



The Asian ESP Journal

Spring Edition

April 2011

Volume 7 Issue 2

Chief Editors: Professor Winnie Cheng

Dr. Paul Robertson

Published by the Asian ESP Journal Press
A Division of Time Taylor International Ltd

<http://www.asian-esp-journal.com>

©Asian ESP Journal Press 2010

This book is in copyright. Subject to statutory exception
no reproduction of any part may take place without
the written permission of the Asian ESP Journal Press.

No unauthorized photocopying

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 the Asian ESP Journal.
editor@asian-efl-journal.com

Publisher: Dr. Paul Robertson
Chief Editor: Dr. Winnie Cheng

ISSN. 1833-3001



Index

Foreword. Professor Winnie Cheng	4
1. Zahra Akbari. Vocabulary comprehension and learning in ESP context: Strategy use and knowledge sources	5- 27
2. Vahid Aryadoust. Validity Arguments of the Speaking and Listening Modules of International English Language Testing System: A Synthesis of Existing Research	28-54
3. Budsaba Kanoksilapatham. Civil engineering research article Introductions : Textual structure and linguistic characterization	55-84
4. Takashi Matsuzawa. English Workshop for Effective Teleconferencing	85-98
5. Matthew Peacock. The Structure of the Methods Section in Research Articles Across Eight Disciplines	99-124
6. Chia-hsiu Tsao. English for Specific Purposes in the EFL Context: A Survey of Student and Faculty Perceptions	125-149

Foreword

Welcome to the Spring 2011 issue of *The Asian ESP Journal* Spring!

It is our pleasure to publish six selected papers that have comprehensively and usefully examined a range of ESP-related topics ranging from vocabulary comprehension and learning (Zahra Akbari), literature review of validity arguments of IELTS's listening modules (Valid Aryadoust), a genre analysis of the introductions in Civil Engineering research articles (Budsaba Kanoksilapatham), effectiveness evaluation of a workshop on teleconferencing (Takashi Matsuzawa), a cross-disciplinary genre analysis of the methods section in research articles (Matthew Peacock), to a comparative survey study of the attitudes of both students and teachers towards English for specific purposes (ESP) and English for general purposes (EGP) (Chia-hsiu Tsao).

I hope you will enjoy reading the papers that report on ESP studies conducted in various countries and cities in Asia, namely Iran, Singapore, Thailand, Japan, Hong Kong, and Taiwan, and recommend the papers to your colleagues and students to enhance the impact of the studies.

Prof. Winnie Cheng, Chief Editor



Title

**Vocabulary comprehension and learning in an ESP context:
Strategy use and knowledge sources**

Author

Zahra Akbari

Abstract

Investigating the cognitive processes used by EFL learners is a revealing enterprise in the domain of L2 vocabulary acquisition. A fundamental question that has remained unanswered in the literature on vocabulary acquisition concerns the relationship between the range of strategies and knowledge sources EFL learners use. In this study, a qualitative approach was designed and carried out in which 103 participants were selected randomly from among medical and paramedical students who enrolled in ESP I at Isfahan University of Medical Sciences. Data on VLSs in an ESP context were elicited by observation, interview and questionnaire. This study has implications for EFL learners, instructors, curriculum developers and materials writers.

Keywords: knowledge sources; strategy use; vocabulary comprehension strategies;
English for specific purposes; specialized and non-specialized vocabulary

1. Introduction

The notion of learning strategies was motivated by two fields, namely cognitive psychology and second language acquisition. In the 1980s, researchers shifted focus from what the good learner does to learn a language to the classification systems of strategies (Griffiths and Parr, 2001). Although taxonomies of a broad range of vocabulary learning strategies (VLSs) do exist, they tend to be incomplete in terms of strategies or factors arguably important for vocabulary learning (Riazi and Alavi, 2004; Riazi et al., 2005). Despite the interesting patterns seen in quantitative studies

(Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999; Schmitt, 1997), these studies do not show how a particular type of strategy is used in the development of vocabulary. Thus, we adopted a qualitative approach in this study in order to have a clearer picture of the process of vocabulary acquisition with a focus on intentional learning of vocabulary especially in ESP contexts.

Research suggests that readers use a variety of strategies when they encounter new words, including ignoring unknown words, consulting a dictionary for their meaning, writing them down for further consultation with a teacher or attempting to infer their meaning from context (Harley & Hart, 2000). Fraser (1999) and Paribakht and Wesche (1999) found that lexical inferencing was the most frequent and preferred strategy used by adult L2 learners to learn the meanings of new words. Conversely, the most frequently used strategy for understanding and learning the meaning of the new words in Kudo's study (1999) was using bilingual dictionaries.

The theoretical foundation of this study is based on Schmitt's study (1997) since in the area of VLS taxonomy development, the most notable effort in terms of range of strategies considered has been that of Schmitt's. However, some critical features make the context of this study significantly different from that in Schmitt's study and these features may influence the kind of strategies and knowledge sources used by ESP students for comprehending and learning specialized and non-specialized vocabulary.

First, in an ESP context, words (mainly specialized vocabulary) are expected to be used both productively (i.e., interactional communication of ESP students with their content teachers, doctors and peers in clinical and academic settings) and receptively (i.e., ESP students' comprehension and/or translation of two main English sources: their English references (printed sources) and information sources from the Internet (online sources)). But "use" was mainly defined by Schmitt (1997) as vocabulary practice rather than interactional communication.

Second, it seems that in Schmitt's taxonomy, what is dealt with are strategies for remembering form-meaning pairs and what is missing are skill-oriented strategies which are characteristic of learning specialized words in an ESP context.

Third, Despite Schmitt's study which was conducted for different age groups at different educational levels and with different purposes, this study was carried out

with a relatively homogeneous group of learners, in terms of their age, language proficiency and their field of study and who are supposed to learn English (i.e. acquire mainly their required vocabulary items) through reading academic texts in ESP courses. ESP students learn English as a means to achieve their subject-specific and ultimately their occupational goals.

Knowledge sources range from knowledge of grammar, morphology, phonology and knowledge of the world to knowledge of punctuation, word association, cognates, L1 knowledge and discourse knowledge (Nassaji, 2003). Paribakht and Wesche (1999) found that their university ESL readers appealed to a variety of linguistic and nonlinguistic knowledge sources when attempting to derive the meaning of new words from context. Specialized words are made up of words that occurred frequently in a specialized text or subject area but did not occur or were of very low frequency in other fields (Nation, 2001; Nation & Chung 2004; Oh et al., 2000) and non-specialized vocabulary are terms that have a specific meaning in a scientific context (Childs & O' Farrell, 2003; Strevens et al., 2000). Taking into account the specific features of ESP contexts in the domain of VLSs and in the light of Schmitt's taxonomy (1997), the following questions are raised:

1. What strategies are used for vocabulary learning and vocabulary comprehension in an ESP context?
2. What knowledge sources are used for vocabulary learning and vocabulary comprehension?

2. The Present Study

2.1. Participants

137 medical and paramedical students, taking ESP I course at Isfahan University of Medical Sciences. (i.e., medicine, pharmacy, dentistry, midwifery, nursing, physiotherapy, management of healthcare services and medical records) were selected to take part in this study. A standardized language proficiency test (Intermediate TOEFL Test Practices by Folse, 1997) was administered to the participants. Although it was a standard test, it was piloted in conditions similar to our main study to ensure its reliability for the context of our study, its calculated test- retest reliability turned

out to be 0.75. Three areas of language proficiency were tested: grammar (20 items), vocabulary (20 items), and reading comprehension (10 items)

Then, 103 learners with mean \pm 1SD were identified. 14% of the participants were male and 86% were female. The participants' mean age was 20 ± 1.12 .

Table 1. The distribution of the participants by sex, degree, field of study, and language proficiency

Field	Total N	Ph.D.	BSc	Assoc. degree	mean	SD	Selected N	Male	Female
Medicine	17	*			28.82	6.47	14	5	9
Dentistry	17	*			28.18	5.07	13	0	13
Pharmacy	21	*			25.43	4.63	16	0	16
Midwifery	14		*		20.07	6.07	11	0	11
Nursing	17		*		21.24	4.20	13	3	10
Physiotherapy	15		*		23.47	6.09	9	2	7
Management of healthcare services	15		*		20.33	5.91	11	2	9
Medical records	21			*	18.24	4.61	16	0	16
Total	137						103	12	91

* =have Ph.D., BSs or Associate degrees

Since the study was qualitative, random sampling was not practically possible in its real sense. On the other hand, it was not feasible to study all of the students in different classes and in all fields of study. For instance, some classes were held simultaneously by different teachers and it was not possible for the researcher to observe them at the same time. Or in some classes, the teacher did not allow the researcher to carry out her research. So, from among the existing ESP I classes for each of the above-mentioned fields, one was selected randomly in order to attempt stratification of the participants. A normality test was performed using EViews software* which ensured the normality of the population in each field (i.e., Jarque_???Bera statistics for all fields was less than 5.99 with confidence interval 95% and degrees of freedom equal to two). Since the status quo was examined (i.e., the existing ESP classes) in this study, it was clear that we did not have any other choice than to accept limitations such as lack of gender distribution or equal number of participants in each group.

2.2. Methodology

In order to elicit data on VLSs and knowledge sources used, three methods were used: a) interviewing the students individually about their vocabulary comprehension and vocabulary learning activities and knowledge sources used while studying their academic texts, b) observing the students in person in the classroom and outside the classroom while studying their academic texts, and c) using a questionnaire based on theoretical considerations of some previous attempts to study VLSs, including that of Schmitt (e.g., Kudo, 1999; Segler et al., 2002; Winke, 2002), to identify VLSs types and based on data collected from interviews and observations. All 103 participants were subject to all three phases of data collection.

First, a semi-structured interview (Mackey & Gass, 2005) was used in the pilot study based on which the final fixed set of questions was developed to be asked in the main study through structured interviews (Cohen, 1998). The purpose of the interviews was to elicit task-specific VLSs, to uncover general vocabulary comprehension strategies and knowledge sources used in this regard. The interviews were conducted in the participants' first language and the students were asked to bring their books to illustrate what they do when they face a new word.

In the main study, the researcher was motivated to find out what VLSs and vocabulary comprehension strategies were used and elicit what knowledge sources were being used to facilitate vocabulary comprehension and learning. Thus, it was decided to observe the students' performance in and outside the classroom before carrying out the questionnaire. The researcher observed the students' VLSs which were used in each classroom while they were asking their ESP teacher questions about vocabulary items, answering their ESP teacher's questions about vocabulary items, reading their ESP texts aloud and translating the translation section of each lesson in their books. The researcher also observed the students' books to check their notes about vocabulary items.

The questionnaire was constructed for the collection of data on what the participants actually do while comprehending and learning the vocabulary items in their ESP texts. The questionnaire, written in learner L1, consisted of two parts: questions to gain demographic information about the participants and questions related to the vocabulary comprehension and learning strategies the participants might

have used. Nonparametric statistics were used to analyze the frequency of ordinal variables in Likert-scale data. Frequency of use was measured as a five-point Likert scale: 1: *never*, 2: *seldom*, 3: *sometimes*, 4: *often* and 5: *always* on a 62-item questionnaire.

The questionnaire consisted of strategies for both comprehension and learning of the new words. The final version of the questionnaire was used in a pilot group. Before administering it to all of the participants, the validity and reliability of the test was examined. , the reliability of the test turned out to be satisfying (Cronbach alpha= 0.82).

To determine the validity of the questionnaire, it was read critically by experts in applied linguistics to clarify its possible problems. Furthermore, in order to measure the degree to which the questionnaire accurately reflects or assesses vocabulary comprehension and vocabulary learning strategies, factor analysis with varimax rotation was also run. The results of the rotated solution revealed the presence of two components showing certain items loading substantially on one component and certain other items loading on another component. Factor analysis showed that the 62 strategies fit into the two main tentative factors as originally hypothesized (i.e., discovery and consolidation strategies). The Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy was 0.607 which was satisfactory where the KMO measures “the sampling adequacy which should be greater than 0.5 for a satisfactory factor analysis to proceed” (University of New Castle upon Tyne, 2002).

3. Results

To answer research question one: "What strategies are used for vocabulary learning and vocabulary comprehension in an ESP context?", the analysis of the questionnaire revealed that using bilingual dictionary for determining the meaning of unknown words whether specialized or non-specialized was the most frequently used comprehension strategy.

Table 2. The most frequent comprehension strategies for specialized and non-specialized vocabulary

Field	Med	Mean	SD	Dent	Mean	SD	Pharm	Mean	SD	Mid	Mean	SD
Comp non-sp	A15	3.21	0.802	A13	3.58	0.793	A13	3.42	0.793	A24	3.55	0.522
	A13	3.14	0.846	A44	3.25	0.965	A15	3.33	0.492	A15	3.18	1.250
	A24	3.00	1.038	A15	2.83	0.937	A19	2.83	0.937	A13	3.00	0.894
Comp Sp	Q15	3.21	0.802	Q61	3.25	1.055	Q13	3.50	0.905	Q24	3.64	0.505
	Q13	3.14	0.864	Q13	2.92	1.379	Q15	3.33	0.651	Q15	3.18	0.982
	Q24	3.00	1.038	Q44	2.92	1.084	Q28	3.17	0.718	Q13	3.09	0.831

Field	Nurse	Mean	SD	Physio	Mean	SD	Manag	Mean	SD	Med	Mean	ds
Com non-sp	A24	3.08	0.954	A13	3.60	0.516	A13	3.73	0.647	A13	3.40	0.910
	A13	3.08	0.760	A15	3.30	0.675	A15	3.00	0.775	A44	2.93	1.335
	A15	3.00	1.155	A24	2.70	1.252	A4	2.91	1.044	A24	2.55	0.743
Comp sp	Q13	3.00	0.577	Q13	3.40	0.699	Q13	3.73	0.647	Q13	3.60	0.737
	Q28	2.92	0.954	Q15	3.30	0.675	Q15	3.09	0.701	Q44	2.93	1.335
	Q24	2.92	0.862	Q24	2.70	1.252	Q4	2.91	0.944	Q24	2.55	0.915

The strategies for non-specialized vocabulary were shown with letter A and strategies for specialized vocabulary were shown with letter Q. The first three most frequent vocabulary comprehension strategies were shown in table 2. Comparing the first three strategies reveals that strategy 13, using bilingual dictionary for determining the meaning of unknown words, was the most frequent one.

Table 3. The most frequent learning strategies for specialized and non-specialized vocabulary

Field	Med	Mean	SD	Dent	Mean	SD	Pharm	Mean	SD	Mid	Mean	SD
Lrn non- sp	A39	2.79	1.122	A43	2.50	1.000	A39	2.83	1.267	A54	2.82	1.250
	A53	2.54	0.967	A33	2.25	1.055	A29	2.67	0.985	A39	2.80	0.919
	A59	2.29	1.069	A39	2.25	1.658	A33	2.50	0.905	A17	2.64	1.433
Lrn	Q39	2.79	1.122	Q43	2.67	0.651	Q47	3.17	0.835	Q39	2.73	1.272

sp	Q53	2.50	0.941	Q54	2.42	1.443	Q41	3.00	0.853	Q54	2.73	1.272
	Q59	2.29	1.069	Q39	2.42	0.996	Q39	2.92	0.996	Q46	2.73	0.905

Field	Nurs	Mean	SD	Physio	Mean	SD	Manag	Mean	SD	Med	Mean	SD
Lrn	A39	3.17	0.937	A39	2.89	1.167	A39	3.27	0.647	A39	2.93	1.163
Non-sp	A59	2.69	0.630	A46	2.60	1.174	A45	3.18	0.982	A54	2.67	0.976
	A29	2.62	1.044	A54	2.56	1.130	A29	2.64	1.27	A43	2.40	0.828
Lrn sp	Q43	3.23	0.927	Q46	3.00	0.943	Q45	3.18	0.982	Q39	3.07	1.033
	Q39	3.15	0.987	Q39	2.70	1.418	Q39	3.18	0.751	Q54	2.73	0.884
	Q11	3.00	0.816	Q54	2.50	1.179	Q29	2.73	1.104	Q11	2.60	1.056

Among the first three most frequent strategies used for learning specialized and non-specialized vocabulary items, the highest mean is repeating the word silently together with its meaning as it is placed in text (A39). The next most frequent strategy is repeating the word together with its meaning aloud (Q54). Choosing between strategies 39 and 54 depended on the students' learning environment (i.e., whether they are at home, in the library or in the dormitory) on the one hand and their learning style (i.e., the role of hearing the pronunciation in learning) on the other.

To answer research question two "What knowledge sources are used for vocabulary learning and vocabulary comprehension?", the integrative summaries of all three data sources revealed that a great majority of the participants in different fields of study did not demonstrate satisfactory levels of EFL vocabulary knowledge. Since comprehension is one of the basic requirements for learning a word, usually made explicit reference to the following knowledge sources in their comprehension and learning strategies depending on their degree of awareness of the existing knowledge sources and their usefulness in vocabulary comprehension.

Table 4. The frequency of knowledge sources used in each field of study

	medicine	pharmacy	dentistry	midwifery	nursing	physiotherapy	management of healthcare services	medical records
L1 knowledge	64%	62%	44%	73%	62%	78%	73%	69%
Grammatical knowledge	50%	46%	38%	64%	54%	56%	56%	50%
Morphological knowledge	43%	38%	31%	55%	46%	44%	36%	44%
Academic knowledge	36%	31%	25%	45%	38%	33%	27%	25%
Knowledge of contextual clues	24%	23%	19%	36%	23%	22%	18%	18%
Knowledge of pronunciation and spelling	21%	15%	13%	18%	15%	11%	9%	6%
Knowledge of Collocational phrases	14%	8%	6%	9%	8%	0%	0%	0%

4. Discussion and Conclusion

As far as vocabulary learning and vocabulary comprehension strategies are concerned, the results are in line with those of Kudo (1999) and Schmitt (1997). With regard to comprehension strategies, the results of this study were rather predictable since the use of bilingual dictionary is a common practice among Iranian EFL learners due to the widespread practice of the Grammar-Translation Method and the lack of necessary study skills (such as how to use different pieces of information in monolingual dictionaries, word analysis and guessing the meaning from context).

According to the participants, the easiest and quickest way to find the meaning of the new words and comprehend the passage is to consult a dictionary. In addition, without a certain amount of vocabulary, it is hard to understand the definition of new words when using a monolingual dictionary. Students would find themselves looking

up word after word as unknown words in one definition lead to other definitions. In sum, dictionary use aids second language learners' vocabulary development because it is the initial step in learning a new word. It provides fast and reliable support for learners who have vocabulary and language limitations.

In fact, the results of the questionnaire were verified by those elicited from interviews and observations of English classes and students' English textbooks. The participants in different groups explained that because of their limited knowledge of vocabulary, they could neither rely on their guesses nor use the monolingual dictionary; so it was better to save time and energy by immediately referring to the dictionary. They did not have the essential comprehension skills including deriving the meaning of new words from the contextual clues and they got used to translating the text word for word.

The role of synonym learning revealed some interesting trends like some students wrote the Persian translation above each of the new words. However, there were students in different groups who, based on their reports in the interviews, attended English classes in private institutes where they had learned how to use monolingual dictionaries. These students had already written the English synonyms of the new words in their English textbooks before they were given the synonyms in the classroom. Many learners fixed on the meaning they had already learned for a given word and found it very difficult to use another one even if the one they knew had no sense in the new context. For instance, it was observed when the teacher asked the synonym of the word "*effectiveness*" in the sentence: "*The nurse evaluates the effectiveness of health care interventions*", the nursing participants' answer was "*influence*". They had, in fact, written the Persian meaning "*tasir*" above it in their textbooks for which they already had the English word "*influence*" in their background knowledge and said it out loud in answer to their teacher.

The results of vocabulary learning strategies were also verified by those elicited from the other two sources. Students in different groups reported that memorization is the easiest and quickest way to learn the meaning of the new words. They agreed that although this strategy leads to short-time retention of the new words, it can allow them to satisfy their immediate need, i.e., passing the exam.

During the term, they underlined the new words in context or highlighted them so as to distinguish them from familiar words. They wrote their meanings (Persian or English if they have any synonyms for them in their mind) above them in the passage and after comprehending the meaning of the text (to make themselves ready to attend each ESP session), at the end of term, they tried to learn the words by repeating them silently as they looked at the word in the text while hiding its meaning. (Because of time constraints, large numbers of new words and high load of course work in each term; students often prepared a bilingual word list at the end of the term for their final exam and tried to learn the new words through repetition. In other words, students focused primarily on short-term comprehension strategies and not on long-term language-learning strategies, and the reason for this is that students were not externally motivated (by, say, a mid-term exam) to learn words before they had to for the final exam.

In interviews, when they were asked: “How do you learn a new word?” they replied: “we memorize it”. For them, learning was synonymous to memorization as they were not aware of other more effective learning strategies. This fact was also revealed when they were asked to answer the questionnaire. The students in different groups said that they did not know that there could be other strategies for learning a new word. However, their reports of the process of memorization in interviews revealed that they did not use pure memorization and it was associated with a kind of land marking or assigning some kind of code to the words to facilitate memorization (i.e., their learning) and to increase their retention period. They mostly related the sounds in the pronunciation of the English word whether specialized or non-specialized to the sounds in the meaning of the same word in Persian. For example, “*chill*” means “*larz*” in Persian; both of them have the sound /l/. Or for learning the word “*pituitary*” which means “*hypofiz*” in Persian, one of the medical students said that he made a code in this way: since “*pituitary gland*” is said to be as small as a “*pea*” and both “*pea*” and “*pituitary*” begin with the sound “p”; learning the meaning of *pituitary* becomes easy. Or in learning the meaning of the word “*ill*”, one of medical records students reported that the word “*ill*” reminded him of the Persian word “*i:l*” (i.e., tribe) who do not have the necessary facilities to observe hygienic rules and ,therefore; they may frequently become “*ill*” (i.e., sick). For learning the word “*discard*” one of the nursing students said that this word consists of the part

“*card*” which in Persian means “*knife*”; so the word acts as a knife that cuts something and the unusable part is “thrown away”. Or the word “*universal*” is like “*university*” which includes all of the fields of study so it reminds her the words “*jame’ v kamel*” in Persian (i.e., comprehensive, including everything).

Since comprehension is one of the basic requirements for learning a word, they usually made explicit reference to the following knowledge sources in their comprehension and learning strategies depending on their degree of awareness of the existing knowledge sources and their usefulness in vocabulary comprehension. The findings in this part of the study are in agreement with Nassaji’s (2003) study of the relationship between strategies and knowledge sources. The knowledge sources are arranged from the most- to the least-frequently used by the participants in different groups.

L1 knowledge. Using L1 knowledge, the learner attempts to figure out the meaning of the new word by translating or finding an equivalent for it in their L1, i.e., word for word translation strategy. From classroom observations, this over-reliance word-for-word translation is illustrated nicely in the following exchange in midwifery classroom. When asked for a synonym of the word "enable" in this sentence: "Enzymes enable microorganisms to make proper use of food." Students answered: "able", "power", "ability", "can", or "capable". Their answers were mainly based on the meaning they derived from the new word itself without paying any attention to its part of speech or its grammatical function in the sentence.

The findings about dictionary use revealed that the first piece of information all participants looked for in a dictionary was the meaning of the word. As far as part of speech of the word was concerned, they also tried to take it into account to some extent since this feature could also help them in the translation/comprehension of the words in a sentence. They usually inferred this feature based on the meaning of the word and/or the suffixes attached to the word and they did not usually pay any attention to the abbreviations used for the parts of speech in each entry in the dictionary. They also used the grammatical function of the words in the sentence to infer its part of speech. It was usually the case when they were good at recognizing grammatical relations between the words.

As one of the most important functions of dictionaries is to provide word meanings, it is understandable that L2 learners consult dictionaries most frequently to check the meaning of new words. Learners value dictionaries because they can improve the reading comprehension of lower proficiency L2 learners and they assist vocabulary learning at all levels of proficiency (Hulstijn et al., 1996; Knight, 1994). Also, their definitions are short and relatively easy to understand. However, they can contribute to over-reliance on one-to-one word translation (Baxter, 1980; Tang, 1997).

Grammatical knowledge. Based on the results elicited from calculating the frequency of items mentioned in question (one of the interview in part one and the questions in part two of the interview and from observation of students' notes), since the participants in different groups highly relied on the translation of the English passages into their L1 as a main device for exact comprehension of the texts, one of the frequently used knowledge sources was their limited grammatical knowledge of L2 words. It is in line with Paribakht (2004) who found that such knowledge can make significant contribution to understanding the meaning of unfamiliar words while reading.

Grammatical knowledge is using knowledge of grammatical functions or syntactic categories such as verbs, adjectives or adverbs. For example, in observing the nursing class, this sentence: "*Surface temperature fluctuates depending on blood flow to the skin and ...*" was translated as "*taghirat-e damay-e sathi bastegi darad be...*" by one of the students who indicated that "*depend*" was considered as the main verb of the sentence and "*fluctuates*" was considered as a plural noun and as the subject of the sentence. In this regard, when the teacher asked the students in the same class to make a question about this sentence, some students (6 students) answered: "*What do surface temperature fluctuates depend on?*"

Morphological knowledge. Due to our participants' limited morphological knowledge in different groups (based on observations and interviews), they usually considered compound words as single units with one meaning for the whole word unless they became familiar with their constituent parts in their Medical Terminology course. The participants also seemed to use their knowledge of the most frequent

affixes to make some overgeneralizations. For example, it was observed that the participants in management of healthcare services took the word "*disappointed*" as the opposite of "*appointed*" or "*discover*" as the opposite of "*cover*". In another example in the nursing group, in the sentence: "*It may be misleading for the nurse to diagnose a problem on the basis of one assessment finding.*", the word "*misleading*" was translated as "*monjar nashodan*" which indicates that "*mislead*" was considered as the opposite of "*lead to*".

When the participants in medicine, dentistry and pharmacy groups became familiar with the medical terminology or the meaning of the constituent parts of the specialized vocabulary, they (based on observations and interviews) mainly relied on word analysis for both comprehension and learning of such words.

Studies show that introducing the morphological method into language teaching helps students to learn quickly, easily and accurately with a minimum time and effort and maximum growth and retention of vocabulary (Xiao and Lu, 2005). Kieffer & Lesaux (in press) found that those students who take unfamiliar words and break them down into smaller parts, or morphemes, have increased success in deciphering and learning unfamiliar vocabulary.

Academic knowledge. Based on the results elicited from observations and interviews, in reading ESP books, the participants used their knowledge of academic (specialized) courses to a great extent to understand ESP words. For instance, it was observed that in a passage about techniques of assessment such as palpation, percussion, auscultation and inspection, since the nursing participants obtained enough knowledge about these techniques in their theoretical specialized courses as well as in their practical training courses, although they had difficulties in understanding some of the sentences word for word, they could understand the essence of the sentences.

Similarly, it was observed in the medical group that almost all participants did not have any problems in understanding this sentence: "common regional anesthesia techniques include spinal (subarachnoid block), epidurals, caudals and major peripheral nerve blocks."; since they had already learned (internalized) the concepts for the specialized terms in their anatomy course so that they did not need to have any translation for the words "caudal" and "epidural" and automatically pronounced them

in Persian pronunciation system (i.e., /kodal/ and /epidural/). In fact, translation of many such specialized words was difficult for the participants in all groups and they preferred to use the word as pronounced in English in Persian system when they learned the concept.

Knowledge of contextual clues. Another related source of knowledge was knowledge of contextual clues. Inferring the meaning of unknown words from the context in which it is used is one of the comprehension strategies used by the learners. Based on observations and interviews, the participants in all groups did not ordinarily pay attention to the contextual clues in the text and they insisted on deriving the meaning from their bilingual dictionaries. For example, it was observed when the teacher asked for the synonym of the word "*deliberate*" as in the following sentence: "*Unlike the immune system, which is slow and deliberate, the effects of inflammation are immediate*", the nursing students gave its literal meaning based on what they found in bilingual dictionaries (i.e., "*agahane*") instead of giving its contextual meaning (i.e., the synonym of the word "*slow*" or the opposite of "*immediate*" in this sentence). This implies that students need consciousness-raising to improve their comprehension of new words based on the contextual clues that are used.

In line with Laufer (2003), it was observed that when a word is noticed as unfamiliar, the reader may try to infer its meaning from context. Yet, not all contexts provide clues for unknown words. Sometimes, clues are ignored when the reader thinks s/he understands the message or when the correct meaning is not compatible with the learner's knowledge of the world. There are also cases when knowledge of the world overrides the use of a linguistic clue. More importantly, clues may appear in words which themselves are unknown to the learner and are therefore unusable.

Knowledge of pronunciation and spelling. Pronunciation was also important for learning the meaning of the new words especially for auditory learners. However, there were some variations in the way the participants used it. It was observed when the teacher asked the participants in different groups randomly to read part of the text; they usually pronounced the new word as it was written/spelt. For example, the nursing students usually pronounced the word "*heart*" as the word "*hurt*" or

pronounced the letter “p” in "*pneumonia*" or the word "patient" as "patiyent". There were few participants (one or two students) in each group especially those attending private English institutes who preferred to refer to monolingual dictionaries as they were taught how to find the meaning together with the pronunciation of the new word there. Based on observations and interviews, the participants in different groups mainly relied on pronunciation of the new words when they had to distinguish words with similar or identical pronunciations (homophones) but with different meanings and they mainly relied on spelling when they had to distinguish words with similar or identical spellings (homonyms) but different meanings.

In fact, knowledge of pronunciation is related not only to the use of L2 words but also to their acquisition. Research findings have shown that when students know how to say a word, it is easier for them to commit the word to memory (Chi et al., 1994; Fan, 1998; Fan et al., 1996). However, it is a fact that most students have difficulty with phonetic scripts (Taylor, 1988; p. 89) and this will continue unremedied until they get help from the teacher.

Correct pronunciation provides the basis for accurate spelling. With correct pronunciation and fair knowledge of spelling rules, correct spelling can be achieved. Clearly, correct pronunciation and proper knowledge of spelling are a great help to memorize words. As a result, spelling rules as well as the relationship between pronunciation and spelling should be expounded thoroughly to the EFL learners.

Knowledge of collocation. In fact, it was observed that the students in different groups were not usually able to recognize most phrases due to their unfamiliarity and the infrequent use of monolingual dictionaries. Therefore, participants translated them literally and tried to stick to the most frequent meaning which they first learned for the same word or the first meaning mentioned in the bilingual dictionary entry. It was observed that since the participants in different groups got used to translating each word in the order in which it occurred in the sentence, they could not recognize the components of a phrase especially when they were separated from each other. For example, in the sentence: "*Despite extremes in environmental conditions and physical activity, temperature-control mechanisms of human beings keep the body's core temperature or temperature of deep tissues relatively constant.*", midwifery students had difficulty to recognize the verb of the sentence as "*keep constant*". Or in the

sentence: "*Precise determinations of the client's health problems can be made*", all midwifery students considered "*can be made*" as the main verb rather than "*make determinations*".

However, there were cases where students in all groups mentioned that learning the words together facilitated their learning. They reported that it is true only when they knew that the words are related in a phrasal group. For instance, midwifery students mentioned that learning "amniotic sac" is much easier for them than learning each of the words by itself. It is also the case for learning the words in the phrase "viral invasion".

McCarthy claims that "in vocabulary teaching there is a high importance of collocation", and describes that 'the relationship of collocation is fundamental in the study of vocabulary, and collocation is an important organising principle in the vocabulary of any language' (1990:12), focusing on collocation acquisition as an appropriate perspective to enrich vocabulary. Students should be encouraged to adopt a more context-based approach by going beyond the word and paying attention to the phrase, clause, sentence and even the paragraph in which the word is located.

