

**Main Grammatical Writing Errors of
Chinese Undergraduate Students in UK Universities:**

An Error Analysis and Contrastive
Analysis Hypothesis Approach.

Richard Bond

Student id s1475975

R.f.bond@umail.leidenuniv.nl

bondsenglish@gmail.com

Supervisor: Dick Smakman

Second Reader: Mike Child

MA Thesis English Language and Linguistics

Leiden University

January 2016

Abstract

Chinese students studying in UK Universities contribute significantly to the UK economy (Leedham, 2011). There are other non-economic potential benefits to both UK universities and Chinese students which are not currently efficiently utilised. These include academic skills, cultural competency and the knowledge base of domestic students (Arthur, 2015). Students' success at university depends on their ability to write academic English however, some UK academics report that Chinese students' writing is not at a sufficient level for academic success. Previous studies have examined Chinese students written errors in their home countries or during their pre-session course, but have not examined the main grammatical errors affecting students' in-course writing. A corpus of Chinese students academic writing in UK universities was accessed and examined to identify the main grammatical writing errors following the procedure for Error Analysis proposed by Corder (1967). The main errors were examined using Lado's 1957 Contrastive Analysis procedures to determine the difficulty of the grammatical forms for Chinese students. Previous strategies for redressing Chinese students' grammatical errors are reviewed and the importance of the link between academic success and cross cultural communication competence in the efficient utilization of the non-economic benefits of Chinese students studying in the UK is highlighted.

Keywords: Second Language Acquisition, Error Analysis, Contrastive Analysis Hypothesis, Chinese Students, Written Grammar Errors.

Table of Contents

Abstract.....	i
Table of Contents.....	ii
Tables.....	iv
Abbreviations.....	v
Chapter 1. Introduction.....	1
1.1 Background.....	1
1.2 Operationalization.....	3
1.3 Summary of Previous Research.....	4
1.4 The Goal of the Study and Expected Outcomes.....	5
Chapter 2. Background Literature.....	7
2.1 An Overview of Chinese Language.....	7
2.2 Error Analysis.	10
2.3 Contrastive Analysis Hypothesis.....	16
2.4 English for Academic Purposes in China.....	25
2.5 Previous Studies.....	27
2.6 Conclusion.....	31
Chapter 3. Methodology.....	32
3.1 The BAWE Data Set.....	35
3.2 Error Correction.....	40
3.3 Contrastive Analysis.....	43
Chapter 4. Results.....	44
4.1 Error Analysis Results.....	44
4.1.1 Categories not Considered Further.....	44
4.1.2 Conflated Categories.....	45
4.2 Major Error Categories.....	48
4.3 Contrastive Analysis Hypothesis Results.....	50
4.3.1 Articles.....	52
4.3.2 Plurals.....	55

Chapter 5. Discussion.....	62
5.1 Research Questions.....	62
5.2 Recommendations.....	73
References.....	76
Appendices.....	90

Tables

Tables

Table 2.1 Distribution of plural markers for English and Spanish.....	20
Table 2.2 Word order change in question formation in English and Spanish.....	21
Table 2.3 Summary of main errors identified in previous studies.....	30
Table 3.1 Breakdown of available data on age, gender and L1 of participants.....	38
Table 3.2 Error Categories.....	41
Table 4.1 Error categories with few tokens.....	45
Table 4.2 Article Errors.....	46
Table 4.3 Number of errors for Singular and Plural Errors across First Year and Third Year essays.....	47
Table 4.4 Revised Error Category ranked by number and percentage for first year and third year essays.....	47
Table 4.5 Summary of English noun plural formation.....	57

Abbreviations

BAWE	British Academic Writing of English
BNC	British National Corpus
CAH	Contrastive Analysis Hypothesis
CNKI	China National Knowledge Infrastructure
EA	Error Analysis
EAC	English Across the Curriculum
EAP	English for Academic Purposes
EFL	English as a Foreign Language
ELF	English as a Lingua Franca
ESL	English as a Second Language
ESOL	English for Speakers of Other Languages
HLTMM	Hospitality, Leisure and Tourism Management
IELTS	International English Language Testing System
IL	Interlanguage
L1	First Language
L2	Second Language
NL	Native Language
PRC	People's Republic of China
TL	Target Language
TOEFL	Test of English as a Foreign Language

Note

The data in this study come from the British Academic Written English (BAWE) corpus, which was developed at the Universities of Warwick, Reading and Oxford Brookes under the directorship of Hilary Nesi and Sheena Gardner (formerly of the Centre for Applied Linguistics [previously called CELTE], Warwick), Paul Thompson (Department of Applied Linguistics, Reading) and Paul Wickens (Westminster Institute of Education, Oxford Brookes), with funding from the ESRC (RES-000-23-0800).

