of gaps in past research that highlight the need for more research. In Move 3, known as ‘presenting the present work’, writers employ seven steps, one of which was ‘presenting research questions or hypotheses’. Assigning the aforementioned labels to all the steps requires us to conduct a coding process to clearly differentiate research questions and hypotheses from other rhetorical steps. Two coders, being the researchers in this study, independently assigned labels to each text segment. The reliability of the coding was measured in terms of “percent agreement”, defined as “the number of agreements per total number of coding decisions” (Biber, Connor, & Upton, 2007: 35). In this research context, the number of agreements was defined as the number of coding decisions (made by Coders 1 and 2) that were the same in each dataset comprising 30 texts. In Dataset A (ES RAs), 432 coding decisions which were made by Coder 1 were the same as 432 coding decisions made by Coder 2, and hence the total number of identical coding decisions made by both coders was 864. When both identical and different coding decisions were considered, 459 coding decisions were made by Coder 1, while 466 decisions were made by Coder 2, and thus the total number of coding decisions was 925 in ES. A similar process was conducted for Dataset B (IR RAs). The calculations for each discipline are indicated as follows:

Dataset A:  
\[
\frac{864 \text{ (i.e. } 432 + 432 \text{) coding decisions}}{459 \text{ coding decisions by Coder 1} + 466 \text{ coding decisions by Coder 2}} \times 100\% = 93.4\%
\]

Dataset B:  
\[
\frac{1224 \text{ (i.e. } 612 + 612 \text{) coding decisions}}{642 \text{ coding decisions by Coder 1} + 673 \text{ coding decisions by Coder 2}} \times 100\% = 93.1\%
\]

In this study, the inter-coder agreement in the first round was 93.4% in ES and 93.1% in IR, thus meeting the requirement that inter-coder reliability should be above 90.0% (Miles & Huberman, 1994). Both coders “discussed their segmentation and labelling” so as to iron out the discrepancies emerging from the
independent coding decisions explained above (Moreno & Swales, 2018: 51). The revision of the labels was then finalized, resulting in a complete inter-coder agreement in the second round.

The analysis of the linguistic resources involved three steps. First, using a word processor, the researchers recorded all text segments for ‘formulating research questions’ and ‘postulating hypotheses’ in two main lists representing the two disciplines. Second, every main list was manually divided into different sub-lists, each of which contained complete clauses that performed the same communicative function (such as a function performed in a factual question or a circumstantial question). Third, repeated attempts were made to ensure that each sub-list contained instances with (i) comparable and recurrent phrases, and/or (ii) similar and recurrent lexico-grammatical structures. This means that the linguistic resources were considered as sufficiently salient on the condition that they (i) appeared in two or more of the 30 research introductions in each discipline, and (ii) performed the same communicative function concerned (Lim, 2017, 2019).

2.3. Procedure for eliciting information from specialist informants

After we had completed the genre-based discourse analysis of rhetorical structure and salient linguistic resources used to realize the communicative functions in Ethnic Studies and Industrial Relations RAIs, face-to-face interviews were conducted with our specialist informants to “triangulate the discourse analysis” (Berkenkotter & Huckin, 1995: 13). This study adopted Bhatia’s (1993) principles for choosing specialist informants given that each specialist informant included was (i) a practicing member of the disciplinary culture in which the genre was routinely used, and (ii) an experienced member who could confirm the analyst’s findings, bring validity to his or her insights, and add psychological reality to the analysis. The interviews involved six specialist informants, with three coming from each discipline. All the informants interviewed in this study were (i) either full or associate professors with doctorates in their respective disciplines, and (ii) had published papers in high-impact journals indexed in the Web of Science (WoS) and Scopus databases. The six specialist informants (SIA to SIF) served at five established universities, three of which were listed in the top 200 universities according to the Quacquarelli Symonds (QS) World University Rankings. In ES, SIA was the Chairperson of the Centre for Policy Studies, while SIB was the Principal Research Fellow at an Institute of Ethnic Studies, and SIC was the Vice President for Academic and Internationalization at the university concerned. In IR, while SID was the Deputy Vice-Chancellor for Academic and International Affairs, SIE was an Editorial Board Member of a WoS-indexed journal, and SIF was the Director of Industrial Linkages and Collaboration Offices. Email messages were written to
invite the informants to participate in semi-structured face-to-face interviews to be conducted in their respective offices.

Following Bhatia’s (1993) suggestions, the interviews were conducted on the basis of the information derived from the results of our initial textual analysis of the frequencies of the rhetorical step. The two sets of questions posed to the informants in the two disciplines are shown as follows:

(i) It has been found that nearly half of the research article introductions in ES incorporate research questions or hypotheses of their studies. Over half of the research article introductions in IR incorporate research questions or hypotheses of their studies. Under what circumstances do you use research questions or under what circumstances do you include hypotheses?

(ii) (a) Why is it that writers in ES use first person pronouns (e.g. ‘I’, ‘we’) to formulate research hypotheses?
    (b) Why is it that writers in IR use first person pronouns to formulate research hypotheses?