5. Implications of the study

The results of this research offer some insights for ESP learners, instructors, curriculum developers and materials writers. One of the crucial characteristics of ESP students that make them fundamentally different from other English as foreign language (EFL) students is that they do not usually need concept-formation, without which the simple retention of word-meaning pairs is meaningless, since they have already acquired the necessary concepts in their L1. In addition, they are expected to learn (i.e., use receptively and productively) both non-specialized and specialized vocabulary in their field of study.

Most of the students, while answering the questionnaire, reported that they did not know that there were so many different strategies to learn vocabulary. If students can find strategies suitable for them and actually use them, this might increase their vocabulary size. ESP learners should learn to recognize different VLSs and be advised to select the appropriate ones depending on the nature of the words they are learning, their purpose of learning and the conditions in which the words are learned. Use of appropriate VLSs enables students to take responsibility for their own learning

by enhancing learner autonomy, independence and self-direction (Lee 2003). These factors are important because learners need to keep on learning when they are no longer in formal classroom setting.

In this study, it was revealed that in fields in which the content of the material in the students' specialized field was presented mainly by using English terminology, the references were mainly in English, the content teachers' lectures were full of English terminology and students also practically experienced the necessity, critical role and usefulness of specialized vocabulary in their training clinical courses, ESP students resorted to more elaborative strategies to learn the new vocabulary for long term purposes while in fields in which the students did not feel the practical usefulness of the vocabulary items in simulated occupational settings, most of their references were translated into Persian and their content teachers used mostly pure Persian, ESP students resorted to repetition and memorization for short term purposes.

Since referring to bilingual dictionaries is the most frequently used comprehension and to some extent learning strategy, dictionary work, including practicing good dictionary skills, is useful as an independent vocabulary acquisition strategy. Students usually come to the language classroom without these study skills, so it is helpful to expose them to a variety of ways to practice words and their definitions and let them choose the manner which is comfortable for them.

The primary lexical objectives are increasing vocabulary breadth, elaborating vocabulary knowledge, developing fluency with known vocabulary and coordinating the use of strategies with various knowledge sources. The means for achieving these objectives are skill-based and include training learners to effectively learn decontextualized lexis, consolidate and elaborate previously met lexis, consult dictionaries, infer from context and engage in reading for meaning.

It is equally important to foster a type of word learning environment that helps students actively use enriched knowledge of words through developing different knowledge sources as students are exposed to words in different contexts.

Teachers should become aware of learners' strategies and learning styles so that they can develop teaching strategies that are compatible with their students' ways of learning. In addition, diversified vocabulary learning activities should be

incorporated in both the classroom and ESP books to accommodate various styles of learning and various types of vocabulary items.

Content-based ESP teaching presents a challenge to ESP instructors (Bell 1996). The problems faced by ESP teachers and the students can be placed at the opposing ends of a bipolar scale; the students struggle with the language of the content and the teachers struggle with the content of the language (Liyanage & Birch 2001). As a result, solving this problem requires collaboration between content teachers and language teachers.

In this regard, the curriculum of presenting courses must also be revised in a way that the necessary specialized courses were held first to familiarize the students with basic specialized concepts and then the ESP courses together with Medical Terminology course would be presented. In this way the schematic knowledge is formed to anchor the new labels to it. On the other hand, the ESP books must be revised in a way that short interesting specialized topics together with attractive pictures are presented to the students to motivate them to read ESP texts as extensively as possible.

In this study it is observed that there is time interval between general English course and ESP courses. During this time interval, students did not have the necessary exposure to English texts (except in few cases in the form of optional translation work). So they mainly forgot the English words they had already learned for two reasons: the main strategy of learning new vocabulary is memorization, the conditions for their recurrent use or frequent reference to them are not provided.

To mitigate this problem, one of the alternatives is to present two-credit English courses in each term to provide enough time for teachers to develop the basic study and reading skills together with different kinds of VLSs to develop the students' reading comprehension and expose students to as many ESP texts as possible to provide conditions to practice and reinforce those skills and strategies. This might remove one of the main hindrances of learning which according to students' reports is learning a large number of new vocabulary in a very short time under the pressure of passing the final exam rather than willingly and with high motivation.

Finally, material designers should first conduct a needs analysis and define objectives of the learners in relation to their field of study before writing materials.

This can lead to providing interesting, informative, short, pictorial passages with few new vocabulary items which are repeated in subsequent passages with enough exercises to practice and reinforce the essential learning strategies.

References

- Baxter, J. (1980). The dictionary and vocabulary behavior: A single word or a handful? *TESOL Quarterly*, 14(3), 325-336.
- Chi, M. L., et al. (1994). Collocational problems amongst ESL learners. A corpus-based study. In L Flowerdew & K K Tong (Eds), *Entering text* (pp 157-65). Hong Kong The Hong Kong University of Science and Technology, Language Centre.
- Childs, P. E. & O'Farrell, F. J. (2003). Learning science through English: an investigation of the vocabulary skills of native and non-native English speakers in international schools. *Chemistry Education: Research and Practice*, 4(3), 233-247.
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. Longman: London.
- Fan, M Y. (1998). An investigation into the problem of recoding technical vocabulary. *Asia-Pacific Journal of Teacher Education & Development*, 1(1), 83-92.
- Fan, M. Y. et al. (1996). *An investigation into the learning of technical vocabulary*. Paper presented at the International Language in Education Conference 1996, Hong Kong.
- Folse, K. (1997). *Intermediate TOEFL Test Practices*. Ann Arbor: University of Michigan Press.
- Fraser, C.A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21, 225-241.
- Griffiths, C. and Parr, J. M. (2001). Language Learning Strategies: Theory and Perception. *ELT Journal*, 55(3), 247-254.
- Gu, Y. & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46, 643-679.
- Harley, B., & Hart, D. (2000). Vocabulary learning in the content-oriented second

- language classroom: Student perceptions and proficiency. *Language Awareness*, 9(2), 78-96.
- Hulstijn, J. H. et al. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use and reoccurrence of unknown words. *MLJ*, 80(3), 327-339.
- Kieffer, Michael J., and Nonie K. Lesaux (in press). *Breaking down to build meaning: Morphology, vocabulary, and reading comprehension in the urban classroom. The Reading Teacher*.
- Knight, S. (1994). Dictionary use while reading: the effects on comprehension and vocabulary acquisition for students of different verbal abilities. *MLJ*, 78(4), 285-299.
- Kojic-Sabo, I. & Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal*, 83, 176-192.
- Kudo, Y. (1999). *L2 vocabulary learning strategies*. Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center. Retrieved from the World Wide Web: <http://www.lll.hawaii.edu/nflrc/NetWorks/NW14/>.
- Laufer, B. (2003). Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59(4), 1-15.
- Lee, K.O. (2003). The relationship of school year, sex and proficiency on the use of learning strategies in learning English of Korean junior high school students. *Asian EFL Journal*, 5(3), 1-36.
- Liyanage, I. & Birch, G. (2001). English for general academic purposes: Catering to discipline-specific needs. *QJER*, 17, 1-12.
- Mackey, A. & Gass, S. M. (2005). *Second Language Research: Methodology and Design*. London: Lawrence Erlbaum Associates.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- Nassaji, H. (2003). L2 vocabulary learning from context: strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, 37(4), 645-667.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I.S.P. and Chung, T.M. (2004). Identifying technical vocabulary. *System*, 32,

251-263.

- Oh, J. et al. (2000). Japanese term extraction using dictionary hierarchy and a machine translation system. *Terminology* 6, 287-311.
- Paribakht, T.S. (2004). The role of grammar in second language lexical processing. *RELC Journal*, 35(2), 149-160.
- Paribakht, T. & **Wesche, M.** (1999). Reading and incidental L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition*, 21(2), 195-224.
- Riazi, A.M. & Alavi, A. (2004). Strategy activation in learning English words. *Academic Exchange Quarterly*, 6(2), 199-203.
- Riazi, A. M. et al. (2005). The effect of EFL students' L2 proficiency and age on their overall pattern of vocabulary learning strategy use. *IJAL*, 8(2), 87-106.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt and M McCarthy (eds.). *Vocabulary, description, acquisition and pedagogy*. Cambridge: Cambridge University Press.
- Segler, T. et al. (2002). Second language vocabulary acquisition and learning strategies in ICALL environments. *Computer Assisted Language Learning*, 15(4), 409-422.
- Stevens, R. A. et al. (2000). Academic language and content assessment: measuring the progress of ELLs. Los Angeles, University of California: National centre for research on evaluation, standards and student testing.
- Taylor, A. (1988). Learners and English dictionaries: Some assumptions and challenges. *Institute of Language in Education Journal*, 4, 88-92.
- Tang, G. M. (1997). Pocket electronic dictionaries for second language learning: Help or hindrance? *TESL Canada Journal*, 15(1), 39-57.
- University of New Castle upon Tyne. (2002, March). Statistics support: How to perform and interpret factor analysis using SPSS. Retrieved March 12, 2002, from http://www.ncl.ac.uk/ucs/statistics/common/specialisttopics/factor_analysis/factoranalysis.html
- Winke, P. (2002). Vocabulary acquisition in the Chinese as a foreign language

classroom. Paper presented at the International Association of Applied Linguistics, Singapore. Strategies for vocabulary acquisition in a Chinese as a foreign language \ classroom. Paper from linguistics 681. Georgetown University. Washington DC.

Xiao, H. and Lu, Na. (2005). Towards a scientific teaching of English. *Sino-US English Teaching*, 2(7), 1-10.

***Note:** (EViews 6 Student Version allows students to analyze datasets whose size is limited only by available computer memory. Among the features students may use via EViews' easy-to-use context sensitive menus and dialogs are: basic descriptive statistics and ANOVA, tabulation, cross-tabulation, covariance and correlation analysis, principal components, **factor analysis**, empirical distribution function tests), time series plots, distribution graphs (histograms, distribution plots, kernel density plots. Instead of this software, we can also use SPSS.



Title

Validity Arguments of the Speaking and Listening Modules of International English Language Testing System: A Synthesis of Existing Research

Author

Vahid Aryadoust

Nanyang Technological University

Biodata

Vahid Aryadoust is a PhD candidate in applied linguistics at the National Institute of Education (NIE) in Singapore. He has written numerous books and articles on English grammar instruction, applied linguistics, and language and psychosociological assessment. As co-principal researcher in the Spaan Fellowship Program of the University of Michigan, he investigated the construct validity of the Michigan English Language Assessment Battery (MELAB) listening test. His current research focuses on listening assessment, and particularly on the validity arguments of standardized listening tests. Correspondences should be sent to Vahid Aryadoust, ELL AG, NIE, 1Nanyang Walk, Singapore, 637616. Email: ng0666907m@stdmail.nie.edu.sg

Abstract

This study reviews research into the IELTS Speaking and Listening modules to build a validity argument for them. Based on Kane's (1992, 2001, 2004) validity argument framework, the researcher postulates seven assumptions to examine the two modules' interpretive arguments, as well as the sufficiency and efficacy of research conducted on them. The Speaking Module has been thoroughly studied in many respects, but its validity argument is nevertheless seriously compromised because IELTS has yet to articulate a constituent theory of second-language speech on which the module's analytic scoring system is based, and because a number of studies have shown very

limited correlations between performance on the module and performance in target language domains. The Listening Module is the least-researched module of the test, and is in urgent need of investigation before a validity argument can even be attempted.

Keywords: IELTS; interpretive argument; listening; speaking; validity argument

1. Introduction

International English Language Testing System (IELTS) was established in 1989. To date, fifteen rounds of research designed to improve the test have been jointly commissioned by the British Council, the University of Cambridge English for Speakers of Other Languages (ESOL) Syndicate, and the International Development Program of Australian Universities and Colleges (IDP). These projects have examined very diverse aspects of the test, from the language testing theory underpinning it to test taker attitudes and classroom training methods. As a result of this research, the test has undergone changes, and the IELTS of 2010 is quite different from the IELTS of 1989.

Today, many universities whose medium of instruction is English have adopted IELTS as an entry requirement for non-native English speakers. Many require a score of at least six out of nine on the test as a threshold of minimum proficiency. Students with lower scores may be refused admission or asked to take supplementary English language programs even if they have already demonstrated other qualifications.

In the present study, I review existing research on the IELTS Speaking and Listening modules, and create a unitary structure of validity arguments that organizes the results of this research. This synthesis will help spot weak areas and direct future attempts to bridge existing gaps.

This article proceeds as follows: It briefly reviews the history of the validity concept and validity arguments; provides an overview of the Speaking Module, including its analytic scoring system, the role of examiner judgment, its structure, the Revision Programs and impact studies, and some miscellaneous topics; highlights the paucity of research on the Listening Module; explains the interpretive arguments of the IELTS Speaking and Listening modules and investigates their plausibility using

the findings of existing IELTS research; and proposes conclusions and guidelines for future research for both the Speaking and Listening modules.

2. Theoretical Framework

The first references to validity arguments were made by Cronbach (1988) and Messick (1988), and several years later were used as a building block of Kane's (1992) validation framework. The concept of validity has continued to evolve since that time: whereas it was once considered a characteristic of measuring tools (see Chapelle, 1999), it is now considered a feature of score uses and interpretations, or, as Messick (1988, p. 3) put it, of the "inferences derived from test scores."

This new definition of validity accepts three types of supporting evidence: construct, content, and criterion (Messick, 1989). Construct-referenced evidence is evidence of the ability of the measuring device to assess what it is intended or claims to measure (Messick, 1989), and is typically based on statistical analysis. Content-referenced evidence is evidence of the degree to which a psychometric tool represents a psychological trait, and is typically based on expert judgment. Criterion-referenced evidence is evidence of the strength of the relationship between test scores and an external criterion, and is divided into concurrent and predictive classes: concurrent evaluations correlate test scores with scores from another standardized test, and predictive evaluations correlate test scores and future academic or work performance.

From this discussion it follows that there are two views of the validity of measuring tools: "theoretical construct" and "pragmatic ascription to ability" (Upshur, 1979, p.85). Predictive validity evaluations, for example, follow the pragmatic ascription to ability: they examine test takers' proficiency in future academic or work environments. Conversely, a study that targets the theoretical construct focuses on whether the test taker has the proficiency to perform the language task at hand. A common truism holds that any attempt to merge these two views would lead to an incoherent definition of validity.

However, Messick (1988) has called for a unitary concept of validation, based around construct-referenced evidence. Messick (1980) argues that construct-referenced evidence is a "unifying concept that integrates criterion and content considerations into a common framework" (p. 1015), and that content- and criterion-referenced evidence do not suffice in validation. Messick argues that content-referenced evidence has an element of subjectivity since it is mainly a function of

expert judgments, and leaves out the psychological processes of test takers, internal structures of the test, and differences in performance across test takers (Messick, 1988, p. 8); and that criterion-referenced evidence's correlation of test scores with future performance on a criterion may compound confusion, because the criterion will need to be validated like the test itself (Messick, 1988, p. 9).

In a later attempt to improve validation, Bachman (2005) argued that the conventional validity approaches are vague and do not provide a clear start and end point in validation. Drawing chiefly on Kane (1992, 2001, 2002, 2004), Bachman introduced a framework to assess the validity argument of any language measuring tool. In Bachman's framework, the researcher systematically examines supporting and attenuating data for a validity argument to reach a final conclusion that supports, weakens, or rejects the main validity claim. Bachman divides validation into two stages, which complement each other. In the formative stage prior to operationalizing the test, an interpretive argument (which Bachman calls an "assessment validity argument") is developed. This argument establishes the intended interpretations of test scores, sets up a web of inferences and assumptions that move from observed scores to the decisions made based on those scores, and "provides guidance as to the type of research needed" (Chapelle, 2008, p. 321). Interpretive arguments are based on assumptions that are temporary, explicit, and "defeasible in the sense that... they can be overturned in a particular case" (Kane, 2004, p. 147). In the summative stage of validation, a validity argument (which Bachman calls an "assessment utility/use argument") is created to back the interpretation. The validity argument comprises qualitative and quantitative data collected by the researcher to support the assumptions and claims made in the interpretive argument.

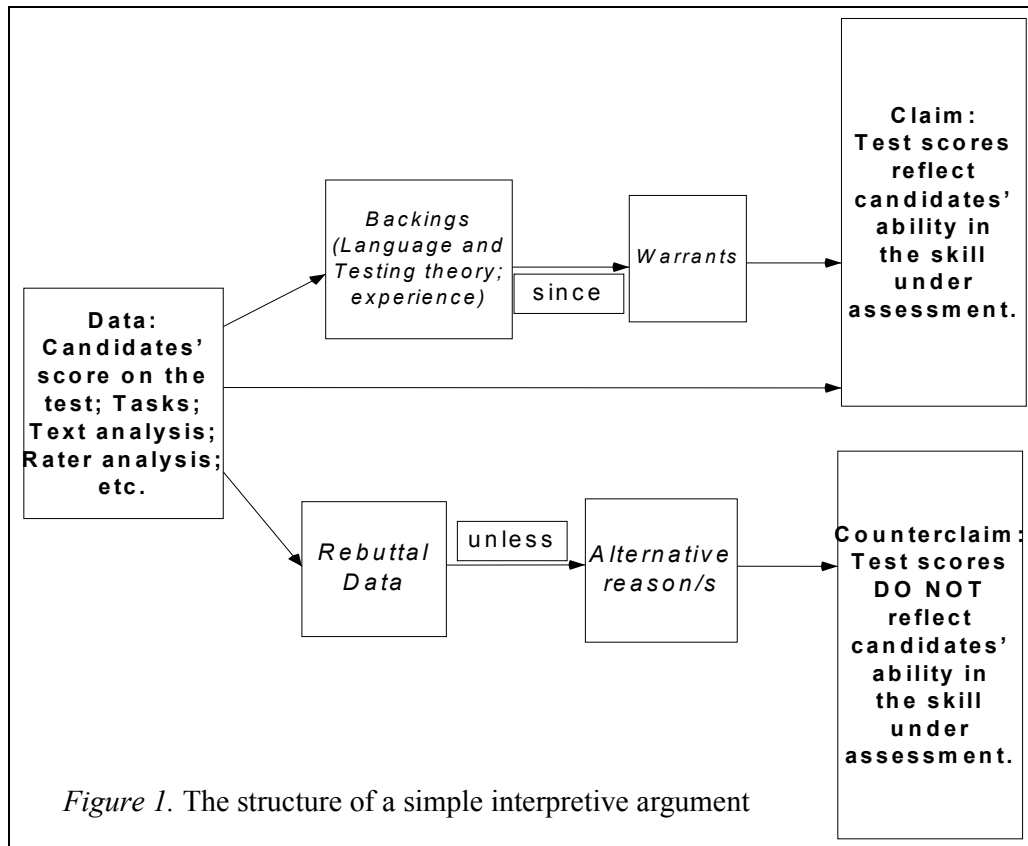
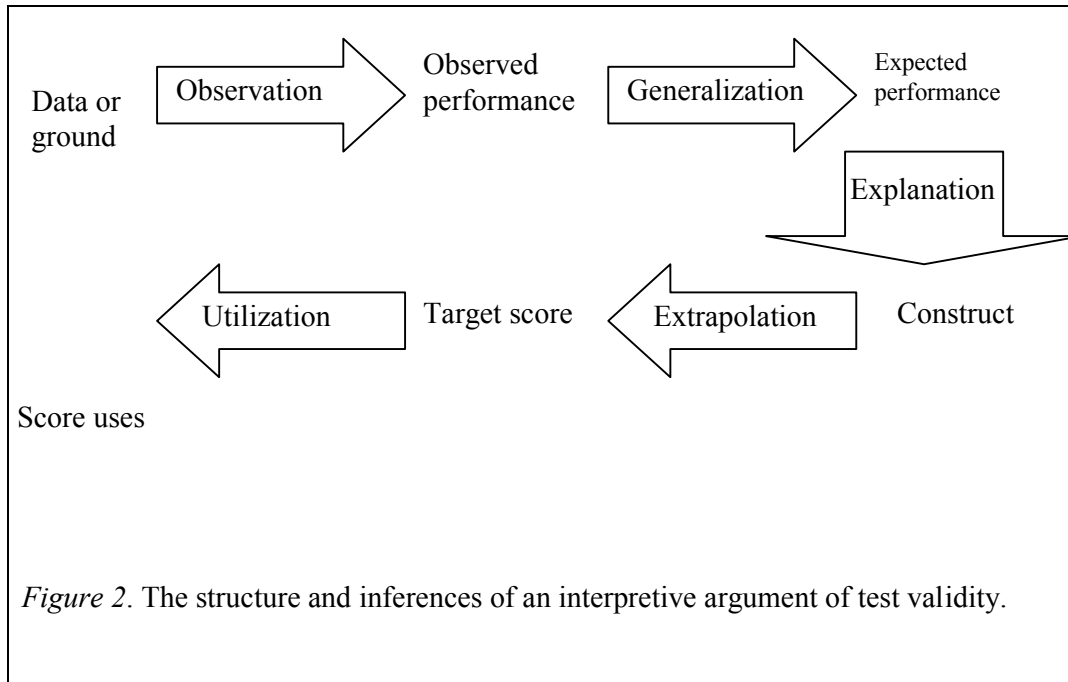


Figure 1 displays an interpretive argument structure comprising a claim, a warrant, data, backing, alternatives, and rebuttals. According to Toulmin (1958), claims are conclusions or interpretations that we intend to draw. A claim relies upon a warrant, one or more general statements that legitimize it (Toulmin, 1958). For example, syllogisms are strong warrants that guarantee a certain outcome. To provide warrants with authority and currency we need backings; these are any form of “assurance,” the “evidence supporting the warrant” (Kane, 2004, p. 148). These backings, in turn, should be drawn from data: expressions of what we have witnessed or observed corresponding to the inference or inferences being made (Mislevy, Steinberg, & Almond, 2003).

To establish a claim as false, a rebuttal is needed. Rebuttals weaken or reject a claim, its backings, or its warrants, unless alternatives are found to resolve them and support the claim.

As noted earlier, an interpretive argument of test validity is the first step in creating a validity argument for a test (Bachman, 2005). Kane (1992) proposed that a properly constructed interpretive argument for a test should follow four inferential

steps: observation, generalization, extrapolation, and utilization. More recently, Kane (2004, 2006), Bachman (2005), Chapelle (2008), and Bachman and Palmer (2010) argued that a fifth inference, called explanation, is needed to bridge the gap between generalization and extrapolation. Figure 2 incorporates this fifth inference, and displays the full structure of a test's interpretive argument.



The observation inference assesses the consistency of scoring methods with measurement processes. The process of measurement should be standardized irrespective of the measuring tool used. In assessment procedures where interpretations and evaluations are performed by an assessor, clear guidelines for the assessor's behavior ought to be stated. Irrelevant and contaminating factors, such as environmental variables (Bachman, 1990), can affect observation inferences seriously.

The generalization inference generalizes the observed scores and interpretations to a broader domain. In this inference, evidence should be collected to assess the degree of variance of test data across the domain. Chapelle (2008) proposed using G-theory and reliability analysis to warrant this inference.

The explanation inference associates test scores to the postulated construct theory. Smith (2005, p. 413) proposed that tests be subjected to a five-stage construct validation program to provide the construct-referenced evidence of validity needed to support this inference.

Because the explanation inference cannot extrapolate scores and observations to non-test behavior, Kane (1992) proposed the extrapolation inference, which involves the use of criterion-referenced evidence. This evidence can be either predictive or concurrent (Chapelle, 2008).

Utilization is the final inference in this framework. It legitimizes the uses to which scores are put. This inference is critical in high-stakes testing, because decision makers may use test scores to admit or reject applicants into their programs.

The present study employs this framework to investigate the validity of scores' use and interpretation in the IELTS Speaking Module, and to suggest the research that remains to be done before such an investigation becomes possible for the Listening Module.

3. Assumptions of the Study

Table 1 lists the assumptions made in this study, and the inferences required to make them. (The term "assumption" here is synonymous with the conventional term "hypothesis": "Assumption" is appropriate to this study because it is an underlying concept in validity arguments.) These assumptions are the basis of the interpretive argument phase of the study.

Table 1. Assumptions of the Study

Category	Definition
Assumption 1	Band scores in the Speaking and Listening Modules are clearly defined and testing conditions are standardized. (Observation) ^a
Assumption 2	Rater judgment does not affect scores in the Speaking Module. (Observation)
Assumption 3	Observed scores mirror expected scores across parallel task versions. (Generalization)
Assumption 4	Listening and Speaking skills have distinct construct definitions which are sufficiently operationalized and represented in the test. (Explanation)
Assumption 5	Construct irrelevancies are not detectable from data analysis. (Explanation)
Assumption 6	IELTS Speaking and Listening tasks portray real-life communication. (Extrapolation)
Assumption 7	Test scores are useful for decision making purposes. (Utilization)

Note. ^a = Every assumption underlies an inference which has been parenthesized.

4. Review of Research on the Speaking Module

Table 2 organizes twenty-eight studies on the Speaking Module into four broad categories, plus a “miscellaneous” category for two studies that did not fit in this framework but rendered relevant information about the module. The “examiners’ judgment” column lists five studies on the behavior of raters; the “analytic scoring and analysis of test and language structures” column includes eight studies that discuss the discourse produced in interviews and the functionality of the module’s analytic scoring system; the “revisions” column includes five studies that report on past and planned revisions to the structure of the Speaking Module; the “impact on future academic performance” column includes eight studies on the relationship of Speaking Module scores to future performance in the target language domain; and the “miscellaneous issues” column includes two relevant studies that fall outside the first four categories.

Table 2. Categorization of Research to Build a Validity Argument for the IELTS Speaking Module

Examiners' judgment	Analytic scoring and analysis of test and language structures	Revisions	Impact on future academic performance	Miscellaneous issues
1. Brown and Hill (1998)	1. Read (2005)	1. Tonkyn (1998)	1. Cotton and Conrow (1998)	1. O'Loughlin (2000)
2. Brown (2000)	2. Read and Nation (2006)	2. Lazaraton (1998)	2. Kerstjen and Nery (2000)	2. Issitt (2008)
3. McDowell (2000)	3. Seedhouse and Egbert (2006)	3. Taylor (2001)	3. Hill, Storch, and Lynch (1999)	
4. O'Sullivan and Lu (2006)	4. Brown (2006a)	4. Taylor and Jones (2001)	4. Allwright and Banerjee (1997)	
5. Merrylees (1999)	5. Brown (2006b)	5. Brooks (2003)	5. Paul (2007)	
	6. Elder and Wigglesworth (2006)		6. Rea-Dickins, Keily, and Yu (2007)	
	7. Weir, O'Sullivan, and Horai (2006)		7. Banerjee (2003)	
	8. Stoyhoff (2009)		8. MacNamara (2006)	

The review presented in this section synthesizes previous research to investigate the validity of test score uses and interpretations in the IELTS Speaking Module. The twenty-eight studies in the five categories presented in Table 2 are the dataset of this section, which I use to provide warrants and backings for the assumptions articulated in Table 1 and build a validity argument for the module.

4.1. Examiners' Judgment

Because examiner judgment is a crucial element of performance tests, several studies on the IELTS Speaking Module have focused on examiner behaviour. Brown and Hill (1998), for example, investigated the effects of examiners' rapport with test takers and the extent to which examiners offered help during the test (for example, by rephrasing test questions) on the structure of the discourse elicited from test takers. They used two discourse levels of interviews, easy and difficult, and observed that examiners who tended to use simpler questions had more problems changing discourse levels.

They also found that examiners substantially influenced test takers' performance. In response, they suggested a behavioural framework for examiners to follow in order to minimize this influence, and recommended that IELTS establish a more restricted code for examiner behaviour. It is worth noting that there was no standardized behaviour set for examiners until that time.

Merrylees (1999) investigated behavior among IELTS examiners in Australia, Honk Kong, Indonesia, Malaysia, the Philippines, and Thailand. At the time, the module had five sections, and Merrylees found that examiners did not elicit enough discourse from test takers in sections four and five. He proposed that "since the object of the IELTS interview is to elicit assessable discourse from the candidates but not to assess their listening comprehension skills, this is clearly a problem" (p. 34). He also reported some discrepancy between prescribed and observed interview durations. This was especially observed of low-proficiency test takers (which is acceptable because they could not produce lengthy stretches of discourse), but Merrylees also found that many test takers assigned a Band 5 score in the test's fourth section were given more time by raters to display their weaknesses than Band 6 scorers. Merrylees found that "these findings call into question the reliability of the test in the hands of examiners who chose not to extend their candidates fully" (p. 34).

Brown (2000) studied examiners' decision making processes to rate them, the relationship between "linguistic and non-linguistic aspects of [test] performance" (p. 49), the salience of the rating criteria, and the probability of a single test taker being assessed differently by different examiners. She used the tape-recorded data she had collected in her previous study on the Speaking Module (Brown & Hill, 1998) and recruited eight raters to assess performance in each interview. She found that topic choice, test condition, and the fact that raters were rating taped interviews affected the scores they assigned to the interviews. Brown's conclusion was almost the same as in the 1998 study: since the rating criteria were holistic and vague, it was not logical to expect consistent examiner behaviour (Brown, 2000).

Similarly, McDowell (2000) examined the efficacy of and discrepancies between IELTS examiners' evaluations of test performance. This study showed that examiners' effectiveness was influenced considerably by their choice of training materials. At the time, examiners could select either self-access to training packs or face-to-face training. Most preferred the self-access method, and McDowell observed behavioural discrepancies based on this choice. McDowell recommended practical

measures to amend the training guidelines and augment the quality of the training packages.

O'Sullivan and Lu (2006) examined effects of examiners' deviations from "the interlocutor frame," and the effects of these deviations on the performance of test takers in the Speaking Module. They found that there were few deviations in the first sections of the module and many in the last part—especially in paraphrasing the questions for test takers—but that these deviations did not influence test takers' performance much.

4.2. Analytic Scoring and Analysis of the Test and Language Structures

Following the 2001 replacement of the holistic scoring system in the Speaking Module by an analytic scale, eight IELTS studies addressed the features and scoring patterns of the new module. Read (2005) used a mini-corpus of 88 speaking performances of IELTS test takers to investigate different aspects of their vocabulary resources. He calculated the token (the total number of words used) and the type of words. More proficient test takers typically used a wider range of vocabulary, but tokens varied widely within each band score. In general, Read found no significant correlation between the use of vocabulary and the band scores assigned to test takers.

Following this study, Read and Nation (2006) attempted to investigate test takers' lexicon and use of "formulaic" language in the module. They found considerable variance in lexicon within band score levels, indicating the limited significance of vocabulary use in assessment, which agreed with Read's (2005) report. They also found that fluency and use of formulaic language and common words were decisive in helping test takers attain high band scores. For instance, among the candidates who were awarded a band score of four (indicating poor performance), the use of formulaic language was not common.

Addressing the interactional make-up of the module, Seedhouse and Egbert (2006) investigated its pragmatic features, such as "turn-taking, sequence, and repair" (p. 164). They found that most examiners adhered to the interview framework. They also contended that the interaction in the Speaking Module is unique in that it resembles both second-language interactions and interactions between students and academic staff at university. They argued that these similarities supported the construct representation of the module on the one hand, and the authenticity of the module and its tasks on the other.

Brown (2006a) made a general assessment of the module's analytic scale. She investigated the degree to which raters were able to differentiate between band scores, their consistency in scoring, and the "salient features" they used to judge test taker performance. Brown found evidence that the analytic scales were effective, but she also highlighted the problems of inference caused by the vagueness of some descriptors, especially in the Fluency and Coherence scale. In addition, she reported many examiners' insistence that the Pronunciation scale was not clear enough for them to confidently assign band scores.

