



THE ASIAN ESP JOURNAL

The Asian ESP Journal

December 2016
Volume 12, Issue 3



Chief Editor - Roger Nunn

Associate Production Editor - Colin Toms



Published by ELE Publishing
(Division of the TESOL Asia Group)

TESOL Asia Group is wholly owned by SITE SKILL TRAINING Pty Ltd (Australia)

All rights reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise, without the prior written permission of ELE Publishing or the Chief Editor of Asian ESP.

No unauthorized photocopying

Publisher: ELE Publishing

Managing Editor of ELE Publishing: Dr. John Adamson

Chief Editor of the Asian ESP Journal: Professor Roger Nunn

Associate Production Editor: Colin Toms

Production Editors: David Young, Dawn Garinger; Enaam Hunaini; Jenny Taylorson; Lala Movsesian; Lynne Robinson; Ronald Jones; Samira Fahmi; Sarah Tatz; Tiffany Cammidge

ISSN 2206-0979 (Online)



Table of Contents

1. Roger Nunn & Colin Toms.....	4 – 6
<i>Foreword</i>	
2. Michael Guest.....	7 – 24
<i>Overcoming Institutional Barriers to Establishing an ESP Programme: A Case Report from Japan</i>	
3. Anie Attan, Abdul Halim Abdul Raof, Masputeriah Hamzah, Noor Abidah Mohd Omar & Masdinah Alaayah Md Yusof.....	25 – 39
<i>Exploring Industry Expectations of Graduating Students' Oral Communicative Ability</i>	
4. Luo Yan & Yingli Yang.....	40 - 69
<i>Examining Business English Majors' Business Vocabulary Knowledge Development</i>	
5. Masoomeh Estaji & Hussein Meihami.....	70 – 93
<i>Revisiting the Topical Knowledge of Iranian ESP Learners in Reading Comprehension: Text Types and Question Types</i>	
6. Seyyed Mohammad Reza Adel & Mostafa Janebi Enayat.....	94 – 119
<i>Gender Representation and Stereotyping in ESP Textbooks</i>	



Foreword

Roger Nunn

The Petroleum Institute, Abu Dhabi

Colin Toms

The Petroleum Institute, Abu Dhabi

In this edition of *The Asian ESP Journal*, we are happy to present another varied set of studies in the Asian context. At *The Asian ESP Journal*, we are keen to publish well-researched studies that make a difference. This may be in the local context but it is also very important that a study has some resonance beyond the local context that is translatable to other setting.

In the first paper, *Overcoming Institutional Barriers to Establishing an ESP Programme: A Case Report from Japan*, Michael Guest provides a very interesting example of research and reflection based on real experience that made a difference. Importantly it also shows how an experienced academic can approach a real-world issue and use research to come up with a solution. His study is therefore a good example of the way that first-person narrative and reflection counterbalanced by triangulated research can provide solid grounds for curriculum improvement. AESP welcome the first-person input of this type, which acknowledges transparently the inevitable presence of the teacher/researcher in the report of his research. It is also interesting to note the process through which this particular study helped overcome initial institutional resistance: something that many ESP specialists can relate to beyond this particular setting.

In the interest of variety, our second study *Exploring Industry Expectations of Graduating Students' Oral Communicative Ability* was conducted by a team of researchers in Malaysia: Anie Attan, Abdul Halim Abdul Raof, Masputeriah Hamzah, Noor Abidah Mohd Omar and Masdinah Alaayah Md. Yusof. Not unlike the first study, this team designs new activities to make curriculum more relevant. They grapple with the complex but very real ESP issue of the relevance of graduate study to industry and the potential mismatch in expectations. This study focuses on spoken language and comes up with a practical oral test that illustrates how assessment and task design needs to go beyond developing language skills into areas such as thinking and interactive ability.

We follow this with a piece by Luo Yang and Yingli Yang, which notes that in China, there has hitherto been little research into the mastery of discipline-specific vocabulary knowledge by Business English majors. Their study examines Business English majors across four grade levels. With a vocabulary test adapted from Hsu's (2011) *Business Academic Word List*, one hundred and twenty seven Business English majors were tested as part of this study. Although students generally demonstrated better mastery of business vocabulary as grade level increases, their overall scores did not show a consistent pattern. Furthermore, a significant difference was found among the four grades in word mastery at different frequency levels. Results of the study are interpreted in terms of both testing and curriculum, with specific implications for vocabulary testing and curriculum development being discussed.

Our focus shifts to Iran for our next piece, *Revisiting the Topical Knowledge of Iranian ESL Learners in Reading Comprehension: Text Types and Question Types* by Masoomah Estaji and Hussein Meihami. Their study considers the effect of topical knowledge on ESP learners' reading comprehension performance, specifically on different types of reading questions. In particular, comprehension, inference, and lexical question types were investigated. The study comprised a total of 46 upper-intermediate students, 26 of them civil engineering students, the remainder being General English students. All participants were in their mid- to late twenties. Two types of reading texts, General English and ESP, were assigned to both groups. Results showed that the ESP students performed better on both the General English reading text and the ESP reading text. In addition, the topical knowledge of the ESP students was deemed to be instrumental in helping them answer different types of questions. The study also revealed that there was no statistically significant difference

between the ESP students' performance on specific reading texts and their performance on General English reading texts. The authors conclude that, despite the limitations of what is a relatively small-scale investigation, their results have implications for both teachers and test designers.

We remain in Iran for our final offering, a topical study by Seyyed Mohammad Reza Adel and Mostafa Janebi Enayat. In *Gender Representation and Stereotyping in ESP Textbooks*, the authors present an exploration of gender equality in current ESP textbooks. Their concern was to investigate the gender representation in the images and texts of the 'Oxford English for Careers' series including *Commerce*, *Nursing*, *Technology*, and *Tourism*. Two separate taxonomies - the first from Goffman's *Gender Advertisements*, the second from Kress and van Leeuwen's *Reading Images* - were converged as an analytical instrument. Subsequently, a systematic quantitative content analysis was conducted to determine gender visibility. The principal findings were threefold. Firstly, men are portrayed as more active and competent in the specific arenas of technology and tourism. Secondly, women appear as objects of both scrutiny and desire in all the titles under scrutiny. Finally, men are more the frequently depicted 'at work' and identified as principal breadwinners in all but the *Tourism* text. The authors conclude that their results reflect extant institutionalized gender discrimination and, consequently, have implications for materials writers, teachers and learners alike.

It might be noted that this edition of *The Asian ESP Journal* departs somewhat from the established norm. Generally speaking, each issue is united, albeit loosely on occasion, by a central governing theme. The manifest heterogeneity of the offerings in this issue, however, defies any such thematic unity. That much said, all the articles presented herein are, as ever, underpinned by our firm belief in the supremacy of the authorial voice. We do not impose limits on the linguistic choices available to authors, nor do we impose a deterministic generic structure or style. In that regard, then, the current edition of *The Asian ESP Journal* remains true to its essential vision, one that is fundamentally Asian in both voice and orientation.



Overcoming Institutional Barriers to Establishing an ESP Programme: A Case Report from Japan

Michael Guest

University of Miyazaki, Japan

Biodata

Michael Guest is Associate Professor of English in the Faculty of Medicine at the University of Miyazaki, Japan. Guest was a long-time columnist on ELT matters in the Daily Yomiuri and Japan News newspapers. His academic interests include ESP, medical English, discourse analysis, testing, and English as a Lingua Franca.

Email: mikeguest59@yahoo.ca; michael@med.miyazaki-u.ac.jp

Abstract

This paper is a Case Report describing a triangulation of 1) action research, 2) the application of my own previous field research, and 3) direct argument, undertaken in a university in Japan in order to inculcate a more informed and accepting attitude towards the implementation of ESP programmes among both peers and administrators. I undertook these actions in response to initial administrative claims that establishing a university-wide ESP-based programme was not feasible due to three stated reasons: 1) that specialist English content would be too difficult for first and second year students, 2) that students needed to master general English before beginning an ESP course, and 3) that ESP teachers should be content experts, not applied linguists. To address the first claim, I carried out in-class action research in which lesser achieving students were given a specialized task usually demanded of more competent learners. To address the second claim, I applied some results of existing field research from within the medical professional discourse community which indicate discourse competence without a mastery of English minutiae. To address the third claim, I presented to programme administrators a combination of established ESP research arguing

that ESP teachers may actually be preferable to content experts when conducting ESP classes or in developing materials. These three actions were carried out in order to augment the case for the inclusion of an ESP approach in the university's English programme and were eventually instrumental in changing and upgrading the university's English curriculum. It is hoped and believed that these responses and activities might have application beyond Japan, throughout the region wherever similar conditions and sentiments may exist.

Keywords: ESP, medical English, Japan, action research, genre analysis, pedagogical policy, non-native English speakers

Introduction

I teach English to medical and nursing students, most of whom are in their first and second years of study, at a national university in Japan. Over time, I had established a successful ESP-based course with these students. However, after being asked to offer recommendations regarding revising existing English education programmes in the other university faculties (education, agriculture, engineering), my initially proposed outline, explanation, and recommendations (to establish university-wide faculty-based ESP programmes) were rebuffed by university English programme authorities.

The reasons offered were threefold. These were: 1) that specialist content in English would be 'too difficult' for first and second year students in other faculties, 2) that students need to master general English (EGP) before beginning an ESP course, and 3) that ESP teachers should be content experts, not applied linguists. In order to erase or minimize these suspicions and doubts, I carried out a combination of action and applied field research – both further supported by existing published research – over two years, after which results were presented to selected senior programme administrators and professors at the author's university. The background of the action and applied field research, and subsequent arguments presented to the university in support the establishment of a wider ESP-based approach, are reported and discussed in this paper. It is hoped that teachers and researchers who have faced similar obstacles when trying to implement ESP programmes may be informed or encouraged by the actions taken or research presented in this case.

Literature Review

Although ESP teaching and learning in Japan has generally followed the same remarkable growth that has been experienced in much of East Asia, certain local factors have led to misunderstandings and suspicions regarding the approach's role and function, particularly in universities. Terauchi (1995) discusses an influential Japanese education policy rift based upon the question as to whether university English should be treated as a practical subject with professional implications or as a general subject as one of the humanities -- as a foundation of a standard liberal arts education. Many university administrators have opted for the latter approach, effectively relegating the adoption of ESP programmes to less academically prestigious vocational schools. The upshot of this debate has led to an emphasis upon specialist English as product, without addressing the value of treating ESP learning as a process (Anthony, 2011). Terauchi et al (2010), in a survey of current ESP practice in Japan, describe how the emphasis upon ESP as a product-centered approach has caused many ESP practitioners to lose confidence, not being content experts in the field, and thus relying upon commercial ESP textbooks rather than developing their own materials based upon needs or genre analysis. This attitude towards ESP persists particularly among many senior teachers and administrators in Japan. This then, is the ideological framework that the author chose to address.

The literature indicates that such perceptions are far from being limited to Japan. The traditional Confucian emphasis upon the educational process, as opposed to tangible outcomes, has clear implications for product-focused ESP pedagogy (Yum, 1994). Resistance towards methods and approaches that aim to go beyond translation and mastery of grammatical/lexical detail is widespread in many Asian societies (Shih, 1999), with criticisms of ESP teaching practice among peers in Taiwan being particularly pronounced (Chen, 2011). Reticence among Hong Kong learners of English is, "strongly influenced by Chinese norms and expectations (and therefore) are apt to find it more difficult to make the transition to interactive, problem-based learning," which "has significant pedagogical implications for...ESP teachers who are charged with preparing them for this mode of learning" (Jackson, 2002, p.78). In Vietnam too, a lack of institutional support for, and understanding of, ESP learning at universities has been noted (Huan, 2013). Clearly, the types of problems I faced are widespread in the region.

As mentioned earlier, in order to both supplement and inform both the action and applied field research (described later) carried out in response to the administration's rejection of the ESP programme I had initially proposed, published research which addressed these claims by 1) supporting the practice of utilizing specialist L2 English content for university freshmen, 2) asserting that English for General Purposes (EGP) and ESP need not be taught successively but could work in a complementary manner, and 3) claiming that non-content expert teachers could successfully create and manage an ESP course, were gathered and subsequently cited to various educational policy authorities at my university at subsequent meetings, informal discussions, and policy-making sessions.

To support points (1) and (2) above, research and subsequent argumentation supporting the suitability of an ESP approach within a university setting, not only for the benefit of already competent in-service speakers, are likely well-known to experienced ESP practitioners. Widdowson (1983) was among the earliest to argue that the generic ESP focus upon contexts is precisely suited to university education. Hutchinson and Waters (1987), in arguing for a more complementary approach between EGP and ESP, contend that learners can make better sense of new ESP information based on their existing language systems, and thus advocate a complementary role between the two approaches. Although Strevens (1988) somewhat controversially divorced EGP from ESP as an 'absolute' quality of ESP teaching, this was done in reference to teaching approaches and materials, not the sequencing of courses or any gradation in student abilities.

Claims that ESP teaching can be effective before EGP has been mastered can often be noted in Dudley-Evans' research, the strongest claim espousing that ESP teaching can even be applied to beginners (Dudley-Evans, 1994). In Dudley-Evans and St. John (1998), it is further argued that while some EGP foundation should be initially established, ESP can be implemented from intermediate levels. A further argument that divisions between EGP and ESP are often variable and determined more by content and approach rather than age or ability, was also put forth (Johns & Dudley-Evans, 1991). Referring to Japanese settings, Anthony (1998) also speaks of the boundaries between EGP and ESP becoming blurred.

Adding to the overwhelming amount of research that supports the successful teaching of ESP to students who have not yet mastered general English, and thus that EGP and ESP can be mutually supportive and are not necessarily sequential approaches, are a number of localized

studies conducted in Asian locales. Lo (2012) successfully established an ESP programme for low-achieving vocational high school students in Taiwan, with similar complementary, overlapping relationships practiced between EGP and ESP also being noted in Far (2008) in Iran and Flowerdew (2000b) in Hong Kong.

In regard to the third administrative claim, that ESP teachers should be content experts and not applied linguists, established research in the pedagogical value of genre analysis was sought out and presented to authorities.

The paramount importance of the teacher's knowledge of genre analysis in order to better understand the communication practices of discourse communities is largely traceable to the pioneering work of Swales (1990). The role of genre analysis as the basis of specialized syllabus and materials design was further established by Bhatia (1993), while Flowerdew (2000a) emphasized how ESP teachers can transform good genre descriptions into syllabus and materials design (although he admits that it may require some field knowledge and expertise). Belcher (2004) maintained that an 'immersion' approach, utilizing a content expert alone, is often unhelpful in terms of learners accessing the target language and calls for support in making the transition to the new discourse community that the ESP teacher can provide.

Flowerdew (2011) has argued further that the various ways in which meanings are realized in language (such as content and identity, as opposed to merely action), is something a teacher/researcher would note, but would often overlooked by learners. Similar arguments maintaining that the role of the ESP teacher is not to be that of the content expert, but to be an individual sufficiently informed by genre and the conducting of needs analyses so as to create viable syllabi and materials are made by Anthony (2007), who envisions the ESP classroom as a setting in which the teacher also serves as a learner.

The teacher's role in developing a language description, particularly in knowing what the features of a given genre are and how to best scaffold them for the learner ESP students, is emphasized by Hutchinson and Waters (1987). ESP learners often being the knowers of the content, as opposed to the teacher, allows the ESP classroom to become a collaborative learning experience in which it is natural that students struggle with the language while teachers may struggle with specific content (Dudley-Evans & St. John, 1998). And since ESP

teaching focuses more on the process of learning rather than the product, the former may present a more viable approach in a rapidly changing workplace (Anthony, 2011). In short, the literature supports the central role of the ESP teacher in teaching specialized discourse to learners in the field.

Many of these prominent studies were among those cited to supplement my position to university administrators and policy makers in response to their claims that ESP courses are too difficult, should be taught only after mastering general English, or should be taught by content experts and not language teaching specialists. This then constituted the first part of my response to the authorities.

Further Actions and Responses

Affecting change in the classroom, or in an entire institutional program, demands identifying a problem and then taking suitable action to define, address or otherwise tackle said problem. The goal of this process is to achieve what is known as ‘catalytic validity’ (Wolcott 1994; Anderson, Herr, & Nihlen, 1994). Catalytic validity is the validation of an action by bringing about some change in the perspective or activities of other participants. The actions I took in response to the initial rebuff of my proposals to establish an ESP-based curriculum across all university faculties aimed at establishing this type of validity. Moreover, in order to convince authorities and administrators of the veracity of my claims I also had to demonstrate interpretive validity, that is, to show that any interpretation of the actions I took is based upon an accurate and reliable description of those participants within the action.

My responses to the first two claims made by administrators, those being (1) that ESP is too difficult for standard Japanese university students and 2) that learners should master English for general purposes before entering ESP courses), involved a two-pronged approach, the first being a truncated form of action research and the second being the direct application of previous field research I had undertaken. These two different methods, approaches, and sources employed are described in separate sections below, with the results of each approach described thereafter.

1.1 Action research: background and method

Action research in EFL typically calls for the researcher to identify a gap or problem in the classroom or teaching practice, make observations to define the problem, take actions, evaluate or reflect upon the actions, and then, possibly, apply the actions to correct or minimize the original gap or problem (Burns, 2009).

The problem identified in this case was the initial dismissal of my ESP proposals and the related claims that such a program would be too difficult for most university students -- the outright rejection of what I considered to be ‘good practice’. In order to demonstrate that ESP materials were not ‘too difficult’ for university majors other than the ‘more academically-proficient’ medical students, I decide to apply a representative classroom lesson item (see Appendix) from my ESP-based medical classes to nursing students to see if the nursing students could adequately manage the task. If they were able to do so, this would be used as support in my next discussion with educational authorities and fellow English Program Committee members.

I chose to have nursing students tackle a task usually reserved for medical students because there is a widely established and acknowledged experience, maturity, and academic ability gap between students in the two faculties. According to in-house statistics noted in the spring of 2011, the average age of 1st year medical students was 21.1, as opposed to 18.5 for nursing students. Of the 108 1st year medical students, 45 admitted to having some experience studying or using English beyond the standard 6 years of formal Japanese education, while only 4 of 63 first-year nurses stated so. English is also a core subject on both national and second-stage university entrance exams for medical students but is not required for nursing students. If nursing students could complete the task, the claim of ‘being too difficult’ made by the programme committee members would be weakened.

The task used to test the nursing students’ ability to manage and/or master ESP materials was one that I had regularly used hitherto only in medical English classes. The task and materials (see Appendix) are based upon a standard hospital patient admission form (obtained from the attached university hospital and translated from Japanese), originally adapted after discussion with in-service doctors and nurses regarding the form’s most salient features, and then modified for the classroom.

Over a series of classes, students learn and practice how to 1) formulate each of the categories on the form as questions to be asked to patients or other health workers, 2) to write short responses in abbreviated forms, and 3) to be able to report the data appropriately to other health workers. Finally, the students' abilities to carry out these skills is assessed in a role-play setting in which the teacher serves as a patient while two students alternately ask, write, and finally report selected items as if the teacher were a fellow health worker.

The grading criteria based upon accuracy, speed, interactive skills, flexibility and extension of certain contents, and the ability to complete and convey the data in a professional manner is made known to the students and judged on a holistic 35-point scoring basis (the overall total of 35 is chosen due to the fact that it is one of the two assessments given to students as a part of the course grade, the other is also graded out of 35, while other classroom factors account for the remaining 30% of the course grade). Under 22 points is deemed unsatisfactory and requires re-testing, 22-26 points is deemed to be adequate, 27-30 is regarded as proficient, and 31-35 as excellent.

1.2. Results from the action research

The average score that had been given to 108 first-year medical students upon completing this role-play task over the previous ten years was 28.4. For the 63 first-year nursing students completing the same task for the first time the average assessment score, using the same assessment criteria, was 27.2. In order to avoid confirmation bias and unduly inflating the scores given to the nurses as well as to increase the interpretative validity of the action, a second adjudicator, who was familiar with the activity, content, and grading criteria, having taught and assessed the same task in the past, was asked to also observe and grade the nursing students' performance. This observer's average score was even higher than my own, at 27.8, with an SD= 1.78, while joint probability of agreement using correlation coefficients (CC) used to measure inter-rater reliability = 0.94.

It must be admitted that the nursing classes took 50% longer than the medical students to practice and prepare for the assessment (six 90-minute classes, as opposed to four for the medical students) but the salient point is that the nursing students were able to achieve an average score deemed 'proficient', only slightly below that of more English-experienced medical students, using an ESP-based approach and materials.

These results appear to lend credence to the notion that ESP materials can be managed by less advanced or proficient English learners and are not therefore the sole province of more advanced or highly-educated learners. In short, a lack of academic proficiency need not preclude a learner's ability to master ESP content.

These results give further voice to the claim that ESP-trained teachers can provide pedagogical and methodological perspectives that might not be available to the field content expert. In the case described above, I had utilized both an analysis of student needs and a key genre within the medical domain (that of completing and reporting data on patient admission forms), transformed these into materials that were motivating to the students, and thereby provided a framework within the syllabus that allowed these new language forms to be consolidated within the students' existing L2 systems.

Augmented by the citations given in the literature section earlier, this argument was also duly presented to university authorities in follow-up meetings in order to provide a rebuttal to their initial objections.

2.1. Utilizing previous field research

During the 2012-2013 academic year, I was awarded a scientific-grant-in-aid by the Japanese Ministry of Education to investigate the performance factors of Japanese doctors in English at international medical conferences (Guest, 2013). One characteristic of this field study was to develop a comparison between Japanese and other East Asian non-native presenters at the same medical conferences, to note if there were any common performance strengths or weaknesses.

Analysis involved observing, monitoring, and notating the various generic moves performed in academic/medical presentations, such as opening and closing gambits, transitions, narrative flow, as well as intonation, and the use of formulaic academic phrases. This collection and analysis loosely followed the formula originally developed by Swales (1990) in which models of speech could be broken down into analyzable moves within a specific genre.

The generic moves and related language choices that speakers made in those conference presentations deemed particularly effective were analyzed for common lexical and strategic patterns. This was done to ascertain what may be considered normative in this particular domain of the medical/academic discourse community as well as note in which way Japanese doctors might be outperforming or under-performing vis-à-vis their peers in East Asia in terms of conference presentation efficacy.

2.2. Results utilized from the field data

One of the most noteworthy features arising from the collection of data from over 140 non-native English medical English presentations was the remarkable number of occasions in which effective and skilled non-native English speakers made repeated, systemic English ‘errors’ in speech. These violated standardized notions of correctness but in no way negatively impacted the overall effectiveness of their presentations. A representative sample of these is reproduced in Table 1 below. The actual utterances appear on the left side in italics with the standard forms appearing on the right. Each basic lexico-grammatical pattern represented below was noted in speech at least five times, from at least three distinct presenters and presentations, with the speakers all being non-native English speakers coming from minimum of three different East Asian countries.

Table 1: Samples of Non-Standard Presentation English from Asian Non-native Speakers of English (adapted from Guest, 2014)

• <i>We placed clamp on X</i> - We placed the clamp on the X
• <i>Three colonoscopy were performed during two separate period</i> – Threecolonoscopies were performed during two separate periods.
• <i>I’d like to show you some case</i> - I’d like to show you a few cases
• <i>In the case with X</i> - In the case of X
• <i>Left side approach we will find X</i> -Using a left side approach we will find X
• <i>We must take care of X</i> - We must be aware of X

• <i>We can well observe X</i> - We can easily/clearly observe X
• <i>There was so significant difference</i> - There was a very significant difference
• <i>Why we chose X is because</i> - The reason we chose X is because...
• <i>May have some advantage to do by endoscope</i> - There may be some advantages in doing it by endoscopy
• <i>How to X?</i> - How can/should we do X?
• <i>First, I present X</i> - First, I will present X
• <i>We want to ask why is this so</i> - We want to ask why this is so.
• <i>How should we do?</i> - How should we do it?
• <i>Next I show you</i> - Next, I'll show you
• <i>We discussed to operate this case or not</i> - We discussed whether to operate in this case or not
• <i>Even we had prepared thoroughly</i> - Even though we had prepared thoroughly
• <i>I'm going to deal with like this</i> - I'm going to deal with it like this.
• <i>Because of no symptom</i> - Because there was no symptom
• <i>It is not clear about the background of x</i> - The background of X is not clear
• <i>In this technique</i> - Using this technique

Many of these samples may be considered representative examples of an emerging English as a Lingua Franca (ELF), as most conform to speech data forms noted in both the Asian Corpus of English (ACE, 2014) and the European-based Vienna International Corpus of English (VOICE, 2013). However, for the purposes of our discussion, the paramount feature is that these non-native English speakers were able to present effectively at international settings in English, within their professional discourse community, utilizing what one might describe as occasionally unpolished or 'incorrect' English.