Chapter 1. Introduction

1.1 Background

Chinese students are an important part of the educational sector of many western economies, not least from the point of view of the economic impact they make to the universities, local businesses and tourism. Their success at university depends on their academic writing skills and the responsiveness of universities' support services to identify their main writing errors and develop strategies to address them. The importance of Chinese students to the economy is evident in the UK as the following statistics indicate.

International students studying in the UK contribute substantially to the UK economy and Chinese students make up a significant number of those international students. International students in total accounted for over 7 billion pounds in the UK economy in 2012-2013 (Universities UK, 2014) and the UK's education exports are estimated to reach 21.5 billion pounds by 2020 (Conlon, Litchfield, & Sadlier, 2011). This is the result of a substantial increase in the number of Chinese students studying in British universities in the last decade (Chuang, 2005; Leedham, 2011). Chinese students made up the largest number of international students in the UK (Universities UK, 2013) and for the UK academic year 2013-2014, students from the People's Republic of China (PRC) totalled 87,895 (UK Council for International Student Affairs, 2015). While various sources on the numbers and the financial impact of international students in the UK differ due to a number of factors such as: rounding up of statistics, different definitions of 'a student', education level under consideration, and the base year prices of predictions, there is nonetheless strong evidence that the Chinese student market is a valuable one for the UK.

The Chinese student market is important to the UK economy, however to succeed at university the students need to have the academic writing skills required for

success and the English language skills which form an integral part of that. In many cases the international students do not have those skills, according to many UK university academics (Parr, 2015). Parr (2015) reports that problems range from claims that students are being set up to fail to perceptions that academic standards are being lowered in UK universities so that international students with low English language skills can keep up. Solutions such as raising English language admission levels may potentially lead to a drop in numbers and therefore income. Greater screening of students to ensure that they have sufficient language levels may have the same effect (Parr, 2015). In order to provide the students with the skills they require to succeed at university and to continue to contribute to the economy, the students require more services tailored to improve their academic writing and English language requirements so that they are not set up to fail nor are academic standards lowered.

International students contribute not just to the economy but also to the internationalisation of academic life on campus (Universities UK, 2014). Despite this, international students remain an underutilised resource for developing the academic skills, cultural competency and knowledge base of domestic students (Arthur, 2015). Language skills are central in exploiting the benefits to the economy and accessing the non-economic benefits which international students bring. One key language skill is in the area of academic writing.

Academic writing is vital for success at university for all students, international and domestic (Leki, & Carson, 1994). For international students the difficulties are often more challenging because the academic writing which is required is in a second language, further complicated by cross-cultural factors (Leki, 1996, as cited in Hu, 2007). Chinese students are among those international students most likely to experience greater difficulty adapting to Western university cultures (Wang, Newton, Matsuo, & Pascoe-Chavez, 2013). Academic writing skills are important not only for academic success but as a part of many students greater objectives. Many students undertake international study as a strategy to increase their chances of employment (Brooks, Waters, & Pimlott-Wilson, 2012). Academic writing skills are especially important for

Chinese students because in China acceptance of academic work in international journals, where the publication language is English, is a necessary requirement for success for Chinese academics, and in some Chinese universities this also applies to PhD students (Cargill, & O'Connor, 2012; Xu, 2012). Universities in the UK are keen to provide services which can assist international students to succeed at university. Language support programs, including academic writing services, are key considerations in Chinese students' success in UK universities as is the motivation to improve Chinese students access to other services which impact on their academic success (Krishnamurthy, & Kosem, 2007).

1.2 Operationalization

Conducting a needs analysis is the first step in developing strategies to address students' writing issues by collecting and analysing information to develop a curriculum (Hyland, 2006). The source of the writing in this study is discipline-specific essays. Discipline-specific essays refers to writing that students have produced as a part of their academic course. This specifically excludes writing produced as part of the EAP (English for Academic Purposes) course, either for diagnostic purposes or for writing practice. Hyland stresses that a needs analysis is a continuous process and the needs analysis should be regularly re-visited to ensure relevance (Hyland, 2006).

Analysing errors as a means of identifying students' needs has been one of the main approaches of second language learning since the early 1970's (Dulay, Burt, & Krashen, 1982). Error Analysis (EA) provides a procedure for collecting information on errors, identifying errors and correcting them (Corder, 1967). EA continues to be a useful tool, although identifying errors is not always easy (Ellis, 1997). Once the errors have been identified, the underlying causes of the errors need to be identified before strategies are developed to address the errors (Corder, 1973).