In another study, Brown (2006b) examined the "Speaking Test band descriptors and criteria key indicators" to find what scoring criteria examiners used. She found that as test takers' overall level of proficiency increased, their awarded scores on individual assessment criteria (such as grammar and vocabulary) also increased, supporting the test's construct validity argument. However, the intervals between performance levels were found to be too large, and the test takers within individual levels formed non-homogenous groups. On most measures, the overlarge intervals between band scores meant that test takers falling near band score boundaries were assessed poorly relative to test takers falling squarely within band scores. Brown concluded that examiners were using the assessment criteria in combination, and not in isolation as intended.

The format of the Speaking Module was changed concurrently with the change in scoring. The module, which had been administered in five consecutive stages, was modified into a three-stage format. Two studies examined this new format. Elder and Wigglesworth (2006) investigated the effect of timing on the second phase of the module, which evaluates test takers' planning skills and language proficiency on a timed speaking task. In this phase, test takers are given a topic card containing a main question and a few sub-questions. They are then given one minute to take notes on the questions and prepare a short response. Elder and Wigglesworth's results did not show any essential difference in performance under timed and untimed conditions. The researchers recommended continuing to allot one minute of planning time in the interest of fairness and "face" (perceived) validity.

Weir, O'Sullivan, and Horai (2006) also investigated the effects of changes in preparation time on the difficulty of the second phase of the module and the "amount of scaffolding" given to test takers. ("Scaffolding" is a metaphorical concept that refers to the visible or audible assistance that a more expert member of a culture can

provide in any social setting; it is often analyzed in classroom discourse [Aryadoust, 2006, p. 149].) Weir et al. (2006) worked “within the socio-cognitive perspective of test validation” (p. 119) and employed both qualitative and quantitative research methods. They chose four tasks of the same difficulty level, three of which were altered in terms of “planning time, response time, and scaffolded support” (Weir et al., 2006, p. 143). They found that the unaltered test version produced the highest band scores, and that test takers of different abilities responded differently to the alterations to the other three tests. Contrary to Elder and Wigglesworth’s (2006) research, Weir et al. found that eliminating preparation time could negatively affect test takers’ performance, depending upon their level of proficiency.

Recently, Stoyhoff (2009) criticized the IELTS scoring system, including scoring of the Speaking Module, on the grounds that IELTS scoring lacks an underlying theoretical construct: “At present, the IELTS does not include an explicit theoretical rationale to support the interpretation and use of test scores and one must infer the conceptualizations underlying the test constructs” (p. 18).

4.3. Revisions

The results of commissioned IELTS studies have been applied in a number of IELTS Revision Programs. Several reports have narrated and analyzed these programs. In particular, Taylor (2001) and Taylor and Jones (2001) reviewed five major phases that the IELTS Speaking Module has gone through.

In phase one, “consultation, initial planning and design,” an extensive study was commissioned in June 1998. Taylor and Jones (2001) mention Tonkyn (1998) and Lazaraton’s (1998) articles in this stage. For example:

...along with findings from earlier studies, [Lazaraton’s work] raised the question of how well the existing holistic IELTS rating scale and its descriptors were able to articulate key features of performance at different levels or bands. It was felt that a clearer specification of performance features at different proficiency levels might enhance standardization of assessment. (p. 8)

In the second phase, “development”, the “prototype of the revised test format” was trialled (Taylor & Jones, 2001, p.1). This phase lasted from January to September

1999. Second and third drafts of the test's assessment criteria were then constructed to help examiners understand the rating scale descriptors.

The third phase, "validation," proceeded from October 1999 to September 2000, and generated the following results, as reported by Taylor and Jones (2001):

- A. Pronunciation is a separate construct, but grammatical and vocabulary scales are highly correlated.
- B. There was no need to collapse the rating scales, as the Rasch model fit the data.
- C. Raters' scoring reflected a largely unified interpretation of performance, although some scores were found to misfit.
- D. The test was highly reliable and generalizable. In particular, its G- and Phi-coefficients increased as the number of subscales increased.

The fourth phase, "implementation," lasted from October 2000 to June 2001. In this phase, retrospective studies further compared the new and old rating methods and showed the revised rating methods to be of good quality, and a worldwide examiner retraining was undertaken to implement the new version of the test. Finally, the fifth or "operational" phase included the full operationalization of the revised IELTS Speaking Module worldwide.

Taylor (2001) also explained the process of retraining examiners worldwide. After the first revisions, some Senior Trainers were retrained again at a regional level. The new materials contained "an IELTS Examiner Induction Pack with accompanying video and worksheet [and] an IELTS Examiner Training Pack, with two accompanying videos and detailed Notes for Trainers" (Taylor, 2001, p. 9). After the training session, ninety-nine percent of examiners and trainers described the package contents as "very good" or "fairly good." Their suggestions were to be incorporated into future versions of the packs. The entire process of retraining IELTS examiners worldwide took between four and five months.

Brooks (2003) explained how an observation checklist (OC) used in the speaking section of the First Certificate in English (FCE) and Cambridge Proficiency in English (CPE) tests was converted into a checklist for the IELTS Speaking Module. Since IELTS speaking tests are audiotaped, Brooks described the advantages of audiotaped samples that did not have any animated visual input to distract the users of the OC.

Although the scholarly literature on IELTS from 2007 and 2008 did not focus on the Speaking Module, its results impacted the module heavily: they convinced the Cambridge ESOL Examinations, the British Council, and IELTS Australia to adopt half-band scores, significantly changing the Speaking Module's scoring system; and they led those organizations to resume study of stakeholder attitudes and test impacts as their main research agenda, some of which focused on the relationship between performance on the Speaking Module and future academic performance.

4.4. Impact on Future Academic Performance

The major backing for the extrapolation inference is future performance, the role of the theoretical construct in the task domain. Predictive validity research can assess this relationship by correlating test results with academic performance, as indicated by grade point average (GPA) in the task domain.

Investigating the predictive power of IELTS at the University of Tasmania, Cotton and Conrow (1998) found no significant correlation between GPA and performance on the Speaking Test. Academic staff ratings also did not significantly correlate with test results. More than half of interviewed students believed that IELTS could not bring into focus their areas of weakness in English, and many students reported language problems in their courses despite having received acceptable IELTS scores. The researchers pointed to the subjectivity of the interviews and questions as problems with the Speaking Module.

Two more similar studies were conducted, one by Hill, Storch, and Lynch (1999), and the other at RMIT University in Australia by Kerstjen and Nery (2000). In these studies, IELTS scores did not predict academic performance strongly, although scores on the Reading and Writing Modules were significantly correlated with it in the latter study. The Speaking Module had no significant correlation with students' academic performance, nor did it account for any portion of the observed variance in academic performance in regression analysis. Kerstjen and Nery further argued that the predictive power of the IELTS scores relied on the test takers' field of study. Allwright and Banerjee (1997) also found that IELTS scores above 7.0 indicated a low risk of failure in academia, but they reported no obvious relationship between IELTS scores and academic performance. Likewise, Paul (2007) suggested that "Language production at a micro level similar to that in IELTS tasks is not necessarily an indicator of overall language adequacy at a macro level or successful task

completion [in the IELTS Speaking Test]” (p. 205). Most directly, Rea-Dickins, Keily, and Yu (2007) found the Speaking Module to be a poor predictor of test takers’ future academic performance.

The studies reviewed above generally reported that academic performance is not solely a matter of language production and proficiency, but depends on other significant variables such as personality, socioeconomic background, nationality, and affective/cognitive factors. Rea-Dickins et al. (2007) found that acceptable scores on the Speaking Module do not guarantee a lack of communication problems in tutorials, since other intervening factors such as stress and the use of specialised academic language may impede students’ ability to communicate effectively.

Further, Banerjee (2003) reported that many university admissions decision makers lack a clear understanding of the meaning and interpretation of IELTS scores. Rea-Dickins et al. (2007) confirmed this finding, and added that even when students’ scores on the sub-skills of the test satisfy a program’s admission requirements, students are found to “lack critical thinking and evaluative skills” (p. 117).

Finally, McNamara (2006) raised the concern that IELTS scores are being interpreted and used in use contexts where the test has not been validated yet, such as “immigration selection in Australia and other countries” (p. 30).

4.5. Miscellaneous Issues

Two issues are discussed here: the effect of gender on examiners, and methods for enhancing candidates’ performance. The former has bearing on the test’s construct-referenced evidence of validity and the latter on its consequential validity—that is, on the consequences of the test on the educational system.

Gender has been a longstanding topic of discussion in performance tests. In high-stakes tests such as IELTS, its effects on both examiners and test takers may cause bias and unfairness. O’Loughlin (2000) investigated the effect of gender on the Speaking Module rating process and the discourse produced during the interview, and found no systematic differences.

Some IELTS research has also focused on the possibility of enhancing candidates’ performance. In an uncommissioned study, Issitt (2008) recommended using three strategies in IELTS preparation courses for the Speaking Module: materials to build candidates’ confidence, materials to encourage candidates to “think

critically,” and “close inspection and utilization” of the rating criteria available from IELTS materials. Issitt argued that these strategies are especially effective for alleviating excessive test taker stress, although he cautioned that they had not been tested on large samples of test takers.

5. Review of Research on the Listening Module

The IELTS Listening Module is the least-researched IELTS module. Until very recently, no published studies focused exclusively on the module.

Coleman and Heap (1998) examined IELTS Listening Module and Reading Module rubrics for evidence of confusion or misunderstanding resulting from the concurrent processing of auditory and visual input. They found the test rubrics to be clear, with “only minor areas which need tightening up” (p. 70), but that some test items were not as clearly written as the rubrics.

In a survey study, Merrylees (2003) showed that 78% of IELTS test takers had no problem understanding rubrics, instructions, and items in the Listening Module. Some individuals had trouble with the subject of the stimuli; 23% believed that the pace of speech delivery was too fast, whereas 44% believed it was not; participants categorically agreed that the more they moved through the test, the more the test became difficult; and some participants had difficulties understanding the British accent.

A number of predictive validity studies have examined the Listening Module alongside the other modules. Kerstjen and Nery (2000) found no correlation between test takers’ Listening Module scores and their academic performance in business courses. They argued that, “It may be that, as the Listening test does not test academic listening skills, it does not accurately predict the kind of listening skills that students in these Business courses are required to develop” (p. 96). Cotton and Conrow (1998) found that Listening Module scores were actually negatively correlated to academic performance for two groups of students: the correlation was weakly negative (-0.19) for a first group of twenty-six test takers after a full academic year, moderately negative (-0.58) for a second group of seventeen students after their first semester, and moderately negative (-0.56) for the same seventeen test takers in their second semester. Cotton and Conrow also interviewed students on the face validity of the Listening Module. Some commented that the module was too general and did not suit academic purposes, others that the test was easy, and others

that performance was affected by test conditions such as the distance of the test taker from the tape player.

Working in Australia, Ingram and Bayliss (2007) reported a weakly positive correlation between test takers' Listening Module scores and their self-evaluation of their lecture comprehension skills, but that "there was not, in fact, any obvious relationship between IELTS Listening scores and the amount of understanding of classes and discussions, according to the participants" (p. 162). Students experienced difficulty when interacting with their peers, although they felt better able to comprehend teaching staff. They further observed that most students did not take many notes in lectures and that their notes did not contain much "substantive content" (p. 179). All students had problems with the speed of speech delivery and the Australian accent, particularly in their first semester.

IELTS has also commissioned four studies on the Listening Module, two of which have only recently been published. Table 3 presents these studies.

Table 3. Unpublished IELTS-Commissioned Studies on the Listening Module

Researcher	Study title	Comments
J. Field	A cognitive validation of the lecture-listening component of the IELTS listening paper.	Round 11, 2005
R. Badger and O. Yan	The use of tactics and strategies by Chinese students in the listening component of IELTS.	Round 12, 2006
R. Breeze and P. Miller	Predictive validity of the IELTS Listening Module as an indicator of student coping ability in English-medium undergraduate courses in Spain.	Round 14, 2008; Unpublished
F. Nakatsuhara	The relationship between test-takers' listening proficiency and their performance on the IELTS Speaking test	Round 15, 2009; Unpublished

Field (2009) found that gap filling and MCQ items in the fourth section of the test engage construct-irrelevant skills. He stated that:

The former [i.e., gap filling] has unfortunate effect on focusing candidate attention at word level and providing *gratis* a great deal of structure of the lecture which it should be the listener's responsibility to construct. The latter [i.e., MCQ] imposes heavy reading demands. Both foster a practice of switching attention away from the recording to the written modality. (Field, 2009, p. 47)

Field (2009) pointed out that IELTS listening test format confines test takers' cognitive processing because they hear the audio input a single time—which has been a long-established tradition in many Cambridge exams—and have to rely heavily on their reading skills while listening to the audio material. Shifting to different modes of input (i.e., listening and reading), test takers can become overanxious if they miss an item. This invites test-wiseness strategies and those who are test-wise stand a better chance to answer the item successfully. In addition, the test structure and the construct irrelevant processes jointly lead test takers to a superficial comprehension of the aural text. The lower level comprehension processes coupled with the limited redundancy of the audio materials and their richness in terms of detailed information make a high cognitive demand of test takers, which is beyond the discourse structure of the university lectures.

The findings of Field's (2009) qualitative research resonate with Aryadoust's (in press) Rasch-based differential item functioning study of the IELTS listening test. Aryadoust found that the listening construct in the IELTS test was underrepresented, which "is probably an important cause of the lack of significant correlation between test results and academic performance observed in previous studies" (Aryadoust, in press, p. ##). He further reported that "short answer items, which feature prominently on the test, are likely to be biased in favor of higher-ability listener subgroups in listening comprehension because of these test takers' ability to apply swiftly what they have understood" (Aryadoust, in press, p. ##). Like gap filling and MCQ items, these items are not effective to measure the listening construct due to the cognitive limitations they impose on listeners. Finally, that low-ability test takers who are test-wise appear to be taking advantage of this fact leads to flawed test results.

In another study, Badger and Yan (2009) found no significant differences between pre-undergraduate and pre-postgraduate students taking the IELTS listening test in using strategies and tactics. Although Badger and Yan believed that this finding provides a piece of evidence supporting the validity arguments of the IELTS listening test, they argued that some of the written text in the examination should be replaced by listening texts, which is in agreement with Field's (2009) finding. They agreed with Field in that an efficient alternative to the present IELTS system is a test system where the test takers hear the audio materials twice. Another solution, as Badger and

Yan argued, would engage maintaining the current practice where test takers listen once without reading the test items. They are encouraged to take notes and use them to answer the test items which are displayed or read to them later. This alternative, which would lead to a major change in the structure of the test and the way the construct is operationalized, is the testing method which has been recently adopted by the Test of English as a Foreign Language (TOEFL).

Because of the paucity of published research on the IELTS Listening Module, and the findings of the studies above which generally do not support its functionality, many outstanding questions preclude the establishment of a comprehensive validity argument for the module.

6. Results and Discussion

IELTS Speaking Module studies have prioritized improving the assessment power of the module and improving the theoretical construct. Findings from these studies have helped test developers devise newer formats of assessing speaking. Now we are in a position to set forth the validity argument for the IELTS Speaking Module. This validity argument is presented in Table 4. All components in this argument, and its division into seven inference classes, are extracted from the main research assumptions that have guided IELTS researchers.

Table 4. Validity Argument for the IELTS Speaking Module, Based on Data from IELTS Studies and Revision Programs

Inference	Warrants	Backings/rebuttals
<i>Observation (A)</i> Band scores in the Speaking Module are clearly defined.	According to IELTS descriptions in Research Notes, band scores are intended to precisely describe speaking ability.	<i>Backings:</i> Although the findings of Brown and Hill (1998) and Brown (2000) questioned the module's scoring system, the later adoption of an analytic rating scale largely resolved these concerns. Tonkyn (1998), Lazaraton (1998), and Taylor (2001), reported on a Revision Program and retraining of examiners; McDowell (2000) recommended ways to improve the training packages; Brown (2006b) concluded that the Speaking Test and its scores are valid and the band scores are clear. Half-band scores were introduced to the scoring system in 2008 to clarify the scoring process.
<i>Observation (B)</i> Rater judgment does not significantly affect	IELTS has made a concerted effort to neutralize the effect of examiner judgment on	<i>Backings:</i> O'Sullivan and Lu (2006) found that examiner deviations did not adversely affect test taker performance. Seedhouse and Egbert (2006) found that most examiners adhered to the

scores.	scores.	interview framework.
<i>Generalization</i> Observed scores approximate expected scores across parallel test/task versions.	Significant G-theory and high reliability coefficients support the generalizability of the module.	<i>Backings:</i> Merrylees's (1999) study, training, and revision programs support the generalization inference. The IELTS official Web site reports an inter-rater correlation of 0.77 and a G-coefficient of 0.86 for the test as a whole.
<i>Explanation (A)</i> Second-language speaking as operationalized in IELTS entails representative components which are measured in the test.	The test assesses speaking as a component of language proficiency using an analytic scoring system comprising vocabulary, grammar, discourse, and cohesion/coherence of ideas.	<i>Backings:</i> Read and Nation (2006) found that fluency and the use of formulaic language and common words affect scores. O'Sullivan and Lu (2006) found that raters mainly followed the module's analytical framework, comprising vocabulary, grammar, discourse, and cohesion/coherence. <i>Rebuttals:</i> Read (2005) found that proficient students often had wider vocabularies, but otherwise found, with Read and Nation (2006), that vocabulary has little relationship to scores. Stoyloff (2009) found that the module's analytical framework is not based on any explicit underlying theory of second-language speaking.
<i>Explanation (B)</i> The current practices in the IELTS Speaking Test do not impede candidates' real performance.	Standardized tests should not expose students to impeding variables: excessive stress and construct irrelevancies should be eliminated.	<i>Backings:</i> Weir, O'Sullivan, and Horai (2006) showed that test takers performed better under the module's standard time and support constraints than under any of three alternatives. Conversely, Elder and Wigglesworth (2006) found that preparation time did not significantly affect test taker performance. The IELTS Revision Programs unified the test's rating system.
<i>Extrapolation</i> IELTS Speaking tasks which operationalize the construct can predict the future performance of test takers.	The IELTS Speaking Module predicts candidates' ability to complete tasks they may encounter in academia and social life.	<i>Backings:</i> Seedhouse and Egbert (2006) concluded that IELTS Speaking tasks were quite similar to both university and second-language interactions. <i>Rebuttals:</i> Most impact studies—by Allwright and Banerjee (1997), Banerjee (2003), Cotton and Conrow (1998), Hill, Storch, and Lynch (1999), Kerstjen and Nery (2000), Paul (2007), Rea-Dickins, Keily, and Yu (2007)—showed no correlation between academic performance and Speaking Module scores.
<i>Utilization</i> Test scores are useful for making admissions and remedial program placement decisions.	Standard setting, available score interpretation materials, and positive attitudes of stakeholders are evidence of the utility of scores.	<i>Backings:</i> Impact studies showed that the majority of candidates, academic staff, and decision makers considered IELTS a fair test. Also, regular Revision Programs, which retrain examiners and rethink benchmarks, represent a policy of standard setting. <i>Rebuttals:</i> Banerjee (2003) and Rea-Dickins et al. (2007) reported that many admissions decision makers are unqualified to interpret IELTS scores, and McNamara (2006) reported the use of IELTS scores in untested domains, such as immigration.

The two observation inferences are attenuated by studies by Brown and Hill (1998), Brown (2000), and MacNamara (1996), but these concerns have been largely

addressed by the new analytic scoring system developed in 2001 and modified in 2008, and by the introduction of a number of guidelines for examiner behaviour.

The generalization inference is made based on observed scores. G-theory, reliability coefficients, and similar analyses provide support for this inference.

The explanation inference is supported by pertinent warrants which are based on the definition of construct-referenced evidence of validity. However, the inference is compromised by the lack of an IELTS-sanctioned model of the constituent structure of second-language speech. The Speaking Module's analytic scoring system assumes that second-language speech comprises vocabulary, grammar, discourse, and cohesion/coherence, but IELTS provides no explicit theoretical model to support this assumption. The lack of construct models is a problem extending to the entire test (Stoynoff, 2009).

The extrapolation inference is also highly problematic for the Speaking Module. Supporting evidence for this inference is admittedly difficult to collect through predictive validity studies, since academic performance is a function of many factors other than language proficiency; nevertheless, tested language proficiency should be able to predict a significant portion of future academic performance. The preponderance of existing literature indicates that the Speaking Module is inadequate in this regard.

The utilization inference also has significant complications. It is warranted by institutionalized standard-setting procedures, available score interpretation materials, and the largely positive attitudes of stakeholders toward IELTS generally, but rebutted by studies that show that admission officers and academic staff are poorly equipped to interpret IELTS scores, and by the misuse of IELTS test results for decision making in unapproved contexts such as immigration and job selection.

7. Conclusion

This article reviews and organizes research on the IELTS Speaking and Listening modules. The existing literature lends some support to the Speaking Module's validity argument, but also presents important evidence against it: the module's scoring system does not refer to any explicit underlying theory of second-language speech; performance on the module does not appear to predict future performance in the target language domains; and test results are often interpreted in unintended contexts, or by decision makers with limited understanding of the test. In the five-stage inferential

framework adopted in this article, the module's explanation, extrapolation, and utilization inferences need improvement. Focusing IELTS studies on these three stages of validation and inference-making could help improve the test.

There is also an urgent need to support the test structure and validity argument of the Listening Module, whose almost total lack of scholarly attention contrasts sharply with its complexity and the complexity of the listening skill.

Acknowledgements

Acknowledgements are gratefully extended to Professor Philippa Mungra and two anonymous reviewers for their comments on an earlier draft of this article.

References

- Allwright, J., & Banerjee, J. (1997). *Investigating the accuracy of admissions criteria: A case study on a British university*. Lancaster: Institute for English Language Education, Lancaster University.
- Aryadoust, S. V. (2006). *A dictionary of sociolinguistics, plus pragmatics and languages*. Shiraz: Faramatn Publication.
- Aryadoust, S. V. (in press). Differential item functioning in while-listening performance tests. *International Journal of Listening*.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1–34.
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in the real world: Designing language assessments and justifying their use*. Oxford: Oxford University Press.
- Badger, R., & Yan, X. (2009). The use of tactics and strategies by Chinese students in the Listening component of IELTS. In Thompson, P. (Ed.), *IELTS research report*, (Vol. 9), (pp. 67-96). UK and Australia: British Council and IELTS Australia.
- Banerjee, J. (2003). *Interpreting and using proficiency test scores*. Unpublished doctoral dissertation, Lancaster University, Lancaster, UK.
- Brooks, L. (2003). Converting an observation checklist for use with the IELTS speaking test. *Research Notes*, 11, 20-21.

- Brown, A. (2000). An investigation of the rating process in the IELTS oral interview. In R. Tulloh (Ed.), *IELTS Research Reports* (Vol. 6) (pp. 49-84). Canberra: IELTS Australia, Pty Ltd.
- Brown, A. (2006a). An examination of the rating process in the revised IELTS speaking test. In P. McGovern & S. Walsh (Eds.), *IELTS Research Report* (Vol. 6) (pp. 41-70). Canberra: IELTS Australia, Pty Ltd & British Council.
- Brown, A. (2006b). Candidate discourse in the revised IELTS Speaking Test. In P. McGovern & S. Walsh (Eds.), *IELTS Research Reports* (Vol. 6) (pp. 71-90). Canberra: IELTS Australia, Pty Ltd & British Council.
- Brown, A., & Hill, K. (1998). Interviewer style and candidate performance in the IELTS oral interview. In S. Wood (Ed.), *IELTS Research Reports* (Vol. 1) (pp. 1-19).
- Chapelle, C. (1999). Validity in language testing. *Annual Review of Applied Linguistics*, 19, 254-274.
- Chapelle, C. (2008). The TOEFL validity argument. In C. Chapelle, M. Enright, & J. Jamieson, (Eds.), *Building a validity argument for the test of English as a foreign language* (pp. 319-350). New York: Routledge.
- Chapelle, C. A., Enright, M. K., & Jamieson, J. M. (2008). Test score interpretation and use. In C.A., Chapelle, M.K., Enright, & J.M., Jamieson (Eds.), *Building a validity argument for the Test of English as a Foreign Language* (pp. 1-25). New York: Routledge.
- Coleman, G., & Heap, S. (1998). The misinterpretation of directions for the questions in the Academic Reading and Listening sub-tests of the IELTS test. In S. Wood (Ed.), *IELTS Research Reports* (Vol. 1) (pp. 38-71). Canberra: IELTS Australia, Pty Ltd.
- Cotton, F., & Conrow, F. (1998). An investigation of the predictive validity of IELTS amongst a sample of international students studying at the University of Tasmania. In S. Wood (Ed.), *IELTS Research Reports* (Vol. 1) (pp. 72-115). Canberra: IELTS Australia, Pty Ltd & British Council.
- Elder, C., & Wigglesworth, G. (2006). *An investigation of the effectiveness and validity of planning time in Part 2 of the IELTS Speaking Test*. In P. McGovern, & S. Walsh (Eds.), *IELTS Research Reports* (Vol. 6) (pp. 13-40). Canberra: IELTS Australia, Pty Ltd & British Council.

- Field, J. (2009). A cognitive validation of the lecture-listening component of the IELTS Listening paper. In Thompson, P. (Ed.), *IELTS research report*, (Vol. 9), (pp. 17-66). UK and Australia: British Council and IELTS Australia.
- Hill, K., Storch, N., & Lynch, B. (1999). A comparison of IELTS and TOEFL as predictors of academic success. In R. Tulloh (Ed.), *IELTS Research Reports*, (Vol. 2), (pp. 52-63). Canberra: IELTS Australia Pty Limited.
- Ingram, D., & Bayliss, A. (2007). IELTS as a predictor of academic language performance, part 1. In L. Taylor (Ed.), *IELTS Research Reports* (Vol. 7) (pp. 137-204). Canberra: IELTS Australia, Pty Ltd & British Council.
- Issitt, S. (2008). Improving Scores on the IELTS Speaking test. *ELT Journal*, 62(2), 131-138.
- Kane, M. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112, 527-535.
- Kane, M. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38(4), 319-342.
- Kane, M. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice*, 21(1), 31-41.
- Kane, M. (2004). Certification testing as an illustration of argument-based validation. *Measurement: Interdisciplinary Research and Perspectives*, 2(3), 135-170.
- Kerstjen, M., & Nery, C. (2000). Predictive validity in the IELTS test. In R. Tulloh (Ed.), *IELTS Research Reports* (Vol. 6), (pp. 85-108). Canberra: IELTS Australia, Pty Ltd.
- Lazaraton, A. (1998). *An analysis of differences in linguistic features of candidates at different levels of the IELTS Speaking Test*. Unpublished study commissioned by UCLES.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- McDowell, C. (2000). *Monitoring IELTS examiner training effectiveness*. In R. Tulloh (Ed.), *IELTS Research Reports* (Vol. 3), (pp. 109-142). Canberra: IELTS Australia, Pty Ltd.
- McNamara, T. (1996). *Measuring second language performance*. Longman: New York.

- McNamara, T. (2006). Validity and values: Inferences and generalizability in language testing. In M. Chalhoub-Deville (Ed.), *Inference and generalizability in applied linguistics: Multiple perspectives* (pp. 27-46). Philadelphia, PA, USA: John Benjamins Publishing Company.
- Merrylees, B. (1999). An investigation of speaking test reliability. In R. Tulloh (Ed.), *IELTS Research Reports*, (Vol. 2) (pp. 1-52). Canberra: IELTS Australia Pty Limited.
- Merrylees, B. (2003). An impact study of two IELTS user groups: Candidates who sit the test for immigration purposes and candidates who sit the test for secondary education purposes. In R. Tulloh (Ed.), *IELTS Research Reports*, (Vol. 4) (pp. 1-58). Canberra: IELTS Australia Pty Limited.
- Messick, S. (1980). Test validity and the ethics of assessment. *American psychologist*, 35, 1012-27.
- Messick, S. (1988). *Meaning and values in test validation: The science and ethics of assessment*. Research Reports. New Jersey: Educational Testing Service, Princeton.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). New York: ACE and Macmillan.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1, 3-62.
- O'Loughlin, K. (2000). The impact of gender in the IELTS oral interview. In R. Tulloh (Ed.), *IELTS Research Reports*, 3, (pp. 1-28). Canberra: IELTS Australia, Pty Ltd.
- O'Sullivan, B., & Lu, Y. (2006). The impact on candidate language of examiner deviation from a set interlocutor frame in the IELTS Speaking Test. In P. McGovern & S. Walsh (Eds.), *IELTS Research Reports*, (Vol. 6) (pp. 119-160). Canberra: IELTS Australia, Pty Ltd.
- Paul, A. (2007). IELTS as a predictor of academic language performance, Part 2. In P. McGovern & S. Walsh, *IELTS Research Reports*, (Vol. 6) (pp. 205-240). Canberra: IELTS Australia, Pty Ltd.
- Read, J. (2005). Applying lexical statistics to the IELTS speaking test. *Research Notes*, 20, 12-16.

- Read, J., & Nation, P. (2006). *An investigation of the lexical dimension of the IELTS Speaking Test*. In P. McGovern & S. Walsh (Eds.), *IELTS Research Reports* (Vol. 6) (pp. 207-231). Canberra: IELTS Australia, Pty Ltd & British Council.
- Rea-Dickins, P., Keily, R., & Yu, G. (2007). Student identity, learning and progression: The affective and academic impact of IELTS on 'successful' candidates. In P. McGovern & S. Walsh, *IELTS Research Reports*, (Vol. 6) (pp. 56-136). Canberra: IELTS Australia, Pty Ltd.
- Seedhouse, P., & Egbert, M. (2006). The interactional organization of the IELTS Speaking Test. In P. McGovern & S. Walsh (Eds.), *IELTS Research Report*, (Vol.6) (pp. 161-206). Canberra: IELTS Australia, Pty Ltd & British Council.
- Stoynoff, S. J. (2009). Recent developments in language assessment and the case of four large-scale tests of ESOL ability. *Language Teaching*, 42(1), 1-40.
- Taylor, L. (2001). Revising the IELTS speaking test: retraining the IELTS examiners worldwide. *Research Notes*, 6, 9-11.
- Taylor, L., & Jones, N. (2001). Revising the IELTS speaking test: Retraining the IELTS examiners worldwide. *Research Notes*, 6, 9-11.
- Tonkyn, A. (1998). *Reading University/UCLES IELTS rating research project*. Unpublished Interim Report.
- Toulmin, S. E. (1958/2003). *The uses of argument*. Cambridge: Cambridge University Press.
- Upshur, J. A. (1979). Functional proficiency theory and a research role for language tests. In E. J. Briere & F. B. Hinofotis (Eds.), *Concepts in language testing: Some recent studies* (pp. 75-100). Washington DC: TESOL.
- Wagner, E. (2010). The effect of the use of video texts on ESL listening test-taker performance. *Language Testing*, 27, 1-21.
- Weir, C., O'Sullivan, B., & Horai, T. (2006). Exploring difficulty in speaking tasks: An intra-task perspective. In P. McGovern & S. Walsh (Eds.), *IELTS Research Reports* (Vol.6) (pp. 119-160). Canberra: IELTS Australia, Pty Ltd & British Council.