The interview questions in this study were only “semi-structured” in the sense that the researcher could rephrase the questions in the process of posing the questions based on the responses provided by the informants (Cohen, Manion, & Morrison, 2000: 146). Each interview was recorded, transcribed and further verified by the informants. The precision of the transcription was considered in three aspects. First, further clarification was sought from the specialist informants during the interview in cases where unclear and ambiguous meanings occurred. Second, notes were taken during the interview in order to serve as the guidelines and back-up for the transcriptions. Third, the transcriptions were done repeatedly to maximize precision, and verbatim transcriptions of the interviews were conducted immediately after each interview.

3. RESULTS AND DISCUSSION

In this section, the quantitative findings relating to the frequencies of research questions and hypotheses in the two social science disciplines will be reported first before attention is focused on the qualitative findings on the linguistic resources used to perform the communicative functions. Our specialist informants’ inputs will be reported if they are associated with each set of findings.
3.1. Analysis of frequencies and percentages of RAIs containing research questions and hypotheses

Our analysis has shown that ‘presenting research questions or hypotheses’ constitutes an important part of the RAIs given that this rhetorical step is included in 43.3% (13/30) of the ES RAIs and 56.7% (17/30) of the IR RAIs. A detailed frequency analysis has revealed that research questions alone are incorporated in 23.3% (7/30) of research articles in Ethnic Studies and 30.0% (9/30) in Industrial Relations, while hypotheses are included in 30.0% (9/30) of the ES RAIs and 36.7% (11/30) of the IR RAIs. We have also found that 10.0% (3/30) of the research articles in ES and IR incorporate both research questions and hypotheses. This means that a vast majority (90.0%) of the research introductions in both disciplines do not simultaneously include both research questions and hypotheses. A further analysis shows that the mean frequencies of research questions in ES and IR are 0.47 (14/30) and 0.63 (19/30) respectively, while the mean frequencies of research hypotheses in ES and IR are 0.67 (20/30) and 1.2 (36/30). The findings suggest that research questions are more likely to be recurring (in the same RAI) in IR RAIs compared to ES RAIs. It also means that on average researchers in IR are likely to incorporate twice as many hypotheses as those in ES.

With regard to Ethnic Studies, Specialist Informant A (SIA) explained that research questions and hypotheses should be made “precise”, “brief”, and “not lengthy” so as to minimize confusion. According to SIA, research questions would normally be formulated when it is a qualitative study, and hypotheses are generally postulated only in a quantitative study. Specialist Informant B (SIB) was of the view that “if our research is conceptual, qualitative or ethnographic, what we should do is to come out with a research question” rather than a hypothesis; however, a “hypothesis normally goes with quantitative research, (so) you set the hypothesis, verify it empirically, (and) then you confirm or disconfirm” it. Specialist Informant C (SIC) was of the view that “if it is exploratory”, she “would use research questions”, because she does not know what assumption she should make, but if she intends to find out “a correlation between (variable) A and (variable) B”, then she would “use a hypothesis to test” it. As she considered herself “a more qualitative person”, she acknowledged that she would usually use a research question, but she pointed out that “quantitative (researchers) tend to use more hypotheses”.

As for Industrial Relations, Specialist Informant D (SID) pointed out that after identifying the gap through literature review, researchers need to formulate a hypothesis for testing. Formulating hypotheses was considered as the “translation of research questions”. He pointed out that many investigations in Ethnic Studies tend to conduct ontological studies, and the researchers are just “accumulating information, mapping certain behaviour”, or attempting to “capture something”. To him, ES studies are “largely inductive” as researchers just collect information (to find out) “why a particular aborigine does that”. However, it was also mentioned
that numerous studies in IR are “epistemological” and “deductive” given that “you extract a hypothesis to test, (and) if it holds, you replicate the theory, but if it doesn’t hold, you reject the theory”. However, Specialist Informant E (SIE) only stated that he used “research questions almost always” and “never used hypotheses”, but pointed out that a “hypothesis has to be novel, plainly obvious and meaningful”. Specialist Informant F’s (SIF’s) view more closely resembles that given by SID, rather than SIE, as he commented that “research questions transfer (convey the meaning of) the purpose or objective, so research questions will guide you all the way around”. According to SIF, IR researchers often “use research questions in qualitative research, but hypotheses are used “in most quantitative research” in their discipline. While the specialist informants provided pertinent information on the context in which research questions and hypotheses are used, it is necessary to report (in the next section) our findings on how these directional determinants are performed using salient linguistic resources in the two disciplines.

3.2. Linguistic resources used in research questions and hypotheses in Ethnic Studies and Industrial Relations

Overall, two types of research questions, namely polar questions and wh-questions, are found in RAIs on Ethnic Studies and Industrial Relations. Polar questions are incorporated in only 6.7% (2/30) of RAIs in Ethnic Studies, and only 3.3% (1/30) in Industrial Relations (see Table 1). The main clause structure of polar questions appears in the form of ‘primary auxiliary verb (e.g. ‘do/does’) + subject + main verb’ structure (VpaSVm structure), which is employed to ask about a regular event (as in ‘Do higher investments in flexicurity policies and practices indeed result in stronger feelings of security?’). An alternative structure appears in the form of ‘modal auxiliary verb (e.g. ‘would’ or ‘may’) + subject + main verb’ structure (VmaSVm structure), which is intended to enquire about a possible action (as in ‘Would the Hijab Effect influence the perceived attractiveness of Muslim women?’). The findings suggest that polar questions are generally not the preferred interrogative structure among writers in both of these social science disciplines.