One of the early theories which aimed to identify the cause of errors was Lado's (1957) Contrastive Analysis Hypothesis (CAH) which was designed to identify the difficulty or ease with which a speaker of one language could learn another language. This was done by comparing the structures of the two languages and assessing their similarity or difference. The theory was greatly criticised in the past for its lack of predictive power and was generally dismissed, however the theory is still useful especially in its explanatory, or *post hoc*, role (Sheen, 1996; Swan, 2007). This will be expanded on in the section on background literature.

1.3 Summary of previous research

There is an emerging issue with the level of Chinese students' English language skills when they study in the UK. In an effort to provide support for international students many UK universities offer a pre-sessional language course in EAP. Despite the effectiveness of these pre-sessional courses, a third of university lecturers surveyed in the UK believe that international students do not have sufficient language skills for their courses (Parr, 2015). In addition to being a gateway to better employment and academic success, language skills are also linked to the ability to adapt to a host country by contributing to the life on campus and in the community generally, which in turn has an effect on international students' perceived abilities in English and consequently their performance in academic writing (Wang, et al., 2013).

Previous research investigating Chinese undergraduate students' academic English writing is varied and has included: foundation program writing in the UK (Chuang, & Nesi, 2006), EAP writing in the UK (Chuang, 2005), EAP in Hong Kong (Evans, & Green, 2007), university students' English writing in China (Li, & Yang, 2010; Zheng, & Park, 2013), analysis of features of coursework essays in the UK (Leedham, 2011), students perceptions of EAP writing in China (Cai, 2013) and Chinese students' improved test scores from a specialised course in Singapore (Hu, 2007).

The results of previous studies do not entirely support each other although many studies agree that the main errors which occur are related to the use of the English article system (Chuang 2005, Chuang, & Nesi, 2006; Zhang, & Xie, 2014), plural suffix –s errors (Zheng, & Park, 2013) and word choice, verb form, missing subject and verb tense (Tan, 2007). Other results will be discussed in more detail in the Literature Review. Of note is that previous research has not focused on analysing grammar errors in Chinese undergraduates' coursework, that is, on discipline-specific texts.

1.4 The Goal of the Study and Expected Outcomes

The goal of studying Chinese undergraduate students' grammatical errors in academic writing in English in UK universities is to discover what the errors are and develop strategies to deal with these errors resulting in improved student access to inclusion in the academic world, greater success in achieving their goals of better employment prospects, the opportunity to contribute and learn from their campus environment and local social setting and the continued benefit to the UK economy (and potentially other countries). This chapter has introduced the setting for the study, outlined the main linguistic theories involved and mentioned the most relevant findings from previous research.

I will review the existing relevant literature in Chapter 2, initially focussing on the theoretical background by outlining the theories of EA by Corder and the CAH by Lado, including a brief overview of what is meant by the term 'Chinese language'. Then I will review the relevant research which has been conducted in relation to Chinese students' academic writing in English in tertiary education with an explanation of EAP in China to include the cultural and educational background of the Chinese students. This will highlight the gap in the research which has not included discipline-specific coursework texts. In Chapter 3 I will detail the rationale for identifying an appropriate database and define the term 'Chinese students'. I will provide an explanation for the choice of error

category and identify reasons to refine the data selection. The results will be reported in Chapter 4 by following the EA guidelines, and using the CAH procedure I will identify previous research which could explain the linguistic causes for the identified errors. I will identify the significant findings for further discussion in Chapter 5. In the final chapter I will also discuss some of the limitations of the current study and the implications of the results for future research, service provision and teaching practice. Throughout I will use the pinyin system with diacritic markers to indicate the tone for Chinese examples. This is explained further in the next section and in more detail in Appendix 1.

Chapter 2. Background Literature

In this chapter I will discuss the main theoretical frameworks relevant to this study. These are Corder's (1967) Error Analysis theory and Lado's (1957) Comparative Analysis Hypothesis. Although Lado developed his theory before Corder developed his framework of analysis, they are presented in this order to be consistent with the order of their application in the study. Then I will outline the key elements of EAP in China to highlight the general educational and English language learning experience of Chinese students. Following this I will review the relevant previous research in this area and the main findings on research into Chinese students' errors in English academic writing. I will begin with a brief overview of what is meant by Chinese language.