Title

**Civil engineering research article Introductions:
Textual structure and linguistic characterization**

Author

Budsaba Kanoksilapatham
Silpakorn University

Biodata

Budsaba Kanoksilapatham is an associate professor at the Department of English, Faculty of Arts, Silpakorn University, Thailand. She completed her PhD in Linguistics (Concentration: Applied Linguistics) at Georgetown University. Her current interests include genre analysis, multidimensional analysis, sociolinguistics and pronunciation teaching.

Abstract

This study reports on the examination of the Introductions of research articles in civil engineering in English. Specifically, this study has the objectives of 1) identifying the structural organization of the Introductions in terms of moves and steps, and 2) capturing linguistic features that are commonly used in each move and step. A dataset of civil engineering research articles was compiled to represent the high quality journals in the field. Subsequently, the dataset was analyzed by using Swales' genre analysis model (2004) to segment the Introduction texts into sub-units called 'moves' and 'steps,' respectively, according to their communicative functions. Finally, text segments representing each move and step were scrutinized to capture a typical set of linguistic features used to perform a given communicative function. While the two-tiered analysis reveals unique features of the language of civil

engineering Introduction discourse, this study elucidates the proclivities of the civil engineering discourse community in constructing its research articles. Pedagogically, the description of both the structure and the linguistic levels contributes to a better understanding of how research article Introductions are constructed in Civil engineering.

Keywords: structural organization; linguistic features; Introductions; genre analysis; civil engineering

1. Introduction

Due to Anglophone supremacy, English has become one of the most dominant international languages in both professional and academic settings. Academically, English is increasingly used in advanced education. Among all academic genres, research articles constitute a key genre used by scientific communities as a medium for the dissemination and ratification of knowledge (Kanoksilapatham, 2005: 270).

Writing a research article is a daunting task to both non-native speakers and native speakers of English (Canagarajah, 1996; Flowerdew, 2001). It requires, for instance, substantial skills in reading research articles published previously in order to know the state of the art of the field, individual scientists' expertise and contribution to the field, and finally writing skills to ensure that scientists can efficiently and successfully express themselves academically. For writing effectively, scientists need to make use of top-down and bottom-up approaches. Based on Cook (1990) and McCarthy (1991), the bottom-up knowledge of the research article genre refers to the scientists' knowledge gained from the smallest chunks of research article discourse (e.g., flexical and grammatical features) moving to more and more general features (e.g., sentences, discourse units, and structural organizations of the discourse units). In contrast, the top-down knowledge of the research article genre starts with general features of a text (e.g., from structural organizations of the discourse units) gradually moving to more specific ones (e.g., lexical and grammatical features).

A multitude of genre-based studies have focused on the rhetorical organizations of different sections of research articles, in particular the four sections of Introduction, Methods, Results, and Discussion (e.g., Brett, 1994; Nwogu, 1997; Kanoksilapatham, 2003; Martin, 2003; Posteguillo, 1999; Varttala, 2002). These studies reveal that, among the four typical sections of research articles, the Introduction section seems to display a relatively rigid and uniformed rhetorical

organization. Nevertheless, other studies have found that rhetorical variations of this particular section are discernible across academic disciplines (e.g., biochemistry and microbiology (Kanoksilapatham, 2007b); applied linguistics (Ozturk, 2007); wildlife behavior and conservation biology (Samraj, 2002). To illustrate, Kanoksilapatham (2007b) compared the rhetorical organizations typically followed in the corpora of 60 biochemistry and 60 microbiology research articles and found that the occurrence of ‘Move 3, Step 5: Announcing principal outcome’ in biochemistry is twice as frequent as that in microbiology (53 vs 26 instances). The study illustrates that each academic discourse community seems to have its own proclivities with regard to style and rhetorical structure for effective and efficient knowledge dissemination among its members.

However, to ultimately become professional genre writers within a discourse community, merely knowing and following the conventions of the genre might not be adequate for researchers. As challenged by Nunn and Adamson (2007) and as remarked by Carter (1995), being trained to follow such conventions can have unpleasant impacts, creating “unreflective writers who will be able to do no more than sustain the genres” (p. 55). As advocated by Harwood and Hadley (2004: 357-358), and subsequently substantiated by Flowerdew (2007), gate keepers of their respective discourse communities (be they journal editors or university tutors) and their roles in maintaining norms of academic writing practice should be challenged. Specifically, novice researchers of a discipline, having been exposed to those conventions, should “question the desirability of reinforcing these predominant norms” (Flowerdew, 2007: 22). This statement corroborates the notion of “conversations”, a term introduced by Bazerman (1980: 657) to refer to the practice when people engaged in actions whose meanings can be negotiated. Since academic researchers are not isolated persons, they are encouraged to get involved in “conversation”.

Whereas a number of scholars cited above emphasized that scholars should not only follow and readily accept the conventions laid down by the community, such a practice is rather intimidating for both novice researchers and for non-native speakers of English. To facilitate the long and arduous process of getting accepted as discourse community members, especially at its initial stage, by preparing them for the academic demands, this study focuses on the analysis of a civil engineering dataset to investigate how research articles are constructed. Civil engineering is known to be an academic discipline dating back to the year 1711, when John Smeaton,

the founder of a Society of Engineers and the builder of the famous Eddystone Lighthouse in Great Britain, claimed to be the first civil engineer (<http://whatiscivilengineering.csce.ca/civill.htm>). With a long history of over 300 years, any academic institution in the world that offers a degree in engineering tends to offer a degree in civil engineering. An insight from the analysis of civil engineering articles will thus be beneficial to a large number of scholars, practitioners, and fellow students in the field.

This study aims to provide a thorough description of the Introductions of civil engineering research articles. A corpus of 60 Introductions taken from top quality civil engineering journals was compiled and subsequently analyzed by Swales' (2004) theoretical framework of genre analysis. The objectives of the present study are twofold: 1) to identify the textual move structure in the Introductions of civil engineering journal articles, and 2) to provide the linguistic characterization of each individual move and step identified. This study elucidates how the Introductions of civil engineering journal articles are constructed, allowing scientists and practitioners in the field to effectively share and update their scientific knowledge, contributing to the growth of the discourse community at large.

2. Swales' analytical framework and linguistic features

Genre analysis was developed by Swales in 1981 to describe the rhetorical organizational patterns of the research article genre. According to Swales (1990), the basic notion of genre analysis is that discourse communities are powerful in shaping the conventions of the genre. For instance, research articles, one of the most important academic genres, represent a good example of such discourse communities where novice writers are indoctrinated into the paper-writing genre in their advanced studies.

The objective of this discursal approach of genre analysis is to describe the communicative purposes of a text by categorizing the various discourse units within the text according to their communicative functions, a.k.a. moves (Swales, 1990, 2004). A 'move' refers to a text segment that performs a specific communicative function (Connor, Upton, & Kanoksilapatham, 2007: 17). In genre analysis, texts typically consist of a series of moves, with moves being functional units contributing to the overall communicative purpose of the genre. Since a move is semantically determined, the size of the move is not syntactically bound. That is, it can vary

ranging from a phrase, a sentence, a paragraph, to multiple paragraphs. But at least, a move contains a proposition. Not all moves occur equally frequently. Some moves occur more frequently than others, and they can be described as obligatory and optional, respectively. A rhetorical move can, in turn, consist of a number of elements that together, or in some combination, contribute to the same function of the move to which they belong. These elements are referred to as ‘steps’ by Swales (1990) or ‘strategy’ by Bhatia (1993).

Swales (1981) analyzed research article Introductions written in English from a wide range of disciplines including, physics, medicine, and social sciences; he observed that although taken from different academic disciplines, these Introductions displayed a common series of moves. In 1990, Swales proposed a three-move schema for article Introductions, the so-called Create a Research Space or CARS model. In a nutshell, the model depicts the preferred sequences of move types and steps which are found in research article Introductions from a wide range of disciplines. Swales’ CARS model for research article Introductions has been widely studied and validated by subsequent studies in other academic disciplines, e.g., medicine (Nwogu, 1997); computer sciences (Posteguillo, 1999) and professional genres (e.g., promotional letter genre (Bhatia, 1993). Focusing specifically on research article Introductions, Swales, incorporating the findings generated from these genre-based studies conducted during the last two decades, modified his 1990 model to accurately reflect the variations of Introductions in diverse research fields (Swales, 2004), as shown in Figure 1:

Figure 1. Swales’ revised model for research article Introductions (2004: 230, 232)

Move 1: Establishing a territory (citations required)* via**

Topic generalizations of increasing specificity

Move 2: Establishing a niche (citations possible)* via**

Step 1A: Indicating a gap or

Step 1B: Adding to what is known

Step 2: Presenting positive justification*

Move 3: Presenting the present work via

Step 1: Announcing present research descriptively and/or purposively (obligatory)

Step 2: Presenting research questions or hypotheses*

Step 3: Definitional clarifications*

Step 4: Summarizing methods*

Step 5: Announcing principal outcomes**

Step 6: Stating the value of the present research**

Step 7: Outlining the structure of the paper**

* Optional and less fixed in order

** Probable in some academic disciplines

*** Possible cyclical patterning of moves particularly in longer Introductions

As shown above, Swales' (2004) model for Introductions in general consists of three moves. They are: 'Move 1: Establishing a territory,' 'Move 2: Establishing a niche,' and 'Move 3: Presenting the present work.' Each move, in turn, consists of a number of steps. At this juncture, it should be noted that there are two additional characteristics of moves in this revised model. First, the status of a move is specified based on its frequency of occurrence, either obligatory or optional. Second, a cyclical fashion of moves within Introductions is determined. By extension and corresponding to Swales, the model proposed is not to prescribe, but to provide guidelines with regard to how ideas are marshalled when writing Introductions.

Rhetorical organizations of texts are helpful, providing the schema of what informational elements or moves are presented. A pertinent question that emerges is: What are the typical linguistic features used to express these moves? In this regard, Swales (1990: 131-132) features the initial attempt to summarize the findings of 40 published studies which described the use of linguistic features in the four major sections of research articles. He concluded that, for instance, Introductions and Discussions have the functions of providing the background of the current study and interpretation of the results, respectively. The common linguistic features associated with these functions of the two sections are *that* complements, use of the present tense, and author's comments.

Swales' (1990) observation of typical linguistic features expressing each communicative function was illuminating, representing an initial attempt to linguistically characterize the language of particular sections of research articles. Meanwhile, his observation implies that a set of linguistic features is used to express a communicative function. However, each of the linguistic features identified by Swales might be considered too crude or a general category. For instance, the feature

of 'that' complement clauses, according to Biber et al. (2007: 267-271), embraces a number of sub-categories including 'that' clauses controlled by a verb, 'that' clauses controlled by an adjective, and 'that' clauses controlled by a noun. In turn, each subcategory can be further differentiated. For instance, a verb can be further categorized as non-factive verbs (e.g., *imply*, *resort*, *suggest*), attitudinal verbs (e.g., *anticipate*, *expect*, *prefer*), factive verbs (e.g., *demonstrate*, *realize*, *show*), and likelihood verbs (e.g., *appear*, *hypothesize*, *predict*). Therefore, the availability of a finer set of lexical and grammatical features would be helpful not only to accurately describe Introduction texts but also to help learners and practitioners in different academic disciplines to choose the appropriate set of linguistic features in crafting their research articles.

Biber's (1988) tagger of linguistic features provides a headstart in this direction. With the continued development of computational programs and taggers, a more elaborate and thorough list of linguistic features has been developed (e.g., Biber, 1988; Biber et al., 2002, 2007). For instance, thanks to the Biber et al. (2002) tagger, Kanoksilapatham (2003, 2007a) demonstrated that scientific texts in biochemistry can be analyzed, in addition to the top-down approach like Swales' genre analysis, by the bottom-up approach like multi-dimensional analysis (for more information, refer to Biber, 1988). The frequency counts of 80 linguistic features were fed into multidimensional analysis, yielding a detailed and accurate linguistic characterization of the biochemistry texts. For example, focusing on biochemistry research articles, Kanoksilapatham (2007a) revealed that in the Introductions, the instances representing 'Move 1: Establishing a territory,' on Dimension 5: Attributed knowledge vs Current study, contain frequent occurrences of present tense verbs, references, and type-token ratio (ratio between the number of different lexical items in a text and the total number of words in that text). In contrast, the instances representing 'Move 2: Establishing a niche,' on Dimension 7: Contradictory proposition, contain high frequent occurrences of concessive markers (e.g., *however*, *nevertheless*, *despite*), pointers, *not* negation, and adverbs.

Studies on the same academic genre of research article Introductions across disciplines reveal that discrepancies in the textual structure of Introductions are discernible (e.g., Kanoksilapatham, 2007b; Ozturk, 2007; Samraj, 2002). This finding implies that even though Introductions of research articles seem, to a certain extent, to follow the model outlined, each discipline has its own preferred rhetorical pattern

which might be identical to, or vary from, the model. In this regard, given the prominent role of civil engineering in the field of engineering, it is interesting to examine the characteristics of research article Introductions in this academic discipline to obtain top-down knowledge. In contingency to this premise, a set of linguistic features identified will equip scholars and practitioners in civil engineering with bottom-up knowledge. Pedagogically, these bodies of knowledge will be useful in training scholars in a focused discipline to write successfully and internationally to disseminate their research discoveries.

This study attempts to qualitatively examine civil engineering research article Introductions written in English to bring into focus the most appropriate academic style of writing in this discipline and genre. The two objectives of the present study are: 1) to identify moves and sub-moves or steps, their occurrence frequency, and their typical sequence, and 2) to examine the linguistic characterization of these moves and steps in this specific discipline.

3. Methodology

3.1. Dataset compilation

A dataset of research articles in civil engineering was compiled. To assure that the Introductions analysed in this study represent the high quality journals in the discipline, the impact factors of the journals published in English on experimental research in the year 2005 (the latest impact factors available at the time of the study) were examined. At this juncture, it should be noted that, recently, the use of impact factors, though pervasively used in sciences, has been challenged, particularly in the humanities. For more details on this, please refer to Rogers, Campbell, Louhiala-Salminen, and Rentz (2007: 413). The five journals with the highest impact factors in civil engineering were 1) *The Journal of Hydraulic Engineering (JHA)*, 2) *The Journal of Water Resources Planning and Management (JWR)*, 3) *Earthquake Spectra (ESA)*, 4) *The Journal of Hydrologic Engineering (JHE)*, and 5) *Transactions on Intelligent Transportation Systems (INT)*. Then, twelve research articles published in each of these five journals during the year 2006 were randomly selected. In short, the dataset compiled for this study consists of 60 civil engineering Introductions. All of the 60 Introductions were referred to, for instance, as JHA1 to JHA12 according to the actual order of their publication.

3.2. Data analysis

Two main stages in analyzing data were used in the present study to address the two objectives. First, following Swales' (2004) theoretical framework, the Introduction dataset was subjected to genre analysis to identify moves and to examine their rhetorical schema. With relevance to this objective, the typical sequence of moves was identified. In addition, the frequencies of the moves identified were recorded to indicate how conventional/optional the moves are. Steps that belong to the same move represent possible options that writers have at their disposal to contribute to the major function of the move. The frequencies of the steps found in each move were also reported.

Second, individual moves and steps identified in the first stage were complemented with detailed linguistic characterization by identifying the lexico-grammatical regularities in each move and step. To address the second objective of the study, Biber's original list of 67 linguistic features developed in 1988 to characterize linguistic variation among a variety of text types provides an initial impetus for this study. Since 1988, the tagger of linguistic features has been incessantly developed, and the tagger has successfully captured major grammatical and lexico-grammatical features in a variety of texts (e.g., Atkinson, 2001; Biber & Finegan, 1994; Conrad, 2001; Reppen, 2001). Nevertheless, given the specificity of the current dataset of the single genre of research articles, the list cannot be exhaustive. That is, some features that are prominent in this single genre might not be captured by Biber's tagger. Therefore, to accommodate the specificity of the corpora and to accurately characterize the texts, additional features might be required. For instance, the *reference* feature, referring to citations (no matter how long the citations are), has been found to be prominent in a number of genres (e.g., Dong, 1966; Hyland, 1999; Kanoksilapatham, 2007a). This feature demonstrates the crucial role of citations in science, for instance, where current scientists need to be aware of relevant previous research and make connections with them. Given the consensus that, in academic settings, previous studies must be acknowledged and credited, this *reference* feature is considered essential in academic texts. The significance of references is clearly substantiated in Swales' (2004) model (Figure 1), in which Swales asserted that citations or references are prevalently used in Move 1 and Move 2. This study does not attempt to use the tagger to quantitatively analyze the linguistic characterization of individual moves. Instead, however, based on the

categories corresponding to those in Biber's (1988) tagger and some additional features mentioned above, this study aims to quantitatively characterize the regularity of the co-occurrence of the linguistic features within individual moves. Given the specific focus on Introductions of this study and consequently the limited size of the dataset and limited number of linguistic features prevalent in texts, quantitative analysis of the linguistic features like multidimensional analysis would not be feasible. For more details on certain requirements of multidimensional analysis, refer to Biber (1988) and Kanoksilapatham (2003).

4. Results and discussion

4.1. Textual structure of civil engineering Introductions

4.1.1. Moves and steps

The results obtained from the application of genre analysis reveal that the Introductions of civil engineering articles, compatible with Swales' model, consist of the same set of three main moves: 'Move 1: Establishing a territory,' 'Move 2: Establishing a niche,' and 'Move 3: Introducing the present work.' The rhetorical pattern of civil engineering Introductions is delineated as follows:

Figure 2. Proposed model for civil engineering Introductions

Move 1: Establishing a territory* via	100.00%
(Topic generalizations of increasing specificity)	
Step 1: Claiming centrality	48.33%
Step 2: Making topic generalization(s)	96.66%
Step 3: Reviewing previous studies	93.33%
Move 2: Establishing a niche* via	71.67%
Indicating a gap	71.67%
Move 3: Introducing the present work* via	100.00%
Step 1: Announcing present research purposively	78.33%
Step 2: Summarizing methods	66.66%
Step 3: Announcing principal outcomes	45.00%
Step 4: Stating research values	38.33%
Step 5: Outlining the structure of the paper	28.33%
Step 6: Justifying procedural decisions	20.00%

Step 7: Describing study sites

15.00%

* Possible cyclical patterning of moves particularly in longer Introductions

As shown, the model generated by this study is similar to Swales' (2004) model at the move level, but not at the step level. It is noted that 'Move 1: Establishing a territory' identified in this civil engineering dataset is different from that in Swales' (2004) model. In this regard, the current model resembles Swales' (1990) model, in which Move 1 features a possibility of three steps. Swales' modification of his 1990 model was triggered by the fact that Swales (2004) and other genre analysts including, for instance, Samraj (2002), argue that the three steps identified in Move 1 are not easy to differentiate, especially from the analysis perspective. As such, Swales (2004) proposes to eliminate the steps of Move 1 and suggests that the accounts of this move are of increasing specificity. Even though this study substantiates Swales' (2004) observation about the nature of Move 1 whose contents feature topic generalization of increasing specificity, from the perspective of pedagogical implications, the identification of the three steps provides novice engineers and practitioners leeway and a more concrete guideline to construct Move 1 in a way that conforms to the conventions of the field.

The current model also deviates from Swales' (2004) model in the use of 'Move 2: Establishing a niche.' According to Swales, this single move can be manifested through two strategies or steps. They are: Step 1: Indicating a gap or adding to what is known, and Step 2: Presenting positive justification. In contrast to Swales' model, only one step of 'indicating a gap' is used in this dataset.

The analysis in general shows that most of the steps found in 'Move 3: Introducing the present work' of civil engineering are closest to Swales' (2004) model, except for two major deviations. First, Swales' steps of 'Presenting research questions or hypotheses' and 'Definitional clarifications' are not found in this dataset of civil engineering. Second, two additional steps of 'Justifying procedural decisions' and 'Describing study sites' are identified in this dataset. The step of 'Justifying procedural decisions' is to convince readers that the procedures employed are well thought-out and carefully selected. To avoid any possible doubt or questions on the methodological issues of the study, this step is used 12 out of 60 Introductions (20.00%) in civil engineering Introductions. Finally, the step of 'Describing study

sites' is deemed necessary in civil engineering (found in 9 Introductions or 15.00% of the entire dataset) because this discipline focuses on harnessing natural forces to improve and protect environments. To enable readers to appreciate the research accounts, a description of the sites is crucial. At this juncture, it should be noted that instances of moves and steps are presented and further scrutinized in Section 4.2.

4.1.2. Move frequency and pattern

As suggested by Swales (2004) and previous genre-based studies (e.g., Brett, 1994; Kanoksilapatham, 2005), the frequencies of occurrence of moves play a critical role in determining the status of the moves. As for the frequencies of occurrence of individual moves, the findings reveal that the three moves occur relatively often. 'Move 1: Establishing a territory' and 'Move 3: Introducing the present work' are the most frequent, occurring in every Introduction (100% of analyzed dataset). 'Move 2: Establishing a niche' is found quite frequently (71.67%), but not as frequent as the other two moves. Despite variations in their frequencies, all of the three moves are considered obligatory and prominent in performing their communicative functions in Introductions.

Based on the results of the move analysis of the 60 civil engineering Introductions, some observations about textual structure can be made with reference to move sequence and cyclical patterning. First, the scrutiny of the move sequence demonstrates that civil engineers vary in their construction of their research article Introductions. A number of move patterns are found but in a small number of Introductions (e.g., the move sequences of 1-3-2, 3-1-2, 3-1, and 2-3-1, in 7, 2, 1, and 1 Introductions, respectively), indicating possible and expected idiosyncrasy of individual writers or groups of writers. Despite minimal variations, the most preferred move sequence is 1-2-3 which is found in 33 Introductions or 55% of the entire dataset. The other move structure which is relatively frequent is 1-3, in 16 Introductions or 26.67%. The most prevalent move structure of civil engineering research article Introductions or the move sequence of 1-2-3 is congruent with those in Swales' model and studies in other disciplines, e.g. humanities (Fakhri, 2004); biochemistry and microbiology (Kanoksilapatham, 2007a, 2007b); applied linguistics (Ozturk, 2007); and wildlife behavior and conservation biology (Samraj, 2002). However, variations across academic disciplines are prevalent, especially in the use of steps. For instance, Samraj (2002) compared the move structures followed in

conservation biology and wildlife behavior Introductions ($N = 24$, 12 from each discipline). She found that the conservation biology Introductions preferred to use the step of centrality claims or Move 1, Step 2 (11 out of 12 Introductions) more frequently than the wildlife behavior Introductions (6 out of 12 Introductions).

In summary, genre analysis applied to civil engineering research article Introductions is elucidating. First, it reveals the construction of texts with regard to their subcomponents or move/step schema. That is 'Move 1: Establishing a territory' can be realized by the use of three steps, whereas 'Move 3: Introducing the present work' by a possible combination of up to seven steps. In contrast, 'Move 2: Establishing a niche' is realized with one possible step in civil engineering. Second, as for the frequencies of occurrence of these moves, Move 1 and Move 3 are invariably present, whereas Move 2 is frequent with 71.67% occurrence. Third, with regard to their typical sequence, the analysis also shows that Move 1 is typically the initial move and Move 3 is the final move. Cyclical patterning is prevalent in this dataset. In fact, all of the three moves are cyclic. Move 1 is the most cyclic, recurring in 52 out of 60 Introductions or 86.67%. Move 2 and Move 3 are cyclic, though to a lesser extent compared with Move 1: Move 2 in 11 Introductions or 18.33%, and Move 3 in 33 Introductions or 55%.

4.2. Linguistic characterization of civil engineering Introductions

The scrutiny of all the instances of each move makes it possible to describe the typical linguistic characterization of each move type, which is the second objective of this study. The following sections illustrate excerpts representing each individual move and step previously identified by genre analysis. Even though the communicative function of each move and step is explicit and self-explanatory from its move label, a brief description of its function is presented. All excerpts illustrating each move and step are taken directly from the dataset, with two modifications. First, citations in the excerpts were replaced by (R). Second, linguistic features that are found to be commonly used across text instances, be they grammatical or lexical, are highlighted in bold.

4.2.1. Move 1: Establishing a territory

The first move, establishing a territory, occurs in all Introductions. It is a typical opening move for the majority of civil engineering Introductions. The communicative

function of this first move type is to introduce the topic of the study. As Swales (2004) noted, this move usually begins with topic generalizations and ends with topic specificities. On this note, Swales' (1990) model seems to be more directly applicable to the textual structure of this dataset, featuring three steps of this move. They are namely, 'Step 1: Claiming centrality,' 'Step 2: Making topic generalization(s),' and 'Step 3: Reviewing previous studies.'

'Move 1, Step 1: Claiming centrality' is illustrated by the following instances. This step is for the author to announce their interest or the importance of the research topic.

- (1) *These instruments **are** now **used extensively** for measuring discharge in rivers throughout the world. [JHE3]*
- (2) *Spectral **acceleration**, SA, **is the most commonly used** intensity measure in practice today for analysis of buildings. [ESA4]*
- (3) ***Erosion** by different forms of jets **is perhaps one of the most studied** scour related problems (R). [JWR6]*
- (4) *The dynamics of internal waves in stratified reservoirs **have been extensively studied** in numerous publications (R). [JHA11]*
- (5) ***Prediction** of solute mixing occurring along a stream **is important** in many applications. [INT12]*
- (6) *Accurate **information** about rate of flow in rivers **is important** for a variety of hydrologic **applications** such as water resources planning and conveyance structures. [JHA5]*

These instances display the use of grammatical features like the non-past tenses (simple present tense, present perfect tense). The choice of active or passive voice is likely to be determined by the semantic categories of verbs. That is, the passive voice usually co-occurs with activity verbs (*used, studied*). Other prominent lexical categories of this step include amplifiers (*extensively, commonly*) and evaluative adjectives (*important, most*). Noticeably, the evaluative adjectives are used as a predicative adjective. Another lexical category includes nominalization or gerund which is formed from the conversion of verbs and adjectives into nouns to refer to abstract concepts (*erode-erosion, accelerate-acceleration, predict-prediction, inform-information, apply-application*). The co-occurrence of these features helps

contribute to highlighting the importance of the topic and on-going interest to investigate a particular concept or area of civil engineering.

‘Move 1, Step 2: Making topic generalization(s)’ represents a statement about the state of the art of a research topic, as illustrated by the following examples.

- (7) ...the **locally averaged instantaneous critical stress** is considered to be a property of the uniform sand particles. [JHA2]
- (8) Site response **includes** the effect of the uppermost several hundred meters of rock and soil and the **surface topography** at the recording site. [ESA1]
- (9) The **headspaces** of sanitary sewer systems are normally prone to pressurization and ventilation. [JHE8]
- (10) Scour of a **sediment bed** around a vertical **cylinder** involves a complex flow field marked by large-scale turbulence structures generated by flow around the pier or abutment.... [JHA1]
- (11) **Acoustic Doppler current profilers (aDcp)** have become standard **instruments** for measurement of spatially distributed water velocities in both freshwater and marine environments. [JWR5]

The use of the non-past tenses (simple present tense, present perfect tense) is prominent in Move 1, Step 2. Similarly, the choice of present simple or present perfect seems to correspond to particular semantic categories of the verbs. For instance, aspectual verbs (*become*) co-occur with the present perfect as shown in (11); whereas in other instances, existence verbs (*involve*, *include*) tend to co-occur with the simple present tense as shown in (8) and (9). In contrast to Step 1, the nouns used in Step 2 are not abstract ideas, but technical and concrete nouns (*acoustic Doppler current profilers*, *sediment bed*, *cylinder*, *instruments*, *headspaces*). These features co-occur to highlight the current state of the art, of knowledge, and of technique. The current accounts of the topic are comparatively more specific (as evidenced by concrete nouns), as opposed to Move 1, Step 1, which is relatively more general (as indicated by frequent use of abstract nouns).

‘Move 1, Step 3: Reviewing previous studies’ is the last step of this move. Through this step, selected previous research is reviewed, as shown by the following examples.

- (12) Several **methods have been proposed**, for a review see (**R**). [ESA1]
- (13) In **recent** years, artificial neural networks (ANNS) **have proved** to be a better alternative for modeling complex and nonlinear processes (**R**). [JHE11]

- (14) *Local scour has been studied primarily by means of laboratory flume experiments (R).* [JHE1]
- (15) *Many studies have been conducted with plane wall jets that interact with non-cohesive sand beds and related details can be found in (R).* [JHA6]
- (16) *Experiments have been performed by, among others, (R), all of which report the impossibility of adequate rough boundary measurements.* [JWR2]

Interestingly, these instances show the distinct use of the present perfect tense, which usually co-occurs with passive voice and activity verbs (*studied, conducted, performed, proved*). A unique feature of this step is the use of citations or reference (R), the feature previously found to be significant in scientific writing (e.g., Dong, 1996, Hyland, 1999, Kanoksilapatham, 2003). It is a tradition in any academic discipline that previous studies, to a certain extent, have provided a foundation for subsequent scholars to build knowledge on. As such, acknowledging the accomplishment and importance of previous research is a strongly encouraged practice. Finally, in congruence with the reference feature, the instances illustrated above reveal the frequent use of common nouns to refer to specific studies or research work (*methods, studies, experiments*). Through this step, civil engineers demonstrate their familiarity with relevant work in the field, build rapport and consensus within the discourse community, and express attribution to previous contributors in the field. In short, to successfully establish the research territory, civil engineers convince the readers by claiming the significance of the research area, by introducing the knowledge of the topic, and by showing familiarity with the topic.

4.2.2. Move 2: Establishing a niche

This move serves the function of identifying the more specific areas of research that require further investigation. In contrast to Swales' (2004) model, which outlines two possible steps of this move, Move 2 of civil engineering research article Introductions is relatively much more homogenous, featuring only the use of 'Step: Indicating a gap,' as shown in the following examples.

- (17) *Despite the numerous studies on hydrologic applications of ANNs, only a few studies have addressed the problem of predicting hysteretic looping stage–discharge relations.* [INT5]
- (18) *Unfortunately, the stage–discharge relationship is not always a simple unique relationship.* [JHA5]

- (19) *The main **drawbacks** of this approach, though, are the **inability** to preserve the monthly correlation between the first month of a year and last month of the previous year. [JHE5]*
- (20) *To date, there **has been little** study of the effect of basin-scale internal waves on the dynamics of selective withdrawal. [JHA11]*
- (21) *No experiments involving the systematic variation of the deflection angle **have been carried out** to date, and therefore it **is not** known how the flow pattern varies. [JWR10]*
- (22) ***However**, the logistics of collecting direct measurement of streamflow on a continuous basis **is costly**.... [JHE5]*
- (23) ***Despite** the numerous studies on hydrologic applications of ANNs, **only a few** studies **have addressed** the problem of predicting hysteretic looping stage–discharge relations. [JHA5]*
- (24) *This model however assumes that there **is no** wastewater in the sewer conduits. This supposition is **a handicap** for an accurate prediction of ventilation dynamics of an operating sewer system. [JHA8]*
- (25) *The writers suggest that laboratory-based equations used for predicting local scour depths at cylinders **inadequately take** into account the similitude considerations [JHA1]*
- (26) ***but** one substantial similitude effect **seems** to have been **overlooked** in prior studies. [JHA1]*

Move 2 performs a communicative function of indicating a gap in previous research. To accomplish its function, the non-past tenses (simple present tense, present perfect tense) are commonly used. The lexical entries used in this move are quite diverse, involving a large set of lexical words ranging from attitudinal verbs (*overlook*), attitudinal nouns (*drawbacks*, *handicap*, *inability*), evaluative adjectives (*costly*), evaluative adverbs (*inadequately*, *unfortunately*), negation devices (*little*, *not*, *no*, *only a few*), and contradiction connectors (*however*, *despite*, *but*, *yet*). Despite their diverse syntactic categories, these lexical words, to a certain extent, share some common semantic category of negative evaluation. This is in congruent with the purpose of this move of identifying gaps in previous studies, pinpointing how inadequate they are. Through these features, civil engineers successfully construct critical comments based on the scrutiny of previous research. At the same time, this

move provides a transition for civil engineers to highlight their own research work in Move 3.