Wh-questions are comparatively far more prevalent than polar questions, so far as research questions are concerned in RAIs in both Ethnic Studies (23.3% or 7/30) and Industrial Relations (40.0% or 12/30), as shown in Table 1.
Wh-questions, as research questions, can be divided into (i) four categories in ES RAIs, and (ii) six categories in IR RAIs (see Table 1). First, factual wh-questions occur in a ‘subject-predicate-complement’ (SPC) interrogative structure, which involves (i) a wh-pronoun and a copular verb, or (ii) a wh-pronoun modifying a noun followed by a main verb (e.g. ‘What are the major challenges and barriers to the implementation of effective trade unions?’). These factual questions are found in 3.3% (1/30) of RAIs in Ethnic Studies and 10.0% (3/30) in Industrial Relations. Second, circumstantial wh-questions involve the use of the primary verb ‘do’ in a VpSaSVm structure to explore how a dependent variable varies due to the effect of an independent variable (e.g. ‘How does hijab and Islamic religiosity affect...')

| DISCIPLINE      | TYPE OF RESEARCH QUESTIONS | INSTANCE                                                                 | FREQ. OF RQS | NO. OF RAIS | % OF RAIS |
|-----------------|---------------------------|--------------------------------------------------------------------------|--------------|-------------|-----------|
| Ethnic Studies  | Polar questions           | Did different newspapers present Muslims’ opinion poll evidence as more or less negative? (RAES 5: 677) | 3            | 2           | 6.7       |
|                 | Factual wh-questions      | What remains of the association of ethnicity with occupation, in a specific service industry sector (global luxury hotel industry) allegedly representing Malaysian modernity? (RAES 1: 816) | 2            | 1           | 3.3       |
|                 | Circumstantial wh-questions | How does hijab and Islamic religiosity affect perceptions of attractiveness for Muslim women in the US and the UAE? (RAES 2: 746) | 4            | 3           | 10        |
|                 | Evaluative wh-questions   | To what extent was it the media reflecting their own preferred narrative to create headlines and sell their newspapers? (RAES 5: 676) | 2            | 1           | 3.3       |
|                 | Determinative wh-questions | Which strategies of boundary making are deployed by the ethnic and migrant groups in order to be recognised as a legitimate part of society? (RAES 10: 415) | 3            | 1           | 3.3       |
|                 | Factual wh-questions      | What factors motivate management to pursue a hybrid model? (RAIR 12: 687) | 5            | 3           | 10        |
|                 | Circumstantial wh-questions | How does minimum wage policy affect enrolments and expenditures on means-tested public assistance programs? (RAIR 2: 668) | 5            | 4           | 13.3      |
|                 | Evaluative wh-questions   | To what extent is the adoption of PRP in companies based in CEE and Western Europe influenced by the MNEs’ centralisation policies, i.e. headquarters determining pay and benefits in the subsidiary, on the one hand, and industrial relations, i.e. collective bargaining coverage and union influence, on the other? (RAIR 17: 292) | 3            | 1           | 3.3       |
|                 | Determinative wh-questions | Which set of FLSA amendments would be most beneficial for the more than 80 million non-exempt workers in the United States, and also perhaps, for some "exempt" workers? (RAIR 1: 719) | 2            | 1           | 3.3       |
|                 | Causative wh-questions    | Why and how did UWRs emerge? (RAIR 18: 273) | 2            | 2           | 6.7       |
|                 | Personal identification wh-questions | Who are the key GPN actors? (RAIR 14: 643) | 1            | 1           | 3.3       |

Table 1. Numbers and percentages of RAIs containing different types of research questions in ES and IR

\[\text{Wh}-\text{questions, as research questions, can be divided into (i) four categories in ES RAIs, and (ii) six categories in IR RAIs (see Table 1). First, factual wh-questions occur in a ‘subject-predicate-complement’ (SPC) interrogative structure, which involves (i) a wh-pronoun and a copular verb, or (ii) a wh-pronoun modifying a noun followed by a main verb (e.g. ‘What are the major challenges and barriers to the implementation of effective trade unions?’). These factual questions are found in 3.3\% (1/30) of RAIs in Ethnic Studies and 10.0\% (3/30) in Industrial Relations. Second, circumstantial wh-questions involve the use of the primary verb ‘do’ in a VpSaSVm structure to explore how a dependent variable varies due to the effect of an independent variable (e.g. ‘How does hijab and Islamic religiosity affect...’)}\]
perceptions of attractiveness for Muslim women in the US and the UAE?). Such circumstantial research questions appear in 10.0% (3/30) of the RAIs in Ethnic Studies and 13.3% (4/30) of those in Industrial Relations.