This then brings into question the notion that efficient ESP learners must have mastered the minutiae of EGP before effectively carrying out interactions within their professional discourse communities. These research findings, based on my own previous field research, were subsequently submitted and explained to the university's English Programmes Committee.

Reflections and Implications

The two actions described above were undertaken not to confirm a research hypothesis but in order to convince authorities that an ESP-based curriculum for university is valid and productive. Current ESP practitioners are likely already familiar with the validity of the arguments presented in order to both substantiate my own teaching practice and advocate this type of approach for other university faculties. The purpose of undertaking these activities, and subsequently reporting them in the current paper, was and is to help convince others of the efficacy of these methods and approaches.

The value for other practitioners lies in the narrative, the story of my response to the initial rebuff by programme administrators. In the earlier literature section I made several references to researchers and practitioners in Asia who had met barriers in establishing ESP programs or resistance to the pedagogy that it entails. This is a story I have also encountered in the region several times in face-to-face discussions and through personal anecdotes on numerous occasions.

English programme administrators may naturally be resistant to what appear to be new methods and approaches. They may assume that advocates are merely cheerleaders for novelty, that they are ungrounded in sound theory, viable practice, or meaningful outcomes. The advocate then must take the initiative to correct this notion, to show that the approach being advocated is grounded in established research (which administrators may be unfamiliar with), that the practice can be carried out within their own institutions, and that they can produce positive outcomes among learners.

Any call for further study then would not be for more research into the viability of ESP programmes at universities in Asia, but rather on the means and methods of gaining understanding and agreement from those holding the reins of power. More tales of advocacy,

grounded in sound pedagogical theory, practice, and research, might establish both a motivational and functional groundwork for those teachers who advocate new, and presumably more suitable, pedagogies and approaches.

As each institution will hold varying degrees of conservatism or resistance to change, have different types of people in charge, as well as differing learners and learning goals, the initial criticisms of ESP programs made at my own institution and the nature of my own responses may not be viable for all cases. But the practice of grounding advocacy in theory, meaningful practice, and positive outcomes in a way that can positively alter the perspective of administrators and others holding power, is something that can, and perhaps should, be applied elsewhere in the region.

Summary and Conclusions

In this study, I have described the responses undertaken to defend the implementation of an ESP programme across all faculties at my university. A combination of published research, classroom action research, and my own previously published field research was gathered in order to lend greater credence to the arguments that 1) university freshmen could manage an ESP programme, 2) ESP need not sequentially follow the mastering of EGP, and 3) trained ESP specialists might be better suited to this task than content specialists.

In the spring of the 2014 academic year, the English Programmes Committee at my university adopted, and began developing, an ESP-based approach for English education within university faculties other than medicine, with one language education ESP expert assigned to each of the university's four faculties, much as per my original recommendation.

I cannot pretend to take full credit for this adoption of what I view as a more progressive, pedagogically sound approach to university English teaching. One factor was likely a change in the membership of the programmes committee, another being a gradual widespread recognition of the validity of ESP approaches that can be noted throughout Asia, as I am far from the only teacher who has advocated a shift towards ESP-based programmes. Nonetheless, having originally asked for my input, and despite the initial rebuff, an understanding of ESP practices has ultimately been established. By subsequently augmenting the original proposal with the combination of action research, applied previous field research,

and cited studies and then presenting these to authorities to further support the proposal, groundwork had been established for change, which was thereafter taken up by newer members of the programme committees.

It remains to be seen what successes (or demerits) this new approach produces, but I feel both comforted and justified by the fact that the combination of results based upon my action research, field research findings, and presentation of published works in the field had an impact upon the adoption of a system almost identical to the one that I had proposed three years previously. In this sense, I believe that catalytic validity has been achieved.

Among both fellow teachers and administrators in the region who are not versed in applied linguistics it may be common to hold some resistance towards the adoption of new teaching approaches, as was noted in the literature. Among those forms of resistance, the claims that ESP requires content experts alone, that such an approach is too difficult for standard university students, and that learners should master EGP in advance of developing ESP skills, may arise. It is hoped that my ultimately gratifying, and pedagogically affirming, experience, based on the results garnered from both classroom action and field research, may serve as an inspiration and guidepost for others who are trying to establish ESP-based programmes elsewhere in Asia, but who are met with peer and/or administrative resistance.

References

- ACE (2014). The Asian Corpus of English. Director: Andy Kirkpatrick; Researchers: Wang Lixun, John Patkin, SophiannSubhan. <http://corpus.ied.edu.hk/ace/> [Accessed Feb. 10, 2016].
- Anderson, G., Herr, K. & Nihlen, A. (1994). *Studying Your Own School: An educator's guide to qualitative practical research*. Thousand Oaks, California: Sage.
- Anthony, L. (1998). Defining English for Specific Purposes and the role of the ESP practitioner. *Center for Language Research 1997 Annual Review*, 115-120.
- Anthony, L. (2007). The teacher as student in ESP course design. *The Proceedings of 2007 International Symposium on ESP & Its Applications in Nursing and Medical English Education*, 70-79.
- Anthony, L. (2011). Products, processes and practitioners: A critical look at the importance of specificity in ESP. *Taiwan International ESP Journal*. 3:2, 11-18.
- Belcher, D. (2004). Trends in teaching English for specific purposes. *Annual Review of Applied Linguistics*, 24, 165-186.
- Bhatia, V.K. (1993). *Analysing Genre*. Harlow: Longman (U.K.).
- Burns, A. (2009). *Action research in second language teacher education*. In A. Burns & J.C. Richards (Eds.), *The Cambridge Guide to Second Language Teacher Education* (pp. 289-297). New York: Cambridge.
- Chen, Y. (2011). The institutional turn and the crisis of ESP pedagogy in Taiwan. *Taiwan International ESP Journal*, 3(1), 17-30.
- Dudley-Evans, T. (1994). Genre analysis: an approach to text analysis for ESP, in M. Coulthard (Ed.), *Advances in Written Text Analysis* (pp.219-228). London: Routledge.

- Dudley-Evans, T. & St John, M. (1998). *Developments in ESP: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Far, M. (2008). On the relationship between ESP and EGP: A general perspective. *English for Specific Purposes World* 1/17(7), 17-27.
- Flowerdew, J. (2000a). Discourse community, legitimate peripheral participation, and the non-native English-speaking scholar. *TESOL Quarterly* 34 (1), 127-150.
- Flowerdew, L. (2000b). Using a genre-based framework to teach organizational structure in academic writing. *ELT Journal* 54/4, 369- 378.
- Flowerdew, L. (2011). ESP and corpus studies. In D. Belcher, A. Johns & B. Paltridge (Eds.) *New Directions in English for Specific Purposes Research* (pp. 119-144). Ann Arbor, MI: University of Michigan Press.
- Guest, M. (2013). Japanese Doctors at International Conferences: Why the worry? *JASMEE (Japan Association of Medical English Education) Journal* 12(3): 47-55.
- Guest, M. (2014). Asian Professional Discourse Communities: Pedagogical Focus on Speech Forms for Oral Performance. *Language Education in Asia Journal*, 5(1), 129-141. <http://dx.doi.org/10.5746/LEiA/14/V5/I1/A10/Guest>
- Huan, N. (2013). Beliefs about support for teacher change in English for specific purpose university classes. *New Zealand Studies in Applied Linguistics*, 19(2), 36-48.
- Hutchinson, T. & Waters, A. (1987). *English for Specific Purposes: A learning-centered approach*. Cambridge: Cambridge University Press.
- Jackson, J. (2002). Reticence in second language case discussions: anxiety and aspirations. *System*. 30(1), 65-84.
- Johns, A. & Dudley-Evans, T. (1991). English for Specific Purposes: International in scope, specific in purpose. *TESOL Quarterly*, 25, 297-314.

- Lo, Y.H. (2012). ESP versus EGP: A Case Report of an ESP programme for vocational high school students of Tourism. *Taiwan International ESP Journal*, Vol. 3(2), 71-100.
- Shih, M. (1999). More than practicing language: Reading and writing for Asian settings. *TESOL Journal*, 8(4), 20-25.
- Strevens, P. (1988). ESP after twenty years: A re-appraisal. In M. Tickoo(Ed.), *ESP: State of the art* (1-13). SEAMEO Regional Language Centre.
- Swales, J.M. (1990). Genre Analysis: *English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Terauchi, H. (1995). Issues in English language teaching in Japanese universities. *The Japanese Learner* 7 (2), 150–171.
- Terauchi H., Yamauchi, H., Noguchi, J. & Sasajima, S. (Eds.), (2010). 21世紀のESP—新しいESP理論の構築と実践 (*ESP in the 21st century – in Japanese*). Tokyo, Japan: Taishukan Press.
- VOICE (2013). *The Vienna-Oxford International Corpus of English*. (version POS Online 2.0) [Accessed Feb. 4, 2016].
- Widdowson, H.G. (1983). *Learning Purpose and Language Use*. Oxford: Oxford University Press.
- Wolcott, H. (1994). *Transforming Qualitative Data: Description analysis and interpretation*. Thousand Oaks, California: Sage
- Yum, J. (1994). The impact of Confucianism on interpersonal relationships and communication patterns in East Asia. In L. Samovar & R. Porter (Eds.) *Intercultural Communication* (pp.75-86). Belmont, California: Wadsworth.

(*This paper represents the written version of the author's presentation 'Overcoming Resistance to ESP Teaching in Japan: A Case Report' given at the International ESP Conference held at Shih Chien University, Taipei, Taiwan on April 25-26, 2015)

Appendix

Getting Patient Information 1

Personal Information:

- Name _____
- DOB _____
- Sex _____
- Marital Status _____
- Nationality _____
- Occupation _____
- Next of Kin (name and relationship) _____

Medical Details:

- Blood Type _____
- Allergies _____
- Current Medications _____
- Underlying Conditions _____
- Hospitalizations _____
- Ht _____
- Wt _____
- Exercise (frequency) _____ (type/length) _____
- Smoking ____ yes ____ no (if yes, frequency: _____)
- Alcohol (frequency) _____ (type/amount) _____

HPI:

- Present Complaint:
- Duration:
- Location:
- Frequency:

O/E:

- P _____
- BP _____
- TEMP. _____
- HS _____
- RS _____



Exploring Industry Expectations of Graduating Students' Oral Communicative Ability

**Anie Attan, Abdul Halim Abdul Raof, Masputeriah Hamzah, Noor Abidah Mohd Omar
& Masdinah Alauyah Md Yusof**

University of Technology Malaysia

Bio data

Anie Attan is Professor of Applied Linguistics at the Language Academy of Universiti Teknologi Malaysia, Kuala Lumpur branch campus. Anie specialises in Language Testing & Evaluation and English for Specific Purposes. Currently she is heavily involved with assessment of students' writing performance on a nationwide scale.

Email: anie.kl@utm.my

Abdul Halim Abdul Raof is Associate Professor at the Language Academy, Universiti Teknologi Malaysia, Skudai Johor. His research interest includes Language Testing & Evaluation, Teaching & Assessment of the Speaking skill and English for Specific Purposes.

Email: m-halim@utm.my

Masputeriah Hamzah is Professor of Applied Linguistics at the Language Academy, Universiti Teknologi Malaysia. Her research interest includes English for Specific Purposes, Computer Mediated Language Learning and Discourse Analysis.

Email: m-mputri@utm.my

Noor Abidah Mohd Omar is Associate Professor at the Language Academy, Universiti Teknologi Malaysia. Her research interest includes Language Testing & Evaluation, ELT Methodology and Written & Oral Communication.

Email: abidah@utm.my

Masdinah Alauyah Md.Yusof is Associate Professor at the Language Academy, Universiti Teknologi Malaysia. She has more than 20 years of experience teaching English proficiency and TESL courses. Her interest includes ELT Methodology & Materials Development, Self-Access/Autonomous Language Learning and Written & Oral Communication.

Email: masdinah@utm.my

Abstract: This paper is our attempt to address the issue of industry expectations of the speaking ability of prospective graduates through a study involving graduating students and workplace professionals. Information regarding oral tasks, minimum standards and quality expected from new graduate employees were gathered through interviews and discussions with the human resource personnel from various industries. Based on the information gathered, a pilot test of group oral interaction was designed and administered to a group of four graduating students. The interactions of the group were video-recorded and were then assessed by professionals from various specialisations based on their respective criteria of assessment. The professionals' responses, comments and suggestions to interview questions posed were noted. The results were analysed to establish the different categories of criteria being applied by the professionals in their assessment of the new graduates' performances. Findings show that thinking ability, interactive ability and professional image, which go beyond language skills, were other equally important criteria of assessment, besides language accuracy. From the analysis, the construct of oral communication ability for an exit oral test was determined.

Keywords: oral language and communication construct, minimum oral standard expected, workplace assessment criteria, industries' expectations.

1. Introduction

English language proficiency has traditionally been one of the determinants for the acceptance of applicants into organisations or countries for purposes of vocation, academic study and/or residence. Applicants' competency score or grade in the language is normally obtained through a language proficiency test, either an entrance or exit. The International English Language Testing System (IELTS) is an example of an established, globally recognised entrance test. Currently, foreign students seeking admission into certain courses and institutions in countries where English is the medium of instruction are expected to show a required minimum grade attained as evidence of their ability to cope with academic study and to perform future tasks in non-test contexts. On a similar note, in the local context, applicants aspiring to pursue tertiary education must show certification of a functional command of the English language, in this instance, a required level on the Malaysian University English Test (MUET), expected by the receiving institutions, as proof of their capacity to cope with future academic study and tasks. An exit test, on the other hand, is one administered to candidates at the end of their academic study. Opinions vary as to the contents in which the exit test should aim to measure. To some, it should not be different from an achievement test, i.e., that the test should aim to measure whether the candidates have shown progress in their acquisition of the skills related to the English language for the duration of their academic study. To others, it should be as close to an employability test as possible (Fugate et al., 2004) aiming to measure whether the candidates have shown acquisition of skills, including English language proficiency, as an indicator of their readiness for employment. Given the above different focus on language test performance, this implies that there exist a theoretical construct and availability of relevant and adequate tests for describing language test performance for each purpose, to correspond to the language abilities needed for expression of intentions in non-test contexts of language use (Bachman, 1991).

In setting up the criteria for describing language test performance and validating the theoretical construct, opinions vary as to whose standards should be more useful. In the case of an exit test for graduating students in which there is none presently, we strongly believe that the gap in exit language and communication certification of these students is addressed and they be informed that a measure of their workplace readiness communicative ability is being developed. In light of this, two issues need to be addressed. First, from whose

perspective should these concerns be addressed? Should it be from the perspective of the academics or the employers? Second, what criteria should be applied to determine the candidates with the required oral communicative ability? In our earlier paper (Attan et al., 2012) we have identified difficulties with determining the exit written construct of the Test of English Communication Skills (TECS) seen from both the Real Life Approach and Components Ability Approach perspectives. From the Real Life Approach perspective, specifically seen from the point of view of domain experts, because the students have yet to join the job market and thus lack the experience and abilities needed for effective performance in the workplace, any measurement of expected abilities may be deemed inaccurate. From the Components Skills Approach perspective, since the expectations are viewed from the academics' point of view, they may be unrealistic and its adoption may not meet actual skills needed for performance of tasks in the workplace, and this again may result in inaccuracy of measurement. Similarly, we believe the same argument applies for determining the exit oral construct of the Test of English Communication Skills.

On that note, we propose to adopt a combination of ideas from the Canale and Swain's (1980) model of communicative competence and the United States Secretary's Commission on Achieving Necessary Skills (SCANS) (CSE Technical Report, 1992) for the conceptual framework of the oral construct. Canale and Swain's model claims that a person is said to be communicatively competent if he has within his language repertoire, a knowledge of the way language is shaped by cultural conventions in particular communities (sociolinguistic competence), control of the linguistic system (linguistic competence), ability to cope communicatively if one's linguistic resources are inadequate (strategic competence) and a knowledge of how texts are organised in speech and writing (strategic competence) to match his specific purposes for communication. The SCANS study found five competencies which are based on a three-part foundation, i.e., basic literacy skills, thinking skills and personal qualities, as requirements for effective participation of workers in today's and tomorrow's workplace. The construct of exit proficiency would thus include components of linguistic, strategic, sociolinguistic and discourse competencies as well as basic literacy ability, thinking ability and personal qualities.

Review of Literature

Despite its increasing prominence, a literature search on validation of the MUET for entry into universities from matriculation and undergraduate test takers' opinion revealed only two. A study by Abu Kassim et al. (2007) comparing the university matriculation students' performance on the MUET and the university's English Placement Test (EPT) showed that the MUET and the EPT appear to be measuring the same construct, although they are not interchangeable. This was supported in the findings of the listening and reading sub-tests of both tests, in which they loaded on the same factor, suggesting that the sub-tests were measuring the same underlying constructs. Similarly the findings of the MUET and EPT speaking sub-tests loaded on the same factor while the writing sub-tests on another. Additionally, findings of the EPT reading sub-test loaded on the same factor as the writing, suggesting that the writing and reading skills were closely related. The explanation for this close relationship was that the summary writing task in the EPT writing test was based on the reading passage of the EPT reading sub-test. In a separate study, Rethinasamy and Chuah (2011) compared first-year students' performance on the MUET and the university Preparatory English 1 (obligatory for students achieving Bands 1, 2 and 3 on the MUET). The findings revealed that there is a significant positive relationship between students' MUET band and Preparatory English 1 grades, in which those who achieved a higher MUET band score were indeed performing well in the university English language proficiency course. Similarly, students who achieved lower MUET band scores tend to obtain lower grades in the Preparatory English 1 course. It was further highlighted that since the benchmark study reported by the Malaysian Examination Council (2005) showed a good correlation between IELTS and MUET bands ($r=0.662$), it goes to show that MUET is a reliable measure of the students' English language ability. In like manner, they asserted that "it can also be inferred that the English Preparatory 1 test results which correlate strongly with MUET band is an indication of its validity" (p.243).

Validation of the oral construct to elicit candidates' speaking communicative readiness for the workplace from domain experts' point of view has shown a mixed response. In Brown's (1993) study of test-takers' response on an oral proficiency test of Japanese for the tourism and hospitality industry, findings confirmed that the test tasks were indeed eliciting the required language response relevant to the industry. Thus the results were treated as evidence that the test was measuring appropriate language skills. In a separate study, Elder (2007)

compares the adequacy and relevance of the International English Language Testing System (IELTS) and the Occupational English Test (OET) as screening device to measure the communicative competence of health professionals for registration. In the study, it was found that while the IELTS could be considered a valid measure of general English proficiency, the OET was, over and above, a better measure of the health professionals' communicative competence. From the above two studies, it can be shown that feedback from domain experts can be considered by responsible bodies in determining the relevance of the test construct and workplace communication readiness of the participants to the target language use domain.

In other contexts, validating the oral construct is a challenge. In Bessette's (2005) study of the implementation of a proficiency testing policy aimed at ensuring adequacy of civil servants' provision of service in both French and English in bilingual regions in Canada, findings showed negative feedback received from the civil servants on many aspects ranging from the training programme to the test itself, which they considered irrelevant to their jobs. Similarly, Murray, Riazi and Cross's (2012) study in New South Wales, Australia, on the relevance of the Professional English Assessment for Teachers (PEAT), a professional screening test, found negative attitudes shown by the qualified overseas-trained teacher test takers towards the test as they felt that the test was irrelevant to their needs. Further negative feedback was found in Kim and Elder's (2014) exploration of pilots' and air traffic controllers' perceptions of tests administered in Korea. An overwhelming majority of the subjects perceived that the test failed to measure their communicative competence in radiotelephony communication as the test items were unsuitable and unrelated to the tasks that they perform and that the test procedures to carry out the tasks were ambiguous. In like manner, Knoch's (2014) study of the criteria used by native English speaking pilots to evaluate speech samples from a number of different aviation English tests, found a wider range of criteria being applied such as technical knowledge, experience and level of training, besides those set by the International Civil Aviation Organisation guidelines. With a mixed response obtained on the above validation studies from the views of domain experts, more research needed to be done to obtain more conclusive results.

For our purpose to determine the workplace oral communicative ability construct, since our graduating students have yet to obtain work experience and to be immersed in actual work context, we believe the appropriate persons for us to obtain relevant information would be the human resource personnel from technical and technology-related industries. Since these

personnel will be responsible for recruitment of deserving applicants, they would be in a better position to assess the quality of potential recruits for their organisation. Based on the above scenario of a gap in the exit oral communicative test construct development, two important considerations have prompted us to conduct the study. First, the critical need to know the employers' expectations of graduating students' oral communicative ability in light of the graduating students' effective participation in the workplace. Second, the urgent need to develop a rating scale that can accurately measure the true oral communicative ability of the graduates when they enter the job market.

In this paper, we attempt to seek answers only to our first concern, i.e., to identify and describe the workplace oral language and communication construct. First, we discuss the process and procedures that we have adopted to determine that oral construct; and next, we present our proposed construct of workplace oral language and communication, based on the findings obtained, for feedback and suggestions. Two research questions have been designed in our attempt to establish the workplace oral language and communication construct. They are:

Research Question 1 : What criteria do employers apply when assessing the quality of oral output?

Research Question 2 : What is the minimum level expected by employers when recruiting potential employees?

3. Research Methodology

3.1 Participants of the Study

In our attempt to find answers to the research questions, three groups of participants were identified. One was a group of five English language practitioners from Universiti Teknologi Malaysia (UTM) with teaching experience ranging between 18 and 25 years. The second group comprised nine workplace professionals from various technical and technology-related industries. These professionals were either Human Resource Managers or personnel responsible for recruitment in their respective organisations in Malaysia. These two groups served as informants or assessors for this study.

The third group of participants was a group of four final year undergraduates of UTM; three male students and one female student (see Table 1). All, with the exception of two male students, had different English language proficiency levels based on their Malaysian University English Test (MUET) band scores.

Table 1: Background of Student Participants

Student	Gender	MUET Band	Faculty
A	Male	Band 5	Education
B	Female	Band 4	Management
C	Male	Band 2	Engineering
D	Male	Band 2	Engineering

3.2 Procedure of the Study

The study involved two phases, adapting the procedures used by Abdul Raof (2002, 2011). Prior to the video recording of the group of students (see Table 1), preliminary input were obtained through informal discussions with a few Human Resource personnel to guide the researchers in designing the oral communicative task. Based on the information gained, a speaking task was designed based on the assumption that new graduate employees would be involved in a discussion with fellow employees. For the pilot study, the students were asked to deliberate, give their views, interact with each other and eventually reach a collective decision on what they think is the most important quality in a worker. The task lasted for about 20 minutes and was moderated by an English language practitioner. The session was video-recorded and used as speech sample in this study.

In the first phase, the video-recording was then shown to all five English language practitioners. Each was asked to view the recording, rank and assess the oral interaction of the four students without referring to any rating scale. After the assessment was done, a group interview cum discussion was conducted with the language practitioners to compare rankings and to know how these were determined. The criteria used were then compiled.

This procedure was repeated in the second phase, this time with the workplace professional group. Each personnel was approached individually and was asked to view the video-recording, rank and assess the students. As with the language practitioners, it was done without the aid of any rating scale. At the end of the assessment, each of them was subjected to a semi-structured interview session. Among the questions posed were:

1. Who have you ranked as the best?
Why have you ranked this student as the best?
What other qualities would you like this student to possess?
2. Why have you ranked X higher than Y?
3. Why have you ranked Z the lowest?
4. Who would you recruit as an employee in your firm?