4.2.3. Move 3: Introducing the present work

While Move 1 and Move 2 contextualize the current study by establishing the significance of the topic and providing accounts of previous research related to the topic, Move 3 focuses on the study being reported. Swales' (2004) model stipulates a possibility of seven steps for Move 3. However, the analysis of civil engineering research articles Introductions reveals that this move type can be realized by seven steps. The two steps delineated in Swales (2004), but not found in this dataset, are the steps of 'Presenting research questions or hypotheses' and 'Definitional clarifications.' The absence of the step 'Presenting research questions or hypotheses' indicates that while scholars from other academic disciplines tend to use either the step of 'Presenting research questions or hypotheses' or the step of 'Announcing present research purposively' as stipulated by Swales (2004), civil engineers possibly prefer to address the goal of the study by 'Announcing present research purposively' (with the frequency of occurrence of 78.33%). Similarly, the absence of the step of 'Definitional clarifications' in this civil engineering dataset, as substantiated by Connor, Upton, and Kanoksilapatham (2007), is usually not found in research article Introductions but in PhD thesis Introductions. The following sections discuss individual steps of Move 3.

'Move 3, Step 1: Announcing present research purposively' typically announces the objective(s) of the research presented.

- (27) *The purpose of the experiments was to investigate the nearfield characteristics of the discharge for a range of port spacings encountered in outfall design practice.* [JHE4]
- (28) *In the present paper, a new comprehensive enumeration model is developed for this "intractable" layout problem.* [JHA9]
- (29) *The objective of this research is to present a new approach for calculating the overall space average and fluid bed shear stresses for uniform-sized sand particles.* [JHA2]
- (30) *The purpose of this paper is to assess the cost-effectiveness of various levels of seismic design and construction quality,* [ESA3]

- (31) **The objectives of the study were:** (1) to compare the accuracy of vDGPS from various DGPS units ...; and (2) to determine which DGPS correction procedures are acceptable [JWR3]
- (32) **The purpose of this paper is** to examine the utility of data-driven computational techniques to model and predict such complex relations that are encountered in low-gradient coastal streams. [INT5]

These instances of Move 3, Step 1 state explicitly the objective(s) of the research. The simple present tense and past tense of verb-to- be are used interchangeably in this step. Also commonly found are common nouns (*research, experiment(s), paper*), cognitive nouns (*purpose, objective(s)*), and deictic references to the present texts, including specific determiners and demonstrative adjectives (*this, the*). In fact, the combination of these grammatical and lexical features forms a typical formulaic expression delineated as follows:

	<i>objective(s)</i>			<i>research</i>		<i>is, are</i>
<i>The</i>	<i>purpose(s)</i>	+	<i>of the/this</i>	+	<i>paper</i>	+
	<i>aim(s)</i>			<i>experiment(s)</i>		<i>was, were</i>
				<i>study</i>		

‘Move 3, Step 2: Summarizing methods’ provides an account of the experimental procedures of the research. However, only major procedural features are selectively presented in the Introductions.

- (33) *Data from four different **DGPS** corrections **were collected** simultaneously during 22 repeated **aDcp** surveys of a single river cross section. The accuracy of the calculated boat velocity from various **DGPS** correction strategies (**vDGPS**) **was estimated** by comparison with boat velocity by bottom tracking (**vBT**) and the resulting influence on discharge measurement **was assessed**.* [ESA3]
- (34) *... the experiments **focused on** discharges driven primarily by buoyancy. The model diffuser **consisted** of T-shaped risers with horizontally opposed ports. Dilutions **were measured** by a precision **microconductivity probe** and the geometrical wastefield characteristics ... **were measured*** [JHA4]
- (35) **The tests were performed** with a single circular pile, two uniform diameter cohesionless **sediments**, and a range of water depths and flow velocities. All but four of the tests **were conducted** in the live-bed scour range. [JWR7]

- (36) *On a seasonal scale, **fluctuations in temperature and turbidity** of withdrawn water **are compared to** the internal wave dynamics. Steady-state formulas of selective withdrawal **are used to** obtain layer thickness in the reservoir where density stratification is unsteady due to internal wave activity; the results **are compared** to measured water quality data. The interaction of long internal waves and selective withdrawal **is studied** in detail on one particular seiching event, where the lake **hydrodynamics is assessed** using direct and inverse methods, followed by the interpretation of observed fluctuations of the water quality. [INT11]*
- (37) *In the current study, the method of **fragments-modified (R) (MFM)** disaggregation model and the nonparametric Markov order 1 model with long term persistence (NPL) **are applied** to generally monthly rainfall sequences [JHE5]*

The grammatical features used to express this step in the discipline of civil engineering include both the present simple tense and past tense. Lexical features include activity verbs (or to be precise research activity verbs) which are prevalent (*collect, estimate, assess, substantiate, perform, measure, compare, study, apply*) and are usually used in passive voice.

‘Move 3, Step 3: Announcing principal outcomes’ showcases significant findings generated by the research.

- (38) ***This paper presents the results** of 24 local sediment scour experiments conducted by the lead writer in the Hydraulics Laboratory at the University of Auckland in Auckland, New Zealand. [JHA7]*
- (39) *Generalized fragility curves are developed and a simple expression **is presented**, which is envisioned to be useful for design. [ESA10]*
- (40) ***The writers present insights** showing that two parameters...substantially **influence** equilibrium scour depth. [JHA1]*
- (41) ***In this paper** a system formulation incorporating system variables and major air energy inputs **is presented** for analyzing air movement in sanitary sewer systems. [JHA8]*
- (42) ***This technical note presents the results** of clear water local scour generated by 3D wall jets in a non-cohesive sand bed at low tailwater depths. [JHA6]*

The linguistic features used to express this step include the simple present tense together with a communication verb (*present*). Prevalent lexical features

include the use of the cognitive nouns (*results, insights*) and deictic elements (*the, this, this paper*) which are followed by common nouns (*paper, technical note*). All these features contribute to highlighting certain discoveries from the current study.

‘Move 3, Step 4: Stating the value of the present research’ claims the significance or contribution of the study being reported.

- (43) *This study brings some new understanding of the physical processes affecting dispersion in meandering channels with a loose bed.* [JHA12]
- (44) *Once validated, this approach can be used to design enhanced anaerobic bioremediation systems.* [JHE6]
- (45) *These models would be a substantial advancement in the field of traffic management due to their potential contribution towards traffic safety as well as freeway operations.* [INT2]
- (46) *This information will be used to comment on the validity of selected models found in the literature.* [JHA12]
- (47) *In the future this data may also be useful for the development and validation of new theoretical and numerical models.* [JWR12]
- (48) *...some implications regarding the prediction of the nature of bed geometry in sine-generated meandering streams are discussed.* [JHA10]

This step goes beyond scientific results, which are objectively generated by the study being reported. As a result, the simple future tense (*will*) or time adverbial indicating the future tense (*in the future*) is used. Other key linguistic features include cognitive nouns (*understanding, contribution*) and likelihood nouns (*implications*), and predictive and possibility modals (*will, would, can*). Deictic elements including specific determiners or demonstrative adjectives (*this, these*) followed by common nouns (*data, study, models*) are highly frequent in these texts. All these linguistic features co-occur to project a possible contribution of the research results.

‘Move 3, Step 5: Outlining the structure of the paper’ familiarizes readers with the structure of the research article being read. The information allows readers to know what to expect while reading the research article.

- (49) *We begin this paper with a brief introduction to relevant aDcp and DGPS theory. We then describe the various DGPS units tested, the field site, and data collection procedures. The analytical techniques developed to process the data are presented. The resulting accuracy estimates for single ping velocity*

and full transect discharge measurements are presented. Finally, the results are generalized through a simulation.... [JHA3]

- (50) *First, a description of the study site and the available stage and discharge measurements is presented. Then, the two modeling methods are briefly explained. Afterward, the setup and the results of each model are presented. The paper closes with conclusions and recommendations.* [JHA5]
- (51) *We begin by providing some background on the project, and then detail our analysis and findings.* [ESA3]

A prominent grammatical feature is the simple present tense. The choice of active or passive voice seems to be determined by the presence or absence of deictic elements (*the/this paper, we*). That is, with the presence of deictic elements, verbs are in active voice; without deictic elements, verbs are usually in passive voice. The presence of deictic elements also seems to determine the verb choices because they usually co-occur with activity verbs (*begin, close*). Otherwise, a communication verb (*present, explain, describe*) is very common. The most distinct lexical feature of this step is sequential connector (*first, then, afterward, finally*). By using these linguistic features, civil engineers are able to direct readers' attention to how a research article is outlined or structured.

'Move 3, Step 6: Justifying procedural decisions' allows civil engineers to assure article readers that a procedure adopted is carefully thought out and justified.

- (52) *We select significant duration because it is relatively stable with respect to the definitions of beginning and end thresholds (R).* [ESA12]
- (53) *DSPM was chosen because it can simulate snow accumulation and....* [JHE 4]
- (54) *To avoid any confusion, in this paper, we define and use the terms acceptable routes and service level in the stochastic scenario.* [INT3]

Common linguistic features found in these instances include cause connectors (*because*) and 'to' clauses controlled by a desire, intention, and decision verb (*avoid*). Both of these linguistic features share the same communicative function of introducing reasons or purposive propositions. The co-occurrence of either of these two features with deictic elements (*we, in this paper*) contributes to the communicative function of announcing the justifications for choosing a procedural step in the research in order to safeguard the authors from possible inquiry.

'Move 3, Step 7: Describing study sites' is a step that provides general information or description of the target site(s) for conducting research.

- (55) *The **study area** is located in **cold** regions. Snow accumulation and spring snowmelt are **important** factors that can cause **significant** spring floods.* [JHE4]
- (56) *Six **sites** have varying amounts of **soft**, clay material over possible Pleistocene till, over bedrock. The seventh site (PCG) is situated on competent bedrock.* [ESA1]
- (57) *The **test site** for this application is **the Snake River**, the **largest** tributary of **the Colombia River**.* [JWR4]

The scrutiny of these instances reveals that the grammatical feature frequently found include the simple present tense (*is, have*). Lexical features include the use of common nouns (*study area, sites, test sites*), attributive adjectives (*cold, soft, significant, important, largest*), and specific place names (*the Snake River, the Colombia River*). These features provide precise and accurate description of the sites.

This step can be claimed to represent a unique feature of the Introductions of civil engineering research articles because it is neither included in Swales' (2004) model, nor found in other disciplines such as biochemistry, microbiology, humanities, and social sciences (Kanoksilapatham, 2005, 2007b; Fakhri, 2004). A question that emerges and is worth investigating is why this step is uniquely found in this discipline. A look at the history of this discipline is enlightening. According to the American Society of Civil Engineers (<http://live.asce.org/hh/index.mxml>), civil engineering

... is the profession in which a knowledge of the mathematical and physical sciences gained by study, experience, and practice is applied with judgment to develop ways to utilize, economically, the materials and forces of nature for the progressive well-being of humanity in creating, improving, and protecting the environment, in providing facilities for community living, industry and transportation, and in providing structures for the use of humanity.

Based on the above definition, civil engineers plan, design, construct, and operate on facilities essential to modern life. Clearly, civil engineering research deals closely with the environment. As a result, description of the environment is an essential element to be presented in Introductions.

This section illustrates that each move and each step belonging to individual moves in research article Introductions perform their communicative functions through different co-occurrence patterns of linguistic features. In conclusion, addressing the two objectives put forward, the study has shown that the Introductions of civil engineering research articles can be analyzed into sub-units of moves and steps. The three moves have proved to be crucial and prevalent due to their high frequencies of occurrence. These moves seem to occur in a specific order or pattern. For some articles, certain moves can occur more than once. The examination of move/step instances also highlights the fact that each move or step seems to use a particular set of linguistic features to help achieve its communicative function.

5. Implications

The genre analysis of the Introductions of civil engineering research articles in English has important pedagogical implications for ESP instruction in general and for civil engineering in particular. As pioneering and influential work in English for Specific Purposes (ESP), Swales' (1981) genre analysis was designed to help non-native speakers of English to read and write research articles. Although the dataset of civil engineering research article Introductions is only a small part of academic English, it is hoped that the contribution of this research should not be limited to the field of civil engineering but should extend to include the scientific community at large.

The results generated from this research show how genre analysis can be applied in the teaching of ESP in resonance with Swales' impetus in creating genre analysis. The study identified the frequency of occurrence of each move and the typical organizational pattern provides civil engineers, especially novice civil engineers, with basic knowledge on how to start crafting their own Introductions.

In today's world, English is one of the most dominant modes of communication in all academic activities. It is thus necessary to prepare ESP learners to meet the needs and challenges of academic settings. In this respect, learners should be exposed to, and should engage in, a variety of academic genres to learn rhetorical variation, not only across genres but also across academic disciplines. They should also be guided to make an appropriate choice of rhetorical or linguistic features to suit the demands of the academic settings in which they are likely to operate. Once the learners can answer why a particular text is used and written the way it is (Bhatia,

1993, 1997), they will fully understand and be able to successfully construct a genre in the discourse community in which it is routinely used.

In the broader scope of the teaching and learning of ESP, this study demonstrates that ‘generic competence’ (Bhatia, 2004: 145) should be considered a primary goal to attain so that learners are equipped with suitable knowledge for communicating in various academic settings. According to Bhatia (2004), generic competence refers to “the ability to identify, construct, interpret, and successfully exploit a specific repertoire of professional disciplinary or workplace genres to participate in the daily activities and to achieve the goals of a specific academic/professional community” (p. 145). Generic competence will enable learners to gain understanding of the practices and conventions of particular communicative tasks embedded within academic and professional settings.

In this regard, the notion of conforming to the norms and conventions laid down by the discourse community should be adopted with great caution. Genre writers should be trained to be sensitive to different move structures of genre usage, especially in response to the demands of changing socio-cultural practices and evolving genres; that is, they should feel free to adopt or subvert the dominant practices as they wish. Additionally, they should be able to manipulate their genre knowledge to serve their own intentions and also meet the demands of today’s ever-changing and complex professional communities. As stated by Rentz (2005: 291), while scholars are concerned with incorporating gatekeepers’ feedback for publication purposes, they should try their best to maintain their “primary ownership”, contributing to the enhancement of research quality (Lave & Wenger, 1991). Although current trends do not seem to favor scholars who are non-English speakers, it is hoped that one day these non-English speakers will be able to get involved in “conversations”, negotiating the genre conventions as recommended by Bazerman (1980). Having achieved this, not only can the role of gatekeepers and researchers be balanced (Rentz, 2005: 91) but also the “primary ownership” be maintained (Rentz, 2005: 291), contributing to the enhancing of research quality (Lave & Wenger, 1991).

The learners of ESP should be guided to be genre explorers themselves. In a class, students may be assigned a piece of writing, for example, abstracts, the methods section, and acknowledgements. Their first task should be a small discussion in which students analyze the text and try to list the stages of those writings. Leading questions from the teachers may be “What is the writer trying to tell us?”, “What kind

of information is included in the writing?”, “Is there any sequence in which the story is written?”, “What kind of information is likely to be found in the beginning and at the end of the story?”, and “What are the typical linguistic features found in texts?”.

In class discussions, when the students are more familiar with a genre, the teachers should ask them to write a text of a particular genre and share it with the class. In this manner, the students are likely to gain a better understanding of genre knowledge rather than teaching them formulaic patterns. At the same time, students should be provided with an opportunity to compare texts across languages, cultures, and disciplines. The comparison will enable students to be aware of the volatile nature of each academic genre and its prevalent variations in order to enhance their sensitivities to different characteristics of various academic texts.

The language used in one academic discipline may be somewhat different from that in another discipline. Therefore, people involved should train novice civil engineers to be sensitive to certain conventions prevailing in their own discipline. Pedagogically, teachers or practitioners should bear in mind that the principal objective of teaching is to help prepare their students to communicate and disseminate their discovery successfully in their research articles. Students should be exposed to different language styles and guided to make appropriate language choice, by examining authentic research articles. Teachers should allow students to analyze texts by themselves. In so doing, students can learn conventions and variations revealed from texts, and so they can become more effective in manipulating other academic genres.

6. Conclusion

This study analyzed a representative dataset of Introductions of civil engineering research articles, using both top-down and bottom-up approaches to discourse analysis. A series of three moves in the Introductions contributes to the overall communicative purposes. These three moves are frequently but not always arranged in the order of Move 1, Move 2, and Move 3. Furthermore, they are cyclic. In addition to the textual organization, the study elucidates the linguistic characterization of each move and step identified. The insights obtained from both the rhetorical structure and the linguistic characterization will help form the pedagogical template for research article Introductions to develop coping strategies in response to the rigorous writing demands of academia encountered by advanced learners of English

in the civil engineering discourse communities. Students of civil engineering and practitioners who are interested in writing research articles can use the description of the textual organization and linguistic characterization as a guideline of how professional research article writers construct their Introductions. On this final note, caveats are in order. The findings are generated from the specialized dataset of civil engineering containing 60 Introductions, and thus may have little capacity for generalizations to be extended to the same genre of other academic disciplines in English and other languages and other genres in general. Finally, without interviews with civil engineers, especially those who are prolific article writers, it is difficult to deeply probe into the contexts of creating research articles.

Acknowledgement

The compilation and analysis of the civil engineering dataset is based upon work supported by the Thailand Research Fund and the Commission on Higher Education, Thailand under Grant No. RSA5080005. The author is also grateful for thought-provoking suggestions from two anonymous reviewers.

References

- American Society of Civil Engineers. (1961). History and heritage of civil engineering. Retrieved October 12, 2008 from <http://live.asce.org/hh/index.mxml>
- Atkinson, D. (2001). Scientific across history: A combined multidimensional/rhetorical analysis of the Philosophical Transactions of the Royal Society of London. In D. Biber & S. Conrad (Eds.), *Variation in English: Multi-dimensional studies* (pp.45-65). Harlow: Pearson.
- Bazerman, C. (1980). A relationship between reading and writing: The conventional model. *College English*, 41, 656-661.
- Bhatia, V. (1993). *Analyzing Genre: Language Use in Professional Settings*. London: Longman.
- Bhatia, V. (1997). Applied genre analysis and ESP. In T. Miller (Ed.), *Functional Approaches to Written Texts: Classroom Applications* (pp. 134-149). Washington, DC: USIA.

- Bhatia, V. (2004). *Worlds of Written Discourse: A Genre-Based View*. New York: Continuum.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. & Finegan, E. (1994). Intra-textual variation within medical research articles. In N. Oostdijk & P. de Haan (Eds.), *Corpus-Based Research into Language* (pp. 201-221). Amsterdam, Netherlands: Rodopi.
- Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A Multidimensional analysis. *TESOL Quarterly*, 36, 9-48.
- Biber, D., Connor, U., & Upton, T. (2007). *Discourse on the Move*. Amsterdam: John Benjamins.
- Brett, P. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, 13, 1, 47-59.
- Canagarajah, A.S. (1996). "Nondiscursive" requirements in academic publishing, material resources of periphery scholars, and the politics of knowledge production. *Written Communication*, 13, 4, 435-472.
- Carter, R. (1995). *Keywords in language and literacy*. London: Routledge.
- Civil engineering. Retrieved October 12, 2008 from <http://whaticivilengineering.csce.ca/civil1.htm>
- Connor, U., Upton, T., & Kanoksilapatham, B. (2007). Introduction to move analysis. In B. Biber, U. Connor, & T. Upton (eds.), *Discourse on the Move* (pp. 17-30). Amsterdam: John Benjamins.
- Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In S. Conrad & D. Biber (eds.), *Variation in English: Multi-dimensional Studies* (pp. 94-107). London: Longman.
- Cook, G. (1990). *Discourse*. Oxford: Oxford University Press.
- Dong, Yu-Ren. (1996). Learning how to use citations for knowledge transformations: Non native doctor students' dissertation writing in science. *Research in the Teaching of English*, 30, 4, 428-457.
- [Fakhri, A.](#) (2004). Rhetorical Properties of Arabic Research Article Introductions. *Journal of Pragmatics*, 36, 6, 1119-1138,

- Flowerdew, J. (2000). Discourse community, legitimate peripheral participation, and the nonnative-English –speaking scholar. *TESOL Quarterly*, 34, 1, 127-150.
- Flowerdew, J. (2001). Attitudes of journal editors to nonnative speaker contributions. *TESOL Quarterly*, 35, 1, 121-150.
- Flowerdew, J. (2007). The non-Anglophone scholar on the periphery of scholarly publication. *AILA Review*, 20, 5-13.
- Harword, N., & Hadly, G. (2004). Demystifying institutional practices: Critical pragmatism and the teaching of academic writing. *English for Specific Purposes*, 23, 355-377.
- Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20, 3, 341-356.
- Kanoksilapatham, B. (2003). A corpus-based investigation of biochemistry research articles: Linking move analysis with multidimensional analysis. Unpublished doctoral dissertation. Georgetown University, Washington DC.
- Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles *English for Specific Purposes*, 24, 3, 269-292.
- Kanoksilapatham, B. (2007a). Rhetorical moves in biochemistry research articles. In D. Biber, U. Connor & T. Upton (eds.), *Discourse on the Move* (pp. 73- 120). Amsterdam: John Benjamins.
- Kanoksilapatham, B. (2007b). Rhetorical organizations of research article introductions in biochemistry and microbiology. *ESP Malaysia*, 13, Dec, 21-37.
- Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.
- Martin, P.M. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, 22, 1, 25-43.
- McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.
- Nunn, R. & Adamson, J. (2007). Towards the development of international criteria for journal paper evaluation. *Asian EFL Journal*, 9, 4, 205-228. Retrieved July 11, 2008 from <http://www.asian-efl-journal.com/Dec 2007 m&ja.php>
- Nwogu, K. (1997). The medical research paper: Structure and functions. *English for Specific Purposes*, 16, 2, 119-138.

- Ozturk, I. (2007). The textual organization of research article Introductions in applied linguistics: Variability within a single discipline. *English for Specific Purposes*, 26, 1, 25-38.
- Posteguillo, S. (1999). The schematic structure of computer science research articles. *English for Specific Purposes*, 18, 2, 139-158.
- Reppen, R. (2001). Register variation in student and adult speech and writing. In D. Biber & S. Conrad (Eds.), *Variation in English: Multi-dimensional Studies*. New York: Longman.
- Rogers, P., Campbell, N, Louhiala-Salminen, L., & Rentz, K. (2007). The impact of perceptions of journal quality on business and management communication academics. *Journal of Business Communication*, 44, 4, 403-426.
- Samraj, B. (2002). Introductions in research articles: Variation across disciplines. *English for Specific Purposes*, 21, 1-17.
- Swales, J. (1981). *Aspects of Article Introductions*. Birmingham AL: University of Aston.
- Swales, J. (1990). *Genre analysis: English for Academic and Research Settings*. Cambridge: Cambridge University Press.
- Swales, J. (2004). *Research Genre: Explorations and Applications*. Cambridge: Cambridge University Press.
- Varttala, T. (2002). Hedging in Scientific Research Articles: A Cross-Disciplinary Study, Linguistic Insights. *Studies in Language and Communication*, 2, 141-174.



Title

English Workshop for Effective Teleconferencing

Author

Takashi Matsuzawa

Biodata

Takashi Matsuzawa received an M.Ed. in TESOL from Temple University. He retired from his company in 2009 and currently teaches Japanese in Shenzhen China. He can be contacted at matsuzaw@siren.ocn.ne.jp.

Abstract

This paper reports on the effectiveness of a one-day workshop developed for business people, aimed at improving English teleconference skills and at understanding the problems teleconference participants have. The workshop focuses on reduced forms for listening to natural speed English, conversation strategies for managing communication breakdown, and pragmatic hints and tips for a teleconferencing schema. According to the survey, the workshop was rated ‘useful’ by the attendees and two listening problems were highlighted as their main issues in teleconferencing: difficulty in following conversation speed and in recognizing pronunciations. Addressing those problems would improve the value of such a workshop.

Keywords: Teleconferencing; conference calls; reduced forms; conversation strategies; business; communication

1. Introduction

Today, audio teleconferencing (or conference calls) has become a key business tool for domestic and international communications. However, audio teleconferencing is difficult for non-native speakers of any language because listening and speaking are

the primary required skills even if some visual materials are available for reference. This report describes a one-day workshop built on a framework aiming at teaching basic skills required for English audio teleconferencing. The workshop was attended by 60 business people who had already participated in audio teleconferencing or might have to in the future. Also, a survey of the attendees was conducted to understand problems recognized while participating in teleconferences. The workshop was evaluated as 'useful' by the attendees and two listening problems were highlighted as common problems from the survey. This paper limits the discussion to audio teleconferences held in English, primarily General American English.

2. Workshop background

2.1. Teleconference

There are two types of teleconferences, video and audio only (conference calls). In reality, audio teleconferences are by far the most common due to the lower cost and ease of setting up and running them. Audio teleconferences only require telephone lines and the service or function to connect those lines into one.

In Japan, the business language used when dealing with other countries is commonly English and it is generally known that Japanese people are less skillful in listening/speaking than in reading/writing. Further, teleconference listening can be hampered by the line quality with various sound strengths and the existence of noise, echo, or cross-talk. Because of all these interferences it can be assumed that the challenge for Japanese participants in English teleconferences is not just a matter of improving listening and speaking. A skills framework is required to address the specific environment of English teleconferencing. However, the skills required for or issues involved in attending teleconferences held in foreign languages have not been adequately addressed in the literature.

There are sites on the Internet where some courses can be found that teach English teleconferencing skills to Japanese business persons, provided by professional development companies.¹ They advertise that they teach required skills for teleconferencing. However, on reading the course descriptions, they seem to primarily provide negotiation skills, not mentioning or addressing the fundamental issue that every dialogue occurs on the phone and seemingly not aware of the special problems this gives rise to.

As a TESOL course graduate, having witnessed many business people struggle with communication in English teleconferences, especially with listening, the author felt the need to develop a workshop for participants and potential future participants in English teleconferences. The aim was for the workshop attendees to understand what the key skills components are and use course materials as a reference for continual skills building as they gain further experience, not primarily to improve skills immediately during the workshop. The department manager approved this workshop to be held during business hours as one of the skills courses offered for the company employees.

The following will draw on suggestions from the literature and develop these into a framework for introducing exercises and a training program specifically aimed at addressing the issues faced by non-native speakers faced with conducting teleconferences in English.

2.2. Teleconferencing skills framework

The author selected three components for elaboration in the course: (a) *Reduced Forms* for improving listening to English spoken at natural speed and supporting bottom-up processing, (b) *Hints and Tips* for understanding the business teleconference environment supporting top-down processing, and (c) *Conversation Strategies* for managing potential communication breakdowns that are expected to occur more frequently than in face to face communications.

Regarding reduced forms, Matsuzawa (2006) studied the comprehension of Japanese business people of English reduced forms, such as assimilation, palatalization, contractions, deletion, and linking (Celce-Murcia, Brinton, & Goodwin, 1996). Matsuzawa (2006) reported general listening difficulties not only in reduced forms but also in weak forms of function words (e.g., *have* in the conversation often pronounced as *'ave* or *'ve* instead of in its strong (citation) form, making it difficult to identify), and giving instructions on these resulted in improvements across the board for people with different levels of English skills.

Regarding hints and tips, Peterson (1991) explained that content schemata (e.g., topic familiarity and previous experience of a field) help the top-down processing of listening comprehension. Tyler (2001) reported using Mental Models, where the availability of topic knowledge could decrease the need for relying on working memory, enabling the non-native speaker to attend to passages with lowered

cognitive load. If the teleconference participants know how to sign on in a teleconference, how the teleconference is conducted, and what the key communication considerations are beforehand, then he/she can better focus on the subject matter.

Regarding conversation strategies, during the teleconference, communication breakdowns can frequently occur because the listener relies only on what is being heard through the ears. For example, if there is an interval while listening, the listener should judge whether it is the end of the speaker's turn or a pause. So it becomes both the speaker's and listener's responsibilities to maintain the flow of communication. Thus, successful communication also requires that the teleconference participant is capable of using conversation strategies.

2.3. The workshop

Based on the framework explained above, the workshop was designed as follows:

Reduced Forms: Referencing Matsuzawa's (2006) study, this session aimed at providing knowledge and experience on reduced forms that are one of the key factors in understanding English spoken at natural speed. The main activity was listening to exercises of short reduced-form sentences either from recorded tapes or spoken by the author. Then the attendees were asked to fill in cloze exercises. Additionally, there were exercises where the attendees were required to predict how reduced forms could occur in the given sentences (e.g., *I miss your phone call*). Some exercises were adapted from Hagen (2000).

Hints and Tips: Selected items that had been important to the author's own experience of English-teleconference participation were stressed here and the selection rationale was objectively supported by theories wherever possible (see below). The emphasis was put on discussion considerations which are listed as a) through i) as follows:

Issues to note in discussions:

- a) Speak to the point (Sakamoto & Natsuoka, 1982)
- b) Be careful to use hedges, e.g., *maybe* (Hatch, 1992)
- c) Answering negative or tag questions
- d) Responding to unexpected requests

- e) Interruptions (Hatch, 1992)
- f) The speaker is unknown (don't be too polite) (Guest, 1998)
- g) Speaking in Japanese among Japanese attendees
- h) Exchanging thanks (Hatch, 1992)
- i) Expressing months by numbers (recent American trend)

These items were introduced in the short lectures with handouts but without exercises.

Conversation strategies: A Dornyei and Thurrell's (1994) set of conversation strategies were chosen for this purpose. The elements of the strategies were message adjustment/avoidance, paraphrase, approximation, appeal for help, asking for repetition, asking for clarification, interpretive summary, and the use of filters/hesitation devices. Exercises used were intended to let the attendees try and practice paraphrasing, approximation, and interpretive summarization in small groups of four to five people. For example, in a paraphrasing exercise (game), a group member picked up a face-down card on the table and had to explain the word on the card to the others without directly using the word. The words selected were *sightseeing, airplane, love, dolphin, and apple*.

As expressed by the use of the word *workshop* in the title, the course was built as exercise-oriented, as far as possible, so many exercises were included, keeping lecture to a minimum (except for the *Hints and Tips*) and providing handouts that would be available for future reference.

3. Aims of study

By conducting this workshop, and by receiving survey feedback from the attendees, the author wanted to (a) confirm the appropriateness and value of the workshop, and (b) get information on English skills issues attendees might have in participating in teleconferences, for future workshop improvement. Two months after the workshop, the attendees who had in the interval actually participated in teleconferences were surveyed again to (c) see the long term effect of the workshop and to investigate if the hands-on experiences had resulted in changes in their views.

4. Participants and the survey method

When the author announced the workshop in the department, enrollment in the workshop required that an applicant (a) currently participate in teleconferences for business or had the potential of participating in teleconferences in the future, and (b) had a TOEIC score of 600 or higher. However, the criterion of the TOEIC score was not strict. As long as the applicant's manager approved, enrollment was accepted. As a result, there were 60 attendees in five workshops held for various groups in the department in five months. Table 1 shows the details of the workshop attendees (hereafter students).