Third, evaluative *wh*-questions occur in *VpSVm* structures involving the primary operator ‘do’ or copular verb ‘be’, which is used to seek data that can ascertain the degree to which a phenomenon is prevalent in a given context (e.g. ‘To what extent did the British public see a picture of Muslim public opinion?’). These evaluative research questions have been identified in only 3.3% (1/30) of RAIs in both Ethnic Studies and Industrial Relations. Fourth, determinative *wh*-questions appear in SPC interrogative structures, which are employed to ask for specific information within the restricted range of answers (e.g. ‘Which ethnic minority is included in the Istrian imagined community and which segments of the ethnic and migrant community stand as negative reference groups?’). Such determinative *wh*-questions occur in only 3.3% (1/30) of the RAIs in both Ethnic Studies and Industrial Relations. Fifth, causative questions involve the use of the auxiliary verb ‘do’ in a *VpSVm* structure that focuses on finding out the reasons for an action (i.e. ‘Why do they respond in that way?’). These causative questions appear in 6.7% (2/30) of the RAIs in Industrial Relations, but not in Ethnic Studies. Sixth, a personal identification question appears in the subject-predicator-complement (SPC) interrogative structure intended to identify the actor(s) involved (i.e. ‘Who are the key GPN actors?’). A question in this category involves a personal identification pronoun ‘who’, but it is used in only 3.3% (1/30) of the Industrial Relations RAIs. Interestingly, causative and personal identification *wh*-questions do not appear in Ethnic Studies at all, thus suggesting that the academic discourse community of IR researchers actually expect a narrower range of conventional research questions, especially conventional polar and circumstantial research questions.

Aside from research questions, it is necessary to examine the linguistic resources employed in the hypotheses in the two disciplines. Hypotheses generally appear either after an announcement of the theoretical framework and literature review, and only occasionally after the presentation of research questions in the RAIs in both disciplines. Writers opt to position research hypotheses in two ways. First, research hypotheses are positioned after a theoretical framework and/or literature review under a separate section with a subtitle (e.g. ‘Theoretical framework and research hypotheses’) that indicates how the process of reasoning via a theoretical framework or literature review eventually leads to research hypotheses. Prepositional phrases are used before the research hypotheses to indicate the grounds on which the research hypotheses are proposed (e.g. ‘From the extant literature, we hypothesize that’). Second, research hypotheses are occasionally positioned after research questions in a one-to-one pattern between each research question and its corresponding research hypothesis. Overall, we have identified different types of hypotheses with reference to linguistic resources used in Ethnic Studies and Industrial Relations, as exemplified in Table 2.
Table 2. Hypotheses in RAIs on Ethnic Studies and Industrial Relations

Some differences and similarities can be identified in the linguistic resources used by writers in both disciplines. In hypothesis postulation, even though writers in both Ethnic Studies and Industrial Relations employ anticipatory verbs (e.g. ‘hypothesize’, ‘predict’), the structures involved do differ to a certain extent. ES researchers use them in combination with nouns denoting variations (e.g. ‘increase’, ‘improvements’) while those in IR use verbs expressing variations (e.g. ‘moderates’, ‘have greater improvements’, ‘have greater improvement’, ‘are positively related to’, ‘have a positive effect on’, ‘retards’).
‘mediates’, ‘retards’) to demonstrate how the values of a dependent variable change as a result of the influence from a strong independent variable. Despite such discipline-specific tendencies, it can be noted that in both disciplines, comparative adjective phrases (e.g. ‘stronger’) are used with noun phrases denoting correlations to depict relationships between variables (e.g. ‘the positive influence of MNEs in CEE countries is stronger than in other market economies’). Another salient characteristic of hypothesis postulation in both ES and IR occurs in the form of verb-preposition structures expressing relationships between variables (e.g. ‘will be partly mediated through’), most of which are modified by adverbs signaling the nature and extent to which the relationships exist (e.g. ‘partly’, ‘closely’).

Our analysis also shows that specific syntactic structures used in postulating research hypotheses differ across the two disciplines in other aspects. Table 3 provides some examples of the different lexico-grammatical structures used by writers to present research hypotheses. We have found that in 13.3% (4/30) of RAI s in Ethnic Studies, research hypotheses begin with a subject-predicate structure ‘we hypothesize’, ensued by a that-clause expressing the content of a hypothesis (e.g. ‘we also hypothesized that patients’ caregivers who viewed ‘Fixing Paco’ would have greater improvement in knowledge scores than patients’ caregivers who did not’). Such a straightforward reference to hypothesis, via the verb ‘hypothesized’, is not found in Industrial Relations RAI s. This is, however, not the case when it comes to other anticipatory verbs, such as ‘assume’ or ‘predict’. Table 3 shows that a larger percentage (i.e. 16.7% or 5/30) of the RAI s in Ethnic Studies incorporate research hypotheses in the structure of ‘we + anticipatory verb (in the simple present) + that-clause’ to indicate an extrapolation. These anticipatory verbs (e.g. ‘predict’) are used in combination with verb phrases in the simple present (e.g. ‘we assume that belonging and inclusion in the ethnic group are not static’) or the simple future (e.g. ‘We expect that the relation between perceived discrimination and religious practice will be stronger’) to describe a possible situation to be experienced by an ethnic group.