Data from both groups were analysed to look for similarities and differences in the rankings made. In addition, data were further studied to elicit the criteria used by the two groups of assessors for assessing and ranking the students.

4. Findings and Discussion

The analysis of both the interviews and criteria used revealed that there is a disparity in the ranking of the students' performances made by the language practitioners as compared to the assessment made by workplace professionals. From the analysis of the interview responses, it was evident that the language practitioner group based their assessment of the students mainly on the criterion of language proficiency, as they were used to assessing students in the context of language learning where emphasis is more on the ability of students to use the language appropriately and accurately. The criteria used were mainly on correctness and variety of language used. The language practitioners group had no dispute over the best speaker, i.e. Student A. However, none of the workplace professionals ranked Student A as the best. Instead, majority of the workplace professionals chose Student B as the one who had performed really well in the task given.

Student A, who demonstrated good control of the English language, was the choice of language practitioners as language proficiency was regarded as key to a student's performance. From the perspective of the language practitioners, a student with such

proficiency is regarded as having achieved the desired standard of oral competency and thus, is expected to do well in the workplace. On the other hand, the workplace professionals were not impressed by the language proficiency of Student A. To them, Student A's overall performance was not substantial as he did not contribute much to the task assigned. He was merely managing the discussion; initiating the discussion, asking opinion from the others, making sure the task gets done and concluding the discussion. Basically, workplace professionals looked at overall competency not just in terms of language ability but the capacity to fulfil the task requirement in terms of expression, elaboration and expansion of thoughts and ideas on the given topic in a convincing way. Student A was seen as too engrossed in ensuring the overall flow of managing the discussion, to the point that his contribution was regarded as minimal and lacking in substance. His 'voice' was not heard and hence, he failed to impress the workplace professionals.

In contrast, Student B was ranked highest by eight out of nine workplace professionals on the basis of her maturity in presenting her arguments and her interactive ability to express her opinion and respond to others. She was seen as able to make significant contribution to the discussion and elaborating her viewpoints. She also listened to what others had to say and considered their alternative opinions, while at the same time defending her arguments in a convincing way. Though her language proficiency was good enough (but not as good as Student A's), this was not the basis why she was chosen as the best by the workplace professionals. What impressed them were clearly elements beyond language performance. The language practitioner group, on the other hand, ranked Student B as second behind Student A as the main criterion from their perspective still revolves around language proficiency in terms of fluency and accuracy in language use.

Another point worth noting is the workplace professional group's impression on other criteria which they regard as important, referred to as the overall 'package'. This is reflected in the comments made of Student B, where the workplace professionals felt that despite her good performance, she could still improve on the image she portrayed. From the interview responses, this criterion could be described as the image one carries or overall presence when communicating with others. Also included is one's confidence when voicing views and opinions, and composure which make up the overall demeanour or 'package'. It was also evident that this aspect of the students' performance was not given much emphasis or importance in the ranking criteria of the language practitioners. In fact, this overall package

in terms of positive image, composure and confidence which are important elements that will leave a good impression of a person's personality are somehow overlooked by the language practitioners, and thus was not included in the assessment or ranking criteria.

With regard to the performance of Students C and D, the language practitioners ranked both as equal, i.e., the lowest due to their lack of proficiency, again a language-based criterion. Despite this, the workplace professionals, however, had a more positive view and were willing to recruit Student C, for instance, as he was seen to have some potential based on the contributions made in carrying out the task. Student C also demonstrated the ability to defend his arguments, though to a lesser extent due to his lower proficiency in the language, as compared to Student B. One Human Resource Manager even thought that Student C was the best candidate compared to the others apparently due to the potential that he had as observed by the personnel. From here, it is clear that language proficiency was not used as the main criterion in determining the choice of candidate to be recruited. There were qualities deemed positive shown by the lower proficiency students, like enthusiasm and determination, qualities which can be instilled and nurtured as part of the recruitment process.

In evaluating the quality of oral interaction based on Research Question 1, it seems that while the language practitioners were paying more attention to language skills (as expected), the workplace professionals were focusing on some other criteria beyond language ability. These were found to relate to attributes deemed important in the workplace as part of one's profession such as thinking ability, interactive ability and professional image. Based on these criteria, rankings given by both groups were therefore different.

With reference to Research Question 2, the discussion held with both groups of assessors was aimed at determining the minimal level of language ability expected by employers for acceptance of potential employees into the workplace. It was found that what merits as 'acceptable level' of oral communicative ability for the workplace includes the ability to express opinions, to respond to viewpoints, to participate in discussions, to be able to analyse problems, to have adequate amount of content and level of maturity to dwell further on the topic or discussion, and to keep the communication going with a fair command of the language though still lacking in confidence and sophistication of language. These abilities were deemed as Functional, that is, equivalent to Level 3 on a 6-Level scale. A graduate at this level would still have many errors in his language and poor range of vocabulary.

Nevertheless, this is regarded as the minimum acceptable level that can sustain the potential employee in the workplace. With exposure and further on-the-job training, the new employee is expected to be able to gain better control of the language in the workplace as time progresses and will eventually be able to use the language more confidently and appropriately.

Our proposed workplace oral language and communication construct, conceived from the findings in the collaborative study between the language practitioners and workplace professionals, is captured in the table as follows:

Table 2: Proposed Construct of Workplace Oral Language and Communication Competence

Criteria	Description
Contribution to Task	Able to analyse topics/issues/problems Shows creativity of ideas Shows maturity of ideas
Interactive Ability	Able to express opinions Able to convince
Language	Accuracy Fluency Wide range of vocabulary Sentence variety
Personality	Professional image Confidence Enthusiasm

5. Conclusion and Suggestions

The approach advocated by this study was empirically-based and utilised feedback and insights of workplace professionals. Similar to the study by Abdul Raof (2002, 2011), this

study proves that engaging workplace specialists in determining the oral ability construct of the graduating students is possible, in fact highly recommended. Although it would benefit from a larger number of workplace professionals involved, nevertheless from the analysis of the data gathered, a rating scale of oral communication ability for the oral exit test could be devised. It comprised a more wholesome criteria of assessment deemed significant as part of workplace expectations which include thinking ability, interactive ability and professional image, apart from language ability. Continuous collaboration with workplace specialists would be the way forward to further validate the draft scale before it can be widely used and accepted. More importantly, the study highlights the importance of engaging workplace professionals in the design of the exit test, most critically in determining their assessment of new graduates' performance, which this study reveals, go beyond mere language skills.

References

- Abdul Raof, A. H. (2002). *The production of a performance rating scale: An alternative methodology*. Unpublished Ph.D thesis. The University of Reading, UK.
- Abdul Raof, A. H. (2011). An alternative approach to rating scale development. In B. O'Sullivan (ed.), *Language Testing: Theories and Practices*. Basingstoke: Palgrave Macmillan.
- Abu Kassim, N., Zubairi, A. & Mat Daud, N. (2007). A construct validation of the Malaysian University English Test and the English Placement Test: Two high-stakes English language proficiency tests. *Malaysian Journal of Language and Instruction*, Vol. 4, p.1-16.
- Attan, A., Abdul Raof, A. H., Mohd Omar, N. A., Abdullah, K. I. & Hamzah, M. (2012). Establishing the construct of workplace written communication. *Procedia Behavioural and Social Sciences*, Vol. 66, 76-85.
- Bachman, L. F. (1991). What does language testing have to offer? *TESOL Quarterly*, Vol. 25 (4), p.671-704.
- Bessette, J. (2005). *Government French language training programs: Statutory civil servants' experiences*. Unpublished Master's thesis, University of Ottawa, Canada.
- Brown, A. (1993). The role of test-taker feedback in the test development process: Test-takers' reactions to a tape-mediated test of proficiency in spoken Japanese. *Language Testing*, Vol. 10 (3), p.277-301.
- Canale M. & Swain M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1:1, 1-47.
- CSE Technical Report 343 (1992). *Measurement of workforce readiness: Review of theoretical frameworks*. National Center for Research on Evaluation, Standards and

Student Testing, Graduate School of Education, University of California, Los Angeles, USA.

Elder, C. (2007). *OET-IELTS benchmarking study report*. Australia: Language Testing Research Centre, University of Melbourne.

Fugate, M., Kinicki, A.J. & Ashforth, E. (2004) Employability: A psycho-social construct, its dimensions and applications. *Journal of Vocational Behaviour*, Vol. 65, p.14-38.

Kim, H. & Elder, C. (2014) Interrogating the construct of aviation English: Feedback from test takers in Korea. *Language Testing*, 1-21.

Knoch, U. (2014). Using subject specialist to validate an ESP rating scale: The case of the International Civil Aviation Organization (ICAO) rating scale. *English for Specific Purposes*, p.33, 77–86.

Malaysian Examination Council. (2005). *Benchmarking report: Correlation between the Malaysian University English Test (MUET) and International English Language Testing System (IELTS)*. Kuala Lumpur: Malaysian Examination Council.

Murray, J. C., Riazi, A. M., & Cross, J. L. (2012). Test candidates' attitudes and their relationship to demographic and experiential variables: The case of overseas trained teachers in NSW, Australia. *Language Testing*, Vol. 29 (4), p.577–595.

Rethinasamy, S. & Chuah, K. M. (2011). The Malaysian University English Test (MUET) and its use for placement purposes: A predictive validity study. *Electronic Journal of Foreign Language Teaching*, Vol. 8 (2), p.234-245.



Examining Business English Majors' Business Vocabulary Knowledge Development

Luo Yan

Teacher's College, Columbia University

Yingli Yang

School of International Studies, University of International Business and Economics, China

Bio data

Luo Yan is a graduate student at Teacher's College, Columbia University.

Email: 942116481@qq.com

Yingli Yang is associate professor at the School of International Studies, University of International Business and Economics, China. Her research interests include second language acquisition and ESP. Her publications appear in *Studies in Second Language Acquisition*, *Journal of Pragmatics*, *The Journal of Asia TEFL*, etc.

Email: yyluibei@126.com

Acknowledgement:

This study is funded by a research grant from China National Natural Science Foundation (31400896). We really appreciate the suggestions and feedback provided by the reviewers and the associate editor Dr. Tanju Deveci.

Abstract

While previous studies have investigated vocabulary size of English and non-English majors in China, little research has been conducted on the mastery of discipline-specific vocabulary knowledge by Business English majors. This study examined knowledge of business academic vocabulary by business English majors across four grade levels. A vocabulary test adapted from Hsu's (2011) *Business Academic Word List* was administered as the measuring instrument. One hundred and twenty seven students majoring in business English participated in this study. Although students generally demonstrated better mastery of business vocabulary as grade level increases, their overall scores of business vocabulary at different word frequency levels did not show a consistent pattern. In addition, according to results of ANOVA and *post-hoc Scheffe* analysis, a significant difference was found among four grades on their mastery of words at different frequency levels. The students' scores on level 1 vocabulary were similar, whereas their scores on level 6 were most widely spread. Results are interpreted in terms of vocabulary test items and the courses in the curriculum and implications on vocabulary test and curriculum development are discussed.

Keywords: Business English, vocabulary list, word frequency

Introduction

In the field of second language acquisition, researchers generally contend that knowledge of essential vocabulary is a prerequisite for fluency in reading and writing (Hu & Nation, 2000; Milton, 2013), the learning of grammar (Ellis, 1997), and for developing listening and speaking proficiency (Milton, Wade, & Hopkins, 2010; Milton, 2013; Zimmerman, 2004). Developing proficiency in the four skills thus requires mastery of a certain amount of vocabulary. According to Nation (2001), a vocabulary of 6000-7000 is required for comprehending general spoken texts while a vocabulary of 8000-9000 vocabulary is required for general written texts. Building on data collected from 745 college students in Israel, Laufer and Ravenhorst-Kalovski (2010) found that a minimal threshold for reading comprehension of academic texts for English as a Foreign Language (EFL) learners is 4,000 to 5,000 words and an optimal threshold for reading academic texts is 6,000 to 8,000 words.

Although English is a required subject in middle and high school education, the vocabulary size of college students in China has not reached the satisfactory level (Cai, 2012; Ma, 2001). This is partly because for non-English majors, English courses for general purposes are only provided in the first and second year and the overall teaching goal of developing competency in English may not allow teachers to focus on vocabulary teaching in the limited time for classroom instruction. For non-English majors, the English classes mainly focus on the comprehension of the text and speaking and listening exercises. In addition, three hours of classroom instruction may not provide enough input and output opportunities to develop adequate vocabulary knowledge. According to *College English Curriculum Requirement 2007*, the basic requirement for recommended vocabulary size is a total of 4,795 words and 700 phrases (Cai, 2012). However, a study of 303 non-English majors in Chinese tertiary level context indicated that only senior students reached the required level (Gu & Li, 2013), due to the limitation of vocabulary offered in their course books as well as the lack of focus on vocabulary acquisition in university level classes. Until now, relatively little is known about Chinese EFL learners' English vocabulary size. To our knowledge, only very few studies have been conducted on the knowledge of business vocabulary of Business English majors.

As a sub-branch of English for specific purposes, Business English is now taught in universities around the world. Learners of Business English need to master core business vocabulary in order to communicate effectively in business contexts. According to Hsu (2009), university-level English-medium level business core textbooks covered a smaller number of lexical items (3000-4000 word levels along BNC) compared to General English textbooks (3000-13000 word levels). As a result, when students finish their university level English classes, they may not have acquired adequate amount of vocabulary in preparation for graduate level readings of academic journal articles and research papers (Hsu, 2011). To our knowledge, Hsu's studies were among the few studies that examined business English vocabulary. However, how students in mainland China develop their vocabulary size over the four years' study at university level and how well they prepare for studies at the graduate level is unknown. The main purpose of this study is therefore to investigate the mastery of business vocabulary by business English majors across four grades in a key university in Beijing, using an abridged form of Business Word List adapted from Hsu's Business Word List (2011).

Literature review

Some consensus has been reached among the definition of vocabulary acquisition for English learners. Ellis (1994) points out that “in the first instance at least, the acquisition of L2 words usually involves a mapping of the new word form onto pre-existing conceptual meanings or onto L1 translation equivalents as approximations” (pp.133-134). Nation (2001) defines what should be taught to the students about vocabulary. He distinguished three overarching aspects of vocabulary knowledge, each with three subcategories: form (spoken form, written form, word parts), meaning (form and meaning, concept and referents, associations), and use (grammatical functions, collocations, constraints on use).

In addition to form, meaning and use, Nation (2001) points out receptive and productive vocabulary knowledge. The former refers to recognizing the meaning of the words and linking them to a translation in the first language while the latter refers to the abilities to pronounce, spell or use words properly in appropriate contexts. The receptive vocabulary is sometimes also referred to as “vocabulary breadth” or “vocabulary size”, in contrast to “vocabulary depth” (Milton, 2010). Vocabulary depth is a more complicated concept which involves several facets of vocabulary knowledge, such as associational meaning, collocational knowledge, inflectional and derivational knowledge, knowledge of concepts and referents, and knowledge of constraints on use (Read, 2000). Later, researchers added a new dimension to vocabulary knowledge, arguing that fluency or the automaticity in recognizing and processing the words is another important aspect in vocabulary knowledge (Daller, Milton, & Treffers-Daller, 2007). This aspect of vocabulary knowledge, however, is more dynamic in nature and points out to the processing or acquisition perspective of vocabulary rather than the declarative vocabulary knowledge *per se*, drawing on the assumption that cognitive resources are limited in speech production (Schmitt, 2010).

Despite the various facets of vocabulary knowledge presented in previous research, the present study mainly examines the receptive knowledge of business academic vocabulary since strong correlations were found between receptive vocabulary size and reading comprehension scores with learners from different proficiency levels, possibly due to the fact that reading comprehension is also a receptive skill (Laufer, 1992; Qian, 1999; Albrechtsen, Haastrup & Henriksen, 2008). Furthermore, empirical evidence shows that there is a link between receptive vocabulary knowledge and language performance in general. For example,

Milton (2010) examined 10,000 learners in Greece taking both recognition tests of vocabulary and formal language exams within the Common European Framework of Reference for Languages (CEFR) and found that learners progress through CEFR levels as their vocabulary knowledge increases.

In the Chinese context, Gao (2015) examined the relationship between vocabulary size and accuracy rate in reading comprehension among college students. Gao found that vocabulary size could significantly predict the accuracy rate in reading comprehension. Wang (2006), on the other hand, investigated the association between vocabulary size and various foreign language skills demonstrated in CET 4¹ test including listening, speaking, vocabulary and structure, reading comprehension and writing. Results showed that vocabulary size had a strong correlation with overall language skills, writing and vocabulary and structure part of the CET 4 test, a moderate correlation with speaking and reading comprehension, and a weak correlation with listening comprehension and cloze test. Examining the process of vocabulary development thus would enhance our understanding of the quality of English language teaching at tertiary level and provide some insights regarding the proficiency level of students.

Academic and business vocabulary list

Building on these definitions of vocabulary acquisition, a number of scholars attempted to compile a list of essential vocabulary in a second language. One of these scholars is West (1953) who compiled *the General Service List of English Words* (GSL) consisting of the most frequently 2,000 word families of English. The list covers approximately 90% of colloquial speech and 80% of written texts. However, because GSL was created in 1953, more recent words were not included in this word list. Nation & Waring (1997) stated that beginners of English learners should first become familiar with the first 2,000 most frequently occurring word families. For further studies for academic purpose, they should acquire specialized academic words.

¹ CET is an acronym for College English Test, a national English proficiency test for all English majors. The subcomponents include listening, speaking, vocabulary and structure, cloze, reading comprehension and writing.

Later, Coxhead (2000) compiled *The Academic Word List* (AWL) based on a corpus of 3.5 million running words of written academic text, and contained 570 word families which exclude the most frequently occurring 2,000 word families. The AWL covered 10% of the total words in a general academic text, but only covered 1.4% of the total words in a fiction with the same vocabulary size. The AWL provided a word list for non-native English learners who wish to further their education academically. This list also motivated further research on the coverage and application of academic vocabulary list in various types of discipline-specific discourse types (Chen & Ge, 2007; Shabani & Tazik, 2014; Friginal, Walker & Randall, 2014). These studies generally show that although AWL covers a significant portion of lexical items used in medical research articles, ESP and EFL research articles, Google Ngram Viewer and the Corpus of Historical American English, it definitely needs to be updated to reflect the latest trend in the use of academic vocabulary in large corpora, since every year, there is a large number of new words emerging in the academic texts.

Following Coxhead, Nation & Heatly (2002) created the RANGE program which combined the 2000 word families from West's (1953) General Service List, 570 words families from Coxhead's (2000) *Academic Word List*, fourteen British National Corpus 1,000 word lists and proper nouns and marginal words. The RANGE program is a free downloadable program that can be used to compare a text against the installed fourteen BNC 1000 base word lists. The program shows the percentage of the vocabulary items in the text that have been covered by the lists. Additionally, it could also be used to compare the vocabulary of many text files at a time. It was used as a tool for the analysis of Hsu's (2011) business word list (BWL). Recently, a new General Service Word list was compiled based on 273 million-word subsection of the 2 billion words Cambridge English Corpus (CEC) (Browne, Culligan, Phillips, 2013). This list is an updated version of the GSL word list developed by West (1953) and has a higher coverage rate (92.34%) than the original list GSL (84.24).

Although these general word lists provide a useful tool for language learners and teachers, it should be acknowledged that there are limitations of these word frequency lists in measuring vocabulary size, since "the accuracy of the estimates of vocabulary size is limited by the representativeness of the corpus counted" (Wesche & Paribakht, 1996, p. 4). In addition, the GSL list and the AWL both need a more recent version to reflect the changing nature of vocabulary use in respective discourse contexts. In order to examine business English majors'

vocabulary size, a more recent and rigorous discipline specific vocabulary list should be used to examine learners' vocabulary development.

Various discipline-specific word lists have been created to cover lexical items in specific discourse genres. Konstantakis (2007) developed a business word list, consisting of 560 business related words. He selected these words from 33 business English course books. The first step Konstantakis took was to examine to what extent GWL and AWL cover the Published Material Corpus (PMC) compiled by Nelson (2000). Initial examination revealed that both lists cover approximately 90% of words in PMC. The second step was to investigate an additional vocabulary list (the BWL) which covers another 5% of the PMC, since 95% coverage means the mastery of these words ensures students' reading comprehension of the material. The words in Konstantakis's (2007) BWL were selected for three criteria. Firstly, the words selected should not be included in the General Service List and Academic Word List; secondly, it should appear in more than 5 course books; thirdly, it should appear more than 10 times. Moreover, names, numerals, Latin words, nationalities and abbreviations are excluded from the list. However, this word list has two limitations. First, the words were selected from the Published Material Corpus (PMC) compiled by Nelson (2000), which was based on 33 published Business English course books published between 1986 and 1996. As a result, the words in the list may not cover the popular topics in the business field in recent years. Second, the word list was not classified into different levels, and therefore cannot be used to differentiate learners' business vocabulary development.

Hsu (2011) compiled a more comprehensive business vocabulary list based on a corpus covering 7.62 million tokens of 2,200 business academic articles from online resources. The research articles covered 20 business subjects including accounting, economics, finance and other subjects. Hsu tested the lexical coverage of BWL in the business research articles through the RANGE software and three selection criteria were adopted: the specialized occurrence, range and frequency of a word family. The word families in the lists include the words outside the top BNC 3000 most frequently occurring words of English and they should occur at least in 10 business subject areas and at least 270 times in the corpus of business research articles. BWL consist of 426 words, which are the most frequently occurring word families in corpus. Those words were beyond the BNC 3000, occurring in more than 10 business subject areas and at least 270 times. The BWL covered 5.66% of the tokens in business research articles from the corpus.

Studies on vocabulary knowledge by university students in China

Students in China usually undergo four years of study at the tertiary level. They are usually offered English lessons two to four times a week throughout their four years' study. Results of previous studies suggest that despite extensive reading and classroom teaching at the tertiary level, students' vocabulary level across four years of study at university level remain rather limited and does not meet the requirement from the national curriculum. For example, Jiang (2009) found that at a vocational college, freshman's average vocabulary size is 1,116. In two studies that examined vocabulary size of first-year and second-year university students, Deng (2001) studied the vocabulary size of freshman and showed that their average vocabulary size was 3,500; Lv (2004) investigated the vocabulary size of freshman and found that their average vocabulary size is 2,145. Ma (2001) tested non-English major's vocabulary size across four grade levels and showed that their average vocabulary size is 3,000, with an increase of 200 words per semester, but it is still far below the required vocabulary size. Similarly, Dai (2013) studied 125 non-English majors' vocabulary size of at university in Nanjing and found that sophomores had an average vocabulary size of 3,934, lower than the requirement (4,795) of the national curriculum.

In reviewing studies on vocabulary size of students at tertiary level, we found that compared to the study of vocabulary size of non-English majored university students, studies about the vocabulary size of English major students in China remain limited. The limited research demonstrate that students generally increase vocabulary size as grade level increases, and the largest increase occur from grade one to grade two.

Gui (1985) showed that the freshman had an average vocabulary size of 1,200 words. Xi (1998) conducted a study on the vocabulary size of 558 students from four grades major in English in Xi'an International Studies University. The results showed vocabulary size at the end of first semester was about 4,825, third semester 6,283, fifth semester 6,804, and seventh semester 6,690. Using Ma (2001)'s vocabulary test, Xu (2010) studied the vocabulary size of 450 English major students in Nanjing Normal University, and found that at the end of the first semester, students' vocabulary size was 4,510, 5,343 at the end of the third semester, 5,872 at the fifth semester, and 6,580 at the seventh semester.

Although these studies showed a generally consistent pattern of vocabulary knowledge development across grade levels, most studies mainly investigated knowledge of English vocabulary for general purposes. To our knowledge, few studies have targeted vocabulary knowledge of a specific discipline, especially business vocabulary knowledge. Business English has become a major in the year 2007, and now there are over 270 universities and colleges in China which has Business English as a major (Chen & Wang, 2009). A number of studies show that vocabulary size explain a large percent of the variance in four language skills (Milton, Wade, & Hopkins, 2010; Schooner, Van Gelderen, Stoel, Hilstijn, & De Glopper, 2011). Second language acquisition theory such as Ellis' (1997) Lexical Learning Hypothesis also posits that vocabulary development is very closely related with other features of language development, such as grammatical development (Milton, 2013). As a result, examining learners' business academic vocabulary knowledge would inform ESP practitioners of the current level of learners' vocabulary knowledge, identify their learning needs, and incorporate vocabulary size into designing the curricula. Therefore, the research questions of this study are as follows:

- (1) Is there a difference in business vocabulary knowledge among students from different grades?
- (2) Does word frequency influence students' business vocabulary knowledge?
- (3) To what extent does word frequency influence students' business vocabulary knowledge across grade levels?