Table 1. Workshop attendees

Number of workshops held	5
Total attendees	60
Attendees having teleconferencing experience (subject to the analyses)	50 (men: 29, women: 21)
Minimum TOEIC score	410
Maximum TOEIC score	965
Average TOEIC score	721.0
Median TOEIC score	727.5

Note: TOEIC scores were for the attendees having teleconference experience only.

At the end of the workshop, each student was asked to fill in a survey on:

- (1) number of teleconferences participated in to date,
- (2) selecting from the predefined list (see Table 2), problems the student found when he/she participated in teleconferences (if a student could not find a problem in the list, to write it in) in priority order,
- (3) evaluations of the workshop's three parts (reduced forms, conversation strategies, and hints and tips) and the course overall, by a five point Likert scale, from *very useful* (5 points) to *very un-useful* (1 point), and
- (4) time allocated to each session.

However, the analysis of item (4) is not discussed in this paper. Then, two months after each workshop, questions (1) through (3) above were asked again to the students who had participated in teleconferences during the two month interval.

Table 2. Predefined problem list

-
- a) Speaking – difficulty in timely thinking of expressions
 - b) Speaking – lack of vocabulary
 - c) Speaking – lack of grammar knowledge
 - d) Speaking – formulating sentences first in Japanese
 - e) Speaking – difficulty in getting understood
 - f) Listening – difficulty in following conversation speed
 - g) Listening – difficulty in recognizing pronunciations
 - h) Listening – lack of vocabulary
 - i) Listening – trying to translate into Japanese
 - j) Lack of teleconferencing knowledge
 - k) Lack of business knowledge
 - l) Cultural differences – greetings and jokes, etc.
-

Note: The original list was in Japanese.

5. Findings

5.1. Right after the workshop

Surveys collected right after the workshop showed that students had various amounts of teleconference experience. Table 3 summarizes the responses.

Table 3. Number of teleconferences

Number of teleconferences participated in	Number of students	Group name (by the author)
1 – 10	16	Beginners
11 – 99	25	Regulars
100 - 350	9	Veterans

Note: The low numbers of teleconferences may be more accurate than the high numbers as each participant could remember each teleconference and count them better with fewer teleconferences. However, both low and high numbers were used as reported in the analyses in this study.

The participants were categorized into three groups based on the teleconference experience: beginners, regulars, and veterans for the detailed analyses in the later sections.

Regarding research question (a), the evaluation of the workshop was analyzed for the total 50 people as a single sample because the sample size is not large. The result is in Table 4.

Table 4. Evaluation: Right after the workshop

Number of Students	Evaluation			Overall
	Reduced Forms	Conv. Strategy	Hints & Tips	
50	4.60	4.52	4.52	4.60

Scales 1: very un-useful, 2: un-useful, 3: neither useful nor un-useful, 4: useful, 5 very useful

The average evaluation by the 50 students was 4.60, between *useful* and *very useful*. Then correlation analyses were conducted using the Microsoft Excel (2001)'s built-in correlation analysis function. The first analysis was about the correlation against the overall evaluation to see who saw the most value in the workshop. The correlation between the overall evaluations and the number of teleconferences participated in was $r = -.08$ so the correlation was negligible meaning that overall evaluation scores did not correlate with the students' past teleconference experience. However, due to the accuracy concerns related to the number of teleconferences participated in mentioned above, it may be too early to consider this conclusion final. Between the overall evaluation scores and the students' TOEIC scores, the correlation was $r = -.24$ so there was a low negative correlation implying that there was a tendency for students with lower TOEIC scores to report a higher overall evaluation. Next, analyses were conducted on the correlation of the workshop's three part scores to the overall evaluation scores to see what component most contributed to the overall evaluation. It turned out that *Reduced Forms* had $r = .74$ to the overall evaluation, *Conversation Strategy* $r = .65$, and *Hints and Tips* $r = .73$ suggesting that the *Reduced Forms* and *Hints and Tips* parts equally contributed to the overall evaluation.

Then, regarding research question (b), English problems reported in the survey were investigated. Table 5 shows the top five problems reported by the total and by each group of students.

Table 5. Reported problems by the total and each group: Right after the workshop

Total (N = 50)	
Listening – difficulty in following the conversation speed	66.0%
Listening – difficulty in recognizing pronunciations	58.0%
Speaking – difficulty in timely thinking of expressions	54.0%
Speaking – lack of vocabulary	48.0%
Listening – lack of vocabulary	24.0%

Beginners (1 – 10 times conference participations: N = 16)	
Listening – difficulty in recognizing pronunciations	68.8%
Listening – difficulty in following the conversation speed	62.5%
Speaking – difficulty in timely thinking of expressions	56.3%
Speaking – lack of vocabulary	43.8%
Listening – lack of vocabulary	37.5%

Regulars (11 – 99 times conference participations: N = 25)	
Speaking – difficulty in timely thinking of expressions	64.0%
Listening – difficulty in following the conversation speed	64.0%
Listening – difficulty in recognizing pronunciations	52.0%
Speaking – lack of vocabulary	44.0%
Speaking – formulating sentences first in Japanese	24.0%
Cultural differences – greetings and jokes, etc.	24.0%

Veterans (100 – 350 times conference participations: N = 9)	
Listening – difficulty in following the conversation speed	77.8%
Speaking – lack of vocabulary	66.7%
Listening – difficulty in recognizing pronunciations	55.6%
Speaking – formulating sentences first in Japanese	33.3%
Listening – trying to translate into Japanese	33.3%

Note: Each student was allowed to report up to five problems. The percentage shows how many students in the group listed a problem.

As shown, two listening problems (*difficulty in following conversation speed* and *difficulty in recognizing pronunciations*), one speaking problem (*difficulty in timely formulating of expressions*), and the *lack of vocabulary* in both listening and speaking were ranked as the top five problems. This tendency was common overall and across groups. However, there were some differences, for example, *Listening - lack of vocabulary*, did not appear in the regular and veteran groups' problem lists; the veteran group reported the use of Japanese for the cognition process of the communication.

5.2. Two months after the workshop

Research question (c) was then analyzed. This question was for the workshop attendees who actually participated in a conference within two months after the workshop. There were 14 respondents here. The overall evaluation of the workshop two months later was 4.07, slightly above *useful*. (Refer to Table 6 below.)

Table 6. Evaluations: Two months after the workshop

Number of Respondents	Evaluation			Overall
	Reduced Forms	Conv. Strategy	Hints & Tips	
14	3.93	4.07	3.79	4.07

Scales 1: very un-useful, 2: un-useful, 3: neither useful nor un-useful, 4: useful, 5 very useful

Then analyses were made on the correlations of the three parts to the overall evaluation. It turned out that *Hints and Tips* part contributed the most to the overall evaluation ($r = .85$). This would mean that the information of *Hints and Tips* had more practical value than was assumed in the workshop.

Table 7 shows the items the respondents now felt were problematic after actually participating in teleconferences after the workshop.

Table 7. Reported problems: Two months after the workshop

N=14	
Listening – difficulty in following the conversation speed	69.2%
Listening – difficulty in recognizing pronunciations	61.5%
Speaking – difficulty in timely thinking of expressions	53.8%
Listening – lack of vocabulary	53.8%
Speaking – lack of vocabulary	38.5%

Note: Each student was allowed to report up to five problems. The percentage shows how many students in the group listed it.

The reported problems remained the same as the ones right after the workshop; especially the top three problems which were in exactly the same order. This means that even if the respondents saw value in the workshop, they could not see rapid improvements in their teleconferencing skills, which were not targeted when the workshop was developed.

6. Discussion

For research question (a), the overall evaluation of 4.60 after the workshop can be interpreted to show that the workshop gave the students valuable information and support for English teleconferencing. And it could be said that the lower TOEIC score holders see the workshop as more beneficial than attendees with higher scores. However, after two months, in answering research question (c), the evaluation is 4.07

vs. 4.60 at the workshop. This lower evaluation may mean that the attendees see the value of learned skills, but, they also see the difficulties in applying those skills during real teleconferences. Yet the average evaluation of 4.07 stills indicates a *useful* level, improvements are required to sustain the level of initial evaluation.

For research question (b), the responses show that the problems the teleconference participants have are fairly common across the groups and the problems continue beyond two months after the workshop. Then, what can be done to improve this workshop or give workshop students suggestions for use in day-to-day business teleconferences that will help them continue to feel increasingly more successful here?

Two listening problems, *difficulty in following conversation speed* and *difficulty in recognizing pronunciations*, are at the top of the problem list. For these problems, support in improving listening comprehension specific to the teleconference would be the answer. Krashen (1996) has been advocating the importance of narrow listening and Caspino (2005) and Dupuy (1999) report the effectiveness of narrow listening. So, it may be useful advice to the student to try recording real teleconferences, with the permission of the other participants, and replay it afterwards until the discourse can be fully understood, or to ask a more fluent person explain what has been communicated. This approach would seem most valuable if teleconferences are held with native English speakers.

Concerning one speaking problem, *difficulty in timely thinking of expressions*, it would be beneficial to list high frequency phrases pertinent to each teleconference purpose, e.g., a review of project status, product sales status, financial business results, and so on. As an example, a quality circle named Hera-hera 5 (2003) at IBM Japan developed a booklet titled *Practical Expressions for Conference Calls: Better English with us!*, available for general reference on its Website, which was based on the experience of communications between the circle members and the US software development organization. The main body of the booklet was about commonly used phrases but it also covered various scenarios of teleconferences from start to finish.

Another aspect of improving teleconferencing skills can be the amount of vocabulary the participant has available to draw on, which also appeared as problematic across groups. Adding one component to the workshop suggesting vocabulary enrichment may help the students, for example, by creating and

maintaining a glossary based on each teleconference objective. There is a high frequency of domain specific vocabulary.

Considering the applicability of the predefined problem list provided in the survey, there are eight problems reported besides the predefined problems. Among them, three problems are the same, resulting in six different non-listed problems. (There are more problems given than the eight, which are judged as part of the predefined problems by the author.) So the predefined problem list could be said to cover and capture the main and common teleconferencing problems for these students.

It may be noteworthy though that three participants write the same problem of *difficulty to interrupt* while a teleconference is going on. At the *Hints and Tips* section of the workshop, the lecture covered the topic on the appropriate interruptions (turn taking) so that Japanese participants would not be considered rude. Still, three students have difficulty in interrupting and talking. Indeed, there are situations where a native English speaker continues speaking and there seems to be very little opportunity of breaking into a monologue and of contributing to the teleconference. It may be necessary to elaborate on the subject of *how to interrupt* while someone is speaking during a teleconference.

7. Conclusions

This workshop was developed based on business needs and the survey responses showed the workshop useful; suggesting that the teleconferencing skills framework has worked. On the other hand, the author acknowledges that this study cannot objectively capture what component or element was effectively acquired as teleconferencing skills or helped in teleconferencing. No doubt, the workshop participants are looking for further ways or suggestions on teleconferencing skills improvement. Based on the feedback analyses, some possible suggestions to the teleconference participants and improvements for the workshop are discussed. Accommodating these improvements would make the workshop more valuable to actual teleconferencing.

Footnote

1. For example, teleconferencing training courses are available at Canning Professional (www.canning.co.jp) and Kurdyla and Associates. (www.kurdyla.com).

Acknowledgements

My sincere thanks go to Torkil Christensen of Hokusei University, Sapporo, Japan for encouragement and continual support in writing this paper.

References

- Caspino, B. K. (2005). Support for narrow listening libraries. *The International Journal of Foreign Language Teaching*, 1(4), 2-8. Retrieved July 13, 2009, from <http://www.tprstories.com/ijflt/IJFLTFall05.pdf>
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- Dornyei, Z. & Thurrell, S. (1994). Teaching conversation skills intensively: Course content and rationale, *ELT Journal*, 48(1), 40-49.
- Dupuy, B. C. (1999). Narrow listening: An alternative way to develop and enhance \ listening comprehension in students of French as a foreign language. *System*, 27, 351-361.
- Guest, M. (1998). Spoken grammar: Easing the transitions, *The Language Teacher Online*, 22(6). Retrieved July 14, 2009, from <http://www.jalt-publications.org/tlt/files/98/jun/guest.html>.
- Hagen, S. A. (2000). *Sound advice: A basis for listening* (2nd ed.). White Plains: Pearson Education.
- Hatch, E. (1992). *Discourse and language education*. Cambridge: Cambridge University Press.
- Hera-hera 5. (2003). *Practical expressions for conference calls: Better English with us!* Tokyo: IBM Japan. Retrieved July 26, 2005, from <http://www-06.ibm.com/jp/ibm/cs/concepts/improve/pdf/tpdf.html>.
- Krashen, S. D. (1996). The case for narrow listening. *System*, 24(1), 97-100.
- Matsuzawa, T. (2006). Comprehension of English reduced forms by Japanese business people and the effectiveness of instruction. In J. D. Brown & K. Kondo-Brown (Eds.), *Perspectives on teaching connected speech to second language speakers* (pp. 59-66). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Microsoft Excel 2002 SP3 [Computer Software]. (2001). Tokyo: Microsoft.

- Peterson, P. W. (1991). A synthesis of methods for interactive listening. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 106-122). New York: Newbury House.
- Sakamoto, N. & Natsuoka, R. (1982). *Polite fictions: Why Japanese and Americans seem rude to each other* (pp. 72-78). Tokyo: Kinseido.
- Tyler, M. D. (2001). Resource Consumption as a function of topic knowledge in nonnative and native comprehension. *Language Learning*, 51, 257-280.



Title

The Structure of the Methods Section in Research Articles Across Eight Disciplines

Author

Matthew Peacock

City University of Hong Kong

Bio data

Matthew Peacock teaches students on a wide variety of programmes and courses in the Department of English. His current research interests include English for Specific Purposes, TESL theory and methodology, second language teacher education, and programme evaluation. He co-edited (with John Flowerdew) a 2001 collection from Cambridge University Press, *Research Perspectives on English for Academic Purposes*.

Abstract

There appears to have been little previous research into the structure of research article (RA) methods sections. This paper reports a corpus-based analysis of the communicative move structure of 288 RA methods sections across eight disciplines: physics, biology, chemistry, environmental science, business, language and linguistics, law, and public and social administration. Preliminary examination of the corpus found seven different moves, though not in all RAs or necessarily in this order: *overview*, *location*, *research aims/questions/hypotheses*, *subjects/materials*, *procedure*, *limitations*, and *data analysis*. A number of interdisciplinary differences were found in moves and move cycles, including clear differences between the sciences and other disciplines. The existence of move cycles in methods sections seems not to have been previously predicted.

Keywords: methods section; research article; genre analysis; research writing; communicative move

1. Introduction

Since Swales's (1981) discussion of communicative moves in research articles (RAs), a lot of attention has been paid to the move structure of introductions in academic writing, for example, Swales (2004) and Samraj (2005). However, little previous research has apparently examined methods sections, and there appears to be no model of its move structure.

The RA was chosen for this research because of its importance for the spread of academic knowledge. Berkenkotter & Huckin (1995) and Hyland (1996) call it the key medium for legitimating findings and disciplines. RAs are the preferred genre through which discourse communities communicate, and their unique language distinguishes different disciplines from each other (Williams, 1998). Corpus-based research is an increasingly popular tool for describing those codes (Williams, 2002).

A move is "a segment of text that is shaped and constrained by a specific communicative function" (Holmes, 1997, p. 325). A move cycle is a reoccurring move pair (or set). An example of a typical discussion section move cycle is *statement of result* □ *reference to previous research* (Dudley-Evans, 1994, p. 225). The standard method for classifying moves is the four-step procedure suggested by moves researchers Dudley-Evans (1994) and Holmes (1997, 2001), which was followed for the purposes of this research:

- (1) look for organisation and patterns and identify moves and boundaries
- (2) use sentence-level analysis
- (3) assign all sentences to a move
- (4) authenticate the classification by using two raters, who work independently and then compare their findings.

The aims of this study were to examine the move structure of RA methods sections and interdisciplinary variation, across eight disciplines: biology, chemistry, physics, environmental science, business, language and linguistics, law, and public and social administration. Our corpus included 288 RAs, 36 from each discipline.

2. The Move Structure of Methods Sections

While only a few authors seem to have investigated the move structure of RA methods sections, their results are important regarding our study because of their focus. Each author looked at a fairly small number of RAs from just one discipline. They are Brett (1994), 20 sociology RAs; Nwogu (1997), 15 medicine RAs; Lim (2006), 20 management; Kanoksilapatham (2005), 12 biochemistry; and Wood (1982), 10 chemistry. Results for the first three disciplines, sociology, medicine, and management, were almost identical: the structure was *data collection* → *experiment* → *data analysis*. Findings for biochemistry and chemistry were also very similar: the structure was *materials* → *equipment* → *procedure*. These similarities among certain disciplines are noteworthy, and led us to wonder if the same applies across a broader range of disciplines, thus the investigation reported in this paper. However, regarding comparing the results of this investigation to the previous research, three of the studies are not recent, and writing standards within disciplines may have changed, making comparison problematic. Therefore a direct comparison with our results is only possible for one discipline, chemistry.

The literature tends to downplay the importance of methods sections, but we feel they may be more important than the current research implies. Research writing guides such as Weissberg & Buker (1990) or Swales & Feak (2000) assert other methods section moves exist, including *overview*, *subjects/materials*, *location*, and *restrictions/limiting conditions*. Somewhat less important, though still useful for our study, is previous theorizing on the function, importance, and length of methods sections. Science methods sections have been described as “enigmatic, swift, presumptive...with little statement of rationale or discussion” (Swales, 1990, p. 170). Swales (2004, p. 220) adds that they are “heavily clipped,” explaining that this may happen because authors assume their readers already possess sufficient knowledge. Regarding the humanities, Swales (1990) claims they have a more careful step-by-step description of method. However, perhaps the usefulness of these suggestions—that science methods sections are becoming de-emphasized and shorter—is limited because they are based on looking at relatively few RAs. They also differ from Gledhill (2000), at least regarding one discipline: he measured the four sections Introduction-Method-Results-Discussion (IMRD) of 150 cancer RAs and found

methods to be 32 percent of the total¹. Gledhill's (2000) findings, followed by a preliminary reading of a few methods sections in our target disciplines, led us to wonder if the assertions that methods sections are of lesser importance really are true.

The above descriptions of the move structure and function of methods sections may lead us to think that writing them is relatively uncomplicated, but this is not the case (Swales, 2004). This may be particularly true for non-native speakers (NNS) (Paltridge, 1993; Wood, 2001), and if so NNS will have real difficulties publishing within their discourse communities research writing is not easy for native speakers (NS) either. These difficulties will be even more acute if, as Vassileva (1997) asserts, NNS research writers tend to be unaware of genre conventions, and these conventions probably differ across disciplines.

While all of the above suggests that NNS and NS research writers need to know the move structure of methods sections, few studies have investigated their structure, and they examined only a small number of RAs from a few disciplines. Other authors describe function or length and importance, though they looked almost exclusively at the sciences. We suggest that the area has not received the attention it deserves, that NNS in particular need help with research writing, and that our research may help by providing profiles of methods sections in a number of disciplines. There have recently been calls for further research; Lim (2006) asserts that methods sections are crucial and that we need to look at their generic structure, which has been neglected in the sciences. Kanoksilapatham, referring to all disciplines, states that there is still no clear model because methods sections have received "scant attention" (2005, p. 287). Also, while move cycles appear in RA discussion sections—Dudley-Evans (1994) describes them as important, and Swales as "well-established" (1990, pp. 172-74)—they seem not to have been observed or predicted in methods sections.

We suggest that further empirical research into methods sections is needed, and that such research should use larger corpora to assess variation across a wide range of disciplines. The findings may tell us much about the true nature of methods and also allow teachers of research writers to inform learners of typical move sequences and cycles.

¹**Notes**

He included titles and abstracts in his count, so we recalculated his data to exclude these and only include the four sections IMRD.

¹

3. Methodology

This study examined communicative moves in 288 RA methods sections across eight disciplines: biology, chemistry, physics, environmental science, business, language and linguistics, law, and public and social administration. As previous models differ, methods sections were searched using the move names and not models proposed by the five previous researchers who investigated method section move structures (Brett, 1994; Nwogu, 1997; Lim, 2006; Kanoksilapatham, 2005; Wood, 1982) or the writing guidebook authors (Weissberg & Buker 1990, Swales & Feak 2000).

The aims of this research were to:

- (1) Examine the move structure of RA methods sections across eight disciplines.
- (2) Examine interdisciplinary variation in RA methods sections.
- (3) Compare the move structure to previous models and provide new descriptions.

4. The RA Corpus

The corpus comprised 288 published RAs, 36 from each discipline. The disciplines were selected because they represent a wide range of academic subjects and also contain large numbers of research writers in this writer's university and around the world. Researching these disciplines may increase the usefulness of this research regarding recommendations for teaching. Also, previous research into method sections, apart from Wood (1982), did not cover these disciplines. Six leading refereed journals were chosen from each discipline. Four were taken from law as it was not possible to find six journals that contained empirical data-driven RAs. We visited the relevant academic departments and asked two informants from each discipline to name ten leading journals in their field. The 46 journals are listed in the Appendix. We note that the journals represent variation within individual disciplines. For example, within language and linguistics there is variation in the objects of study between *English for Specific Purposes* and the *Journal of Neurolinguistics*. Another example is public and social administration and the different research strands represented by *Evaluation and Program Planning* and *Social Science & Medicine*; this is probably also true for other disciplines. This variation was unintentional. That these sub-discipline variations could influence the structure of methods sections may be a factor in this research.

Six RAs from 2000/2003 were randomly selected from each journal. This was done by giving each RA a number and drawing six numbers for each journal from a box. Only empirical data-driven RAs with an explicit Introduction-Method-Results-Discussion format were selected, to allow comparison of moves. Review essays, discussions, and RAs by writers already selected were not used. The corpus size was 288 RAs, much larger than in previous studies. We suggest that the disciplinary corpora are acceptably representative because of their size and because of the use of expert informants to select journals. We realise that our corpus is heterogeneous in that it covers eight disciplines, yet we could not avoid this in order to examine interdisciplinary variation. Also, we found considerable parallels in the function of moves across disciplines, as we shall see below. For example, the common move *procedure* functioned in every discipline to describe the data-collection actions taken by the researcher/s.

We thus satisfied Sinclair's (2005) requirement for corpus building, which is to build corpora according to communicative function.

5. Classifying Moves in the Corpus

Two raters were involved in this procedure—the researcher himself and a local university lecturer who has an MA in Applied Linguistics and knowledge of RA move structure. Classification was done using the standard methods described above (Dudley-Evans, 1994; Holmes, 1997, 2001, p 1): (1) Look for organisation and patterns, identify moves and boundaries; (2) use sentence-level analysis; (3) assign all sentences to a move; (4) authenticate the classification by using two raters, who work independently, and then compare their findings. Sentence-level analysis means looking at individual sentences and to be aware that a single sentence can constitute one move. Using steps 1, 2, and 3 to identify and classify moves (step 4 is an added safeguard) was recommended or used by several previous theorists on and researchers into moves (Brett, 1994; Kanoksilapatham, 2005; Samraj, 2005).

The first stage of our analysis was a thorough examination of the corpus. As noted above, this search used the move names, but not models, suggested previously. We found that seven different moves occur, though not in all RAs or in this order: *overview*, *location*, *research aims/questions/hypotheses*, *subjects/materials*, *procedure*, *limitations*, and *data analysis*. Our assertion that seven different moves occur means that we found seven functions within method sections in the entire

corpus. We will now define the function of and exemplify all these moves. Typical examples, all quoted from our corpus, are included below.

Overview. This move gives a brief overview of the research method, at the start of or early in the methods section. For example:

This study examined the portrayals of police officers and criminals in this context, paying specific attention to racial representations and the use of force by police officers. (Law)

Research aims/questions/hypotheses. This move describes the goals of the research and the questions to be answered, and outlines the hypotheses if any. For example:

The aims of the evaluation study were to determine whether "enhanced" serves as a bridge to substance abuse treatment, care for HIV disease, and other services. (Public and social administration)

Subjects/materials. This move describes *subjects* in business, language and linguistics, law, and public and social administration, and *materials* in the sciences. *Subjects* are the people or organizations from whom data were collected. *Materials* are the equipment and supplies used during the research. An example of *subjects* follows:

The subjects involved in this study were 12 Japanese university students (six men and six women, average age approximately 20 years). (Language and linguistics)

An example of *materials* follows:

All solutions were prepared with de-ionized water (18 M cm⁻¹) from a Milli-Q analytical reagent-grade water purification system (Millipore, Bedford, MA). (Chemistry)

Location. This move describes the research site, the geographical location where the research took place, occasionally with reasons why that site was chosen. For example:

The 59.8 km² catchment of the Aar River in the south-eastern Lahn-Dill hill-country was chosen as the test area because it reflects a representative sample of the region. (Environmental science)

Procedure. This move describes the data-collection actions taken by the researcher/s. For example:

Colonies were then lysed on the membrane by freeze-thaw, then incubation with lysozyme (4 mg/ml). (Biology)

Limitations. This move describes the ways in which the research was restricted or limited, sometimes with reasons for the limitation/s. For example:

However, modelling such a complex structure is difficult because the variables are nested within one another. (Business)

Data analysis. This move describes how the data were analysed, the analysis method. For example:

All performances were scored at the time of the interview by the interviewer and the observer, using the analytic rating scale developed for the Foreign Services Institute (FSI) speaking test. (Language and linguistics)

To put steps 1, 2, and 3 into practice, we identified sentences which matched the above definitions. For example, to identify *procedure*, we searched for sentences which describe the data-collection actions taken by the researcher/s; to identify *data analysis* we looked for sentences which described how the data were analysed. We classified all moves by examining the move structure of all 288 RAs in terms of the above seven moves. First, the researcher classified all moves, and second, to measure inter-rater agreement, the second coder independently evaluated 60% of methods sections. The third procedure measured intra-rater agreement. Two weeks after the initial classification, the researcher randomly selected six RAs from each discipline and classified them again. Inter-rater and intra-rater agreement scores were then calculated by measuring correlations between the results using Statistical Package for the Social Sciences. Both were over 90 percent: the former was 91%, the latter 97%. We found little inter-rater or intra-rater disagreement once we had agreed on the definitions and functions of the seven moves. Steps 1, 2, and 3 had to be done using a purely manual search of the corpus, as the seven moves were expressed using a great variety of words. There appear to be no keywords or collocations for the expression of moves which would have allowed us to identify moves through an automatic search.

6. Results

The average number of moves per methods section was 6. This varied by discipline: biology 7, chemistry 6, physics 4, environmental science 8, business 7, language and linguistics 5, law 6, and public and social administration 6.

A number of marked differences were found in the moves used, move cycles, and methods section structure between three of the sciences, biology, chemistry, and physics and the four non-sciences. The differences between the two groups were fairly striking, and we will therefore present the results separately in order to highlight this dissimilarity between the three sciences listed above and the four non-sciences. Environmental science differed in several ways from all the other disciplines, and therefore those results will be presented independently.

Some representative examples of the typical moves from each discipline were given above. Those which were not will be presented below.

6.1. Overall Frequency of Individual Moves

We noted above that sub-discipline variation might affect the structure of methods sections. However, we found greater variation in moves and move structure *between* individual disciplines than *within* them. Table 1 below shows interdisciplinary differences in the overall move frequency: that is, the percentage of each move across the three sciences and the four non-science disciplines. For example, *procedure* made up 41% of all the moves found in biology, chemistry, and physics. The left column shows move names:

Table 1. Overall Move Frequency: Interdisciplinary Differences (%)

Moves	Biology, chemistry, and physics	Interdisciplinary differences (%) for business, language and linguistics, law, and public and social administration
Subjects / materials	29	20
Location	3	11
Procedure	41	35
Data analysis	25	12
Limitations	0	5
Research aims / questions / hypotheses	0	7
Overview	0	9

Biology, chemistry, and physics. Table 1 shows that the three most frequent moves were *procedure*, 41% of all moves, *materials* (29%), and *data analysis* (25%). These made up 96% of all moves. *Limitations*, *research aims/questions/hypotheses*, and *overview* were very rare. *Location* was rare, making up only 3% of all moves. Most were in biology, none in physics.

Business, language and linguistics, law, public and social administration. We see that the three most frequent moves were *procedure*, 35% of all moves, *subjects* (20%), and *data analysis*, 12%, less than half as frequent as in the three sciences above although in contrast with them, these made up only 68% of all moves. However, the variety of moves was greater than in the three sciences. *Location* was much more common, making up 11% of all moves. *Limitations*, *research aims/questions/hypotheses*, and *overview* were also much more prevalent, making up 5%, 7%, and 9% of all moves respectively.

6.2. Frequency of Appearance of Individual Moves

Only *procedure* appeared in all 288 RAs. Table 2 shows interdisciplinary differences in move frequency: that is, in how many percent of RAs the move appeared. For example, *materials* appeared in 97% of biology RAs:

Table 2. Frequency of Appearance of Individual Moves: Interdisciplinary Differences (%)

Moves	Biology	Chem.	Physics	Environ. Science	Business	Lang. & Ling.	Law	Public & Social
Subjects / materials	97	100	75	31	92	94	86	86
Location	36	8	0	67	47	58	58	75
Procedure	100	100	100	100	100	100	100	100
Data analysis	86	100	67	78	72	67	56	50
Limitations	6	0	0	69	44	19	28	6
Research aims / questions / hypotheses	3	0	6	11	36	22	58	67
Overview	3	0	0	50	25	19	42	54

Biology, chemistry, and physics. Table 2 shows that *materials*, *procedure*, and *data analysis* were very frequent, *materials* and *data analysis* less so in physics. *Location* appeared more often in biology.

Environmental science. Compared to the other three sciences, *materials* appeared far less frequently, 31%, while *location*, *limitations*, and *overview* were much more common at 67%, 69%, and 50% respectively.

Business, language and linguistics, law, public and social administration. Table 2 clearly shows that the variety of moves was greater than in the three sciences. *Data analysis* appeared in only 61% of RAs, compared to 84% of RAs in the three sciences. However, *location* was found in 60% of RAs. *Limitations*, *research*

aims/questions/hypotheses, and *overview* appeared in 24%, 46%, and 38% of RAs respectively.

There were individual discipline differences. *Location* was in only 47% of business RAs, whereas *data analysis* and *limitations* were in 72% and 44% respectively. *Research aims/questions* appeared in only 22% of language and linguistics RAs, and *overview* in 19%. Public and social administration used *location*, *research aims/questions*, and *overview* more frequently (75%, 67%, and 54%), and *data analysis* and *limitations* less frequently, in 50% and 6% of RAs respectively.

6.3. Move Cycles and Move Structure

Move cycles were common. The cycles, and overall structure, differed between the three sciences and other disciplines.

Biology, chemistry, and physics. Only two cycles were found, *materials* → *procedure* and *procedure* → *data analysis*. These made up 53% and 47% of all cycles, respectively. Repeated cycles were common: 50% of RAs had two or more of the cycle *materials* → *procedure*, and 37% had two or more of *procedure* → *data analysis*. 69% of RAs opened with *materials* → *procedure*. Some biology RAs opened with *location*, and 75% closed with *procedure* → *data analysis*. Some physics RAs closed with *materials* → *procedure*.

The typical move structure was *materials* → *procedure* → *materials* → *procedure* → *procedure* → *data analysis*.

Business, language and linguistics, law, public and social administration. Move cycle structure was much more complex, and a greater number of cycles were found. The most common cycle was *procedure* → *data analysis*, which made up 24% of all cycles and appeared in 53% of RAs. The second commonest was *subjects* → *procedure*, making up 23% of all cycles and appearing in 50% of RAs. Two other common cycles were *subjects* → *location* and *location* → *procedure*: these made up 17% and 16% of all cycles, and appeared in 44% and 41% of RAs respectively. Other less common cycles were *research aims/questions/hypotheses* → *procedure*, *procedure* → *limitations*, and *overview* → *procedure*; these appeared in 17%, 16%, and 12% of RAs respectively. Repeated cycles were rare; the commonest was *subjects* → *procedure*, but this cycle was repeated in only 8% of RAs.