Likewise, hypotheses in Industrial Relations also employ such anticipatory verbs after first-person pronouns and before that-clauses. Notably, the anticipatory verb ‘assume’ is ensued by verbs in either the simple present or the simple future, but other anticipatory verbs, such as ‘predict’, are consistently followed by the simple future. Nonetheless, despite such cross-disciplinary similarities, some lexico-grammatical structures for hypothesis postulation are employed only in Industrial Relations, but not in Ethnic Studies. For instance, phrasal verbs signaling explicit epistemic modality or tentativeness (e.g. ‘may lead to’) only appear in Industrial Relations RAI s (albeit in only 3.3% of the IR research introductions). In addition, hypotheses are explicitly differentiated from each other using cardinal numerals (e.g. ‘Hypothesis 1a’), but interestingly, such explicit numerical markings for hypotheses are evident in only IR RAI s, but not in ES RAI s. However, in a slightly higher percentage (i.e. 10.0%) of the RAI s in Industrial Relations, writers use noun phrases denoting evidence or theoretical models before anticipatory verbs or nouns
in sentence-object positions (e.g. ‘This model generates a number of predictions that’), thus showing that a broader range of linguistic resources are employed by IR researchers in hypothesis postulation as compared to ES researchers.

| DISCIpline         | TYPE OF HYPOTHESIS                                                                 | INSTANCE                                                                                                                                       | FREQ. OF HYPOTHESIS | NO. OF RAIs | % OF RAIs |
|--------------------|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|------------|-----------|
| Ethnic Studies     | Plural first-person pronouns and hypothetical verbs (in the simple present or simple past) ensued by that-clauses | We hypothesize that there has been a general increase in anti-Semitism due to… (RAES 14: 2710) We hypothesized that patients who viewed the telenovela would have greater improvements in knowledge and intentions to adopt proactive health-care behaviors than patients receiving… (RAES 23: 60-61) | 6                   | 4          | 13.3      |
|                    | Plural first-person pronouns and anticipatory verbs (in the simple present) ensued by that-clauses | Inspired by the above theoretical perspectives, we assume that belonging and inclusion in the ethnic group are not static…(RAES 10: 415) We expect that the relation between perceived discrimination and religious practice will be stronger in more established groups (H1b)… (RAES 17: 2640-2641) | 14                  | 5          | 16.7      |
| Industrial Relations| First-person pronouns and anticipatory verbs (in the simple present) ensued by that-clauses | I therefore assume that working-time flexibility and autonomy are positively related to women’s time adequacy, whereas men experience time squeeze with these arrangements (Hypothesis 1). (RAIR 22: 262) | 10                  | 6          | 20.0      |
|                    | Noun phrase denoting a theoretical model ensued by an anticipatory verb | This model generates a number of predictions that are relevant to the current study, which focuses on issues relating to the provisioning of family care needs… The models also predict that market (as opposed to time-based) alternatives for the provisioning of family care needs will be more heavily utilised in households… (RAIR 13: 669-670) | 8                   | 3          | 10.0      |
|                    | Demonstrative nouns ensued by cardinal numbers used for explicit marking of hypotheses | Hypothesis 1a Subsidiaries of foreign MNEs will have a higher incidence of broadbased individual performance-related pay than domestic firms. (RAIR 17: 294) | 14                  | 2          | 6.7       |
|                    | Modal verbs employed to indicate hypotheses | Because of the strong male breadwinner model, flexible and autonomous working time may lead to time inadequacy primarily for men (Hypothesis 2). (RAIR 22: 262) | 4                   | 1          | 3.3       |

Table 3. Numbers and percentages of RAIs containing different lexico-grammatical structures used in hypotheses in ES and IR.
Overall, what merits attention is that even though a wide range of linguistic resources are used in hypothesis postulation in Industrial Relations, writers in Ethnic Studies consistently use first-person pronouns before an anticipatory verb in practically all their hypotheses. In this regard, SIA pointed out that many researchers in Ethnic Studies are often influenced by “qualitative studies in Anthropology and Sociology”, thus explaining why ES researchers demonstrate a relatively greater propensity to use first-person pronouns in their ethnographic qualitative studies. SIB, however, did not comment on the use of first-person pronouns, but instead only highlighted his tendency to use a hypothetical verb in the simple present after the first-person pronoun even though it might “create confusion” with regard to whether the researchers have completed their field work. In contrast, SIC’s comment resembled that of SIA as it dealt more directly with pronoun usage. According to SIC, researchers in Ethnic Studies are so actively involved that they often get “immersed into the situation they studied”. She was of the view that the use of the first-person pronouns in postulating research hypotheses suggests that ES researchers are “more active, more real, more involved or more engaged, like (as if you (the researchers) have become part of what you (the researchers) study”. This implies that many researchers are no longer “independent outsiders looking at something weird or interesting”, but would “join them” and become “engaged”. Highlighting the significance of self-mention, SIC pointed out that “other pronouns will make your research appear distant, (or) make the community very far from you”. Overall, the use of the first-person pronouns in hypothesis postulation seems to function like an overt signal of self-reference among ES researchers acting as participants in their qualitative studies.