Methodology

Participants

A hundred and twenty-seven undergraduate students majoring in business English from a key university of business and economics in Beijing participated in the study on a voluntary basis. Six of them were male, and 121 of the participants were female. Their ages varied between 18 to 22 years, with a mean age of 20. There were 34 participants from grade 1, 33 from grade 2, 30 from grade 3 and grade 4 respectively. Moreover, for ethical concerns, they were told that the results would be used only for academic purpose and the tests were completely anonymous. Moreover, they were ensured that the results would not affect any of their scores

in any of their courses. All the students took the tests in classrooms in the researcher's presence. The participants were given 15 minutes to complete the tests. No dictionaries and other references were allowed during the tests.

In the present study, as shown in Table 1, the curriculum offered in this university first provides students with wide-angled courses which aim to train their generic skills in the field of business. And in the third and fourth year, many narrow-angled courses were added in the schedule, reflecting more discipline- specific discourse practices.

Table 1: Course schedule of business English major students

YEAR	COURSES	
First Year	Integrated Skills of Business English I&II	
Second Year	Integrated Skills of Business English III & IV	Listening and Speaking in Business English I & II
	Business English Writing I&II	Introduction to Business
Third Year	Import and Export Practices	Introduction to Business Law
	Principles of Economics	International Trade
Fourth Year	Business Communication Practice	

Vocabulary list

The present study adapted Hsu's (2011) Business Word List to test students' business vocabulary knowledge. The list included 426 most frequently occurring words chosen from 2,200 business research articles. The BWL categorized the words by their occurring frequencies in a descending order by the British National Corpus fourteen thousand word levels. As for the present study, a subset of the Business Word list was selected, including 30 most frequently occurring words at six levels from BNC 4th 1000 to BNC 9th 1000. Words that appear in BNC 1st to BNC 3rd 1000 level were not included in the test since the pilot study showed that students already passed these three levels, indicating their vocabulary size were greater than 3000 level. Moreover, the words that had appeared in the word lists of *Integrated Skills of Business English* textbooks were eliminated to avoid the possible effect of classroom teaching. In addition, in order to make the research result concise and clear, new labels were used to represent each frequency levels (See Table 2). For example BNC 4th 1000

was labeled as Level 1, which was followed by Level 2 to Level 6. All of the participants took part in the same vocabulary test composed of 30 business words divided into 6 levels according to their occurrence frequency in BNC.

Thirty words were selected for the following three reasons: 1) In Nation & Beglar (2007) vocabulary size test, 10 items were provided at each level. In order to comply to the original standard and also to accommodate for the time constraints of the present study, we decided to select 5 words at each level. Altogether, we selected 30 words from Hsu's list across 6 different frequency levels. We were consistent in our selection criteria in that at each level, we selected the five most frequently occurring vocabulary items. 2) We selected these 30 items based on a pilot study conducted among students of similar proficiency levels so that these items were within the appropriate range of their receptive vocabulary. 3) Our main purpose of research was to examine students' vocabulary knowledge across grade levels and how this may be mediated by word frequency factor rather than to estimate the learners' vocabulary size. Therefore, only a representative section of Hsu's vocabulary list was selected.

Table 2: Target word list according to ascending order by BNC

	BNC 4 th	BNC 5 th	BNC 6 th	BNC 7 th	BNC 8 th	BNC 9 th
	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Target words	asset audit statistic review transact	journal abstract equity portfolio ratio	cite default affiliate deviate compliance	empirical disclosure parameter classification setup	copyright volatile violate denote dependence	underwrite inventory collateral mitigate transparency

Measurement Instrument

The present study employed part of Hsu's (2011) Business Word List and adapted it into a vocabulary size test similar to Nation & Beglar (2007) to evaluate the mastery of business vocabulary by students majoring in Business English across four grade levels. The test mainly aims to measure participants' receptive vocabulary size and therefore provides little indication of how well these words could be used in speaking and writing. According to

Nation & Beglar (2007), the Vocabulary Size Test was developed to provide a reliable and comprehensive measure of a learner's vocabulary size. It can also compare learner's growth of vocabulary across grade levels. Since the vocabulary items were divided into different levels according to their frequency in BNC, the greatest value of the test will be in measuring students' progress in vocabulary learning (Nation, 2001). The target words were provided in complete sentences² and were highlighted in bold. The sentences were designed to be as concise as possible (See example 1. Part of the test is provided in Appendix I).

Example 1:

Target words: **asset**

Sentence: I need to categorize the **asset**.

Options: a. 资产 b. 形态 c. 装置 d. 物品

Correct Answer: a. 资产

Extra care was taken to avoid redundant words or show any implications of the word's meaning. Four different Chinese words were displayed below as options, but only one of the options was the accurate translation of the target word. The options were designed to have the same part of speech, and all of the options could make sense in the sentence. The students were required to choose one correct answer that corresponds to the semantic meaning in the sentence. The test was piloted among students who are of the same proficiency level as the participants in the same university (Mackey & Gass, 2005).

The first author marked the entire set of the test papers based on the following criteria: if the learner correctly chose one option of the meaning of the word from the multiple choices, one score was awarded. If the choice was wrong, the learner would get a score of zero for this test item (Laufer & Ravenhorst-Kalovski, 2010). The students' total scores were then calculated and recorded on an *Excel* spreadsheet. The second author checked all the test papers and the inter-rater reliability as percentage agreement was 98%³.

² Since the purpose of the test is to examine students' knowledge of business vocabulary, the sentence does not need to provide the clue with regards to the meaning of the word. Instead, it provides clues regarding the part of speech of the word in use.

³ Since this test includes multiple choice items, the two raters scored according to the standard answer sheet provided by the researcher. And the discrepancy was later resolved by two raters checking the answer together.

Data Analysis

Students' performance in the six categories was scored and calculated separately. Question 1-5 which included target words in the BNC 4th 1000 were calculated together and the composite score was labeled as the overall score for level 1. Scores for question 6-10 (BNC 5th 1000) were labeled as scores for level 2, and scores for question 11-15 (BNC 6th 1000) were labeled as level 3 scores. The same rule applies to all frequency levels. The average score and standard deviation were calculated for each student at each frequency level. ANOVAs and *post-hoc Scheffe* tests were also conducted to measure the differences in students' scores among different grade levels at various frequency levels.

Results

For descriptive statistics, the tables and graphs of average score of each level and each grade are displayed. Inferential statistics show the ANOVA results regarding whether grade had an effect on students' different performance in the tests.

Descriptive Statistics

Table 3: Total score of vocabulary test across grade levels

	Grade 1	Grade 2	Grade 3	Grade 4
Total	19.59	21.97	22.9	24.27

As shown in Table 3, the average score of participants in grade 1 was 19.59. The average score of grade 2 students was 21.97; whereas the average score of grade 3 students was 22.90, and the average score of grade 4 was 24.27.

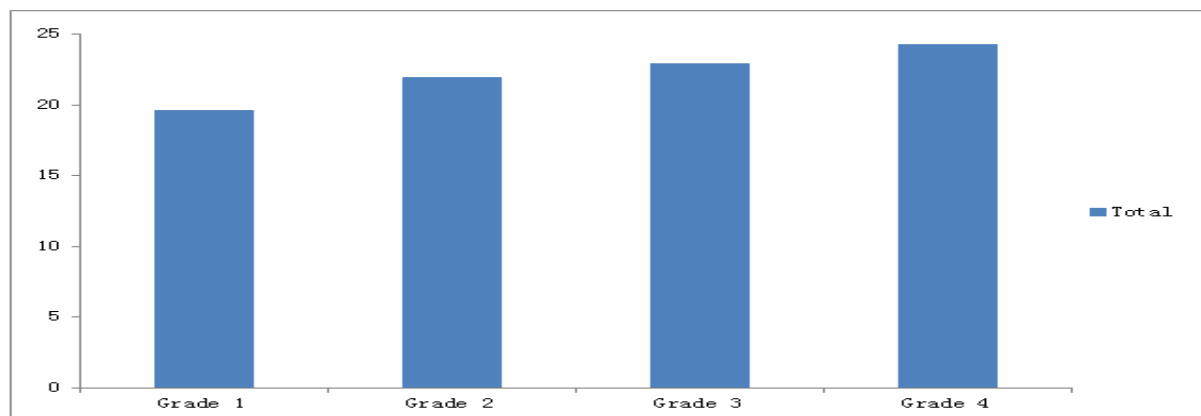


Figure 1: The total vocabulary test score across each grade level (Max 30 points)

According to Figure 1, for the total score, senior students obtained the highest score with an average of 24.27 points, which was followed by the juniors with 22.90 points, sophomores 21.97 points, and freshman 19.59 points. In the vocabulary tests, although grade 4 obtained higher score than the grade 3, grade 2 and grade 1, the biggest improvement occurred from freshman year to sophomore year, and the second largest improvement occurred from juniors to seniors.

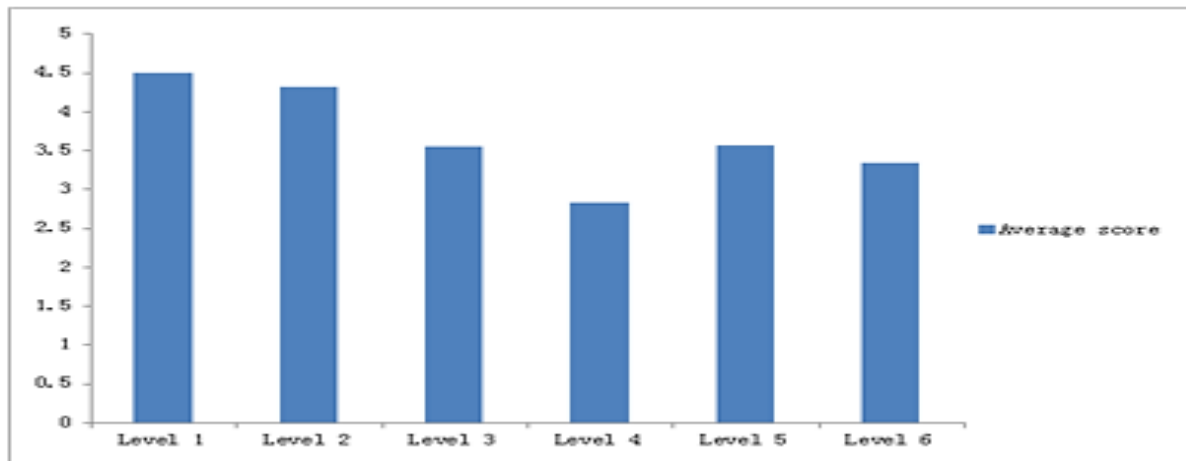


Figure 2: The average score at each word frequency level. (Max 5 points per cell)

Figure 2 shows the average score of all the participants at each word frequency level. Generally, the average score decreases from Level 1 to Level 4, but slightly increases from Level 4 to Level 5, and decreases again from Level 5 to Level 6. These results show that generally, the less frequently the words occur in business texts, the less likely students would know the meaning of the word. However, the results of Level 5 slightly deviate from this trend. Students' score at 3.56 at level 5 and then drop to 3.34 at level 6.

Table 4: The average score of students at each grade across each frequency level

LEVEL	GRADE							
	Grade 1		Grade 2		Grade 3		Grade 4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Level 1	4.32	0.72	4.63	0.65	4.60	0.67	4.40	0.81
Level 2	4.00	1.02	4.24	1.03	4.57	0.73	4.53	0.82
Level 3	3.26	0.86	3.21	1.17	3.77	1.07	4.03	1.00

Level 4	2.32	0.88	3.03	1.15	2.87	1.04	3.17	0.53
Level 5	3.24	0.96	3.58	1.23	3.47	0.97	4.00	0.91
Level 6	2.44	1.21	3.27	1.28	3.63	1.22	4.13	0.90
TOTAL	19.59	3.41	21.97	4.25	22.90	3.90	24.27	2.88

As shown in Table 4, students at grade four have the highest mean score at Level 4, level 5 and level 6. Interestingly, at level 4, where students at each grade level have the lowest mean score, students at grade four not only have the highest mean score, but also the lowest standard deviation. This means that students at grade four had similar performance at this level, whereas students' scores at other grade levels are more varied.

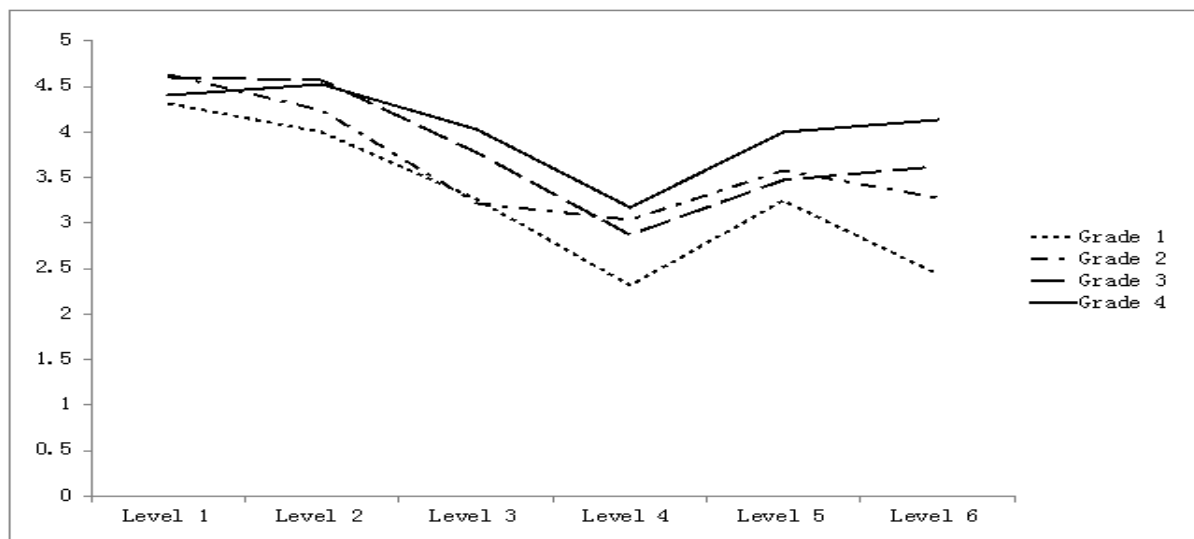


Figure 3: The mean score of students at each grade across each frequency level

It is shown in Figure 3 that the participants' mean scores are highest at level 1 and then gradually decrease from level 1 to level 4. Students at all the grades have the lowest score at Level 4. However, at Level 5, the score of all the grades increase and then decrease again at level 6. At Level 1, there is no obvious difference in the score between grades. At Level 2, students at Grade 4 and Grade 3 have nearly the same score while those at the other two grades have lower scores. At Level 3, the scores of Grade 1 and Grade 2 are approximately the same while Grade 3 and Grade 4 have higher scores. At Level 4, Grade 1 has the lowest score, and it is much lower than the other three grades. There is an increase in scores in all the grades at level 5. The score of Grade 4 is much higher than that of Grade 2 and Grade 3, while score of Grade 1 is much lower than other grades. At Level 6, the differences in scores

between four grades become large with Grade 1 the lowest and Grade 4 the highest. Interestingly, it is shown that freshmen are most sensitive to level of the word frequency in that the accuracy rates drop dramatically at Level 4 and Level 6 which are the two levels with the lowest average score. An unexpected finding is that students' scores at level 5 are higher than that at level 4. A closer examination of participants' score for each word and the accuracy of four words shows that the scores of the words "parameter" (level 4), "setup" (level 4), "copyright" (level 5), "dependence" (level 5) are shown in the following table.

Table 5: Accuracy rate of four grades on the four words (Accuracy Rate, %)

	LEVEL 4		LEVEL 5	
GRADE/WORD	Parameter	Setup	Copyright	Dependence
Grade4	26.67	6.67	93.33	86.67
Grade3	26.67	16.67	93.33	63.33
Grade2	27.27	33.33	87.88	72.73
Grade1	29.41	17.65	88.24	50.00
Average	27.50	18.58	90.70	68.18

Table 5 shows that the average accuracy rate of "parameter" was 27.5%, "setup", 18.58%, "copyright", 90.70%, and "dependence", 68.18%. The accuracy rate of the words "parameter" and "setup" was far lower than the words "copyright" and "dependence". Those two words at level 4 are probably more likely to occur in an academic paper; however, "copyright" and "dependence" are often encountered in the general textbooks and other materials. Thus, participants are much familiar with the latter ones. Therefore, students' scores for the two words "copyright" and "dependence" were higher than the scores for "parameter" and "setup". This may also explain the deviance in statistics.

Inferential Statistics

The statistical software SPSS (16.0) was used to analyze statistics. ANOVA and *post-hoc* analysis were conducted to test whether there were statistically significant differences between grades in different levels and where these differences lie. Significance level was set at 0.05.

Table 6: The ANOVA results of the effects of grades at each frequency level

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Total Score
F	1.456	2.737	4.699	5.117	3.056	12.024	9.344
Sig.	0.230	0.046*	0.004**	0.002**	0.031*	0.000**	0.000**

* The mean difference is significant at the .05 level.

** The mean difference is significant at the .01 level

In Table 6, results of one-way ANOVA show that difference in grades have no significant effect at level 1 ($F=1.456$, $p > .05$), while the difference in grades have a significant effect at level 2 ($F=2.737$, $p = .046$), level 3 ($F=4.699$, $p = .004$), level 4 ($F=5.117$, $p = .002$), level 5 ($F=3.056$, $p = .031$), level 6 ($F=12.024$, $p < .001$) and total score ($F=9.344$, $p < .001$). This indicates that except for the first level (BNC 4th 1000), students' scores are significantly different across different grade levels. It is important to note that at the last level, level 6, the p value is close to zero, which showed that different grades vary dramatically in terms of vocabulary knowledge.

Apart from the ANOVA results, *post-hoc Scheffe* Analysis was conducted to determine which grade levels demonstrated significant difference at each frequency level. The results are shown in Table 7.

Table 7: Post-hoc Scheffe Analysis of grade effects at each frequency level

Multiple Comparisons			
Scheffe			
Dependent Variable	(I) Grade	(J) Grade	Sig.
Level 3	4	2	0.022*
		1	0.035*
Level 4	1	4	0.006*
		2	0.027*
Level 5	1	4	0.036*
Level 6	4	2	0.040*
		1	0.000**
	1	3	0.001**
		2	0.041*

* The mean difference is significant at the .05 level.

**** The mean difference is significant at the .01 level**

According to the results presented in Table 7, there are no significant differences between grade 1, grade 2, grade 3 and grade 4 at frequency level 1 and level 2 ($p>.05$). At level 3, there are statistical differences at a significant level between two groups: grade 4 and grade 1 ($p=.035$), as well as between grade 4 and grade 2 ($p=.022$). At level 4, grade 1 has a significant difference from grade 4 ($p=.006$) and grade 2 ($p=.027$). At level 5, there is only a statistical significance between grade 4 and grade 1 ($p=.036$). However, at level 6, there are significant differences between group 4 and group 2 ($p=.040$), between group 1 and group 3 ($p=.001$), between group 1 and 2 ($p=.041$), and between grade 1 and grade 4 ($p<.001$). This suggests that students of grade 4 significantly outperformed students of grade one for words at lower frequency levels. This is also graphically demonstrated in Figure 3. Students' scores are quite similar at level 1, and they are dispersed most widely at level 6.

Discussion

Research question 1 asks whether there is any difference in business vocabulary knowledge among students from different grade levels. Our results showed that students' average score increased as grade level increased. With regards to research question 2 regarding the influence of frequency on students' vocabulary knowledge, our results showed that students' average scores decreased from level 1 to level 4, but increased at level 5 and gradually dropped at level 6 again. With regards to research question 3, our results showed that the four groups performed similarly at level 1, but their scores became most divergent at level 6.

The results of the study generally confirmed previous studies that revealed English majors' vocabulary size increase as grade level increases, and the most prominent growth occur at grade two (Xi, 1998; Xu, 2010). However, these studies examined general vocabulary only, and our study added to existing literature in ESP by examining the mastery of business vocabulary knowledge across four grade levels by Business English majors. Another interesting finding is that frequency is an important factor in differentiating students' vocabulary knowledge across grade levels. Our study shows that the least frequent level (level 6) showed the most obvious variation among different grades, whereas the lowest frequency level words did not differentiate learners at different grade levels. As argued by Schmitt (2010, p.63), "frequency in which a word occurs in language ... is the single most

important characteristic of lexis...since it is related to the acquisition, processing and use of vocabulary”.

In addition, the present study tests the applicability of Hsu's (2011) *Business Academic Word List* in measuring the variations in Chinese business English majors' receptive vocabulary knowledge. Furthermore, this study contributes to our understanding of how the business English curriculum may enrich business English majors' vocabulary knowledge across different grades.

The difference in vocabulary test scores and the steady increase in the overall scores across four grade levels are consistent with studies on vocabulary size of English majors in the Chinese contexts (Xi, 1998; Xu, 2010). The steady increase in students' business vocabulary knowledge can be attributed to courses offered during their four years' study. As argued by Basturkmen (2014), ESP courses can range from more wide-angled courses such as Business English or Academic English to more narrow-angled courses that are tailored to the specific needs of a particular group of learners. The wide-angled courses provide description of language use in a particular field of discipline, whereas the narrow-angled courses include more specific genres in the target discourse community.

Students in this particular university took a number of wide-angled courses during the first and second year of their undergraduate study and these courses offer some frequently occurring vocabulary items in the business field to the students. For example, the course *Integrated Skills of Business English* is the required business English course for freshmen and sophomores. The texts in the course books were mainly selected from popular magazines published in English-speaking countries. Although the new words appeared in the vocabulary lists in the series of *Integrated Skills of Business English* Book I to Book IV had been deleted from the test, other words appearing in the texts may increase students' vocabulary size in general, especially vocabulary items related to business and economics. As students take more disciplinary specific courses in grade three, such as *Import and Export Practices*, *Introduction to Business Law*, *Principles of Economics*, and *International Trade*, their knowledge of vocabulary in specific sub-fields in business practice may develop. In addition, these courses allow them to develop their business skills using the second language and gradually acquire the vocabulary in the professional speech community.

Apart from the courses in the curriculum, preparation for tests may explain the sharp increase in the overall vocabulary score from grade 1 to grade 2. In the second year of university study, students majoring in English are required to take the Test for English Majors-4(TEM-4). In consequence, many students may recite the vocabulary lists in order to pass the examinations. Moreover, students preparing to further their education abroad are required to pass the examination including TOEFL, IELTS, GMAT and GRE in the second year. These exams all require a large amount of vocabulary. Students who prepare for those tests would be expected to recite the related vocabulary lists which may have profound effects on their vocabulary size.

In terms of the relationship between vocabulary knowledge and word frequency levels, results show that average scores generally decrease as word frequency levels increased. The average score of participants decreases from level 1 to level 4. However, the average score of word level 5 was higher than the former levels. There may be two reasons for this unexpected phenomenon. Firstly, the word frequencies may not show the process of acquisition precisely. The words from higher level categories may be more familiar to participants than the lower level categories. This phenomenon is most noticeable in the level 4 and level 5.

Although different grades have similar trends of average score variation at different words frequencies; the standard deviations are different among grades, especially at Level 4, which is the section that most of the participants have the lowest score. The senior participants have the smallest standard deviations, showing that most participants have similar accuracy rate in the section. This suggests that during four-year study in college, students' knowledge of business vocabulary are gradually becoming similar, especially at level 4.