92% of RAs opened with either *subjects*, *procedure*, *research aims/questions/hypotheses* or *overview*, but none of these dominated. The most frequent opening cycle was *subjects* → *location* in 15% of RAs. Another 13% opened with *procedure* → *subjects*, and 12% with *overview* → *procedure*. 83% closed with either *procedure* or *data analysis*; common closing cycles were *procedure* → *data analysis*, 47% of RAs, and *subjects* → *procedure*, 20%.

The typical move structure was *subjects* → *procedure* → *location* → *procedure* → *data analysis*.

6.4. Individual Disciplines

Representative examples of the typical moves from each disciplinary corpus are given (i.e., those which were not given above, in section 2.2). The variations found in individual disciplines will be described below:

Biology and chemistry. Methods sections were remarkably similar within and across these two disciplines. They were very uniform and standardised, and authors gave very precise and detailed descriptions of *materials*, including their origin. Authors often referred back to previous studies: e.g., “we used their method,” and sometimes highlighted innovations. Chemistry RAs, however, contained far more details of *data analysis* than did biology RAs.

The typical move structure in biology and chemistry was *materials* → *procedure* → *materials* → *procedure* → *materials* → *procedure* → *data analysis*.

Example moves from biology follow:

Materials:

PEG 4000 (molecular weight distribution of 3500-4500), NaH₂PO₄ and Na₂HPO₄ were obtained from Merck, Germany.

Data analysis:

The data sets were tabulated as a time series (distance between pairs of GFP dots as a function of time).

An example from chemistry follows:

Procedure:

Elution of mercury(II) content in foam column was achieved quantitatively by percolating 100 cm³ acetone at 5 cm³/min flow rate.

Physics. Methods sections were very similar to those in biology and chemistry, though shorter, and were also standardised across the discipline. Authors also tended to refer back to previous studies regarding method. However, while the majority opened with *materials* → *procedure* and closed with *procedure* → *data analysis*, fewer did than in biology and chemistry. Physics methods sections contained much more on physics theory, and on the mathematical models that the authors employed, than did the other sciences.

The typical move structure was *materials* → *procedure* → *procedure* → *data analysis*.

Examples from physics follow:

Materials:

The alloy Ti-6Al-4V was used in the form of cylindrical billets, fabricated by cold- and hot-isostatic pressing of elemental powders by Dynamet Technology (Burlington, MA), as described in [28].

Data analysis:

The energies of the ground and core-ionized states were calculated at the Hartree-Fock level, using the norm-extended quadratically convergent SCF [17 and 18] method.

Environmental science. Methods sections were longer and more complex than in any other discipline, and also differed in structure as noted above in 3.2. However, as in the other sciences, they were precise and detailed.

The typical move structure was *location* → *overview* → *procedure* → *limitations* → *procedure* → *data analysis* → *procedure* → *data analysis*.

Examples from environmental science follow:

Overview:

In December 1991 a 3-way factorial experiment was set up to examine the relationship between original stem size, cutting height, and post-harvesting pruning on resultant coppice yield.

Procedure:

The bales were stored in three stores located 8, 10, and 12 km away from the heating plant, respectively.

Data analysis:

All data were tested for normality (Proc Normal; SAS Institute, 1994) and homogeneity of variances (Levines test; Steel & Torrie, 1980).

Business. More explanation/description of the data-collection method was given than in the sciences and also more justification. *Hypotheses* were sometimes given, and also, less often, the cycle *data analysis* → *limitations*, found only in this discipline.

The typical move structure was *procedure* → *subjects* → *location* → *procedure* → *limitations* → *procedure* → *data analysis*.

Examples from business follow:

Procedure:

The subjects, who are managers from one selected industry (the plastic processing industry), were each assigned to one of the eight treatments.

Subjects:

The main survey was conducted among firms which had some (however limited) experience of supply chain partnering.

Data analysis:

Data was processed with LISREL (Joreskog & Sorbom, 1993), which is a structural equation modelling technique frequently used in marketing research.

Language and linguistics. This discipline had the shortest and simplest methods sections among the four non-science disciplines. There was some emphasis on innovation.

The typical move structure was *subjects* → *location* → *procedure* → *data analysis*.

Examples from language and linguistics follow:

Location:

...at the Department of Applied Psychology, Okayama University, Japan.

Procedure:

The interviews were videotaped and audio taped. The interview format was structured, with two parts, the first part being designed to elicit short answers, while in the second part the subject was expected to produce longer responses.

Law. Methods sections often opened with *overview*. Two cycles, *hypotheses* → *subjects* and *overview* → *hypotheses*, were found only in this discipline. They appeared in about 30% of RAs.

The typical move structure was *overview* → *subjects* → *location* → *procedure* → *procedure* → *data analysis*.

Examples from law follow:

Subjects:

A two-week sample of primetime television (8-11 p.m., EST) was constructed in the fall of 1997.

Procedure:

Analyses were conducted at the character level. Every police officer ($n=393$) was coded, whether in plainclothes or uniform.

Data analysis:

The reaction to or initiation of force was measured as either retaliatory force or initiating force.

Public and social administration. *Procedure* was often long, and involved descriptions of multiple methods, described sequentially. And as noted above in 3.2, the discipline showed a greater use of *location*, *research aims/questions*, and *overview* and lower use of *data analysis* and *limitations*.

The typical move structure was *overview* → *research aims/questions* → *subjects* → *location* → *procedure* → *data analysis*.

Examples from public and social administration follow:

Subjects:

The sample consisted of 194 child protection officers from all over Israel.

Location:

Such officers are appointed by the Ministry of Labor and Welfare, and are employed by the social service departments of the country's municipalities.

Procedure:

The child protection workers were asked to fill out all three questionnaires on two children, between 3 and 13 years old, with whom they had dealt in the previous 6 months.

Data Analysis:

MANOVAs and Univariate ANOVAs were carried out to examine the continuous variables.

Finally, while testing the assertions of Swales (1990, 2004) that RA methodology sections are becoming shorter would require investigation of changes over time, we did measure their length. They made up 25% of RA length². Table 3 below shows interdisciplinary differences.

Table 3. Length of Method Sections as a Percentage of the Whole RA: Interdisciplinary Differences

Biology	Chemistry	Physics	Environ. Science	Business	Language and Linguistics	Law	Public and Social Admin.
23	18	21	44	28	24	19	22

Environmental science had much longer than average methods sections, and chemistry shorter. There was little variation within individual disciplines, though we found a few exceptions; the *International Journal of Research in Marketing* had much longer methods sections (42%), while *International Business Review* and *Acta Materialia* had much shorter methods sections, 19% and 10%, respectively.

7. Discussion

Only one move, *procedure*, appeared in all 288 RAs. However, two other moves were extremely common in biology, chemistry, and physics: *materials* and *data analysis*. *Subjects* and *data analysis* were important in business, language and linguistics, law, and public and social administration, *data analysis* somewhat less so. Also, RAs from these disciplines had a greater range of moves: *location*, *limitations*, *research aims/questions/hypotheses*, and *overview* appeared often. Additionally we found move cycles to be almost universal, and repeated cycles fairly important in three of the sciences, although the type and order of cycles differed sharply between the three sciences and the four non-science disciplines, and to a lesser extent among the latter. None of these seems to have been predicted previously, and the discovery of these

The lengths of other sections were: Introduction 20%, results 29%, discussion section 26%.²

clear differences allowed us to present discipline-specific descriptions. We suggest that our finding of greater variation in moves and move structure between individual disciplines than within them indicates some disciplinary conformity and increases the value of the descriptions.

Regarding the *structure* of the methods sections, our findings differ from previous work. They do not match Wood (1982), who did not mention *data analysis* or move cycles, perhaps because our corpus was larger, or Brett (1994), Nwogu (1997), Lim (2006), or Kanoksilapatham (2005), who investigated different disciplines. However, regarding the *function* of methods sections, our findings do confirm the usefulness of some previous descriptions. We did find Swales's (1990) hypotheses to be partially accurate for biology, chemistry, and physics methods sections in that they tend to rely on shared knowledge and have "little statement of rationale or discussion" (1990, p 170); and for business, language and linguistics, law, and public and social administration in that they are a more careful step-by-step description of method. On the other hand, we did not find that methods sections are short. In our corpus they were usually complex and occupied an average of 25% of RA length. This indicates that they remain important to readers.

It is not easy to explain the differences across disciplines. Our findings indicate a number of clear differences in move structure between the three sciences and the four non-science disciplines. In biology, chemistry, and physics the three moves *materials*, *procedure*, and *data analysis* were all-important, and only two move cycles were found, sometimes repeated: *materials* □ *procedure* and *procedure* □ *data analysis*. Business, language and linguistics, law, and public and social administration had a greater variety of moves, including *location*, *limitations*, *research aims/questions/hypotheses*, and *overview*; and move cycle structure was different across articles and more complex. There were also less important differences between these four disciplines, and environmental science differed from all other disciplines. We believe that the disciplinary differences in moves, move cycles, and methods section structure that we found inform us about disciplinary norms, and that the patterns revealed are accepted within the relevant discipline as being the recognized way for writers to present their methods. Based on the results from our corpus, it is evident that in biology, chemistry, and physics it is very important for authors to let their readers know full details of *materials*, *procedures*, and *data analysis*. We went back to our corpus to try to find out why these and only these details are covered in

methods sections in these three sciences. It became apparent that the research very often involved detailed investigations of extremely complex materials and substances, the origin of which was important for readers to know. In environmental science, on the other hand, we found that it is important to describe *location*. On further investigation of this corpus we found that a thorough description of the site where the research took place (which was very often farmland, river systems, wetlands, or industrial areas) is of prime importance to understand the research context and results, and conclude that this information is certainly needed by readers. All these are examples supporting Hyland's (2000) suggestion that RA authors need to "project an insider ethos" (p.78). This is similar to his earlier proposal that discipline differences reflect rhetorical constraints within a discipline (Hyland, 1999). No doubt RA authors, who face considerable pressure to publish, must therefore look for acceptance among editors and readers, and face sanctions such as rejection if they step outside discipline conventions. If much depends on publication, sanctions must strongly motivate authors to follow discipline conventions. The need to gain this acceptance may mean there is a discrepancy between laboratory facts and facts as presented in RAs. This discrepancy, and the idea of writing as a social act, is also addressed in social construction literature; Berkenkotter & Huckin call writers "social actors" (1995, p. 24), and Latour & Woolgar (1986) say scientists present claims as fact and construct reality to increase credibility. This leads to questions about research writing: What is this insider ethos, what are these rhetorical constraints and discipline conventions, and do authors follow them? Our answer based on the results, especially our finding of much greater variation between disciplines than within them, is that among these conventions is the structure of methods sections, and that authors do follow these conventions. We suggest that the writers in our corpus are trying to maximize objectivity with their standardized presentation of facts.

In business, language and linguistics, law, and public and social administration, along with *materials*, *procedures*, and *analysis methods*, these discipline constraints mean that details about the *subjects* are also required, and sometimes also *limitations*, *research aims/questions/hypotheses*, and/or a research *overview*. Other individual discipline differences seem relevant. In business (and sometimes in language and linguistics), it is presumably important to let readers know the *limitations* of the research. We returned to our business corpus to try to find out why this is the case, and also why *location* appears less often and *data analysis* more often. It became

evident that the business RAs often explored mathematical models where the origin of the model (and thus location) is a less important factor, whereas analysing problems with the model, and the output of the model, are more central to the argument. We also revisited our public and social administration corpus to investigate why *location* is more frequent, and *procedure* longer and more complex, than in the other non-science disciplines. We found that the research often took place in unique communities and that it is important to describe the differences in these communities in detail, and also that research *procedures* were sometimes extensively modified for these communities. In public and social administration (and sometimes in language, linguistics and law), it is also a discipline convention to state *research aims/questions* or *hypotheses*; and in law and public and social administration to open with *overview*. We suggest that these individual discipline rules (as Swales, 1990 suggests) reflect differences between individual discourse communities, as well as agreement about the appropriate structure of methods sections inside individual disciplines.

8. Conclusion

The discipline-specific descriptions of methods sections may be useful in that they are based on a large corpus, and one such use may be for teaching. Implications for teaching fall into two different areas, the learning of move structures and the acculturation process of student socialisation into their chosen discipline (Ferenz, 2005). Regarding the first area, we propose knowledge of move structure is important for the teaching of research writing, particularly since teaching a subject through a focus on moves is an appropriate approach according to Dudley-Evans (1997). The findings of this study indicate move structures in methods sections vary between disciplines, and therefore discipline-specific teaching of method section structure may be appropriate. Additionally, sensitivity to interdisciplinary variation is required for teachers of research writing. While little has been written on teaching move structures, a first step to teaching them is to understand how communicative functions are linked to linguistic features (Lim, 2006).

Regarding the teaching of move structures to students involved in research writing, the following steps may prove useful to classroom teachers:

1. Introduce students to the concept of communicative moves and check if they know what they are. If not, at this stage use examples from step 4 below for

- illustration. Discuss how common moves are, why they are necessary, and how they function
2. Tell students the names of the three moves found in their discipline's methods sections, such as *materials*, *procedure*, and *data analysis*.
 3. Tell students the function of each move; *materials* describes the equipment and supplies used during the research, *procedure* describes the data-collection actions taken by the researcher/s, and *data analysis* describes how the data were analysed.
 4. Provide examples of moves for *materials*, *procedure*, and *data analysis*
 5. Tell students what move cycles are and note those that are common to their discipline, such as *materials* □ *procedure* and *procedure* □ *data analysis* in chemistry, and share the common patterns that RAs open and close with.
 6. Show students sample methods sections from their disciplines to act as models. Swales (1990) calls this a necessity, and we agree³. Mark the moves for students in the examples.
 7. Ask students to read other (unmarked) methods sections and ask them to mark and name the moves.
 8. Tell students the methods used in another RA from their discipline, listed in point form, and ask them to write the methods section, in prose using appropriate subheadings. Regarding the second area, the literature on academic writing contains a number of suggestions for facilitating student acculturation into a discipline, which we suggest making part of research writing courses. One such suggestion is appointing mentors for students who can encourage them to participate in disciplinary activities and who the students can identify with (Tardy, 2005). Another option is to promote interactions between students and the academic discourse community (Ferenz, 2005; Lea & Street, 2000), such as through participating in conferences.

Swales further suggests we look at these and see where they fall on the scale between³ “heavily clipped” and “highly elaborated” (2004: 224).

However, for students to be successful we need to provide them with the necessary linguistic skills to effectively join the discourse of their discipline (Stierer, 2000). Since all writing occurs within a social context, it is our responsibility as educators to make our students aware of that context (Koutsantoni, 2006).

9. Finally, we suggest that this exploratory study is limited in that we did not conduct informant interviews to investigate why these discipline differences occur. Future studies may check this, and also add additional disciplines to their investigations.

Summarising, analysis of our large corpus found a number of differences in the structure of RA methods sections across eight disciplines. In particular, we found differences between three of the sciences, environmental science, and the other four disciplines, and put forward methods section descriptions for eight individual disciplines. We hope that this research has added to the knowledge of genre conventions in academic writing and that these findings improve our understanding of RAs. We also hope that the findings have relevance for the teaching of research writing, and help teachers and course designers prepare discipline-specific courses for students.

References

- Berkenkotter, C. & Huckin, T.N. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale, NJ: Lawrence Erlbaum.
- Brett, P. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, 13(1), 47-59.
- 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. (1997). Genre: How far can we, should we go? *World Englishes*, 16(3), 351-358.
- Ferenz, O. (2005). EFL writers' social networks: impact on advanced academic literacy development. *Journal of English for Academic Purposes*, 4(4), 339-351.
- Gledhill, C. (2000). The discourse function of collocation in research article introductions. *English for Specific Purposes*, 19(2), 115-135.

- Holmes, R. (1997). Genre analysis and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes*, 16(4), 321-337.
- Holmes, R. (2001). Variation and text structure: The discussion section in economics research articles. *ITL Review of Applied Linguistics*, 131-132, 107-135.
- Hyland, K. (1996). Talking to the academy: Forms of hedging in science research articles. *Written Communication*, 13(2), 251-281.
- Hyland, K. (1999). Disciplinary discourses: Writer stance in research articles. In C.N. Candlin & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 99-121). London: Longman.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Harlow, UK: Longman.
- Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, 24(3), 269-292.
- Koutsantoni, D. (2006). Rhetorical strategies in engineering research articles and research theses: Advanced academic literacy and relations of power. *Journal of English for Academic Purposes*, 5(1), 19-36.
- Latour, B. & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. (2nd Ed.). Princeton, NJ: Princeton University Press
- Lea, M.R. & Street, B.V. (2000). Student writing and staff feedback in higher education: An academic literacies approach. In M.R. Lea & B. Stierer (Eds.), *Student writing in higher education: New contexts* (pp. 32-46). Buckingham, UK: The Society for Research into Higher Education and Open University Press.
- Lim, J.M.H. (2006). Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25(3), 282-309.
- Nwogu, K.N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes*, 16(2), 119-138.
- Paltridge, B. (1993). Writing up research: A systemic functional perspective. *System*, 21(2), 175-192.
- Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes*, 24(2), 141-156.
- Sinclair, J. (2005). Corpus and text – basic principles. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice* (pp. 1-16). Oxford: Oxbow Books.

- Stierer, B. (2000). Schoolteachers as students: Academic literacy and the construction of professional knowledge within Master's courses in education. In M.R. Lea & B. Stierer (Eds.), *Student writing in higher education: New contexts* (pp. 179-195). Buckingham, UK: The Society for Research into Higher Education and Open University Press.
- Swales, J.M. (1981). *Aspects of article introductions. Aston ESP research reports, no 1*. Birmingham: Language Studies Unit, the University of Aston at Birmingham.
- Swales, J.M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J.M. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.
- Swales, J.M. & Feak, C.B. (2000). *English in today's research world: A writing guide*. Ann Arbor: University of Michigan Press.
- Tardy, C.M. (2005). "It's like a story": Rhetorical knowledge development in advanced academic literacy. *Journal of English for Academic Purposes*, 4(4), 325-338.
- Vassileva, I. (1997). Hedging in English and Bulgarian academic writing. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 203-221). Berlin: Mouton de Gruyter.
- Weissberg, R. & Buker, S. (1990). *Writing up research*. Englewood Cliffs, NJ: Prentice Hall.
- Williams, G.C. (1998). Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics*, 3(1), 151-171.
- Williams, G. (2002). In search of representativity in specialised corpora: Categorisation through collocation. *International Journal of Corpus Linguistics*, 7(1), 43-64.
- Wood, A.S. (1982). An examination of the rhetorical structures of authentic chemical texts. *Applied Linguistics*, 3(2), 121-143.
- Wood, A. (2001). International scientific English: The language of research scientists around the world. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 71-83). Cambridge: Cambridge University Press.

Appendix A. Journals in the Corpus

Biology

Applied Soil Ecology

Biochimica et Biophysica Acta

Biomass and Bioenergy

Chemistry and Biology

Current Biology

Journal of Biotechnology

Business

Industrial Marketing Management

International Business Review

International Journal of Project Management

International Journal of Research in Marketing

Journal of Business Venturing

Journal of Operations Management

Chemistry

Analytical Biochemistry

Analytica Chimica Acta

Corrosion Science

International Journal of Inorganic Materials

Journal of Chemical Thermodynamics

Journal of Solid State Chemistry

Environmental Science

Applied Energy

Atmospheric Environment

Biomass and Bioenergy

Ecological Modelling

Environmental Pollution

Global Environmental Change

Language and Linguistics

English for Specific Purposes

Journal of Neurolinguistics

Language and Communication

Language Sciences

Speech Communication

System

Law

California Law Review

Canadian Journal of Criminology

International Review of Law and Economics

Journal of Criminal Justice

Physics and Material Science

Acta Materialia

Chemical Physics

International Journal of Fatigue

Journal of Luminescence

Journal of the Mechanics and Physics of Solids

Physica C: Superconductivity

Public and Social Administration

Child Abuse & Neglect

Evaluation and Program Planning

Habitat International

International Journal of Public Sector Management

Social Science & Medicine

World development



Title

**English for Specific Purposes in the EFL Context:
A Survey of Student and Faculty Perceptions**

Author

Chia-hsiu Tsao

Fooyin University

Biodata

Chia-hsiu Tsao is a full-time lecturer at Fooyin University of Technology, where she has been teaching various English courses to EFL learners for nearly twenty years. Her research interests include language-learning motivation, learning strategies, needs analysis, and factors that affect foreign language learning.

Abstract

This study reports on a questionnaire survey carried out with 351 students and 23 instructors in a technological university in Taiwan about their attitudes toward English for Specific Purposes (ESP) and related issues. Major findings of this study include: (1) Students favored ESP more than English for general purposes (EGP) while teachers were more reserved about the idea of replacing EGP with ESP; (2) Although both faculty and students recognized the importance of ESP, neither considered students' English proficiency up to the level needed to cope with the ESP course requirements; (3) Both sides agreed that although ESP courses should differ from EGP in their objectives, materials and approaches, they should still focus on the training of language skills while integrating specialized terms and discipline content into the course; (4) The two sides disagreed on whether English should be the only medium of instruction, to which the students give stronger support than the teachers; and (5) Regarding the factors that may affect the effectiveness of an ESP course, there is a notable difference of viewpoints between the two sides. The top-ranked factors, which were all student-centered, are as follows: student needs analysis, students'

learning capacity, and students' learning motivation. In contrast, the teachers emphasize more the course itself, placing teaching materials and methods as the top concerns, course objectives and design the second, and student needs the third. Based on these findings, this paper concludes with pedagogical implications and suggestions for future research.

Keywords: ESP (English for specific purposes); EGP (English for general purposes); specialized English; needs assessment

1. Introduction

As English continues to dominate as the lingua franca of business, media, technology, medicine, education, and research, the demand for ESP is growing rapidly, particularly in EFL countries where English is mainly used for instrumental purposes (Tsao, 2008; Xu, 2008). People in these countries learn English in order to fulfill the school curriculum requirement, to pass standardized English proficiency tests, or to obtain promotion or professional development at work. In response to the great demand for English in academic, vocational, and professional contexts, more and more universities in Taiwan are offering ESP courses to meet the global trend as well as to meet students' future career needs.

The demand for ESP has led some higher education authorities and administrators in Taiwan to claim that ESP should replace EGP, the long-existing practice of English language teaching in Taiwan's universities, and thus become the mainstream of college English education (Chen, 2008; Hsiao & Lin, 2002). The rationale behind their claim is presumably as follows: EGP is basic language learning to be studied before, but not during college; College English should be more advanced, more specialized, and match students' majors of study, particularly in technological universities where students are trained to perform on-the-job; and finally, compared with EGP, ESP is more effective in increasing students' learning motivation because it relates to their fields of study and caters to their needs (Chen, 1993).

Since students' overall English proficiency is generally believed to be an indicator of the success or failure of English education, the value of general English courses offered in local universities has been questioned ([The China Post Editorial, 2007](#)). The extremely low percentage of students in Taiwan, particularly those from

technical universities, who are able to reach the CEF⁴ A2 level of proficiency by the time of graduation, further justifies the inefficacy of general English education. Since the merit of EGP has been questioned, a shift from EGP to the more focused and career-oriented ESP seems worth an attempt. Many in Taiwan, including educators and administrators, consider ESP a potential ‘panacea’ for students’ poor English performance. They believe that students will be more motivated and benefit more by a program geared to their needs and directly relevant to their fields of study, thus resulting in higher learning efficacy. However, before implementing ESP instruction, some necessary preparation needs to be made, such as conducting needs analysis to identify students’ needs and expectations. In addition, many factors need to be taken into consideration, including teacher qualification, learner proficiency, teaching materials, large classes, and limited teaching hours. Moreover, teachers’ and students’ attitudes toward ESP require closer examination, because they are the ones who are directly involved in the course and who also play determinant roles in shaping the outcome of the course. As such, their opinions can provide valuable information for program directors, curriculum planners, and others who are responsible⁵.

Therefore, this study, by taking a private technological university in Taiwan as a case of investigation, attempted to compare and contrast teacher-student perceptions while obtaining answers to the following questions:

1. How do students and faculty view ESP as compared with EGP?
2. What do students and faculty think of student readiness for ESP?
3. What is required of ESP instruction?
4. What are the potential problems facing ESP in the EFL context?
5. What factors may affect the success of an ESP course?

⁴ CEF is the short form for Common European Framework of Reference for Languages: Learning, Teaching, and Assessment. It is a guideline used to describe achievements of learners of foreign languages and is comprised of six scales, moving from beginner to expert: A1, A2, B1, B2, C1, and C2. It has been adopted by the Ministry of Education of the Taiwanese government as the island-wide benchmark for measuring English proficiency of local college students.

⁵ In recent years several ESP centers or programs have been launched in various colleges or universities in Taiwan with a purpose to bridge the gap between students’ English proficiency and the requirement for their major field of study (Tsou, 2009).

2. Literature Review

2.1. Definition of ESP

Differences exist in how people interpret the meaning of ESP. Hutchinson and Waters (1987) defined ESP as "an approach to language teaching in which all decisions as to content and method are based on the learner's reason for learning" (p. 19). Strevens (1988) described it as English language teaching which is designed to meet specified needs of the learner. In the view of Chen (1993), ESP is "a major specialization within the discipline of English language teaching" (p. 80). Still, others specified ESP as the teaching of English for academic studies, or for vocational or professional purposes, as opposed to EGP, English for general knowledge and skills (Brunton, 2009; Carver, 1983; Hyland, 2006). Hence, we have such acronyms as EAP (English for academic purposes), EOP (English for occupational purposes), EMP (English for medical purposes), EBP (English for business purposes), and EST (English for science and technology). All of these are part of the ELT (English Language Teaching) repertoire. Whatever name it assumes, ESP is now a term connoting promise for more effective and more useful English language instruction (Tsou, 2009; Yogman & Kaylani, 1996).

2.2. Implementation of ESP

Sysoyve (2000) introduced a framework for the development of an ESP course. The framework started with student analysis, followed by formulation of goals and objectives, content design, selection of teaching materials, course planning, and course evaluation. In addition, he suggested that course development be viewed as an on-going process, with necessary alteration by the teacher to suit student interests and needs even while the course is in progress. Thus an ESP course should be a customized program, which caters to a certain group of learners with a specific purpose and enables them to prepare for professional communication at future workplaces.

In addition to students' specified needs for English, their learning capacity is also a very important factor that needs to be considered when developing an ESP course. As suggested by Wong (2005) and Huang (2007), students should have a reasonably good command of general English in order to learn ESP successfully. One important question that should always be addressed before implementing an ESP course is: Are the course objectives attainable with the students' current language

levels? In other words, how can ESP instructors design a course based on a more realistic student analysis, so that the course matches students' English ability and arouses their interest?.

According to Dudley-Evans and St. John (1998), the ESP instructor has as many as five key roles to perform: teacher, course designer and material provider, collaborator, researcher, and evaluator. Prior to teaching, the ESP instructors need to formulate the goals and objectives before they determine the content and select appropriate teaching materials for an ESP course. It often happens that the ESP practitioners may need to work with and even team teach with the subject specialists. They should also engage in classroom action research to solve instructional problems and improve teaching practice. And they should evaluate their course regularly to identify students' learning problems and to make proper adjustments. As we can see, to be an effective and competent ESP teacher is absolutely no easy task.

2.3. Empirical Studies of ESP

A multitude of studies have been dedicated to ESP-related issues in Taiwan. Issues discussed ranged from course designs (Chang, 1992), types of ESP courses or materials (Huang, 1997; Yang, Chang & Kao, 1994), learners' needs (Lee, 1998; Yang & Su, 2003), to reading strategies (Hsu, 2008) and vocabulary-learning strategies (Yang, 2005). In order to provide an insight into the ESP movement in Taiwan, we will review some of the more recent empirical research relevant to the interest of this study in more detail.

Huang (1997) conducted a large-scale survey to probe into the attitudes of five different populations—program coordinators, English language teachers, subject teachers, university freshmen, and graduates from 18 universities in Taiwan—toward the need to offer ESP courses under the general English curriculum. Her study found that program coordinators were divided in their attitudes towards the issue, with 7 out of 18 holding positive attitudes, 3 unsure, and 8 negative. The coordinators who supported the practice recognized the value of ESP and believed that students would benefit from ESP, while those who opposed cited reasons that included: (1) Language teachers lacked qualifications for teaching ESP; (2) Students' own departments should be responsible for offering ESP; (3) General language skills were more important for students; and (4) Students could acquire ESP naturally while studying their major subjects. In contrast, the attitudes of the remaining four groups were quite positive;

that is, they generally supported the idea of integrating ESP into the general English program. Huang also found that at the universities where ESP was integrated, it was only offered to a small extent and that the variety of ESP courses was very limited, with most of them related to news and business. Obviously, students' need for ESP at the time was not well met.

Assessing the English needs of nursing professionals, Lee (1998) compared in-service students with full-time students in a junior college of nursing and found that the former were more concerned about using English to communicate, such as describing nursing procedures, expressing comforting and greeting, and conversing in outpatient departments, while the latter expressed a stronger need for more fundamental tasks like learning medical terminology. Tsai (2001) conducted a similar but larger study, using a more refined questionnaire. Her study results demonstrated consistency between the "expectation needs" of nursing students and the "in-service needs" of nurses: both groups placed the need for developing communication ability above the need for learning language components and the need for professional training.

Unlike most ENP studies which chose students and in-service professionals as the subjects, Lee and Joe (2006) surveyed 169 nursing supervisors from local hospitals. They found that linguistic competence needs were deemed as important as communicative competence needs, and that either in EGP or ENP courses, listening and speaking abilities were regarded as the top two priorities, and that ENP courses should be given more emphasis than EGP courses.

A comparison of needs between medical school students and interns revealed that medical students most wanted to comprehend medical textbooks or journals written in English, while interns hoped to write and publish English medical reports or attend international medical seminars (Shen, 1996). Following Shen, Chia et al. (1999) investigated how medical students and faculty perceived the English needs for medical studies and derived the following findings: (1) Among basic language skills, listening was considered the skill most needed to be improved by freshmen; (2) Reading was ranked as most important for medical studies, followed by listening, writing and speaking; (3) The best pattern of curriculum design perceived by both was general English in the freshman year, followed by three years of specific elective English; and (4) Limited vocabulary and slow reading speed were singled out as being the most common problem.

Liang (2007) integrated ESP components into her Freshman English for Leisure Management students by combining subject matter (travel and tourism) and language learning. She adopted a variety of communicative activities to engage her students in the learning process, particularly role plays and oral presentation that allow real communication to happen. As a result, her students reacted favorably to the course. The successful implementation of ESP in her class implied that ESP can be integrated into college language courses as long as students' English proficiency and motivation are taken into account in the first place.

From the studies reviewed above, we can see that abundant research has been devoted to the field of ESP, and yet very few studies have been conducted to investigate how English faculty and students perceive ESP learning and teaching in an EFL setting, thus making the undertaking of the current study significant.

3. Methodology

3.1. Participants

This study involved 24 English teachers and 351 students from a technological university in Taiwan. The teachers surveyed included 19 lecturers, 4 assistant professors, and 1 associate professor, aged between 31 and 55. Eleven teachers had teaching experience of over 15 years, 6 between 11 and 15 years, the other 5 between 6 and 10 years, and 1 within 5 years.