In Industrial Relations, however, SIF explained this linguistic phenomenon from a different perspective. To SIF, the first-person pronoun ‘I’ is occasionally used by British authors, but ‘we’ is “more often used by American writers as an inclusive first-person pronoun as ‘we’ tends to encompass everybody as a whole and appears to be more generous”. SIE was of the view that “in the United States and UK”, and “in Japan and other places, (such as) Malaysia (and) Indonesia”, some of the early Industrial Relations writers were workers”, and this explains why some of them tend to use the “first-person language”. To SIE, these earlier researchers might also have some influence over today’s researchers. However, in the present era, some “academics who were previously workers” may be “more likely to remove the ‘I’ and ‘we’ politely to make (help) them conform to academic standards”. He explained that with respect “to theory, it is supposed to be separated from ‘I’ or ‘we’” and “there’s definitely a movement of becoming more intellectual about it”. SIF, however, considered that the use of first-person pronouns is less common in hypothesis postulation in Industrial Relations even though such pronouns may be used in some “qualitative research” in the discipline, particularly “ethnographic studies”, in which the researcher “also becomes a respondent (participant)”. Overall, it can be seen that even though hypothetical
and anticipatory verbs are commonly used for hypothesis postulation in both disciplines, the use of the first-person pronouns are far more prevalent in Ethnic Studies than Industrial Relations because of the alleged influence of an ontological approach and the discourse community’s expectation about ethnographic research in the discipline.

4. CONCLUSION AND IMPLICATIONS FOR RESEARCH WRITING INSTRUCTIONS

In conclusion, with regard to our first research question about the extent to which research questions and hypotheses are incorporated in RAIs in Ethnic Studies and Industrial Relations, we have found that ‘presenting research questions or hypotheses’, considered collectively as a rhetorical step, is incorporated in less than half (43.3%) of the RAIs in Ethnic Studies, but in the majority (56.7%) of those in Industrial Relations. The percentages of RAIs containing research questions and hypotheses in Industrial Relations are lower than those in Educational Psychology (75.0%) (Loi, 2010), Computer Science (75.0%) (Shehzad, 2010), Clinical and Health Psychology (65.0%) (Martín & Pérez, 2014), but higher than those in Dermatology (45.0%) (Martín & Pérez, 2014), Applied Linguistics (33.30%) (Sheldon, 2011), Political Science (30%), and Political Philosophy (15.0%) (Martín & Pérez, 2014). Given that IR has been considered a field that often adopts epistemological approaches bearing the features of various disciplines, it is understandable that the degrees of prevalence of research questions and hypotheses are higher than those in disciplines which incorporate a broader range of quantitative techniques. In contrast, Ethnic Studies (43.3%), with research that is largely based on an ontological approach involving a larger proportion of qualitative studies, more closely resembles a spectrum of other social science disciplines with numerous qualitative research papers, particularly Political Science and Political Philosophy, in which both research questions and hypotheses are less explicitly stated (see also Martín & Pérez, 2014; Sheldon, 2011).

More precisely, our study has found that research questions alone are incorporated in 23.3% of the Ethnic Studies RAIs and 30.0% of the Industrial Relations RAIs. Research hypotheses, however, are found in 30.0% of the research articles in Ethnic Studies and 36.7% of those in Industrial Relations. A possible explanation for such occurrences is that writers in Industrial Relations have a greater tendency to employ an epistemological research approach in which writers tend to extract a hypothesis to test, and when they find that it holds, they are likely to further enhance the credibility of the theory; otherwise, they would make an attempt to explain why they need to reject the theory. Overall, the aforementioned findings show that hypotheses are comparatively more predominant in Industrial Relations than in Ethnic Studies; likewise, research questions are also relatively more prevalent in Industrial Relations compared to Ethnic Studies.
explanation is that Ethnic Studies constitutes a social science discipline in which writers would prefer to state the direction of their studies through “purpose statements” (Lim, Loi, Hashim, & Liu, 2015: 70) rather than via research questions or hypotheses.

In regard to our second research question about how writers use linguistic resources to present research questions and hypotheses in the two fields, we can conclude that writers in both Ethnic Studies and Industrial Relations demonstrate a propensity to use wh-questions rather than polar questions. This means that polar questions are rarely employed in both Ethnic Studies and Industrial Relations, while wh-questions are much more frequently used by writers in Ethnic Studies and Industrial Relations. Our study has also shown that all the three types of wh-questions found in doctoral dissertations, namely factual wh-questions, evaluative wh-questions and circumstantial wh-questions, which begin respectively with ‘what’, ‘to what extent’ and ‘how’ (Lim, 2014a), do appear in RAs in both Ethnic Studies and Industrial Relations. However, causative and personal identification research questions (beginning with ‘why’ and ‘who’), which were not reported in past research, are found in Industrial Relations RAs, but not in Ethnic Studies RAs. Additionally, determinative wh-questions (beginning with ‘which’) occur in both disciplines despite their low frequencies of occurrence. In brief, determinative, causative and personal identification research questions, which constitute three categories of research questions that were not reported in previous research (Lim, 2014a), are found in this study; however, they are far less prevalent than factual, evaluative and circumstantial research questions.