According to the results from the one-way ANOVA, differences in grades have significant effect on the scores of each frequency section except Level 1 (BNC 4th 1000). The four grades do not show a statistically significant difference in the mastery of the lowest level of words, perhaps because most of the words at this level were acquired before college studies. Those words may be frequently seen in their high school textbooks and reading materials, and they are often used in daily lives. Thus, the college studies have little impact on participants' acquisition of these business words.

However, as for the words with higher levels and lower frequencies, they are more likely to be learned during college studies. Words of higher level contain relatively more morphemes than lower level words, which makes them difficult to memorize. Their meanings are sometimes less transparent so that they may be difficult for students to understand, such as the word “parameter”. They may be used less frequently in daily life but occur regularly in discipline specific academic articles. As a result, students generally obtained lower scores for words at higher levels.

Conclusions and implications

Results of the present study show that there was a significant difference in the knowledge of business academic vocabulary among students of different grades. Participants in grade 4 have the highest average score of vocabulary tests and grade 1 have the lowest average score. Moreover, there is a significant difference in students’ knowledge of business vocabulary at different frequency levels. In addition, there is a significant difference among four grades on their mastery of words at different frequency levels. At Level 1, the four grades do not differentiate with each other in the average score. However, the differences among groups become more noticeable for the words that appear less frequently in academic texts.

Nevertheless, it is important to note that the level of difficulty does not necessarily correspond to level of frequency. As argued by Wesche and Paribakht (1996), there is a “lack of evidence for a clear link between the frequency of a word in a given corpus and the likelihood of its being known by an individual” (p.4). The words that appear less frequently in business academic articles may not be more difficult to learn in other discourse contexts. For example, the word “copyright” may be a core vocabulary item in legal discourse, but it may not be a highly frequent vocabulary item in business discourse. Students may still have access to these words with low frequency through other channels. Vocabulary test developers may take this factor into consideration when they design a vocabulary test.

Although students of higher grades master business vocabulary better than lower grades, their average score for the business academic vocabulary test remains to be improved. We compared the word lists in the textbook series *Integrated Skills of Business English* and Business Academic Word List developed by Hsu (2011). Results showed that vocabulary lists of the textbook only covered 49 words of the Business Academic Word List, accounting

for only 11.5% of the total words in the Business Academic Word List. As most researchers contend that ESP course should focus on specific learning needs of the learners (Hutchinson & Waters, 1987; Dudley-Evans & St John, 1998; Hyland, 2008), ESP course design may also take learners' vocabulary size and learning needs into consideration. In this process, teachers identify the language and skills that learners will use in the target professional community and evaluate the learners' knowledge *status quo* (Basturkmen, 2014). In order to help undergraduate students prepare for their future study in business areas, textbook developers may use a robust discipline-specific vocabulary test to examine the learners' mastery of target vocabulary, and then develop textbooks that contain the vocabulary necessary for communicating in the professional workplace. Moreover, students should be encouraged to do business related research activities and extensive reading in academic literature on current business topics in order to better master business vocabulary.

Limitations and future directions

The present study has certain limitations, one of which is related to the limited number of participants and the limited number of target words. The limited number of participants may render the research results not comprehensive enough. Moreover, as the university in which the study was conducted is one of the top universities focusing on business and international trade, the performance of students in this university may vary from students from the other universities in China. Studies on Business English majors in other universities may provide support for the findings in the present study. Secondly, due to the space and time limit, only thirty words were listed in the tests. As the number of words in each level was limited, the results of the present study need to be confirmed by future studies that use a word list containing more items. Another limitation is that the present test only examines the receptive knowledge of the business academic vocabulary, future studies may investigate the productive knowledge of business academic vocabulary by using various test forms (Laufer & Nation, 1999).

As business English has become a heated topic in the research field of ESP, research on business vocabulary development by business English majors will provide implications for course developers and teaching practitioners. Specifically, more research is needed with regards to individual difference and business vocabulary development, task demand and business vocabulary acquisition, as well as the difference in the acquisition of general

English vocabulary and discipline-specific vocabulary. In addition, with the development in corpus linguistics and studies on business English discourse, we will undoubtedly gain more understanding on the compilation and validation of business vocabulary tests.

References

- Albrechtsen, D., Haastrup, L., Henriksen, B. (2008). *Vocabulary and writing in a first and second language: Processes and development*. Basingstoke: Palgrave Macmillan.
- Basturkmen, H. (2014). *Developing courses in English for Specific Purposes*. Shanghai: Shanghai Foreign Language Education Press.
- Browne, C. Culligan, B. & Phillips, J. (2013). The New General Service List. Retrieved from: <http://www.newgeneralservicelist.org>.
- Cai, J. (2012). The constraints on and necessity for upgrading college English vocabulary requirement. *Journal of PLA University of Foreign Languages*, 35(1), 48-53.
- Chen, Q. & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs), *English for Specific Purposes*, 26(4), 502-514.
- Chen, Z. & Wang, L. (2009). Developing national curriculum for BA program in Business English of China. *Foreign Languages in China*, 30, 4-11.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238.
- Dai, J. (2013). A Study on the Vocabulary Size of Non-English Majors at CET-4 Level. *Journal of Chongqing University of Technology(Social Science)*, 27 (1), 118-122.
- Daller, H., Milton, J., & Treffers-Daller, J. (2007). Editor's introduction: conventions, terminology and an overview of the book. In H. Daller, J., Milton, & J.Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp.1-32). Cambridge: Cambridge University Press.
- Deng, Z. (2001). Studies on survey questions in vocabulary size tests and discussions on a vocabulary size test. *Foreign Language teaching and research*, 33(1), 57-62.

- Dudley-Evans, T. & St. John, M. (1998). *Developments in English for Specific Purposes*. Cambridge: Cambridge University Press.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford, England: Oxford University Press.
- Ellis, N. (1997). Vocabulary acquisition: Word structure, collocations, word-class, and meaning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description acquisition and pedagogy* (pp. 122-139). Cambridge: Cambridge University Press.
- Friginal, E. Walker, M., & Randall, J. (2014). Exploring mega-corpora: Google Ngram Viewer and the Corpus of Historical American English. *EuroAmerican Journal of Applied Linguistics and Languages*, 1(1), 48-68.
- Gao, Y. (2015). An empirical study on the correlation between vocabulary size and reading comprehension. *Course Education Research*, (16), 199-200.
- Gu, Q., & Li, X. (2013). Receptive vocabulary size for non-English majors. *Overseas English*, 2, 25-27.
- Gui, S. (1985). A Survey and analysis of vocabulary size of English majors. *Modern Foreign Languages*, 1, 1-6.
- Hu, M. & Nation, I. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Hutchinson, T. & Waters, A. (1987). *English for Specific Purposes*. Cambridge: Cambridge University Press.
- Hsu, W. (2009). Measuring the vocabulary of college general English textbooks and English-medium textbooks of business core courses. *Electronic Journal of Foreign Language Teaching*, 6(2), 126-149.

- Hsu, W. (2011). A business word list for prospective EFL business postgraduates. *The Asian ESP Journal*. 7(4), 63-99
- Hyland, K. (2008). The author's replies. *TESOL Quarterly*, 42, 113-114.
- Jiang, Y. (2009). An investigation on vocabulary size of non-English major college freshmen and its implications on vocabulary teaching. *Journal of Harbin University*, 30(2), 124-127.
- Konstantakis, N. (2007). Creating a business word list for teaching business English. *ELIA*, 7(1), 79-102.
- Laufer, B. (1992). Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability. *Journal of Research on Reading*, 15, 2, 95-103.
- Laufer, B. & Nation, I. (1999). A vocabulary size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- Laufer, B. & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Lu, C. (2004). Vocabulary size and its influence on English achievement as well as its relationship to depth of lexical knowledge. *Foreign Language Teaching and Research*, 36(2), 116-123.
- Ma, G. (2001). A study on vocabulary size of non-English majors in universities of science and technology. *Foreign language teaching*, 3, 48-52.
- Mackey, A. & Gass, S. (2005) Second language research: Methodology and design.
- Mahwah, NJ: Lawrence Erlbaum Associates.

- Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. In I. Vedder, I. Bartning, & M. Martin (Eds.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research* (pp. 211-232). Second Language Acquisition and Testing in Europe Monographs Series 1.
- Milton, J., Wade, J. & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca- López (Eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.
- Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 57-78). Eurosla Monographs Series 2, European Second Language Association. Retrieved June 25, 2015, from <http://www.diva-portal.org/smash/get/diva2:682928/FULLTEXT01.pdf>.
- Nation, I. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. & Heatley, A. (2002). RANGE [Computer software]. Retrieved March 5th, 2010, from <http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx>
- Nation, I.S.P. and Beglar, D. (2007) A vocabulary size test. *The Language Teacher* 31 (7), 9-13.
- Nation, I. & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
- Nelson, M. (2000). *A corpus-based study of Business English and Business English teaching materials*. Unpublished PhD Thesis. Manchester: University of Manchester.
- Qian, D. (1999). Assessing the role of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review*, 56 (2), 282-307.

- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Hampshire, England, Palgrave Macmillan.
- Schooner, R. Van Gelderen, A., Stoel, R., Hilstijn, J. & De Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary-school students. *Language Learning*, 61(2), 31-79.
- Shabani, M. & Tazik, K. (2014). Coxhead's AWL across ESP and Asian EFL Journal Research Articles (RAs): A Corpus-Based Lexical Study, *Procedia - Social and Behavioral Sciences*, 98, 1722 – 1728.
- Wang, Z. (2006). A study of the correlation between English vocabulary, language skills, and comprehensive language proficiency, (5), 92-97.
- West, M. (1953). *A general service list of English words*. London: Longman, Green, & Co.
- Wesche, M. & Paribakht, T. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, 53(1), 13-40.
- Xi, Z. (1998). A survey of vocabulary development of English majors. *Foreign Language Teaching*, (2), 72-77.
- Xu, D. (2010). A Survey and analysis of vocabulary size for English major students at key normal universities. *Kejiaowenhui*, 10, 115-117.
- Zimmerman, K. (2004). The role of vocabulary size in assessing second language proficiency. Unpublished MA dissertation, Brigham Young University.

Appendix I Vocabulary Test

【特别声明】：本测试仅用于商务英语学科研究目的，测试成绩与其他无关，务必请参与者保证测试的真实性，避免在测试期间使用字典、手机、电脑等进行辅助，请将答案写在题。时间控制于10分钟内，有关测试者的个人信息，科研者承诺对其隐私保护，敬请监督，感谢同学们的参与。

测试号：_____

2. 年级：a. 2010级 b. 2011级 c. 2012级 d. 2013级

3. 背过的词汇表：a. 大学四级 b. 大学六级 c. 专业四级 d. TOEFL e. IELTS f. GRE g. GMAT

4. 英语成绩(分数)：大学四级_____ 大学六级_____ 专业四级_____ TOEFL_____ IELTS_____ GRE_____ GMAT_____

asset: I need to categorize the **asset**.

资产

形态

装置

物品

包装

条款

证券投资基金

全景

audit: The **audit** process queries this key.

扩大

编辑

重复

审计

ratio: This **ratio** gives you the solution.

比率

数据

贷款

整体

statistic: Let me repeat one **statistic** from the Commission.

字母

数据

图案

经纬

cite: Can you **cite** another case?

引导

导致

理解

引用

review: He is **reviewing** his options.

重复

回顾

观赏

修改

default: He **defaulted** this arrangement.

犯错

改正

违背

批评

transact: We will **transact** with them.

交易

转换

移动

分散

affiliate: The majority of people questioned **affiliated** themselves with a religious group.

使附属

使偏离

使敌对

使流动

journal: I like this **journal** very much.

记者

季度

期刊

旅程

deviate: Our leaders will not **deviate** from a market-oriented global economic system.

脱离

转变

产生

出现

abstract: Legal questions rarely exist in the **abstract**.

伸展

收缩

资料

摘要

compliance: Inspectors should verify **compliance** with the treaty.

联合

和睦

服从

合谋

equity: He sold his **equity** in the company last year.

期货

股权

负债

收益

empirical: This part is an **empirical** research.

实证的

君权的

绝对的

空洞的

portfolio: I want to learn more about the **portfolio**.

disclosure: Other firms have struck similar off-balance-sheet deals, but poor **disclosure** makes them difficult to track.

关闭
拒绝
披露
否认

parameter: It is the research of the **parameter** estimation.

段落
参数
周长
数据

classification: It is based on ODS **classification**.

古典
阶级
经典文学
类别

setup: This **setup** is very attractive.

装潢
住房
机构
引擎

copyright: The **copyright** vests in your company.

复制品
版权
打印
抄写

volatile: Perishable goods have **volatile** markets.

不确定的
反常的
反抗的
狭小的

violate: She **violated** the terms of contract repeatedly.

暴力
背叛
违反
蔑视

denote: It **denotes** nothing.

记录
显示
否认
意味着

dependence: It analyses the energy **dependence** on the import.

独立
注射
来源
依赖

underwrite: The company **underwrites** part of the offering.

强调
提示
承保
弱化

inventory: We are short of **inventory**.

发明
投资
存货
虚构

collateral: Their house is a source of **collateral**.

收入
腐败
合资
抵押物

mitigate: The new measures helps **mitigate** the effects of the recession.

强化
减轻
同化
消除

transparency: The first principle in supervision is **transparency**.

转化
传送
交通
透明度



Revisiting the Topical Knowledge of Iranian ESP Learners in Reading Comprehension: Text Types and Question Types

Masoomeh Estaji

Assistant Professor of Applied Linguistics, Allameh Tabataba'i University, Iran

Hussein Meihami

PhD Candidate of Applied Linguistics, Shiraz University, Iran

Bio data

Masoomeh Estaji is an assistant professor of TEFL at Allameh Tabataba'i University. She is also the head of the English Department and an adjunct professor of TEFL at Khatam University. She has published numerous books and papers on ESP, testing, and teacher education. Her research interests include Language Testing and Assessment, ESP, and Teacher Education.

Email: mestaji74@gmail.com

Hussein Meihami is currently doing his PhD in Applied Linguistics (TEFL) at Shiraz University, Iran. He has published research papers in various journals and presented many papers in different conferences. His research interests include second language writing, ESP, language assessment, and discourse studies.

Email: s.hmeihami@rose.shirazu.ac.ir

Abstract

There are disparate findings with regard to the use of the learners' topical knowledge and its impact on their reading comprehension. Therefore, this study was an attempt to investigate the effect of topical knowledge of English for specific purpose (ESP) learners on their

reading comprehension performance. Moreover, in this research the effect of topical knowledge on different types of reading questions namely comprehension, inference, and lexical questions were investigated. To this end, a total of 46 upper-intermediate students, including 26 civil engineering students and 20 general English students, with age range between 23 and 29 who were selected through convenience sampling participated in this study. Two types of reading texts, general and ESP, were assigned to both groups. The results of this study showed that the ESP students had a better performance on both general English reading text ($p=.001$) and ESP reading text ($p=.001$) than the general English students. In addition, the topical knowledge of ESP students was significantly constructive in helping them to answer different types of reading comprehension questions. The study also revealed that there was no statistically significant difference between the ESP students' performance on specific reading texts and their performance on general English reading texts. This investigation bears some implications for teachers and test designers.

Keywords: ESP; general English; lexical question; reading comprehension; text types

1. Introduction

Topical knowledge, as an important factor in teaching and testing L2, is defined by Alexander, Schallert and Hare (1991), as “the interaction between one’s prior knowledge and the content of a specific passage” (p. 334). It finds its way in the realm of L2 teaching/learning when Bachman and Palmer (1996) emphasize the importance of this interaction as a leading criterion for making inferences about the language ability of the test takers. They point out that the interaction between topical knowledge and language use is among the principles that we may stick to find out about real L2 learning. Moreover, as He and Shi (2012) claim topical knowledge is required in different skills, highlighting the role of topical knowledge in the reading skill.

The undeniable role of L2 reading comprehension in acquiring L2 is well-shown through the formidable number of research studies conducted on the various aspects of this skill. For instance, Uso-Juan (2007) states that “it is now widely accepted that reading ability in English is of vital importance to academic studies and future professional success” (p. 210). Reading comprehension is seen by L2 researchers (Cook, 2008; Keene & Zimmerman, 1997; Pressley, Block & Gambrell, 2002) as a complex skill in terms of teaching/learning compared to other L2 skills. According to Zonghong (2005), reading complexity is due to some factors

such as L1/L2 relationships, reader linguistic proficiency, and topical knowledge. In other words, in the realm of second language reading, topical knowledge plays an important role. Likewise, some language scholars have claimed that the lack of related topical and background knowledge can be the source of complexity in reading comprehension (Cook, 2008; Smagorinsky, 2001). The discussions on the roles of topical knowledge are also approached with much attention in the field of ESP. That said, researchers (e.g. Chen & Donin, 1997; He & Shi, 2012; Tan, 1990), declare that this area of language acquisition has more interaction with topical knowledge. However, opponents (e.g. Clapham, 1996; Koh, 1985; Ridgway, 1997) have claimed that it is not always true for the ESP students to have their best performance in their fields.

The ESP learners go through a lot of readings during their academic courses which can help them mutually with both their general English and their academic courses. This happens when ESP learners accrue topical knowledge of their domain of study through numerous courses accomplished during their academic studying period. Hence, this topical knowledge may be constructed due to instructions and interactions of ESP learners with the real-world or pedagogical contexts.

There are several theories proposed to address different aspects of topical and background knowledge and their effects on reading comprehension. In particular, schema theory is the most cited among the theories of topical knowledge which has been defined by Carrell and Eisterhold (1983, p. 73) as “different kinds of background knowledge that facilitate our understanding.” Later on, Anderson and Pearson (1984), point out that by relating the background knowledge to what comes as new knowledge, the learners create a “home” for comprehending a new text. The strength of schema theory, then, is established by this interaction (Alderson, 2000). Although the importance of schema theory for reading comprehension is still doubted, researchers agree that by activating some mechanisms, schema can benefit readers to understand and comprehend the texts (Alderson, 2000; Rumelhart, 1980).

That said, the point that draws ESP researchers’ attention is whether ESP learners who create their specific schema through their academic courses can benefit from the use of that specific knowledge in their ESP reading comprehension courses or not. The important point is whether specificity of the academic knowledge – defined by Hyland (2002) as the

transferability of knowledge and skills among different disciplines and occupations – is transferable to develop the learners' reading comprehension. Accordingly, the main thrust of this investigation was to explore the effects of topical knowledge that ESP learners gain through their academic courses on their reading comprehension performance. Moreover, this investigation aimed to investigate whether the ESP students compared to EGP students are more capable to transfer their topical knowledge as schema to their reading comprehension. This research, on top of these, attempted to measure the participants' discipline knowledge and its impact on their reading comprehension performance considering text type and question type.

2. Literature Review

There are a lot of parameters which potentially affect second language (L2) reading. In fact, the topics which are dealt with in various passages demand particular attention since they commence and direct the act of reading. Bachman and Palmer (1996) referred to topical knowledge as one of the characteristics of an individual that is of interest in language testing. They asserted that this characteristic can exert great influences on language use. Moreover, they mentioned that topical knowledge can be a facilitative factor rather than a preventive one in the test takers' performance. It is sometimes referred to as knowledge schema and real-world knowledge by Bachman and Palmer (1996), which can be considered as what L2 learners have as knowledge structures in their long-term memory. Topical knowledge that each individual possesses should "be considered in the description of language use because this provides the information base that enables them to use language with reference to the world in which they live, and hence is involved in all language use" (Bachman & Palmer, 1996; p. 65). To assess the topical knowledge and language knowledge that each L2 learner has, one should determine the extent to which relevant topical knowledge and areas of language knowledge are accessible, and which of these availabilities help to accomplish an L2 task.

Fisher and Frey (2009) asserted that an essential element of acquiring new knowledge is the background and the topical knowledge about that subject. Cromley and Azevedo (2007), in a similar vein, declared that background and vocabulary knowledge are among the strong indicators of success in reading comprehension that indirectly show whether the learners apply problem-solving – one of the major L2 learning strategies – when they need to reach

meaning. Moreover, Stahl, Hare, Sinatra and Gregory (1991) stated that the most obvious effect of topical knowledge on reading comprehension is to influence *directly* the understanding and the comprehending of what is read. Consequently, the more your topical knowledge about a subject matter is, the more you understand the text.

Research about topical knowledge and its effects on reading comprehension dates back to 1970s. A multitude of studies (Callender, 2008; Carrell, 1987; Roller & Matambo, 1992) have revealed that prior knowledge plays a substantial role in improving reading comprehension performance on comprehension tests and recall tasks. In particular, Krekeler (2006) reported that prior knowledge strongly affects performance on the language for specific academic purposes (LSAP) reading tests. Hereupon, Hudson (2007) has explicated that if prior knowledge can affect the recall and comprehension of readers within a particular culture, there would be cross-cultural and cross-linguistic effects as well. In a study conducted by Steffensen, Joag-Devv, & Anderson (1979), the effect of EFL/ESL background knowledge on cultural issues was investigated. The results of their study, similar to the one conducted by Carrell (1983), showed that the learners responded better to the culture test on which they had topical or background knowledge. However, conflicting evidence surfaced in Peretz and Shoham's (1990) study in language for specific purposes (LSP) testing, showing a non-significant relationship between topic familiarity and EFL reading comprehension test performance.

Regarding content schemata, Liu (2011) investigated the effects of major field of study and cultural familiarity on TOEFL iBT reading performance of 8692 participants. The study employed six reading passages with three passages covering the topic of physical science and three others focusing on the subjects of art and history. The research results revealed the impact of content schemata on reading performance of all the test-takers at individual item and bundle levels. Similarly, Tarchi's (2009) investigation showed that learners who possessed topical knowledge about history had a better understanding of reading comprehension texts including historical subjects. Tarchi stated that the "discipline-specific" aspects of historical reading require readers to establish a causal relationship between different phenomena to form the totality of the text. He stated that the more facts a reader knows about a topic, a better understanding of the text that learner achieves.

Reisman and Wineburg (2008), in a different study, examined the effect of contextualization of the materials in two situations: When the learners possess topical knowledge and when they do not. They suggested that contextualization of the previous thoughts is hard to achieve without background knowledge. They also stated “background knowledge allows the students to decipher unfamiliar terms and create accurate mental images as they read” (p. 203).

However, topical knowledge and its effects on reading comprehension in ESP context did not enjoy a well number of studies. The research conducted by Alderson and Urquhart (1985) focused on this aspect. In this investigation, which was a review of the combination of three studies, Alderson and Urquhart (1985) reported on the effect of background knowledge on ESP students’ ability in reading comprehension. Four groups of ESP students from different majors such as Development Administration and Finance (DAF), Engineers (ENG), Liberal Arts (LA), and Science and Mathematics (SM) participated. The researchers examined the effect of discipline related text on the participants’ reading comprehension. Moreover, they examined the students of different majors through different texts of different disciplines. For instance, they examined the reading comprehension ability of DAF students with texts of ENG major. The results revealed that the students of different discipline performed significantly better in their discipline-related reading comprehension. Moreover, each group outperformed the other group when they were provided with their own major discipline-related texts to read and comprehend. As a whole, Alderson and Urquhart’s (1985) study indicated the influence of background knowledge of discipline-related major on the reading comprehension performance of ESP students.

Regarding the role of language proficiency, some investigations were aimed at studying the effect of background knowledge among different language proficiency levels. In a study conducted by Krekeler (2006), the effect of background knowledge was researched. In addition to background knowledge, Krekeler (2006) investigated the effect of different language proficiency levels and their relationship with background knowledge. In his study, Krekeler (2006) used linguistic threshold “where L2 readers may be able to draw on their background knowledge” (p. 103). In particular, he attempted to determine whether background knowledge of ESP students would be beneficial just as some specific levels of L2 proficiency. He used a c-test to assess the students’ language proficiency level. The

research results showed that the effect of background knowledge is strong; however, this effectiveness varies among different levels.