2.1 An Overview of Chinese Language

'Chinese' is a generic term for languages spoken in China but it can cause confusion due to the several different terms used to refer to Chinese language by different people (Hannas, 1997). According to the Chinese National People's Congress in 2000,

"For purposes of this Law, the standard spoken and written Chinese language means Putonghua (a common speech with pronunciation based on the Beijing dialect) and the standardized Chinese characters." (Article 2. Law of the People's Republic of China on the Chinese Language, 2000).

This definition avoids using the Western term 'Mandarin', which used to refer to the language used by scholars and officials at the imperial court (Sanders, 1987) but has now generally undergone semantic shift and means the same as 'Putonghua' (Mair, 1991). To use the term dialect raises the point that Chinese languages are often called

dialects in English when in fact they do not meet the definition of a dialect (Crystal, 2013). This is because many of them are not mutually intelligible. Other terms are *hànyǔ*, *guóyǔ* and *zhōngwén*. Although my instinct is that these terms are largely interchangeable, Mair (1991) suggests that there may be intricate discrepancies, although he does not explain them further in any detail.

In spoken communication with people from different regions, Chinese people speak with varying levels of proficiency in Putonghua, often heavily accented by their local variety. The difference in spoken languages in China can be highlighted by Mandarin which has 5 tones and Cantonese which has 9 tones and are mutually unintelligible. The differences in pronunciation and the differences in tone do not have a direct relevancy on the main focus of this thesis, which focuses on writing. The differences in Mandarin and Cantonese may be of relevance to the relationship between being able to communicate in spoken English and accessing support services and interacting with the local community, both in academic and in daily life, which is related to a students' confidence and subsequently with academic performance. However, any effect of differences in spoken dialect of Chinese and its affect of academic performance is not the subject of this thesis. It is English academic writing which is being considered.

In writing, Chinese characters (*hànzì*) are generally considered the same for all Chinese speakers but are distinguished by Traditional Chinese and Simplified Chinese. Traditional Chinese characters are used in Taiwan and Hong Kong. Simplified characters are used in Mainland China, although they are making more of an appearance in Hong Kong daily life and there are occasions when traditional characters are still used in Mainland China. The situation is more varied in Singapore, Malaysia and among migrant communities around the world. The difference is that the simplified characters use fewer strokes. There is on-going debate about whether pinyin will replace characters and there are currently some calls for the re-introduction of traditional characters to Mainland China, which use arguments about maintaining traditional culture, building ties with Taiwan, and aesthetics. In the context of this study, character use does not

significantly affect the inclusion of students from either a traditional or simplified character use background because the system is not different between the two, only the number of strokes required to write.

The inability to communicate in spoken Chinese, but having a similar writing system, results in some people resorting to ‘hand writing’ in areas which share a character writing system, that is tracing the mutually intelligible characters onto the hand of their interlocutor when the spoken language barriers prove insurmountable. Native Chinese speakers have informed me that with the adoption of technology such as mobile phones, this is less likely to occur in modern times because people will type the character into their mobile phone. Where characters are not used, pinyin is the most widespread system for using the Roman alphabet to represent Chinese characters both in China and for Western learners of Chinese (Bassetti, 2007). Although sometimes represented without tone markers, pinyin usually includes the tone to fully represent the pronunciation. A brief outline of pinyin and tone markers is included in Appendix 1.

Although Leedham (2011) and others generally consider Chinese to be an ideographic language, DeFrancis (1984, as cited in Wu, 1991) refutes this and provides evidence to demonstrate it is more accurately considered as a phonetic language. I will use the term Chinese to refer broadly to the Standard Chinese, or *putonghua* referred to in the Chinese law, and include varieties of Chinese some of which may not be mutually intelligible (such as Cantonese, or Yue, sometimes called *Guǎngdōng huà* by Chinese).

There is on-going debate about what to call Chinese in English given that there are many languages and dialects in China and throughout the world which are all called ‘Chinese’ but which may be mutually unintelligible. Although there have been some moves to call it ‘Putonghua’ or ‘Mandarin’ to distinguish between ‘Cantonese’ or other varieties, ‘Chinese’ still seems to be the generic default term used, with options for clarification depending on the context. This is seen in the data used in this study where the majority of the respondents stated their L1 as ‘Chinese’ without further specification.