As for the postulation of hypotheses, the salient linguistic resources can be unpacked meaningfully with reference to some previous research. It was reported in a past study (Lim et al., 2014) that hypotheses could be presented using a (i) noun phrase denoting postulation in an SPC structure containing past-tense copular verb preceded by a noun phrase denoting a hypothesis and ensued by a that-clause indicating a prediction (e.g. ‘the research hypothesis was that...’), and (ii) the passive anticipatory it-clause in an extraposed subject beginning with a that-clause (e.g. ‘it was hypothesized that...’); nevertheless such linguistic resources are not used by the writers in these two social science disciplines. Instead, active anticipatory verbs ensued by that-clauses are often employed by writers in both Ethnic Studies and Industrial Relations to postulate hypotheses. The use of the first-person pronouns in hypothesis postulation is, however, far more prevalent in Ethnic Studies than in Industrial Relations. In brief, it can be concluded that that-clauses in active voice constitute a generally unavoidable structure in hypothesis postulation in both disciplines, but the use of personal pronouns appears to be a common practice in Ethnic Studies, but not in Industrial Relations.

The findings reported in this paper can be flexibly adapted in designing EAP instructional materials after considering the degrees of prevalence and range of salient linguistic resources used to present research questions and hypotheses.
Given that ‘presenting research questions and hypotheses’, as a rhetorical step, is used by writers in over two-fifths and the majority of the research introductions in ES and IR respectively, it is reasonable to devote some attention to this major step in our preparation of teaching materials for novice writers in the disciplines. So far as research questions in both ES and IR are concerned, instructors may have to ascertain the nature of the research conducted by novice writers before offering any pertinent advice to learners. If learners’ papers focus on exploratory qualitative descriptions and inductive interpretations (based on the instances or phenomena observed), more time needs to be spent on guiding them to write research questions, rather than hypotheses. On the surface, *wh*-questions, compared to polar questions, appear to merit more attention in instructional sessions given that they are far more prevalent than polar questions in both disciplines. To be precise, sufficient instances of factual questions need to be introduced to learners in both disciplines as they were among the more prevalent categories of *wh*-questions in ES and IR. We may see, at a glance, that circumstantial research questions deserve considerable attention only in Industrial Relations, but not Ethnic Studies. However, when it comes to actually helping second language learners to master the construction of research questions, the apparently lower frequency of polar questions (requiring a ‘yes/no’ answer) does not mean that second language learners do not need to first familiarize themselves with the writing of polar questions before mastering alternative categories of research questions (as shown in Table 1). This is because even the most prominent circumstantial questions in Industrial Relations require an initial construction of a polar question (e.g. ‘Does minimum wage policy affect enrollments and expenditures on means-tested public assistance programs?’) before the circumstantial *wh*-word ‘how’ is added to the inceptive part of the question (e.g. ‘How does minimum wage policy affect enrollments and expenditures on means-tested public assistance programs?’). This suggests that guiding second language learners to construct *wh*-questions often requires instructors to first familiarize learners with polar questions (as exemplified in Table 1).

In cases where the study conducted is essentially based on quantitative data in both disciplines, learners can be introduced to the writing of hypotheses, which deductively delineate a possible situation or phenomenon before quantitative data are obtained to refute or support the hypotheses. While guiding learners in hypothesis postulation (as exemplified in Tables 2 and 3), instructors may introduce learners in both disciplines a range of (i) anticipatory verbs used in combination with *that*-clauses, and (ii) comparative adjective phrases and verb phrases with modifying adverbs that denote relationships between variables, given that they constitute the central linguistic components of the directional determinants that distinctly guide their subsequent work right from the onset of their research. To raise learners’ consciousness of the use of anticipatory verbs mentioned above, it is recommended that some gap completion items be prepared in an exercise that requires learners in both disciplines to consistently match
anticipatory verbs in the simple present with *that*-clauses containing verb phrases in the simple present or the simple future. The difference across ES and IR is that instructors might allow novice writers in Ethnic Studies to use first-person pronouns in hypothesis postulation frequently, especially if the research concerned is based on an ethnographic approach which signals researchers’ overt self-reference, thus reflecting their active engagement as both researchers and participants. Overall, the major types of lexi-co-grammatical resources identified for both research questions and hypotheses can be included in exercises aimed at training learners to construct research questions and statements that guide them in developing their research reports. The range of lexi-co-grammatical resources to be highlighted and emphasized in EAP lessons is largely contingent upon the prevalence of these directional determinants in the social science disciplines, and EAP instructors can make informed decisions if they take into consideration the results of our genre-based textual analysis and the valuable inputs provided by the specialist informants concerned.