All in all, although studies on the effect of topical knowledge on different L2 skills such as speaking are done recently (Khabbazzashi, 2015), research conducted on topical knowledge and its effects on reading comprehension did not lend itself to the fair number of investigations. The case is also true in the realm of ESP studies. Moreover, due to some reasons, such as ignorance of the level of participants' existing discipline-related knowledge and their English proficiency level, the previous studies done on the relationship between topical knowledge and L2 reading comprehension had controversial findings (Chen & Donin, 1997; Tan, 1990; Uso-Juan, 2006). All in all, reviewing the literature and specifying the purpose of the study, this investigation intended to answer the following research questions:

1. Does the topical knowledge of ESP learners have any statistically significant effect on their reading comprehension performance?
2. Is there any statistically significant difference between the effect of topical knowledge on the performance of ESP learners compared with EGP learners while they respond to reading comprehension questions of comprehension, lexical, and inference?
3. Do ESP students perform better in ESP reading texts in comparison with general reading texts?

3. Methodology

3.1. Participants and Research Setting

This research study was conducted in two sites. In order to collect data from the ESP students, Technical Institute for Engineering (TIE) in Tehran, Iran was chosen. In this institute, there are some technical and ESP courses which are taught to the students of various disciplines. Moreover, the institute of Sharif Language Center was considered to obtain data regarding the students of general English. A total of 46 students, all senior students and selected through convenience sampling, including 26 civil engineering students and 20 students of general English participated in this study. It is highly important to mention that for removing the methodological problems, such as research contaminating results, the ESP participants were

selected from among 40 civil engineering students who have already passed the Loading course, which is a mandatory two-credit course assigned for the students of Civil engineering in their BA program, based on the cut score of 15. The rationale behind this selection was that, the knowledge of the ESP students has to be measured since different levels of English proficiency might have an impact on the results of the studies (Chen & Donin, 1997; Tan, 1990; Uso-Juan, 2006). The same assumption might be true about the effect of topical knowledge on ESP students' reading comprehension. Given that, because of selecting reading materials for the purpose of assessing the reading ability of the ESP students from the aforementioned course, Loading, a cut score of 15 was set. Hence, all the 26 civil engineering students, participated in this study, had obtained 15 and above in the Loading course ($M = 17.27$, $SD = 1.36$). Table 1 indicates the descriptive statistics of the ESP students' performance in the Loading course.

Table 1: Descriptive Statistics of the ESP Students' Knowledge Test

	N	Min. Score	Max. Score	Mean	SD
Students' knowledge test performance	26	15	19.50	17.27	1.36

There were 12 female and 14 male students among the students of civil engineering while there were 11 female and 9 male students among general English course ($N=20$). As stated in advance, the EGP students, who were selected through convenience sampling, were English language learners at Sharif language center. It is worthy of note that, in order to ensure the participants' homogeneity in terms of their level of proficiency, all the participants took part in a modular TOEFL proficiency test. Based on the obtained results, all the 46 participants, in the age range of 23 to 29, were specified to be at the level of upper-intermediate.

3.2. Instrumentation

3.2.1. English Proficiency Test

For the purpose of this study, to ensure the level of proficiency and homogeneity of the participants, a modular TOEFL proficiency test designed by *ETS* (Educational Testing Service, 2014, *The Official Guide to the TOEFL iBT* (4th edition). New York: McGraw-Hill) was administrated. To be in line with the purpose of this study, the reading, writing, and listening sections of this modular test were administrated. There were two readings in this

proficiency test, each containing 15 questions. It is important to state that we examined the readings for their topical knowledge inclusion. Since our purpose was to administrate a proficiency test, we selected the readings which did not require the topical knowledge. The participants of this study were required to answer these reading comprehension questions in no more than 30 minutes. Since these reading comprehension questions included different types of questions namely inference questions, comprehension, and lexical, they contributed to the micro level aim of this study which was to assess the effect of topical knowledge that ESP and EGP students possess on the way they respond to different types of reading comprehension questions. Regarding the two other sections, writing and listening, the students were required to listen to an audio and write a paragraph about it. Two rating scales were used to determine the overall performance of the students. First, for assessing their reading comprehension, the Right/Wrong (RW) rating scale was used to count their right and wrong answers to the questions. The second rating scale was the one which was extracted from He and Shi (2012) that includes different writing components such as language, content, and organization. The participants' overall scores were put in the TOEFL scale to specify the participants' proficiency level. It should be stated that the entire test took about 45 minutes.

3.2.2. *Knowledge Test*

As mentioned earlier, one of the problems of the previously done investigations on the relationship between topical knowledge and L2 reading comprehension was that these studies ignored measuring the discipline-related knowledge and took it for granted (Chen & Donin, 1997; McNeil, 2011; Tan, 1990; Uso-Juan, 2006). Owing to this, to assure that the ESP participants were in the same level of topical knowledge, their scores in one of their course exams, Loading, were gathered. These scores were obtained by the students in their final exam. The reason for which this course, Loading, was selected was that the technical reading text of this investigation included Loading topic. In this exam, the students had to provide responses to ten open questions related to Loading. For the purpose of scoring, each right answer was given two scores, each half right answer was given one score, and wrong answers were given no score. Finally, by using the results of this test and setting a cut score criterion, each ESP student whose knowledge of Loading was between the range of 15 to 20 was selected to participate in this study (for results see Table 1, section 3.1).

3.2.3. Reading Comprehension Test

The aim of this study was to determine the effect of ESP students' topical knowledge on their reading comprehension, if any. Consequently, for accomplishing this aim, two types of reading texts were required. First, a specific type of text in accordance to what ESP students are taught in their content classes and a general type of reading which they may encounter in their everyday life and during their general English courses. After consulting this issue with three language experts, the researchers decided to choose specific texts whose topics were related to Learning from MA entrance exams. There were two reasons for this selection. The first and at the same time the most important reason was that these reading texts were designed by a panel of experts, so the reliability and validity of them had been already determined. The second reason was that the participants of this study were in their last semester of their major; therefore, there was a balance between content they have been taught in their classes and the one they have been tested.

As for the general reading text, the researchers selected a text from TOEFL iBT reading comprehension texts by *ETS* publications (2014). The selected reading passages did not include topical knowledge. The readings of this study included ten questions (ten for specific reading text, and ten for general one). These questions varied in types and they were comprehension, inference, and lexical questions. Table 2 shows the number of each type of questions.

Table 2: Number of Each Type of Question

Discipline	Overall number of Words	Types of Questions	Number of each type of question
ESP	482	Comprehension	4
		Inference	4
		Lexical	2
EGP	431	Comprehension	4
		Inference	4
		Lexical	2

The comprehension questions are defined by Koivukari (1987, p. 103) as “those [questions] calling for the reproduction of content and those calling for the generation of the new content.” This definition emphasizes the fact that the students need to focus on the

components of the texts and select the response to answer these questions. Below are two examples of comprehension questions extracted from ESP and general reading texts:

A) *The “image analysis machine” was.... The MAD equipment. (ESP Reading Text Sample)*

1. *As fast as* 2. *Slower than* 3. *A lot faster than* 4. *A lot slower than*

B) *To account for the ‘surprising’ number of animals in a ‘country producing so little food’, Darwin suggests all of the following as partial explanations except...(General English Reading Text Sample)*

1. *food which is a concentrated source of nutrients*
2. *rapid regrowth of plant material*
3. *large area for animals to forage in*
4. *mainly carnivorous animals*

Inference questions, on the other hand, refer to the questions that require the students to use their own knowledge associated with what the reading texts mention. Two examples of inference questions are shown below:

A) *it is inferred that before the lunch of ERTS image analysis method had been ...(ESP Reading Text Sample)*

1. *Very complex* 2. *Quite sophisticated*
3. *Very well-developed* 4. *In their early development stage*

B) *Darwin’s parenthetical remark (line 83-84) infers that..... (General English Reading Text Sample)*

1. *Burchell’s data are not reliable*
2. *Burchell’s ideas are not to be given much weight*
3. *Comparison of the weights of herbivores is largely speculative*
4. *Darwin’s views differ from Burchell’s*

The third type of question was lexical questions. For these questions, the students were required to provide synonymous, antonymous, or the closest meaning relationship among the given questions. Two examples are provided below:

A) *A LOP is a ... (ESP Reading Text Sample)*

1. *Fix* 2. *Position* 3. *Line* 4. *Distance*

B) The word 'vitiated' (line 4) most nearly means... (**General English Reading Text Sample**)

1. *Infiltrated* 2. *Occupied* 3. *Impaired* 4. *Invigorated*

3.3. Data Collection Procedure

For collecting data, the researchers visited the institutes during two days (one day for administrating the English and knowledge proficiency test and another for administrating the reading texts). The reading texts, including both general and specific ones, were administrated to the civil engineering students after examining their proficiency levels. They were asked to read the texts in no more than 45 minutes and provide response to them. The order in which the texts were assigned on the paper was Specific-General (SG). The same procedure with the same time limitation and text ordering was assigned to the students of general English.

3.4. Data Analysis

As it was stated earlier, the reading comprehension questions were all multiple choice. Consequently, a Right/Wrong scale was used to score the students' overall and componential performance on the reading texts. In this regard, for each correct answer the students were given one score and for each wrong answer they were given no score. The students' answers were examined twice: Once for the total score and the other time for determining the students' score on different types of reading comprehension questions namely, comprehension, inference, and lexical questions. For data analysis of the obtained information, a series of descriptive and inferential statistics were run employing SPSS IBM version 21.

4. Results

The first research question of this study was probing the effect of topical knowledge on ESP students' reading comprehension. Table 1 shows the results of descriptive statistics about the

ability of the participants in answering reading text questions of the ESP and the general texts.

Table 3: Descriptive Statistics of ESP and General Responses

	Group	N	Mean	Std. Deviation	Std. Error Mean
ESP Text	ESP	26	6.69	1.46	.28
	GENERAL	20	2.95	1.39	.31
GENERAL Text	ESP	26	5.46	1.81	.35
	GENERAL	20	5.60	1.72	.38

As can be seen, the ESP group has a higher mean score ($M=6.69$) in the ESP reading texts whereas this is 2.95 for the students of general English. Interesting enough, the mean scores of both ESP and general English students on answering general texts questions are so close to each other and are 5.46 and 5.60, respectively. Since one cannot decide on the better performance of ESP students based on the results of descriptive statistics, the inferential statistics was computed for the comparison reasons. As to the normality of the data, the ratios of skewness and kurtosis over their respective standard errors were within ± 1.96 . Table 4 illustrates the results of paired samples t-test on the performance of ESP students in reading comprehension.

Table 4: Inferential Statistics on the Overall Performance of ESP Students

	Levene's <i>F</i>	<i>t</i>	<i>df</i>	<i>Sig.</i>
Overall performance	.117	8.77	44	.001

Based on the findings, it is clear that the topical knowledge which ESP students possessed contributed to their better performance on ESP reading texts. Table 4 indicates that the mean score difference is statistically significant between ESP students and general English students on their performance on ESP texts ($p<0.05$).

The second research question of this study had a micro level analysis orientation in which the reading texts were analyzed through the type of questions they contained. The research question was whether there is any statistically significant difference between the effect of topical knowledge on the performance of ESP learners compared with EGP learners while they respond to reading comprehension questions of comprehension, lexical, and inference.

Table 5 indicates the descriptive statistics computed for both groups of ESP students and general English students in two types of readings including the general and ESP texts.

Table 5: Descriptive Statistics for Both Groups

Group	Question Types	Text Types	N	Mean	SD
ESP	Comprehension Q	ESP text	26	3.03	.77
		General text		3.00	.89
General	Comprehension Q	ESP text	20	1.20	1.22
		General text		3.00	.91
ESP	Inference Q	ESP text	26	3.11	.86
		General text		3.34	.74
General	Inference Q	ESP text	20	1.10	.96
		General text		2.90	.96
ESP	Lexical Q	ESP text	26	1.15	.88
		General text		1.46	.58
General	Lexical Q	ESP text	20	.70	.65
		General text		1.50	.60

Table 5 indicates that with regard to the comprehension questions, ESP students obtained a mean score of 3.03 while it was 1.20 for the students of general English. Nonetheless, both groups' mean scores for general English texts were 3 with a slight difference in the SD (.89 for the ESP students; .91 for the general English students). Besides, the data illustrated in Table 3 indicates that with regard to the second types of question, namely the lexical one, the ESP students achieved a mean score of 1.15 in the ESP text while this was .70 for general English students. For the general texts, however, the mean scores of ESP students and general English students were close to each other; 1.46 and 1.50, respectively. The third type of reading text questions was inference ones. The mean score of the ESP students in such questions was 3.11 for ESP reading texts and 1.10 for the general English texts. With regard to the general reading texts, the ESP students obtained a mean score of 3.34 while the general English students obtained the mean score of 2.94. To figure out whether or not the differences among these three types of questions were statistically significant, an independent samples t-test was run. In addition to meeting the normal assumptions (the ratios of skewness and kurtosis over their respective standard errors being within +/- 1.96), the assumption of homogeneity of variances was also met ($p > .05$). Table 6 illustrates the results for this quest.

Table 6: Independent Samples T-Test of Both Groups with regard to Different Reading Questions

Group	Type of Question	Levene's <i>f</i>	<i>t</i>	<i>df</i>	<i>Sig.</i>
ESP*General	Comprehension	2.80	7.01	44	.001

Inference	.003	7.44	44	.001
Lexical	.104	1.92	44	.04

The results of Table 6 indicate that the mean scores, for the three types of reading text questions, namely comprehension, lexical, and inference are statistically significant ($p=0.001$, $p=.04$, and $p=.001$ respectively). However, the results for the lexical question types of reading should be interpreted with caution.

The third research question of this study was to investigate “whether the ESP students perform better in ESP reading text in comparison with the general reading texts?” Table 7 shows the results of descriptive statistics for ESP students’ performance in general and specific reading texts.

Table 7: Descriptive Statistics for ESP Students’ Performance in General and Specific Reading Texts

		Mean	N	Std. Deviation	Std. Error Mean
Pair	ESP	6.69	26	1.46	.28
	General	5.46	26	1.81	.35

As can be seen, ESP students obtained a higher mean score in reading comprehension of ESP texts ($M=6.69$), compared to general English reading texts ($M=5.46$). To see whether the mean scores in both groups are statistically significant, a paired sample t-test was run. Table 8 shows the results of such analysis.

Table 8: Paired Sample T-Test of the ESP Students’ Performance in Both General and ESP

Reading Texts			
	<i>t</i>	<i>Df</i>	<i>Sig.</i>
Overall performance	2.93	25	.07

As can be seen in the results of Table 8, the difference in the mean scores, for the performance of the ESP students on the ESP and the general reading texts, is not statistically significant ($p=0.07$).

5. Discussion

This study was probing to investigate the effect of topical knowledge that ESP students obtained during their academic courses on their English reading comprehension performance. The first research question of this study was *"Dose topical knowledge of ESP students have any statistically significant effect on their reading comprehension?"* The empirical evidence that was achieved in this investigation showed that the ESP students' performance on the reading comprehension texts was statistically significant ($p=.001$). These results are compatible with that of Alderson and Urquhart (1985) who declared that ESP students had a better performance on their major related readings. Moreover, they showed that when providing the students with non-related discipline reading texts, it would be difficult for them to perform on them since they did not possess the necessary topical knowledge for processing the reading texts. Likewise, another compatible study with the current study is that of Tarchi (2009). In his study with the history students, he found that the students of history had a better performance on the historical reading texts. He came to the conclusion that possessing topical knowledge in the academic major was the reason for their good performance on the ESP reading texts.

It can be claimed that the reason for the ESP students' better performance on the ESP reading texts is that "acquiring new knowledge, the topical knowledge, or background knowledge has a pivotal role" (Fisher & Frey, 2009; p. 12). Associating this fact to the schema theory, one can mention an analogy about everyday life and individual's learning. In our everyday life, our experiences cause us to learn about new things. Consequently, it is the combination of the real world knowledge and the experiences obtained in the real world that create the new knowledge. One more reason that can be mentioned is the role of "genuine text". According to Widdowson (2000), the genuine text refers to the relationship between the text and the readership. The higher the sharing area between the characteristics of the texts and the readers, the higher the readers' level of achievement. Stahl *et al.* (1991) declare that the most obvious effect of topical knowledge on the reading comprehension is its influence on directing the understanding and comprehending what is read. The directness in this relationship causes faster comprehension and finally faster understanding that may end in a better performance of the ESP learners on the ESP related texts.

Having a micro-level orientation, the second research question was probing the effect of the topical knowledge of the ESP learners on the different types of reading questions. The question was *"does topical knowledge of ESP learners have any statistically significant effect*

on their reading comprehension with regard to comprehension, lexical, and inference questions?" The obtained results indicated that the ESP students outperformed while they read the texts containing topical knowledge with regard to the three types of questions. However, none of the groups showed any statistically significant performance on these questions on the general reading texts. As defined earlier, a schema "is a conception abstraction containing slots to be instantiated in various ways" (Zonghong, 2005). Consequently, as Schank (1982) stated the total schema is made up of the different aspects. In reading comprehension, these parts are comprehension, lexical, and inference. ESP students participating in this study showed that they had good topical knowledge, responding well to comprehension, lexical, and inference items. This is true to claim that what they obtained through their subject matter courses was converted into slots of knowledge (Alderson & Urquhart, 1983), and finally to background knowledge. Compared with general English students, in these three aspects, the ESP students had a better performance since they possessed more topical knowledge.

The third research question of this study was *"do ESP students perform better in responding to ESP reading texts in comparison to general English texts?"* The achieved results indicated that there is no difference between the performances of the ESP students in the ESP related reading text and the general English one (Table 8). The rationale behind it can be mentioned as the threshold level (Dudley-Evans & St. John, 1998). Since the ESP students have already passed their general English courses before they started to study their ESP courses, they obtained more background knowledge in terms of general English. Therefore, their background knowledge helped them to answer the general English text with the same competency as they did the ESP related reading text. One can make a reasonable claim to say that ESP students are privileged with their ESP knowledge over the general English students who do not possess it.

Now based on the empirical evidence achieved in this study, we can come back to the question posed in the introduction section. It was stated earlier that Hyland (2002) questioned the transferability of the skills from general English to specific one. The findings of this study revealed that there exists such a possibility. The ESP students showed a better performance in comprehending the ESP and the general English texts (Table 8). This can be referred to as the continuum of ELT course types. In accordance to these types of courses, different steps are arranged through the continuum from the general to the specific English. By passing each

step, an ESP learner can decrease his/her distance from specific knowledge. This shows that the students have the knowledge of general English along with the knowledge of specific English. Consequently, the results of this investigation are supporting the fact that ESP students can transfer their general English knowledge to their knowledge of specific English. By revisiting the specificity, transferring topical knowledge to the language knowledge between general and ESP reading text, another reason may be specified for ESP students' good performance on general English courses. ESP students learn to use a text as TAVI (text as vehicle of information) rather than TALO (text as linguistic objectives) in their L2 classes (Dudley-Evans & St. John, 1998). TAVI is focusing on the comprehension of the text more than TAVO; hence, ESP learners can highly take advantage from TAVI and transfer it to general English text.

6. Conclusions and Implications

This study was an attempt to explore the effect of topical knowledge that ESP students possess on their reading comprehension performance. The results of the present study support the hypothesis that the topical knowledge that ESP students obtained through their subject matter courses has a significant effect on their performance in reading comprehension, thus supporting the findings of the previous investigations (Fisher & Frey, 2009; Cromley & Azevedo, 2007; Stahl *et al.*, 1991). Lumley and O'Sullivan (2005) have also concluded that topical knowledge has a great impact on test performance. Moreover, this study examined the effect of ESP students' topical knowledge on responding to three different types of reading comprehension questions, namely comprehension, lexical, and inference. The obtained results showed that ESP students outperformed the general English students in the ESP related texts with regard to these questions; however, no difference was found between the performance of these two groups on the general English reading comprehension texts. As a whole, the content knowledge that ESP students enjoy on various subjects was found to be significantly effective and constructive in their performance on ESP related reading comprehension.

The implication of this study can be for test designers working in the realm of ESP and general English test designers. The outcome of this study showed that there was no difference in the performance of the students on general English reading texts. Therefore, it is fair if the test designers and the ESP teachers provide the ESP students with the ESP related reading

texts in their exams; however, it is not that much fair to do the same procedure for the general English students. Despite the results obtained in this study, test designers should consider the purpose for the ESP students to either provide them with the ESP related reading texts or the general English reading texts. Based on this finding, L2 test developers and testing researchers should be more conscientious regarding the topic and text selection process in order to diminish the effects of test bias resulting from the influence of topical knowledge.

The current study suffered from several limitations, affecting the generalizability of its findings. For instance, the study was carried out on a small sample of participants, using a non-probability sampling method, due to the constraints which include the time, financial, and human factors. Besides, the participants were selected from two different settings. Hence, if the study was conducted in other institutions, other regions of Tehran and provinces, the results would be more generalizable.

Owing to the limitations of this study, further studies are required to be carried out with larger number of participants and with ESP students of different majors. Besides, the effect of discipline-related and non-related reading comprehension texts on ESP students' performance is another venue which requires further investigation.

References

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alderson, J. C., & Urquhart, A. H. (1983). The effect of student background discipline on comprehension: A pilot study. In A. Hughes & D. Porter (Eds.), *Current development in language testing* (pp. 121-127). London: Academic Press.
- Alderson, J. C. & Urquhart, A. H. (1985). The effects of students' academic discipline on their performance in ESP reading tests. *Language Testing*, 2(2), 192-204.
- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 61(3), 315-343.
- Anderson, R. C., & Pearson, D. P. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Eds.), *Handbook of reading research* (pp. 255-291). New York: Longman.
- Bachman, L., & Palmer, A. (1996) *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Callender, A. A. (2008). *The effects of prior knowledge relevance and organization on text comprehension and memory*. Unpublished doctoral dissertation, Washington University, Saint Louis, St. Louis.
- Carrell, P. L. (1983). Some issues in studying the role of schemata, or background knowledge, in second language comprehension. *Reading in a Foreign Language*, 1(2), 81-92.
- Carrell, P. L. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly*, 21(3), 461-481.

- Carrell, P. L. & Eisterhold, J. C. (1983). Schema theory and ESL reading pedagogy. *TESOL Quarterly*, 17(4), 553-569.
- Chen, Q., & Donin, J. (1997). Discourse processing of first and second language biology texts: Effects of language proficiency and domain-specific knowledge. *Modern Language Journal*, 81(2), 209-227.
- Clapham, C. (1996). *Studies in language testing 4. The development of IELTS: A study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press/UCLES.
- Cook, V. (2008). *Second language learning and language teaching (4th Ed.)*. New York: Arnold Publication.
- Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2), 311-325.
- Dudley-Evans, A. & St. John, M. J. (1998). *Developments in English for Specific Purposes*: Cambridge University Press.
- Fisher, D., & Frey, N. (2009). *Background knowledge: The missing piece of the comprehension puzzle*. Portsmouth, NH: Heinemann.
- He, L., & Shi, L. (2012). Topical knowledge and ESL writing. *Language Testing*, 29(3), 443-464.
- Hudson, T. (2007). *Teaching second language reading*. New York: Oxford University Press.
- Hyland, K. (2002). Language, context and literacy. *Annual Review of Applied Linguistics*, 22(1), 113-135.
- Khabbazzbashi, N. (2015). Topic and background knowledge effects on performance in speaking assessment. *Language Testing*, 32(1), 1-26.

- Keene, E. O., & Zimmerman, S. (1997). *Mosaic of thought: Teaching comprehension in a reader's workshop*. Portsmouth, NH: Heinemann.
- Koh, M. Y. (1985). The role of prior knowledge in reading comprehension. *Reading in a Foreign Language*, 3(1), 275-380.
- Koivukari, A. (1987). Question level and cognitive processing: Psycholinguistic dimensions of questions and answers. *Applied Psycholinguistics*, 8(2), 101-120.
- Krekeler, C. (2006). Language for special academic purposes (LSAP) testing: the effect of background knowledge revisited. *Language Testing*, 23(1), 99-130.
- Liu, O. L. (2011). Do major field of study and cultural familiarity affect TOEFL ® iBT reading performance? A confirmatory approach to differential item functioning. *Applied Measurement in Education*, 24(3), 235-255.
- Lumley, T., & O'Sullivan, B. (2005). The effect of test-taker gender, audience and topic on task performance in tape-mediated assessment of speaking. *Language Testing*, 22(4), 415-437.
- McNeil, L. (2011). Investigating the contributions of background knowledge and reading comprehension strategies to L2 reading comprehension: An exploratory study. *Reading and Writing*, 28(2), 883-902.
- Peretz, A. S., & Shoham, M. (1990). Testing reading comprehension in LSP: Does topic familiarity affect assessed difficulty and actual performance? *Reading in a Foreign Language*, 7(1), 447-455.
- Pressley, M., Block, C. C., & Gambrell, L. (2002). *Improving comprehension instruction: Rethinking research, theory, and classroom practice*. Indianapolis, IN: Jossey-Bass.
- Reisman, A., & Wineburg, S. (2008). Teaching the skill of contextualizing in history. *The Social Studies*, 99(5), 202-207.