The students surveyed were chosen based on stratified sampling (Gall, Borg & Gall, 2003)—a sampling technique in which the entire population is divided into groups, and a random sample of these groups is selected to make sure that the sample is representative of the larger population under study.

3.2. Procedures

The questionnaires to the teachers were either given to them in person or placed in their mail boxes. A total of 29 questionnaires were distributed, and 24 returned.

The survey to the students was administered by the researcher herself, who first obtained the permission of the instructor whose students had been chosen to answer the questionnaires. Before the survey began, the researcher explained the content and purpose of each questionnaire section to make sure the students knew how to answer it. Among 400 questionnaires that were issued, 351 valid responses were obtained, after incomplete ones were discarded. The sample, which had an

average age of 20.7 and an average time of 7.8 years in learning English, comprised students from various departments and in different years of study.

3.3. Instrument

A questionnaire written in Chinese and entitled "A Survey on ESP Instruction" (see Appendix 1 for the English translation) was developed by the researcher to serve the research purpose. The same questionnaire was adopted for both the students and the faculty with a view to comparing their perceptions. A pilot survey was conducted on 10 students and 3 English teachers before it was edited into the present format, which consists of three sections: three questions on the subjects' background information; 21 questions that probe into the subjects' attitudes toward ESP instruction, with responses rated on a 5-point Likert scale, ranging from "strongly agree" to "strongly disagree", and one multiple-choice question concerning the most important factors for the success of ESP. The 21 questions in part two were further broken down into four sub-sections, with 6 questions inquiring about how students and faculty view ESP as compared with EGP, 7 about whether students are ready for ESP instruction, 4 about what is required of ESP courses, and the last 4 about the potential problems facing ESP. The Cronbach's alpha coefficient for this scale was 0.85 for the students and 0.71 for the teachers, indicating that the scale measured responses with satisfactory internal consistency and accuracy.

3.4. Data Analysis

SPSS for Windows was used to perform the following analyses: First, an internal consistency reliability test using Cronbach alpha was conducted to obtain the reliability of the scale. Next, descriptive statistics, including frequencies, mean, and standard deviation, were estimated. Finally, independent t-tests were run to determine if there were any significant differences between students' and teachers' responses.

4. Results & Discussion

4.1. How Is ESP Compared with EGP?

As compared with EGP, ESP is obviously more appealing for both the students and the faculty according to Table 1. Both parties are quite consistent in their responses to items 2, 4, 5, and 6 of the questionnaire (t values are .94, .99, .94, and 1.29 respectively, none reaching the significant level of $p < .05$). More subjects agreed than

disagreed that ESP is superior to EGP in such aspects as having more specific goals, being more effective in promoting students' learning motivation, learning efficacy, and employment competitiveness.

The only significant differences between the two parties are found in items 1 and 3. When asked whether ESP should replace EGP to become the mainstream of college English education (item 1), the teachers were mostly opposed (58.3%, including those who disagreed and strongly disagreed) while the students were largely indecisive (48.4%). But still, there were more students who agreed (29.5%) than disagreed (21.8%). The result implies that students in general favor ESP more than EGP while teachers hold a more dubious attitude toward ESP becoming the core of English teaching at the tertiary level. A possible explanation for this discrepancy of opinions between the two parties is that students are not satisfied with the existing EGP instruction and hope for a change, while teachers either consider EGP a foundation for ESP or do not think such a substitution would lead to a better result.

The significant discrepancy of opinions in item 3 depicts a different story. The means of 3.66 for students and 4.13 for teachers suggest that both agree that tailor-made ESP courses should be offered to meet the needs of different departments, only that teachers are more supportive than students. To put it in a nutshell, the teachers embrace ESP instruction with the same zeal as the students, but not to the extent of substituting it for EGP.

Table 1. ESP vs. EGP (N=351 for Ss; 24 for Ts)

Queries	Responses on a Likert scale ^a							
	Subject	1	2(%)	3(%)	4(%)	5(%)	Mean	t
1. ESP should replace EGP to become the core of English teaching at vocational universities.	S ^b	6.8	15.0	48.4	20.4	9.1	3.10	2.65**
	T	12.5	45.8	20.8	16.7	4.2	2.54	
2. Compared with EGP, ESP has more specific goals.	S	2.5	6.5	39.1	43.3	8.2	3.40	0.94
	T	0	8.3	16.7	62.5	12.5	3.21	
3. Tailor-made ESP courses should be offered to students of various majors.	S	3.1	5.1	32.9	40.5	18.4	3.66	2.39*
	T	0	0	12.5	62.5	25.0	4.13	
4. ESP courses help enhance students' employment competitiveness.	S	2.0	6.5	28.9	45.0	17.0	3.69	0.99
	T	0	0	33.3	45.8	20.8	3.88	
5. ESP is more effective than EGP in improving students' learning efficacy.	S	4.5	9.6	38.0	37.4	10.5	3.40	0.94
	T	0	16.7	54.2	20.8	8.3	3.21	
6. ESP is more effective than EGP in increasing students' learning motivation.	S	4.8	12.5	42.5	33.1	6.8	3.25	1.29
	T	0	12.5	37.5	37.5	12.5	3.50	

*p<.05 **p<.01 ***p<.001

^a Responses on the Likert scale: 5=strongly agree 4=agree 3=unsure 2=disagree 1=strongly disagree^b S = students; T = teachers

4.2. Are Technological University Students Ready for ESP?

As can be seen in Table 2, neither the students nor the teachers have confidence in students' readiness for ESP instruction. The students consider themselves not competent enough to handle ESP learning, and the teachers feel exactly the same (M=3.70 and 3.71 respectively in item 5. The reasons lie in the students' lack of learning motivation, lack of adequate fundamental English ability and lack of sufficient vocabulary, about which the teachers are even more affirmative than the students ($t=2.66$, 2.22 and 2.58 in items 1-3, all reaching the significant level of p

< .05). The teachers' replies indicate that ESP is standing on a shaky ground. In fact, many ESP scholars and researchers (e.g. Chen, 2006; Gatehouse, 2001; Oladejo, 2004) have asserted that ESP must lay its foundation in general English proficiency, only through which higher level of professional and communicative competence can be achieved. If students are not able to read a simple passage or sustain an everyday conversation in English due to the lack of general vocabulary, they are likely to encounter great difficulty in understanding a content-rich journal article or express themselves in a specific occupational context. Echoing the contention of these scholars and researchers, the great majority of the teachers and the students alike in this study agree that students need to have a satisfactory grounding in basic English skills before they advance to ESP learning (item 6). As for the difficulty of ESP learning, there exist significant discrepancy of opinions between the two parties ($t=3.79$, $p<.001$). Students generally thought that "ESP is harder and more challenging than EGP", while teachers were divided in their responses. Around 42% of the teachers were indecisive, and the remaining 58% were equally divided into two opposite camps (half agreed and half disagreed). Indeed, it is not easy to decide which is harder, because ESP and EGP are equally insurmountable for most students in Taiwan.

Table 2. Whether students are ready for ESP (N=351 for Ss; 24 for Ts)

Queries	Responses on a Likert scale ^a							
	Subject	1(%)	2(%)	3(%)	4(%)	5(%)	Mean	t
1. I (Students at Fooyin) don't have any motivation to learn English.	S ^b T	3.1 0	18.4 8.3	16.1 4.2	45.0 66.7	17.3 20.8	3.55 4.00	2.66*
2. I am (Students at Fooyin are) very weak in English foundation.	S T	1.4 0	14.7 8.3	23.2 0	35.7 70.8	24.9 20.8	3.68 4.04	2.22*
3. I am (Students at Fooyin are) very weak in English vocabulary.	S T	0.6 0	10.8 4.2	18.7 0	47.3 70.8	22.7 25.0	3.81 4.17	2.58*
4. With little English vocabulary, I (Students at Fooyin) cannot learn ESP well.	S T	1.1 0	6.8 20.8	24.4 8.3	47.3 54.2	19.3 16.7	3.78 3.67	0.59
5. My (Fooyin students') English ability is not good enough to handle ESP course requirements.	S T	1.1 0	7.1 12.5	30.6 20.8	42.5 50.0	18.4 16.7	3.70 3.71	0.35
6. Students need to have a satisfactory grounding in basic English skills before they advance to ESP learning.	S T	1.1 0	5.9 16.7	22.4 16.7	47.0 45.8	23.2 20.8	3.86 3.71	0.78
7. ESP learning is harder and more challenging than EGP learning.	S T	1.7 0	3.7 29.2	32.6 41.7	42.5 25.0	18.7 4.2	3.73 3.04	3.79***

*p<.05 **p<.01 ***p<.001

^a 5=strongly agree 4=agree 3=unsure 2=disagree 1=strongly disagree^b S stands for students; T stands for teachers

4.3. What Is Required of ESP Instruction?

Table 3 reveals a consistency of opinions between the students and the teachers regarding what ESP courses should focus on and what an ideal ESP instructor should be like. Most of the subjects from the two groups (58.9% of the students and 65.0% of the teachers) agreed that ESP teaching should focus on language training while integrating terminology and discipline content into the course to meet the learners' specific needs. The majority of either group also support the statement that ESP should differ from EGP in its objectives, teaching materials, and teaching approaches. The result implies that when teaching ESP, language teachers should not use the same approach that is used in teaching general English, because these two are apparently different in their goals and learning content.

For item 3, an agreement of opinions is again found between the students and the teachers. The great majority of the respondents from both groups agree that ESP instructors should possess both English-teaching competency and subject content knowledge. However, to find instructors who are experienced and capable of teaching English may be easy, whereas it is certainly not easy to find someone who is at once a competent language teacher and a knowledgeable specialist. To solve this problem, previous research has suggested team-teaching as a coping strategy (Adams-Smith, 1980; Chang, 2005; Jackson & Price, 1981). Yet, while co-teaching may be an ideal way to deal with the shortage of qualified ESP instructors, it is not widely feasible when taking into account the cost and time spent on making co-teaching work and the difficulty of coordinating language and subject teachers. In comparison, the suggestion of seeking advice from subject specialists or obtaining subject area information from students seems to be more plausible (Dudley-Evans & St. John, 1998; Robinson, 1991). As a matter of fact, the lack of content-area knowledge is probably not an issue for most of the ESP courses, because it is general English communication ability, not subject-content knowledge, that is most desired by the learners (Chen, 2006; Oledajo, 2005; Wang, 2004).

Since communication ability is the core of learning in most ESP classes, should English be the only medium of instruction? According to the statistical figures in item 4, 25.7% of the student respondents agreed on the use of English as the only medium of instruction, while only 4.2% of the surveyed teachers did, resulting in a significant discrepancy of opinions between the two parties ($t=3.28$, $p<.001$). Apparently, many teachers (70.8%) would resort to bilingual teaching or even speak more Chinese than English in the ESP classroom. The reasons are either that most of the local English teachers in Taiwan, as claimed by Oladejo (2005), are not communicatively competent enough to teach their students in English, or that the students' English ability is too inadequate to benefit from an "English-only" instruction, as observed by the researcher of this study.

Table 3. What is required of ESP instruction (N=351 for Ss; 24 for Ts)

Queries	Responses on a Likert scale ^a							Mean	t
	Subject	1(%)	2(%)	3(%)	4(%)	5(%)			
1. The focus of ESP teaching should be placed on language training, while integrating specialized content into the course.	S ^b T	2.0 0	4.8 12.5	34.3 12.5	47.3 58.3	11.6 16.7	3.62 3.79	0.99	
2. ESP should differ from EGP in its goal, material, and approach.	S T	2.0 0	4.8 25.0	36.8 8.3	45.9 62.5	9.9 4.2	3.57 3.46	0.66	
3. ESP instructors should possess both English-teaching competency and content-specific knowledge.	S T	0.8 0	3.1 8.3	26.3 12.5	48.7 66.7	20.7 12.5	3.86 3.83	0.13	
4. English should be the only medium of instruction for ESP courses.	S T	9.6 12.5	22.1 58.3	41.9 25.0	19.5 4.2	6.2 0	2.91 2.21	3.28***	

*p<.05 **p<.01 ***p<.001

^a 5=strongly agree 4=agree 3=unsure 2=disagree 1=strongly disagree^b S stands for students; T stands for teachers

4.4. What Are the Potential Problems Facing ESP?

The same problems that have occurred in EGP will also confront ESP. To name a few, there is a large group of uninterested students with poor English foundation and scanty English vocabulary. They have no English environment to be exposed to, once out of the English classroom, nor are they required to or given the chance to use English in their daily life. English, for them, is nothing but a school subject that they have to pass in order to graduate. In addition, large class teaching and limited hours of instruction are also negative factors that will impact the teaching and learning of ESP.

Table 4 reveals the subjects' concerns about the problems with ESP. An overwhelming majority of the students and the teachers alike agree that limited hours of instruction and lack of opportunities to apply English will diminish the effects of ESP instruction, although the teachers feel much more strongly than the students in both cases ($t=4.44$ in item 1, $p<.001$; $t=2.32$ in item 2, $p<.01$). The two parties also generally agree that lack of qualified teachers is an urgent problem of ESP (item 3). When asked if ESP courses are likely to become limited to the learning of specific lexicon and the translation of content-specific texts, over half of the subjects from the

two groups (50.7% of the students and 58.3% of the teachers) gave positive responses, a result making the future of ESP rather pessimistic. In summary, unless the problems affecting the development of ESP education are addressed, little can be expected of the future of ESP.

Table 4. Potential problems of ESP courses (N=351 for Ss; 24 for Ts)

Queries	Responses on a Likert scale ^a							
	Subject	1(%)	2(%)	3(%)	4(%)	5(%)	Mean	t
1. Limited hours of instruction will weaken the effects of ESP instruction.	S ^b T	2.8 0	9.1 0	42.8 16.7	33.1 66.7	11.9 16.7	3.42 4.00	4.44***
2. Lack of opportunities to use English in daily life or the workplace will undermine the effects of ESP instruction.	S T	0.8 0	5.4 0	28.6 12.5	46.7 70.8	18.1 16.7	3.76 4.04	2.32**
3. Shortage of qualified teachers is a potential problem for ESP.	S T	6.3 16.3	19.8 38.1	33.1 30.4	24.0 10.4	16.7 4.8	3.25 2.49	3.58***
4. ESP courses will likely become limited to the learning of specific lexicon and the translation of specialized texts.	S T	2.3 4.2	5.4 8.3	41.1 29.2	39.1 45.8	11.6 12.5	3.53 3.54	0.80

*p<.05 **p<.01 ***p<.001

^a 5=strongly agree 4=agree 3=unsure 2=disagree 1=strongly disagree

^b S stands for students; T stands for teachers

4.5. What May Affect the Success of ESP Practice?

Among the various factors that are likely to affect the success of an ESP course, as displayed in Table 5, the top five for the students are: needs analysis (56.1%), students' learning capacity (51.6%), students' learning motivation (49.3%), specific course objectives and proper planning (48.7%), and appropriate teaching materials and methods (38.5%), while no other factors reached 30%. The top three factors all concern students, which correspond to ESP scholars' assertion that establishing the needs of students and subsequently the goals of a class is the foundation of any effective ESP course. This is in agreement with Chen (1994) and Hutchison & Waters (1987). Robinson (1991) also pointed out that an ESP course, in theory, is goal-oriented and is based on needs analysis.

In contrast with the students' responses, the teachers emphasized more the course itself, placing teaching materials and methods as the top concern (79.2%),

course objectives and design as the second (70.8%), and students' needs as the third (41.7%). Obviously, there is a notable difference of viewpoints between the two parties that needs to be reconciled. The result here provides the teachers with valuable information as to how their views contrast with their students', consequently allowing the teachers to make adjustments accordingly.

Table 5. Factors that may determine the success of ESP (N=351 for Ss; 24 for Ts)

Items	Ss' responses (n=351)			Ts' responses (n=24)		
	N	%	Rank	N	%	Rank
1. Survey students' needs before the class	198	56.1	1	10	41.7	3
2. Specific course objectives and proper planning	172	48.7	4	17	70.8	2
3. Appropriate teaching materials and methods	136	38.5	5	19	79.2	1
4. Appropriate number of credit hours offered	7	2.0		4	16.7	
5. Students' learning capacity	182	51.6	2	5	20.8	
6. Students' learning motivation	174	49.3	3	7	29.2	
7. Teachers' content-specific knowledge	58	16.4		3	12.5	
8. Teachers' teaching competency	67	19.0		6	25.0	
9. Teachers' English ability	28	7.9		0	0	
10. Others	4	1.1		1	4.2	

5. Conclusion

This study investigated the perceptual similarities and differences between students and teachers regarding the demand for ESP in vocational universities. Major findings indicated that both faculty and students agreed on the following: 1) ESP is important and necessary for technological students; 2) Students need to have a satisfactory grounding in basic English skills before they advance to ESP learning; 3) ESP instruction should focus on the training of language communication skills while integrating terminology and subject content into the course; 4) ESP instructors should possess both English-teaching competency and subject content knowledge, and 5) Problems that have affected the effectiveness of general English curricula will also affect ESP negatively. By way of contrast, students' and teachers' perceptions differ in these areas: 1) The majority of students agreed on the idea of substituting ESP for EGP, but it was the opposite case with the faculty; 2) Students showed stronger

support than the teachers about the idea of using English as the only medium of instruction, and 3) Students considered needs analysis, students' learning capacity, and students' learning motivation as the most crucial factors that contribute to the success of an ESP course, while the faculty regarded teaching materials and methods, course objectives and design, and student needs as the most important factors.

6. Pedagogical Implications

Based on the findings of this study and the review of literature, several implications are discussed. First, as ESP today is increasingly taught to large classes of demotivated students with a poor command of English, it seems particularly necessary to base ESP teaching on some competence in EGP, for no students of ESP could go far without basic English skills (Chen, 1994). Moreover, the complexity of the subject content must be controlled and kept within manageable limits of the learners' ability – the content knowledge should be something familiar or not too difficult for the students to handle. Otherwise, learning both language and specialized content could become overwhelming for the students who, according to the survey results of this study, were not well equipped enough to deal with the dual complicity of ESP courses. In other words, ESP teachers should select materials that are less specialized in content knowledge but still related to the learners' fields of study or work.

Second, ESP should be regarded as a complementary language-learning requirement, rather than a substitution for EGP. Many researchers (e.g. Gilmour & Marshal, 1993; Spack, 1988) pointed out that students' problems in comprehending specialized texts are mostly caused by lack of general English words, rather than by the technical terminology of their subject. This finding that general English skills are much more needed than technical English knowledge in the workplace is in agreement with previous studies (Chen, 2006; Oladejo, 2005; Wang, 2004) among in-service professionals such as nurses and engineers for communicative purposes. As Chen (2006, p.26) nicely put it, ESP is like the leaves and branches on a tree of language. "Without tree trunks and roots, leaves or branches cannot grow because they do not have the necessary underlying language support. The same is true of ESP, since content-related specific language cannot stand alone without General English syntax, lexis, and functions."

Third, ESP course planning should begin with the analysis of learners' needs and wants. Based on learners' needs and their future language use, objectives of the course can then be determined, and evaluation measurements can be integrated to ensure that these objectives are achieved. While four skills are stressed equally in an EGP class, in ESP, it is needs analysis that determines which language skills are most needed and characterize the syllabus design (Gatehouse, 2001).

Finally, an ESP teacher does not have to be an expert in the target field, but has to remain flexible and is always engaged in professional development in ESP teaching. Dudley-Evans and St. John (1998) suggested that the ESP instructor can consult with the subject expert when developing materials or encountering problems with the subject area. Anthony (2007) proposed the "teacher as student" approach, by which he meant the ESP teacher, by acting like a student of the target field, can learn a lot by listening to the views of his/her students and can also contribute to discussions by using his/her vast knowledge of English. In addition to specialist informants, self-education such as reading ESP journals, content-area textbooks, and media reports, is also an effective way to obtain specialist knowledge. Acquisition of specialist information and continuing education should thus be a priority for ESP teachers.

7. Limitations of the Study

This study is limited in several respects. First, its subjects were drawn from the same university, i.e. a private technological university in southern Taiwan, and therefore its results may not be generalized well to other educational settings or other population with different backgrounds. Future research involving a larger sample across the nation would help validate the findings of this study. Another limitation of this study concerns its methodological design - only a questionnaire survey was adopted to collect research data. Other types of survey instruments, such as individual or focus group interviews, should be included in further studies to obtain more in-depth information pertaining to the issues researched. Finally, the list of items in the questionnaire of this study is by no means adequate and complete; it is subject to further confirmation and modification through more empirical studies.

References

- Adams-Smith, D. E. (1980). Co-operative teaching: Bridging the gap between E and SP. In British Council (Eds.), *Team Teaching in ESP*. ELT Documents 106. London: British Council English Teaching Information Centre.
- Anthony, L. (2007). The teacher as student in ESP course design. Keynote address given at *2007 International Symposium on ESP & Its Applications in Nursing and Medical English Education*. October 5-6, 2007. Kaohsiung, Taiwan.
- Brunton, M. (2009). An account of ESP – with possible future directions. *English for Specific Purposes Issue*, 3(24), Vol. 8. Retrieved September 30, 2009 from http://www.esp-world.info/Articles_24/An%20account%20of%20ESP.pdf
- Carver, D. (1983). Some propositions about ESP. *The ESP Journal*, 2, 131-137. Retrieved September 30, 2009 from http://www.esp-world.info/Articles_22/esp%20essay%20for%20publication.htm.
- Chang, B. N. (2005). A discussion of team teaching in English for specific purposes (ESP) courses. *Far-Eastern Journal*, 22(4), 531-546.
- Chang, B. Y. (1992). Dahsueh bishou waiwen kecheng shechi chih tantao. [University curriculum design for required foreign language courses]. *The Proceedings of the 8th Conference on English Teaching and Learning in the Republic of China* (pp. 161-170). Taipei: Crane Publishing Co. Ltd.
- Chen, H. C. (1994). Team-teaching English for special purposes: A new perspective. *Selected papers from the second international symposium on English teaching* (pp. 79-95). Taipei: Crane Publishing Co. Ltd.
- Chen, Y. (2006). From common core to specific. *Asian ESP Journal*, 1(3). Retrieved June 18, 2008 from http://www.asian-esp-journal.com/June_2006_vc.php
- Chen, Y. H. (2008). Chuanchouhwa shih tai yingyu de weilai—yingyu tsowei dieryuyen de hsuehhsi tseluei. [The future of English in the globalized age—learning strategies for English as the second language]. *English Career*, 27, 21-27.
- Chia, H. U., Johnson, R., Chia, H. L. & Olive, F. (1999). English for college students in Taiwan: A study of perceptions of English needs in a medial context. *English for Specific Purposes*, 18(2), pp. 107-119.
- Dudley-Evans, T. & St. John, M. (1998). *Developments in ESP: A multi-disciplinary approach*. Cambridge: Cambridge University Press.

- Gall, M. D., Borg, W. R. & Gall, J. P. (2003). *Educational research: An introduction* (7th Edition). White Plains, New York: Longman.
- Gatehouse, K. (2001). Key issues in English for specific purposes (ESP) curriculum development. *The Internet TESL Journal*, 7(10). Retrieved June 20, 2008 from <http://iteslj.org/Articles/Gatehouse-ESP.html>
- Gilmour, M. & Marshal, S. (1993). Lexical knowledge and reading in Papua New Guinea. *English for Specific Purposes*, 12(2), 145-157.
- Hsiao, C. & Lin, C. (2002). Keda nanhai tselueh shindenfan: Fooyin chuanli tuidon esp yifu yingwen chiayu. [The new model of blue sea strategies for technical universities: Fooyin vigorously promotes ESP medical English education]. *English Career*, 22, 28-32.
- Hsu, L. W. (2008). Taiwanese hospitality college students' English reading strategies in English for specific purpose courses. *Journal of Hospitality and Home Economics*, 5(1), 53-67.
- Huang, P. (2007). Metacognitive reflections on teaching business English writing skills to evening program college students. *The Proceedings of 2007 International Symposium on ESP & Its Application in Nursing & Medical English Education* (pp. 267-285). Kaohsiung, Taiwan: Fooyin University.
- Huang, S. Y. (1997). The ESP component in English programs for non-English majors at Universities in Taiwan. *The Proceedings of the 14th Conference on English Teaching and Learning in the Republic of China* (pp. 367-378). Taipei, Taiwan: Crane Publishing Co. Ltd.
- Hutchinson, T. & Waters, A. (1987). *English for Specific Purposes: A learning-centered Approach*. Cambridge: Cambridge University Press.
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. London: Routledge.
- Jackson, M. & Price, J. (1981). A way forward: A fusion of the two cultures or teaching communication to engineers. In British Council (Eds.), *The ESP Teacher: Role, Development and Prospects*. ELT Documents 112. London: British Council English Teaching Information Centre.
- Lee, C. Y. (1998). English for nursing purposes: A needs assessment for professional oriented curriculum design. *The Proceedings of the 7th International Symposium on English Teaching* (pp. 615-625). Taipei: Crane Publishing Co. Ltd.

- Lee, S. C. & Joe, S. G. (2006). A needs analysis of English for nursing professionals in southern Taiwan: A study on junior college graduates. *The Proceedings of 2006 International Symposium on ESP & its Application in Nursing and Medical English Education* (pp. 162-77). Kaohsiung, Taiwan: Fooyin University.
- Liang, Ling-Ling (2007). An ESP course for leisure management students of technological colleges. *The Proceedings of the 16th International Symposium on English Teaching* (pp. 467-473). Taipei: Crane Publishing Co. Ltd.
- Oladejo, J. (2004). Too little, too late: ESP in EFL communicative competence in the era of globalization. *The Proceedings of 2004 International Conference and Workshop on TEFL & Applied Linguistics* (pp. 421-427). Taipei, Taiwan: Ming Chuan University.
- Robinson, P. (1991). *ESP today: A practitioner's guide*. Hemel Hempstead, UK: Prentice Hall.
- Spack, R. (1988). Initiating ESL students into the academic discourse community: How far should we go? *TESOL Quarterly*, 22(1), 48-62.
- Shen, Y. M. (1996). Tantaobien bichao chongyehjenyuan yu tsaihsiaosen duai shuehhsi yishuehyinwen che yitong. [Comparing medical school students and doctors in their needs for English. *The Proceedings of the 13th Conference on English Teaching and Learning in the Republic of China* (pp. 31-40). Taipei, Taiwan: Crane Publishing Co. Ltd.
- Stevens, P. (1988). ESP after twenty years: A re-appraisal. In M. Tickoo (Ed.), *ESP: State of the art* (pp. 1-13). SEAMEO Regional Language Centre.
- Sysoyev, P. V. (2000). Developing an English for specific purposes course using a learner centered approach: a Russian experience. *The Internet TESL Journal*, 6(3). Retrieved June 18, 2008 from: <http://iteslj.org/Techniques/Sysoyev-ESP.html>
- The China Post [Editorial](#) (August 24, 2007). [TEFL scores in Taiwan: a humiliating failure](#). Retrieved October 1, 2009, from <http://www.chinapost.com.tw/editorial/2007/08/24/119749/TEFL-scores.htm>
- Tsai, P. (2001). An assessment of English needs of nursing students and professionals. *The Proceedings of the 10th International symposium on English Teaching* (pp. 608-620). Taipei: The Crane Publishing Co, Ltd.
- Tsao, C. H. (2008). English-learning motivation and needs analysis: a case study of technological university students in Taiwan. Paper presented at the Basic

- Research Conference of Chinese Military Academy on its 84th school anniversary. May 6, 2008. Fongshang, Taiwan.
- Tsou, W. (2009). ESP (English for Specific Purposes) makes college EFL learning effective. Paper presented at TESOL 2009 Annual Conference. March 26-28, 2009. Denver, Colorado, USA.
- Wang, B. R. (2004). An investigation of ESP instruction in Tongji University. *Foreign Language World*, 1, 35-42.
- Wong, C. S. (2005). Refocusing on the learner: A principled remedial language acquisition strategy for ESP students. *The Proceedings of 2005 International Symposium on ESP & its Application in Nursing and Medical English Education* (pp. 63-64). Kaohsiung, Taiwan: Fooyin University.
- Xu, X. (2008). Influence of instrumental motivation on EFL learners in China and its implication on TEFL instructional design. *Educational Communications and Technology*, University of Saskatchewan. Retrieved August 20, 2009 from <http://www.usask.ca/education/coursework/802papers/xu/index.htm>
- Yang, C. H., Chang, H. & Kao, M. Y. (1994). A study on the use of field-specific authentic English texts in a junior college. *The Proceedings of the 3rd International symposium on English Teaching* (pp.171-189.). Taipei: The Crane Publishing Co. Ltd.
- Yang, M. N. (2005). Nursing pre-professionals' medical terminology learning strategies. *Asian ESP Journal*, 1(1). Retrieved June 1, 2008 from http://www.asian-esp-journal.com/June_2006_yc.php
- Yang, M. N. & Su, S. M. (2003). A study of Taiwanese nursing students' and in-service nursing professionals' English needs. *Journal of Chang Gung Institute of Technology*, 2, 269-284.
- Yogman, J. & Kaylani, C. T. (1996). ESP program design for mixed level students. *English for Specific Purposes*, 15(4), 311-324.

Appendix 1. Questionnaire

Part One: Personal Information

1. Major: _____ Year of Study: _____

2. Time spent on English study per week :

- ① more than 3 hours ② 2-3 hours ③ 1-2 hours ④ less than one hour

3. Major ways of learning English (multiple choices) :

- ① classroom teaching ② self-study ③ cramming schools or tutors
 ④ English programs on the air ⑤ English songs or movies
 ⑥ English-media (newspaper, the Internet, etc.)
 ⑦ practicing English with other people ⑧ others

Part Two: Please choose from 1 (strongly disagree) to 5 (strongly agree) for each of the following statements.

1. ESP should replace EGP to become the core of English teaching at vocational universities.
2. Compared with EGP, ESP has more specific goals.
3. Tailor-made ESP courses should be offered to students of various majors.
4. ESP courses can help enhance students' employment competitiveness.
5. ESP is more effective than EGP in improving students' learning efficacy.
6. ESP is more effective than EGP in increasing students' learning motivation.
7. I don't have any motivation to learn English.
8. I am very weak in English foundation.
9. I am very weak in English vocabulary.
10. With little English vocabulary, I don't think I can learn ESP well.
11. My English ability is not good enough to handle ESP course requirements.
12. Students need to have a satisfactory grounding in basic English skills before they advance to ESP learning.
13. ESP learning is harder and more demanding than EGP learning.
14. The focus of ESP teaching should be placed on language training, while integrating specialized content into the course.
15. English should be the only medium of instruction for ESP courses.
16. ESP instructors should possess both English-teaching competency and content-specific knowledge.

17. ESP should differ from EGP in its goal, teaching material, and approach.
18. Lack of opportunities to use English in daily life or the workplace will debilitate the effects of ESP instruction.
19. ESP courses are likely to become the teaching of specific lexicon and the translation of specialized texts.
20. Limited hours of instruction may weaken the effects of ESP instruction.
21. Shortage of qualified teachers is a potential problem for ESP.

Part Three:

For an ESP course to succeed, what factors are influential? (Please choose five that you consider more important than the others)

1. Students' needs analysis
2. Specific course objectives and proper planning
3. Appropriate teaching materials and methods
4. Number of credit hours offered
5. Students' learning capacity
6. Students' learning motivation
7. Teachers' content-specific knowledge
8. Teachers' teaching competency
9. Teachers' English ability
10. Others (please specify): _____