[Paper submitted 23 May 2020]
[Revised version received 1 Sep 2020]
[Revised version accepted for publication 15 Sep 2020]

**Acknowledgements**

Many thanks to the University of Malaysia Sabah for providing a research grant (SBK-0377-2018) that has helped us obtain related references for this investigation into research questions and hypotheses in Ethnic Studies and Industrial Relations. We are indebted to our specialist informants who have given us enlightening and useful inputs. Let us thank Professor Nadežda Silaški and two reviewers for their useful suggestions and insightful comments.

**References**

Berkenkotter, C., & Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bhatia, V. J. (1993). *Analyzing genre: Language use in professional settings*. New York: Longman Publishing.

Biber, D., Connor, D., & Upton, T. A. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. Amsterdam: John Benjamins.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). Oxford: Routledge Falmer Publisher.

Feak, C. B., & Swales, J. M. (2011). *Creating contexts: Writing introductions across genres*. Ann Arbor, MI: University of Michigan Press.

Hyland, K. (2005). *Metadiscourse*. London/New York: Continuum.
Lim, J. M. H. (2011). Delineating sampling procedures: Pedagogical significance of analysing sampling descriptions and their justifications in TESL experimental research reports. *Ibérica*, 21, 71-92.

Lim, J. M. H. (2014a). Formulating research questions in experimental doctoral dissertations on Applied Linguistics. *English for Specific Purposes*, 35, 66-88. https://doi.org/10.1016/j.esp.2014.02.003

Lim, J. M. H. (2014b). Expert writers’ descriptions and justifications of research instruments and materials: Using communicative resources to generate promotional effects. *Journal of Applied Linguistics and Professional Practice*, 11(2), 154-176. https://doi.org/10.1558/japl.34704

Lim, J. M. H. (2017). Writing descriptions of experimental procedures in language education: Implications for the teaching of English for academic purposes. *English for Specific Purposes*, 47, 61-80. https://doi.org/10.1016/j.esp.2017.05.001

Lim, J. M. H. (2018). Directional determinants in U.S. experimental doctoral dissertations: Writers’ communicative resources in language education. Singapore: Partridge.

Lim, J. M. H. (2019). Explicit and implicit justifications of experimental procedures in language education: Pedagogical implications of studying expert writers’ communicative resources. *Journal of English for Academic Purposes*, 37, 34-51. https://doi.org/10.1016/j.jeap.2018.10.006

Lim, J. M. H., Loi, C. K., & Hashim, A. (2014). Postulating hypotheses in experimental doctoral dissertations on Applied Linguistics: A qualitative investigation into rhetorical shifts and linguistic mechanisms. *Ibérica*, 27, 121-142.

Lim, J. M. H., Loi, C. K., Hashim, A., & Liu, M. S. M. (2015). Purpose statements in experimental doctoral dissertations submitted to U.S. universities: An inquiry into doctoral students’ communicative resources in language education. *Journal of English for Academic Purposes*, 20, 69-89. https://doi.org/10.1016/j.jeap.2015.06.002

Loi, C. K. (2010). Research article introductions in Chinese and English: A comparative genre-based study. *Journal of English for Academic Purposes*, 9(4), 267-279. https://doi.org/10.1016/j.jeap.2010.09.004

Loi, C. K., & Evans, M. S. (2010). Cultural differences in the organization of research article introductions from the field of Educational Psychology: English and Chinese. *Journal of Pragmatics*, 42(10), 2814-2815. https://doi.org/10.1016/j.pragma.2010.03.010

Martín, P., & Pérez, I. K. L. (2014). Convincing peers of the value of one’s research: A genre analysis of rhetorical promotion in academic texts. *English for Specific Purposes*, 34, 1-13. https://doi.org/10.1016/j.esp.2013.09.002

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. London: Sage.

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

Ozturk, I. (2007). The textual organisation of research article introductions in applied linguistics: Variability within a single discipline. *English for Specific Purposes*, 26(1), 25-38. https://doi.org/10.1016/j.esp.2005.12.003

Shehzad, W. (2010). Announcement of the principal findings and value addition in Computer Science research papers. *Ibérica*, 19, 97-118. Retrieved from http://www.aelfe.org/documents/05_19_Shehzad.pdf
Sheldon, E. (2011). Rhetorical differences in RA introductions written by English L1 and L2 and Castilian Spanish L1 writers. *Journal of English for Academic Purposes, 10*(4), 238-251. https://doi.org/10.1016/j.jeap.2011.08.004

Soler-Monreal, C., Carbonell-Olivares, M., & Gil-Salom, L. (2011). A contrastive study of the rhetorical organization of English and Spanish PhD thesis introductions. *English for Specific Purposes, 30*(1), 4-17. https://doi.org/10.1016/j.esp.2010.04.005

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

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

Wang, W., & Yang, C. (2015). Claiming centrality as promotion in applied linguistics research article introductions. *Journal of English for Academic Purposes, 20*, 162-175. https://doi.org/10.1016/j.jeap.2015.05.002

Wong, V. F., & Lim, J. M. H. (2014). Linking communicative functions with linguistic resources in short stories: Implications of a narrative analysis for second language writing instruction. *System, 45*, 147-162. https://doi.org/10.1016/j.system.2014.05.008