- Ridgway, T. (1997). Thresholds of background knowledge effect in foreign language reading. *Reading in a Foreign Language*, 11(1), 151-166.
- Roller, C. M., & Matambo, A. R. (1992). Bilingual readers' use of background knowledge in learning from text. *TESOL Quarterly*, 26(1), 129-141.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, N.J.: Erlbaum.
- Schank, R. C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. New York: Cambridge University Press.
- Smagorinsky, P. (2001). If meaning is constructed, what is it made from? Toward a cultural theory of reading. *Review of Educational Research*, 71(1), 133-169.
- Stahl, S. A., Hare, V. C., Sinatra, R., & Gregory, J. F. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension. *Journal of Reading Behavior*, 23(4), 487-508.
- Steffensen, M. S., Joag-Devv C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, 15(1), 10-29.
- Tan, S. H. (1990). The role of prior knowledge and language proficiency as predictors of reading comprehension among undergraduates. In J. H. A. L. de Jong & D. K. Stevenson (Eds.), *Individualizing the assessment of language abilities* (pp. 214-224). Clevedon, UK: Multilingual Matters.
- Tarchi, C. (2009). Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge. *Learning and Individual Differences*, 20(5), 415-420.

- Uso-Juan, E. (2006). The compensatory nature of discipline-related knowledge and English language proficiency in reading English for academic purposes. *The Modern Language Journal*, 90(2), 210-227.
- Uso-Juan, E. (2007). The presentation and practice of the communicative act of requesting in textbooks: Focusing on modifiers. In E. A. Soler and M. P. S. Jorda (Eds.), *Intercultural language use and language learning*. Netherlands: Springer.
- Widdowson, H. G. (2000). On the limitation of linguistics applied. *Applied Linguistics*, 21(1), 3-25.
- Zonghong, L. (2005). A critical review of Hossein Nassaji's article in language learning. *CELEA Journal*, 28(5), 63-96.



Gender Representation and Stereotyping in ESP Textbooks

Seyyed Mohammad Reza Adel

Mostafa Janebi Enayat

Hakim Sabzevari University, Sabzevar, Iran

Bio data

Seyyed Mohammad Reza Adel holds a PhD in TEFL from Ferdowsi University in Mashhad, Iran and is currently an Assistant Professor at Hakim Sabzevari University in Iran. His main research interest includes socio-applied linguistics. He has taught graduate and postgraduate courses for more than a decade, and presented and published papers in many national and international conferences and journals.

Email: sm.adel@hsu.ac.ir

Mostafa Janebi Enayat is a PhD candidate in Applied Linguistics at Hakim Sabzevari University, Sabzevar, Iran. His main research interests include language testing, vocabulary assessment, discourse analysis and critical approaches to materials evaluation. Currently, he is a university lecturer at Golestan University of Medical Sciences and Hakim Jorjani Institute of Higher Education, Gorgan, Iran.

Email: mostafa.enayat@gmail.com

Abstract

The present research was an attempt to explore gender equality in current ESP textbooks as they play a key role in meeting the specific needs of the learners (Hyland, 2002). This study was conducted to investigate the positioning of gender in the images and texts of Oxford English for Careers series including *Commerce*, *Nursing*, *Technology*, and *Tourism*. First, dimensions identified in Goffman's *Gender Advertisements* (1976) were converged with the

image semiotic categories of Kress and van Leeuwen's *Reading Images* (2006) to analyze the images of these ESP textbooks. Second, a systematic quantitative content analysis was carried out with reference to gender visibility in texts. The results indicated that (a) men are more active and competent in technology and tourism; (b) women appear as objects to be scrutinized and objects of desire in all the subjects studied; and (c) men are more frequently presented in workplaces and shown as the breadwinners in commerce, nursing and technology. The results, which mirror institutionalized unfair gender discrimination to the disadvantage of women in society, have implications for materials writers, language instructors, and learners which are fully discussed.

Keywords: gender representation, gender stereotyping, ESP textbooks, image analysis, critical discourse analysis

1. Introduction

One of the most important social issues that has been the focus of many studies and has invoked much debate is "gender", which is often thought of as a bipolar category with males and females as opposite to each other either linguistically or socially (Talbot, 2003). Consequently, people are perceived through a lens of gender polarization and assign certain norms and expectations about verbal and social behavior to males and females (Bem, 1993). There is thus a strong tendency for gender stereotyping to set in. Gender stereotypes are closely related to gender ideologies and can support each other, and many studies have been conducted to find the representation of gender in different domains (Cameron, 1998; Cunningham, 2008; Del-Teso-Craviotto, 2006; Erden, 2009; Kendall, 2004; Kuo, 2003; Ladegaard, 2010; Ohara & Saft, 2003; Pizzini, 1991; Tanaka & Fukushima, 2002).

2. Literature review

2.1. Gender representation in media

A number of studies have examined gender inequality in visual media such as images, dramas, and TV programs (e.g. Giaschi, 2000; Goffman, 1976; Ohara & Saft, 2003). Print media has also gained remarkable attention (Danova, 2006; Del-Teso-Craviotto, 2006; Isanovic, 2006) as a significant domain in which gender inequality can be exercised.

Advertisements were among the first types of media which were analyzed by Goffman (1976). Using an analytical framework in which several dimensions such as gaze, body display, and modality were utilized to decode the ideologies behind the use of images in these advertisements, the researcher found that women were presented as submissive and subordinate, having less power than men. Masse and Rosenblum (1988) also studied the images of advertisements in six traditional magazines and using a similar framework, they found a sex-segregated world of images.

Language and gender researchers have devoted so much attention to different popular media such as newspapers (Litosseliti, 2002; Morrison, 1996), women's magazines (Caldas-Coulthard, 1996; Del-Teso-Craviotto, 2006), or shopping channels (Bucholtz, 1999). A couple of studies investigated the representation of women in south east Europe, among which Isanovic (2006) examined gender representation in the daily newspapers of Bosnia and Herzegovina. The author selected a few types of newspapers with different themes including world, economy and business, entertainment, culture and art, sport, and politics. The framework of analysis was a set of modes of positioning of six techniques: (1) presence/absence; (2) themes; (3) voices/sources; (4) social relations/occupations; (5) visual representation; and (6) gender-sensitive language. The results indicated that women were presented as submissive and weak in comparison with the image of men. In a similar article,

Danova (2006) studied the representation of women in two political Bulgarian newspapers namely *Ataka* and *Trud* and found that women were presented in stereotypical roles. They were presented as the objects of the male gaze, wives and mothers, or as violent and unruly persons. Kress and van Leeuwen (2006) explained that "the gaze of represented participants directly addresses the viewers and so establishes an imaginary relation with them, while more schematic analytical pictures invite impersonal, detached scrutiny" (p. 89).

2.2. Gender Representation in ELT

Critical Discourse Analysis (CDA) is the analysis of linguistic and semiotic aspects of social processes and problems (Fairclough & Wodak, 2010). It is an interdisciplinary field which makes use of other studies to deal with social and cultural problems. Consequently, the field of ELT has not remained an exception and has made use of CDA approaches to deal with its

socio-cultural problems. Many social issues have been raised in the domain of ELT, one of which has been the significant issue of gender equality. Some of the studies which have been concerned with this issue are briefly reviewed here.

In an attempt to investigate gender positioning in ESL textbooks, Giaschi (2000) did a critical image analysis to find the stereotypical ideologies of gender. The images selected for his study were drawn from the award-winning *Headway* series and *4th Dimension*. The results of Giaschi's study revealed that men were given the active role in the images while women were mostly passive in the images, and they were presented as submissive, weak and dependent. Liao and Huang (2003) also examined primary EFL textbooks to see how they observe gender equality. The content analysis of the dialogues provided evidence that men and women were equally presented; however, the analysis of third-person singular pronouns showed that men were more dominant. Sano, Lida, and Hardy (2001) investigated gender positioning in five series of high school English textbooks in Japan. While the linguistic features showed no evidence, some textbooks presented gender-biased messages. In a similar study, Nakamura (2002) focused on the literary stories of Japanese high school EFL textbooks to examine the images of women. Using both quantitative and qualitative analyses, the results indicated that at least 80% of the stories presented women as emotional, sad, guilty, and foolish while men were shown as warm-hearted, having self-control and greatness.

In Hong Kong, Lee (2014) investigated the status of women as reflected in English-language textbooks of their primary schools. The author compared the representation of women in textbooks published in 2005 with the same series published in 1988 and found that gender equity was more in the new textbook series. However, the stereotypical portrayal of women, as the author found, was still evident as they had more limited range of social roles. Stockdale (2006) tried to examine gender representation in the EFL textbook called *Impact Values*. Several categories were analyzed by the researcher and the results indicated that men were more visible in all the categories of this study and they outperformed women in almost all the categories. Levine and O'Sullivan (2010) examined gender bias and female invisibility in images and illustrations of EFL textbooks written for Japanese university students. Using quantitative and qualitative analyses, the results of their analyses showed that males occupied a great number of social roles while females were pictured as schoolgirls in low-status occupations.

Although women were mostly portrayed having less social power than men, as reviewed above, some other studies found that the two genders were equally presented in English-language textbooks. Healy (2009), for instance, studied the representation of men and women in modern EFL textbooks. The overall analysis of the study revealed that gender was equally presented in the textbooks and no bias was found by the researcher. In other words, both men and women had the same opportunity to have a voice in the textbook. Similarly, Mineshima (2008) investigated how an EFL textbook, used in upper secondary English classes in Japan, represented the two genders. The results indicated that both genders were represented equally in the number of male and female characters, their utterances, number of occupations, and family roles. Lee and Collins (2009) also examined the extent of gender stereotyping in Australian ELT textbooks and came to the conclusion that men and women were equally represented in the content and images of the textbooks.

2.3. Gender positioning in Iran

The status of sexism in EFL textbooks was studied by Ansary and Babaii (2003) in the context of Iran. Trying to examine the issue of sexist language and attitudes in EFL materials, the authors used both quantitative and qualitative analyses to analyze two textbooks (*Right Path to English I & II*). The results indicated that the textbooks could be considered sexist as they presented students with an unfair and inexcusable picture of women. Bahman and Rahimi (2010) also studied gender representation in EFL textbooks of Iranian high schools. Three volumes of English textbooks were analyzed by the researchers using linguistic and content analyses and the representation of men and women was found to be unfair. A similar study was conducted to examine the presentation of gender in high school and pre-university ELT textbooks by Fatemi, Pishghadam, and Heidarian (2011). They found evidence for the imbalanced representation of males and females and the failure of the textbooks to reflect the wide range of roles played by women in the society. In a more recent study, Kordjazi (2012) investigated gender positioning in the images of two English-learning software programs, namely *Tell Me More* and *English at Home*. The author found that men appeared as powerful, dominant, competent, and active while women appeared as powerless, subordinate, objects to be scrutinized, and reactive.

3. The present study

As seen in the above-mentioned studies, gender representation was mainly examined in ESL/EFL textbooks which target general English. Although ESP has been characterized as “centered on the language and activities appropriate to particular disciplines, occupations and activities and required by particular learners” (Hyland, 2002, p. 386), the content of ESP textbooks has gained little attention in the literature of CDA in general, and gender representation in particular. Given the importance of ESP in the field of ELT and the fact that ESP textbooks play a vital role in distributing certain ideologies, the present study aimed at finding the representation of gender in the images and content of Oxford English for Careers series, as widely published ESP textbooks. Image and content analyses were conducted to answer the following research questions:

1. Who is active and who is passive in the images?
2. How is the gaze directed?
3. What does the clothing imply?
4. What are the frequent spaces in which the two genders are presented?
5. Who is more visible in the content of the textbooks?

3.1. Corpus

The images and content of four ESP textbooks developed by Oxford University Press for Careers were used as the corpus of this study: (1) *Oxford English for Careers: Nursing* (Grice, 2009); (2) *Oxford English for Careers: Commerce* (Hobbs & Keddl, 2009); (3) *Oxford English for Careers: Technology* (Gledinning & Pohl, 2009); and (4) *Oxford English for Careers: Tourism* (Walker & Harding, 2009).

3.2. Analytical Framework

3.2.1. Image Analysis

Fairclough (1995) argued that texts are now “increasingly multi-semiotic” (p. 4) and cannot be limited merely to a written or transcribed spoken discourse. Kress and van Leeuwen (1996) also mentioned that visual elements of texts are gaining greater importance in the present world of multiculturalism and globalization. According to Kress and van Leeuwen (1996), meanings contained by pictures apply to deeper levels of human perception and

transfer more comprehensive messages than a written text. In the present research, Kress and van Leeuwen's semiological analysis (2006) was converged with Goffman's (1976) content analysis (1979) to investigate gender positioning of images included in the textbooks. The criteria that were used in this study are as follows.

Four values were used to identify the actual role of each participant (*narrative presentation*) based on Kordjazi (2012). First, participants could play the role of an actor who do something to another participant; second, they could be the goals of another participant's action; third, the participants could be behaviors performing a non-transitive action such as smiling or running; and finally, participants could be reacting to a model, object or situation. Two kinds of *Gaze* (eye-direction) were identified: at the viewer and away from the viewer. The clothing of males and females was examined under three categories: sparsely clothed (e.g., shorts or bikini) lightly clothed (e.g., T-shirt or top), and fully clothed (e.g., business suit and long skirt). One of the variables that is very important in analyzing images is the variable of *Space*. Five values were set for space, namely home, workplace, street and neighborhood, leisure areas, and shop.

3.2.2. Content Analysis

In this phase, occurrences of female and male characters in terms of names, titles, pronouns and special nouns were examined. Female and male characters were examined in terms of gender related pronouns, nouns, and names. Total instances of gender related pronouns (she/her/hers/herself and he/his/him/himself), special female and male nouns (madam, sir, mother, father, wife, husband, bride, bridegroom) and nouns with unmarked gender (e.g. neighbor, cousin) naming female/male characters in the text, and titled (e.g. Mrs. Clark, Mr. Sun) names related to female and male characters were counted in the content of the textbook.

4. Results and discussion

4.1. Who is active and who is passive in the images?

The following table shows the narrative representation of the participants shown in the images of "*Oxford English for Careers: Commerce* (Hobbs & Keddl, 2009), *Oxford English*

for *Careers: Nursing* (Grice, 2009), *Oxford English for Careers: Technology* (Gledinning & Pohl, 2009), and *Oxford English for Careers: Tourism* (Walker & Harding, 2009)". This analysis shows the active participants of the image by showing the role of each gender in the images.

Table 1: Narrative role and gender of the represented participants (%)

Textbook	Commerce		Nursing		Technology		Tourism	
Factor	M	F	M	F	M	F	M	F
Actor	4 (8.8)	3 (5.7)	20 (27)	8 (12.1)	10 (21)	1 (6.2)	10 (15.6)	1 (2)
Goal	0 (0)	3 (5.7)	9 (12.1)	9 (13.6)	5 (10.4)	6 (37.5)	7 (11)	5 (10.2)
Behavior	19 (42.2)	18 (34.6)	18 (24.3)	23 (34.8)	3 (6.2)	5 (31.2)	11 (17.2)	19 (38.7)
Reactor	22 (49)	28 (53.8)	27 (36.4)	26 (39.3)	30 (62.5)	4 (25)	36 (56.2)	24 (49)
Total	45 (100)	52 (100)	74 (100)	66 (100)	48 (100)	16 (100)	64 (100)	49 (100)

Table 2: Results of Chi-Square Test

Pearson Chi-Square	Value	df	Sig (2-sided)
Textbook			
English for Commerce	3.402	3	.334
English for Nursing	5.332	3	.149
English for Technology	15.783	3	.001**
English for Tourism	10.423	3	.015*

** Significant at $p = .01$

* Significant at $p = .05$

Table 1 shows that in *English for Commerce*, males were more active in the images than females; however, this difference was not noticeable and could be, therefore, disregarded. For the second value, it was indicated that men were not shown as the goal of another participant while women were presented as the goal of another participant in 5.7% of the cases. While women outnumbered men for being the goal of another participant, men were presented more than women as playing a non-transitive role. The difference between males and females for reacting to a model, object or situation was not significant and showed women as more reacting than men. The results of Chi-Square analysis also indicated that there was not a

significant difference between males and females as for the narrative representation with $\chi^2 = 3.402, p > .05$.

In *English for Nursing*, men were shown as more active participants compared with women. They were the active participants in 27% of the images while women had the active role in only 12.1% of the images. In addition, women were more likely to have non-transitive role in the images than men with 34.8% of the instances. Although a slight difference was found between the two genders, it was not statistically significant with $\chi^2 = 5.332, p > .05$.

The difference between the two genders was highly significant in *English for Technology*. Men had the active role in 21% of the images while women were the active participants in only 6.2% of the cases. Furthermore, females were the goal of another participant in 37.5% of the cases which depict them as powerless participants who are likely to be the goal of another active participant. They also had non-transitive roles in 31.2% of the cases, while men performed non-transitive actions in only 6.2% of the images. The results were also statistically significant with $\chi^2 = 15.783, p < .01$. Consequently, it was found that men were shown as active and powerful participants of the images in *English for Technology* which is considered as a hard science.

Similarly, in *English for Tourism*, men were presented as active participants more than women. In addition, women had a non-transitive role in 38.7% of the cases, while men had such roles in 17.2% of the cases. The Chi-Square analysis also showed that this difference was statistically significant ($\chi^2 = 10.423, p < .05$). Consequently, men were presented as more active and powerful participants in the images of *English for Tourism* which falls into soft science.

4.2. How is the gaze directed?

Kress and van Leeuwen (2006) believe that one of the key features which can uncover the ideology behind the presentation of gender in the images is the gaze or the eye direction of the participants. The participants may look at the viewer or look away from the viewer. The following table shows the gaze of each gender in the images of ESP textbooks.

Table 3: Gaze and gender of the represented participants (%)

Textbook	Commerce		Nursing		Technology		Tourism	
Factor	M	F	M	F	M	F	M	F
At the viewer	14 (35.8)	18 (42.8)	15 (22)	17 (30)	7 (15.2)	6 (35.3)	10 (16.4)	20 (36.4)
Away from the viewer	25 (64.2)	24 (57.2)	53 (78)	41 (70)	39 (84.8)	11 (64.7)	51 (83.6)	35 (63.6)
Total	39 (100)	42 (100)	68 (100)	58 (100)	46 (100)	17 (100)	61 (100)	55 (100)

Table 4: Results of Chi-Square Tests

	Pearson Chi-Square	Value	df	Sig (2-sided)
Textbook				
English for Commerce		1.041	1	.308
English for Nursing		1.388	1	.239
English for Technology		10.703	1	.001 [*]
English for Tourism		10.432	1	.001 [*]

*Significant at $p = .01$

The analysis of gaze direction, as shown in Table 3, indicates that in *English for Commerce*, males gazed away from the viewer in 64.2% of the cases while females gazed away from the viewer in 57.2% of the cases. Dyer (1992) believed that looking at the viewer can represent the weakness of the participant and men do not, therefore, tend to do that. Women on the other hand, looked at the viewer in 42.8% of the images which is more than that of men with 35.8% of the cases. However, the difference was not statistically significant ($\chi^2 = 1.041$, $p > .05$).

In *English for Nursing*, men gazed away from the viewer in 78% of the cases while women gazed away from the viewer in 70% of the cases. Females looked at the viewer in more cases compared with males. This difference was not statistically significant at $p > .05$. Although the difference between men and women in *English for Commerce* and *English for Nursing* is in favor of men, it is not significant and can be thus disregarded.

In contrast, the results of *English for Technology* indicated that not only did men outperform women in the images, but also they were presented as more powerful participants as far as their gaze in the images was concerned. They gazed away from the viewer in 84.8% of the

cases, while women gazed away in 64.7% of the images. Moreover, women's gaze at the viewer was more than twice the size of men's gaze. The results were also statistically significant ($\chi^2 = 10.703$, $p < .01$). The results clearly indicated that men were presented as more powerful participants than women.

Similarly, the analysis of the images in *English for Tourism* indicated that women gazed at the viewer more than men. In addition, men also gazed away from the viewer more than women with more than 83% of cases for men and 63.6% for women. The analysis of Chi-Square also showed a statistically significant difference ($\chi^2 = 10.432$, $p < .01$). Consequently, both *English for Technology* and *English for Tourism* presented a less powerful image of women in comparison with men.

4.3. What does the clothing imply?

One of the main features of an image that can affect the representation of the participants is clothing which shows the participant's beliefs (van Leeuwen, 2005). Clothing can communicate messages about the participants' religion, occupations, values, and attitudes (Danesi, 2004). Clothes are significant means by which certain stereotypes can be made out of people. The following table shows the basic modes of presentation for clothing in which the two genders were presented in Oxford English for Careers series.

Table 5: Body display and gender of the represented participants (%)

Textbook	Commerce		Nursing		Technology		Tourism	
Factor	M	F	M	F	M	F	M	F
Sparsely clothed	0 (0)	1 (2)	2 (2.5)	0 (0)	0 (0)	0 (0)	4 (5.9)	10 (18.5)
Lightly clothed	5 (9.8)	18 (35.2)	23 (28.7)	33 (49.2)	9 (16.6)	10 (58.8)	14 (20.6)	16 (29.6)
Fully clothed	46 (90.2)	32 (62.7)	55 (68.7)	34 (50.7)	45 (83.4)	7 (41.2)	50 (73.5)	28 (51.8)
Total	51 (100)	51 (100)	80 (100)	67 (100)	54 (100)	17 (100)	68 (100)	54 (100)

Table 6: Results of Chi-Square Tests

Pearson Chi-Square	Value	df	Sig (2-sided)
Textbook			
English for Commerce	10.861	2	.004**
English for Nursing	7.651	2	.022*
English for Technology	11.724	1	.001**
English for Tourism	7.401	2	.025*

**Significant at $p = .01$

*Significant at $p = .05$

The results of Table 5 indicated that in *English for Commerce*, men were presented in almost all of the cases fully clothed. In contrast, women were presented less fully clothed in comparison with men. While men were not presented as sparsely clothed at all, women were shown as sparsely clothed in 2% of the cases. Although this difference is not significant, women were lightly clothed in 35.2% of the images while men were lightly clothed in only 9.8% of the cases. The results of Chi-Square in Table 6 also showed that the differences between the two genders were statistically significant ($\chi^2 = 10.861, p < .01$).

The analysis of *English for Nursing* also indicated that women were presented lightly clothed in nearly 50% of the images while men were shown lightly clothed in 28.7% of the cases. Furthermore, men were presented more fully clothed than women. Table 6 also showed a significant difference between the two genders ($\chi^2 = 7.651, p < .05$). As a result, women were more likely to be shown as objects of desire.

The differences were more significant in *English for Technology* in which women were presented lightly clothed in more than 58% of the cases while men were shown lightly clothed in only 16.6% of the images. Moreover, men were fully clothed in 83.4% of the images, whereas women were fully clothed in 41.2% of the cases. Table 6 also shows that the difference was highly significant with $\chi^2 = 11.724, p < .01$. The results clearly revealed that gender stereotyping was prevalent in the images of this textbook.

Similarly, gender differences were found in the images of *English for Tourism* where women were sparsely clothed in 18.5% of the images, while men were sparsely clothed in only 5.9% of the images. They were also more lightly clothed than men. In addition, men were more fully clothed than women in the images of this textbook. The same result was found by Chi-

Square analysis ($\chi^2 = 7.401, p < .05$). Consequently, the results for all the textbooks indicated a clear gender difference in favor of men who were presented as visually more powerful which in turn represents their real-life social roles.

The presentation of women in lightly clothed ways, which shows parts of their body, position women as objects to be observed (Kress & van Leeuwen, 2006). This kind of presentation may prove the claim put forward by Berger (1972) which stated that women are usually portrayed as objects to be desired and possessed and men as subjects of envy. The results can also confirm the results of Giaschi (2000) which indicated that men are usually portrayed in three basic modes: casual, professional, and power-dressed. Moreover, the results are in line with Kordjazi (2012) which found that women are shown as objects to be scrutinized in the images of the English-learning software programs.

4.4. What are the frequent spaces in which the two genders are presented?

Space or the situation in which people are presented is a key element that can affect gender positioning in images. By space we mean the different locations that people are usually presented in daily life. It can be indoor spaces (like home, shop, or work places) or outdoor spaces such as street, neighborhood and work places. The following table shows the different spaces in which the participants were presented in the images of ESP textbooks.

Table 7: Space and gender of the represented participants (%)

Textbook	Commerce		Nursing		Technology		Tourism	
	M	F	M	F	M	F	M	F
Factor								
Home	0 (0)	3 (7.6)	4 (7)	11 (23.4)	2 (4.3)	3 (21.4)	0 (0)	0 (0)
Workplace	24 (55.8)	16 (41)	36 (63)	24 (51)	39 (84.8)	9 (64.2)	20 (32.2)	14 (28)
St./Neigh.	4 (9.3)	2 (5.1)	8 (14)	4 (8.5)	3 (6.5)	1 (7.1)	2 (3.2)	3 (6)
Leisure	13 (30.2)	16 (41)	9 (15.7)	8 (17)	2 (4.3)	1 (7.1)	39 (62.9)	32 (64)
Shop	2 (4.7)	2 (5.1)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1.6)	1 (2)
Total	43 (100)	39 (100)	57 (100)	47 (100)	46 (100)	14 (100)	62 (100)	50 (100)

Table 8: Results of Chi-Square Tests

Pearson Chi-Square	Value	df	Sig (2-sided)
Textbook			
English for Commerce	11.995	4	.017 [*]
English for Nursing	11.558	3	.009 ^{**}
English for Technology	15.153	3	.002 ^{**}
English for Tourism	1.612	3	.657

**Significant at $p = .01$

*Significant at $p = .05$

The results of Table 7 revealed that in *English for Commerce*, women were presented in more indoor spaces such as home with 7.6% of occurrence. Males, on the other hand, were not presented in home space at all. Furthermore, males were more presented in workplace than females. This clearly indicates that men play the role of bread winners who work outside to make a living. Women, on the other hand, are more likely to be inside and work less than men. In addition, men were presented in outdoor spaces more than women. The results of Chi-Square also indicated that the difference was statistically significant with $\chi^2 = 11.995$, $p < .05$.

In *English for Nursing*, women were more shown at home than men. Moreover, men were more presented in workplace spaces than women. As a result, men were shown as more powerful participants than women who were more likely to be presented in indoor spaces like home. Moreover, women were more likely to be presented in leisure spaces and less in streets and outdoor spaces. The results of Chi-Square in Table 8 also indicated that this difference was significant ($\chi^2 = 11.558$, $p < .01$).

The analysis of *English for Technology* also proved a more socially powerful position for men. Women were presented at home spaces in more than 21% of the cases while men were shown at home spaces in only 4.3% of the cases. Furthermore, men were more likely to be presented in workplace than women with more than 84% of the cases. The result was statistically significant ($\chi^2 = 15.153$, $p < .01$). As a result, the number of men's appearance in the images alongside the spaces in which they were shown clearly indicated that they were represented as significantly more powerful than women.

In contrast, the two genders were equally presented in *English for Tourism* as far as the spaces in which they were shown were concerned. The percentages of their appearance in all the spaces were approximately the same with no significant difference. Similarly, the Chi-Square analysis indicated no significant difference between the two genders ($\chi^2 = 1.612$, $p > .05$).

The findings indicated that women were more likely to be represented in less important spaces such as home and leisure. In addition, they were less presented in outdoor spaces which have more power and social status such as workplace and neighborhood. Women were stereotyped as individuals who should work in the kitchen and play the role of bread takers while men were presented as the ones who should work and manage the family affairs and were presented in more workplaces and outdoor areas which deal with serious jobs and businesses.

4.5. Who is more visible in the content of the textbooks?

This part of the study focused on the presence of females and males in the content of ESP textbooks. In the initial stage, the general visibility of both genders was studied. Firstly, the words naming female and male characters in reading passages and exercises were counted.

Analysis was done using the following set of criteria: examination of female and male occurrences in texts, calculation of frequencies of female and male pronouns, nouns and names, and investigation of gendered presentations in ESP textbooks.

Table 9: Gender Pronouns, Nouns, Proper names and Titled Names (%)

Textbook	Commerce		Nursing		Technology		Tourism	
	M	F	M	F	M	F	M	F
Pronoun	76 (24)	49 (21)	156 (51)	144 (47)	112 (49)	24 (52)	58 (53)	34 (33)
Noun	11 (3)	13 (5)	30 (10)	42 (14)	5 (2)	3 (6)	1 (1)	3 (3)
Proper N.	216 (68)	152 (66)	87 (28)	91 (30)	101 (44)	18 (39)	40 (36)	59 (57)
Titled N.	12 (4)	16 (7)	33 (10)	25 (8)	9 (4)	1 (2)	10 (9)	6 (6)

Table 10: Results of Chi-Square Tests

Pearson Chi-Square Textbook	Value	df	Sig (2-sided)
English for Commerce	4.554	3	.207
English for Nursing	3.647	3	.302
English for Technology	3.084	3	.379
English for Tourism	11.688	3	.009*

*Significant at $p = .01$

Table 9 demonstrates the number of gender pronouns, nouns, proper nouns, and titled nouns. The results of content analysis indicated that in *English for Commerce*, the visibility of men and women was almost the same. The percentage of occurrence for each category was so much different between males and females. The results of Chi-Square also showed that this difference was not significant with $\chi^2 = 4.554$, $p > .05$. However, the total instances of occurrence were more for males (315 > 230). This can show gender inequality in *English for Commerce* dealing with business and trading. In other words, men were presented as the dominant group in the field of business and trading and women were shown as subordinate and powerless.

The results of content analysis for *English for Nursing* were quite different though. The frequency of gendered nouns and pronouns showed that males and females were equally presented in almost all categories of content analysis. The percentage of male and female pronouns was 51% and 47%, respectively. In addition, the percentage of male and female proper names was not different which indicates an equal presentation of both genders. Chi-Square analysis also showed no significant difference ($\chi^2 = 3.647$, $p > .05$).

The results of Table 9 showed that the content of *English for Technology* was greatly imbalanced for the two genders. Although the one-by-one comparison of each category showed no statistically significant difference ($\chi^2 = 3.084$, $p > .05$), the overall results of content analysis indicated that the number of occurrences for males was 227 while this was only 46 for females. This clearly shows that men were the dominant group in this textbook with a number almost five times bigger than that of women. Consequently, the results provided evidence for a highly significant difference between the presentation of men and women in the content of *English for Technology*.

The content analysis of *English for Tourism* showed that males and females were presented equally. The overall results indicated that the total instances of occurrence for males were 109 and 102 for females. While the results of Chi-Square indicated a significant difference ($\chi^2 = 11.688, p < .01$), this was in favor of males for pronouns and titled nouns. For other categories, this difference was in favor of females. As a result, content analysis showed that *English for Tourism* is not biased towards any of the two genders as far as the textual

5. Conclusion and implications

The term language ideology generally refers to “sets of representations through which language is imbued with cultural meaning for a certain community” (Cameron, 2003, p. 447). Ideology is in contrast with “truth” and usually refers to the representation of certain beliefs or attitudes in a language. Fairclough (1989) relates ideology to power relations and defines it as a phenomenon which “sustains unequal relations of power” (p. 84). This ideology is not merely represented in the linguistic mode of presentation since all images, as argued by Kress and van Leeuwen (2006), are "entirely in the realm of ideology" (p. 12).

The use of images can be used ideologically to present a group of individuals as weak and less socially powerful (Kress & van Leeuwen, 2006). As Wodak (2006) mentioned, the construction of meanings by the media relates to several dimensions of semiotics including texts and images and “images have been and are being used today for political and media interests in various ways” (p. 3). The term ‘ideology’ is a significant aspect of creating and establishing power relations (Wodak, 2002). Images, as one of the codes for communication, can thus bear ideological meanings which can sustain unequal power relations. They can “stand for different points of view, for meta-narratives, for ideologies” (Wodak, 2006, p. 3).

The current study aimed at finding the positioning of gender in ESP textbooks and the possible gender inequalities that may be exercised in the images and texts of these textbooks. As it was put forward before, the power of media cannot be denied. ESP textbooks, as instruments used by so many people throughout the world, are therefore identified as powerful means of media and their hidden ideologies are worth studying. When a group of individuals like women are portrayed as weak and socially less important in certain types of media, this ideology is more likely to be established in real life. In other words, there is a

close relationship between the presentation of individuals in media and what they are going to be thought of in the society.

In the images of the textbooks analyzed for the purpose of this study, the results indicated that women were frequently shown as the reactive participants of the image. They were more likely to be the goal of the male party than men. In contrast, men were portrayed as the active participants of the images holding more social power and prominence. The results of this part confirmed the findings of previous studies which aimed at investigating the positioning of gender in ELT textbooks (e.g. Bahman & Rahimi, 2010; Giaschi, 2000; Lee, 2014; Liao & Huang, 2003; Sano et al., 2001; Nakamura, 2002; Stockdale, 2006).

The present study focused on four ESP textbooks known as “Oxford English for Careers” series. The overall analysis showed that *Oxford English for Careers: Technology* was the most stereotyped textbook. Males and females were unequally presented in the images of this textbook which deals with technology as a hard science. *Oxford English for Careers: Commerce* was the second most stereotyped textbook which deals with business, a field known to be more masculine. These two textbooks depicted a more unequal depiction of gender than the other textbooks. *Oxford English for Careers: Tourism* and *Oxford English for Careers: Nursing*, which deal with rather less serious subjects, were less stereotyped than the other two textbooks. This clearly shows that women are more likely to be discriminated against in contexts and fields where men are considered to be the dominant party.

Kress and van Leeuwen (2006) argued that all images are ideological and no single image is free from stereotypical power. They mentioned that particular discourses are privileged in visual representations while others may be downplayed or even silenced. Negative and unfair positioning of gender may have destructive consequences (Wodak, 2006). When genders are portrayed as different and unequal, such positioning may narrow the range of roles women may play in society (Milburn, Carney, & Ramirez, 2001). Learning a foreign language through ESP textbooks is a mean to learn another culture and inappropriate positioning of gender may fossilize certain stereotypes of women in the mind of the learners (Mineshima, 2008). It has been argued that ELT materials writers should represent an unprejudiced role of women in the content and images of textbooks to maintain social justice (Laakkonen, 2007).

From the perspective of social cognitive theory (Bandura, 2003; Burr, 1998), schools and educational materials play an important role in gender development. Lee (2014) pointed out that based on the framework of social cognitive framework, language learners "may learn about gender role expectations that are associated with particular settings and the typical actions associated with those places" (p. 369). In other words, stereotypical depictions of the two genders can reinforce learners' gendered schemas. Viewing stereotypical positioning of women in media can affect the users' perceptions and form their future expectations (Smith & Granados, 2009). Constant exposure to gender bias in textbooks may leave stereotypical effects on learners' minds and cause them to maintain inequitable perceptions.

Conducting critical studies for analyzing images and content of ESL/EFL/ESP textbooks for finding gendered discursive practices is a must since they function at a stereotypical level (Mills, 1995) and can have certain implications. This study also has particular implications for gender identity in education with reference to their representation in ESP textbooks as means for teaching and learning a foreign language. The results of this study have implications for materials writers, teachers, and learners.

Showing the hidden gender stereotypes in ESP textbooks can raise the awareness of the stakeholders involved in education in general, and language education in particular. Materials writers need to be "enlightened in terms of gender sensitivity and must be further informed of the guidelines directed to them" (Kordjazi, 2012, p. 71). The findings of this study showed that materials writers should be made aware of the potential social impact of the texts and images they use in ESP textbooks. They have to take this into account and try to maintain gender equity and avoid perpetuating stereotypical presentation of the two genders.

The findings of this study also have implications for teacher education, including for critical linguistic awareness in relation to textbook and text selection, for textbook use and adaptation. ESP instructors need to be meticulous in choosing the right teaching materials and use the textbooks and materials that have the least degree of gender bias (Mineshima, 2008). Giaschi (2000) also believes that studies which investigate gender equality in ELT textbooks can provide support to ESL teachers to avoid reinforcing gender inequity on a global scale. Mineshima (2008) further points out that the teacher can prevent gender

discrimination from taking place in the classroom and can even adapt the texts and/or images before bringing them to the class.

This study can have implications for language learners as well. The ESP learners could be made aware of the possible gender inequity they might face in the texts and images of the textbooks they use in the classroom. Filak (2002) believes that conducting such studies can change the worldviews and attitudes of language learners.

6. Limitations and suggestions for further research

The present study investigated the positioning of gender in four ESP textbooks published by Oxford University Press selected from Oxford English for Careers series. There are a number of other ESP textbooks available on the market published by other authors and publishers which were not examined in this study. Consequently, the results of this study may not be generalized to all ESP textbooks. Analyzing more samples and textbooks from different publishers could have improved the generalizability of the findings of the present study. In addition, focusing on textbooks alone, as argued by some researchers, may not be sufficient since the way a text or image is interpreted "depends, to some extent, on the readers' responses to the text" (Lee & Collins, 2009, p. 366). In other words, gender equity or inequity may not be recognized by some learners or teachers as the users of ESP textbooks.

Future studies can investigate gender representation in other ESP textbooks and/or materials such as software programs. In addition, the methodology can shift from a textbook-based analysis to an audience-oriented one. To put it simply, the positioning of gender could be investigated through the lenses of the teachers and learners who use the textbooks. This could be conducted qualitatively by using observation, interviews and questionnaires.

References

- Ansary, H., & Babaii, E. (2003). Subliminal Sexism in Current ESL/EFL Textbooks. *Asian EFL Journal*, 5 (1), 200-241.
- Bahman, M., & Rahimi, A. (2010). Gender representation in EFL materials: an analysis of English textbooks of Iranian high schools. *Procedia Social and Behavioral Sciences* 9, 273–277.
- Bandura, A. 2003. *Bandura's Social Cognitive Theory: An Introduction (electronic resource)*. San Luis Obispo, CA: Davidson Films, Inc.
- Bem, Sandra L. (1993). *The Lenses of Gender: Transforming the Debate on Sexual Inequality*. New Haven, CT: Yale University Press.
- Berger, J. (1972). *Ways of seeing*. London: BBC/Harmondsworth: Penguin.
- Bucholtz, M. (1999). Purchasing power: the gender and class imaginary on the shopping channel. In Bucholtz, A. C. Liang & L. A. Sutton (Eds.), *Reinventing identities: The gendered self in discourse* (pp. 348-368). Oxford: Oxford University Press.
- Burr, V. (1998). *Gender and Social Psychology*. London: Routledge.
- Caldas-Coulthard, C.R.(1996). Women who pay for sex. And enjoy it': Transgression versus morality in women's magazines. In Caldas-Coulthard, C.R, & Coulthard, M. (Eds), *Texts and Practices: Readings in Critical Discourse Analysis* (pp. 250-270). London: Routledge.
- Cameron, D. (1998). Is There Any Ketchup, Vera: Gender, Power and Pragmatics. *Discourse & Society*, 9(4), 437-455.
- Cameron, D. (2003). Gender and Language Ideologies. In Holmes, J. & Meyerhoff, M. (Eds), *The Handbook of Language and Gender* (pp. 447-467). Blackwell Publishing Ltd.

- Cunningham, M. (2008). Influences of gender ideology and housework allocation on women's employment over the life course. *Social Science Research*, 37, 254-267.
- Danesi, M. (2004). *Messages, signs, and meanings: A basic semiotic textbook in semiotics and communication theory*. Toronto: Canadian Scholars' Press.
- Danova, M. (2006). Women in politics in Bulgarian newspapers: Post-feminism in a post-totalitarian society. In Bamburak, N. M., Jusic, T., & Isanovic, A. (Eds), *Stereotyping: Representation of Women in Print Media in South East Europe* (pp. 111-132). Sarajevo: MediaCentra.
- Del-Teso-Craviotto, M. (2006). Words that matter: Lexical choice and gender ideologies in women's magazines. *Journal of Pragmatics*, 38 (11), 2003-2021.
- Dyer, R. (1992). *Only Entertainment*. London: Routledge.
- Erden, F. T. (2009). A course on gender equity in education: Does it affect gender role attitudes of preservice teachers? *Teaching and Teacher Education*, 25 (3), 409-414.
- Fairclough, N. (1989). *Language and Power*. London: Longman.
- Fairclough, N. (1995). *Critical Discourse Analysis: The Critical Study of Language*. London: Longman.
- Fairclough, N., & Wodak, R. (2010). Critical Discourse Analysis in Action. In Coffin, C., Lillis, T., and O'Halloran, K. (Eds), *Applied Linguistics Methods: A Reader* (pp. 98-111). London: Routledge.
- Fatemi, A.H., Pishghadam, R., & Heidarian, Z. (2011). Gender Delineation in High school and Pre-university ELT Textbooks: A Criterion-Oriented Approach to Text Analysis. *The Iranian EFL Journal*, 7 (3), 32-47.
- Filak, A. (2002). Men and women between cultures: Gender stereotypes in EFL coursebooks in Poland. *Neolit On-line*, 1(3).

- Giaschi, P. (2000). Gender Positioning in Education: A Critical Image Analysis of ESL Texts. *TESL Canada Journal*, 18 (1), 32-46.
- Gledinning, E.H., & Pohl, A. (2009). *Oxford English for Careers: Technology*. Oxford: Oxford University Press.
- Goffman, E. (1976). *Gender Advertisements*. New York: Harper & Row Publishers.
- Grice, T. (2009). *Oxford English for Careers: Nursing*. Oxford: Oxford University Press.
- Healy, D. (2009). The representation of women and men in a modern EFL textbook: Are popular textbooks gender biased? *Memoirs of the Osaka Institute of Technology, Series B*, 54 (2), 91-100.
- Hobbs, M., & Keddl, J.S. (2009). *Oxford English for Careers: Commerce*. Oxford: Oxford University Press.
- Hyland, K. (2002). Specificity revisited: how far should we go now? *English for Specific Purposes*, 21, 385-395.
- Isanovic, A. (2006). Media discourse as a male domain: Gender representation in the daily newspapers of Bosnia and Herzegovina, Croatia and Serbia. In Bamburak, N. M., Jusic, T., & Isanovic, A. (Eds), *Stereotyping: Representation of Women in Print Media in South East Europe* (pp. 43-79). Sarajevo: MediaCentra.
- Kendall, S. (2004). Framing authority: gender, face, and mitigation at a radio network. *Discourse & Society*, 15 (1), 55-79.
- Kordjazi, Z. (2012). Images matter: A semiological content analysis of gender positioning in contemporary English-learning software applications. *Novitas-ROYAL (Research on Youth and Language)*, 6 (1), 59-80.
- Kress, G., & van Leeuwen, T. (1996). *Reading Images*. London: Routledge.

- Kress, G., & van Leeuwen, T. (2006). *Reading Images: The grammar of visual design* (2nd edition). London: Routledge.
- Kuo, S. H. (2003). Involvement vs. detachment: gender differences in the use of personal pronouns in televised sports in Taiwan. *Discourse Studies*, 5 (4), 479- 494.
- Laakkonen, I. (2007). *Breadwinners, sportsmen and conforming girls: Representations of gender in Finnish EFL textbooks*. Unpublished Pro Gradu Thesis. University of Jyväskylä, Finland. Retrieved from <https://jyx.jyu.fi/dspace/handle/123456789/18295>.
- Ladegaard, J. H. (2010). ‘Doing power’ at work: Responding to male and female management styles in a global business corporation. *Journal of Pragmatics*, 43 (1), 4-19.
- Lee, J.F.K. (2014). Gender representation in Hong Kong primary school ELT textbooks – a comparative study. *Gender and Education*, 26(4), 356-376.
- Lee, J.F.K., & Collins, P. (2009). Australian English-language textbooks: the gender issues. *Gender and Education*, 21(4), 353-370.
- Levine, D., & O’ Sullivan, M. (2010). Gender and images in the EFL textbook Talk a lot, Starter Book, Second Edition. *The Journal and Proceedings of GALE*, 3, 33-42.
- Liao, Y.S., & Huang, S.H. (2003). Gender presentation in EFL textbooks children. *Selected Papers from the Twelfth International Symposium on English Teaching*, 2, 124-134.
- Litosseliti, L. (2002). ‘Head to Head’: gendered repertoires in newspaper arguments. In L. Litosseliti J. Sunderland (Eds.), *Gender identity and discourse analysis* (pp. 129-148). Amsterdam: John Benjamins.
- Masse, M., & Rosenblum, K. (1988). Male and female created them: The depiction of gender in the advertising of traditional women’s and men’s magazines. *Women’s Studies Forum*, 11 (2), 127-144.

- Milburn, S., Carney, D. R., & Ramirez, A. M. (2001). Even in modern media the picture is still the same: A content analysis of clipart images. *Sex Roles*, 44, 277-294.
- Mills, S. (1995). *Feminist Stylistics*. London: Routledge.
- Mineshima, M. (2008). Gender representations in an EFL textbook. Retrieved from www.niit.ac.jp/lib/contents/kiyo/genko/13/14_MINESHIMA.pdf.
- Morrison, A. (1996). Barking up the wrong tree? Male hegemony, discrimination against women and the reporting of bestiality in the Zimbabwean press. In C. R. Caldas-Coulthard & M. Coulthard (Eds.), *Texts and practices: Readings in critical discourse analysis* (pp. 231-249). New York: Routledge.
- Nakamura, N. (2002). Images of women in high school EFL textbooks. *JALT 2002 at Shizuoka, Conference Proceedings*.
- Ohara, Y., & Saft, S. (2003). Using conversation analysis to track gender ideologies in social interaction: toward a feminist analysis of a Japanese phone-in consultation TV program. *Discourse & Society*, 14 (2), 153-172.
- Pizzini, F. (1991). Communication Hierarchies in Humour: Gender Differences in the Obstetrical/Gynaecological Setting. *Discourse & Society*, 2 (4), 477- 488.
- Sano, F., Lida, M., & Hardy, T. (2001). Gender representation in Japanese EFL textbooks. *PAC3 at JALT 2001, Conference Proceedings*.
- Smith, S. L., and A. D. Granados. (2009). Content Patterns and Effects Surrounding Sex-role Stereotyping on Television and Film. In Bryant, J., & Oliver M.B. (Eds), *Media Effects: Advances in Theory and Research* (3rd ed., pp. 342–361). New York: Routledge.

- Stockdale, D. A. (2006). Gender representation in an EFL textbook. *Paper submitted to the School of Humanities of the University of Birmingham, UK in part fulfillment of the requirements for the degree of Master of Arts in Teaching English as a Foreign or Second Language (TEFL/TESL).*
- Talbot, M. (2003). Gender Stereotypes: Reproduction and Challenge. In Holmes, J. & Meyerhoff, M. (Eds), *The Handbook of Language and Gender*. Blackwell Publishing Ltd.
- Tanaka, H., & Fukushima, M. (2002). Gender orientations to outward appearance in Japanese conversation: a study in grammar and interaction. *Discourse & Society*, 13(6), 749-765.
- van Leeuwen, T. (2005). *Introducing Social Semiotics*. London and New York: Routledge.
- Walker, R., & Harding, K. (2009). *Oxford English for Careers: Tourism*. Oxford: Oxford University Press.
- Wodak, R. (2002). Aspects of Critical Discourse Analysis. *ZfAL*, 36, 5-13.
- Wodak, R. (2006). Images in/and news in a globalised world: Introductory thoughts. In Lassen, I., Strunck, J., & Vestergaard, T. (Eds), *Mediating Ideology in Text and Image* (pp. 1-16). Amsterdam/Philadelphia: John Benjamins Publishing Company.