2.2 Error Analysis

The practical reality of time constraints means that teachers need to focus on the problem areas which students are facing in order to be efficient and effective as teachers (Evans and Green, 2007). Identifying what these problem areas are in a Needs Analysis is an important first step and is an integral part of EAP (Hyland, 2006). Corder (1967) identified two essentially different kinds of errors and describes them as the difference between errors of performance and errors in knowledge. Errors of performance he called *mistakes*, which may be due to a number of reasons, such as distraction, stress or momentary forgetfulness, while errors of knowledge are systematic and these are called *errors*. Native speakers may also make statements which could be considered to be inappropriate to the situation, or unacceptable, but these mistakes, or lapses, are not due to a 'breach of the code' as Corder calls it, and so they are different to the errors which language learners may make due to incomplete knowledge (Corder, 1973, p.259). In practice, it is not always easy to determine whether an error has been made because of the difficulty in deciding whether the statement is acceptable or not. This depends in some part on being unable to determine what the writer actually meant as opposed to what they have expressed, unless the writer is available and able to explain. Once a writing sample has been assessed as incorrect, it is still sometimes difficult to determine whether an *error* or a *mistake* has been produced (Ellis, 1997).

Focusing on errors is more useful than focussing on correctly produced language because it is difficult to determine whether the correct production of a language form reflects an accurate representation of what the learner knows or whether the learner is only repeating set phrases (Corder, 1967). The study of errors is further useful in three ways. Firstly it helps the teacher to assess what remains to be learned, secondly it assists linguistic researchers in determining how language is learned and thirdly it is a tool for language learners to test hypotheses about the language they are learning (Corder, 1967). Whether language learners do in fact form hypotheses about language

learning is further discussed by Corder (1967) and he claims that the main hypothesis that learners test is whether the L2 systems they are learning are different from their L1 or not and, if so, how they are different. This means that language learners are 'testing' their hypothesis in an informal way. An English speaker learning Spanish might, for example, attempt a statement in Spanish with the adjective in front of the noun and notice that it is not correct and that the adjective comes after the noun. The hypothesis is that the two languages do not share the same placement of the adjective in relation to the noun, ('do they differ?', - yes they do) and that the adjective appears after the noun ('how do they differ?' - adjective after noun).

Corder used the term 'transitional competence' to describe the state of the learners' language learning at one point in time, (Corder, 1967, p. 166) which he developed into the idea of an 'idiosyncratic dialect'. This represents an individual's state of knowledge of a language and is distinct from the concept of *idiolects*, which have characteristics which are similar to those of other members of a closely related social dialect (Corder, 1981). His idea is similar to Selinker's (1972) concept of interlanguage (IL). Interlanguage represents the state of being between the learner's Native Language (NL) and the Target Language (TL), while drawing from both, to create the interlanguage (IL). Interlanguage has been a useful way of referring to the speech acts a speaker makes when learning a foreign language but Henderson (1985) states that the concept does not predict anything and therefore is not useful as a hypothesis. Nonetheless interlanguage continues to be used as a way of conceptualising the internal processing of language learning (Ellis, 1997).

Ellis, Loewen, and Erlam (2006, p. 351) point out that the teacher does not always only analyse errors after a lesson, but that often the language feature being taught has been taught several times before and feedback offered. They use the term 'partially mastered' to reflect that the error items are not always being taught, or assessed, for the first time. This is especially true for Chinese students once they have reached a UK university EAP program.

Corder (1974) outlined three stages in the process of EA which are: Recognition, Description, and Explanation. Lennon (2008) points out that Corder actually also included the selection of the corpus as a preliminary step in his 1974 publication. Also implicit in Corder's 1974 work is that one of the aims of an EA is for the correction of the errors once they have been identified. So Corder's EA framework can be summarised as having 5 steps which follow a logical order:

1. Selection of Corpus
2. Recognition (or Identification) of Errors
3. Description (or Classification) of Errors
4. Explanation of the Causes of Errors
5. Correction (or Remediation) of Errors

Corder recognised that there are errors which are obvious errors, 'overt errors' and these are easily recognised as being errors. There are also errors which occur in complete and acceptable sentences but do not mean what the writer intended, and these are called 'covert errors'. Corder gives the example of a student who produced the sentence: "I want to know the English". This is a grammatical sentence if the meaning relates to getting to know the English people, however if the student intended to convey that the underlying desire was to get to know more about the English language, then there is a grammatical error (Corder, 1973, p. 272-273).

Criticism of EA

Lennon (2008) outlines issues with 4 of the steps in EA, which are briefly summarised here. Firstly, the 'Recognition of Errors' is problematic for Lennon because the distinction between errors and mistakes is often difficult to determine and is dependent on external factors such as the situation and student anxiety levels. Furthermore, native

speakers' judgements of what is acceptable or erroneous are not consistent (Lennon, 2008). This is recognised by Corder and is one of the main points he makes as an assumption to be borne in mind when conducting an EA.

Secondly, the 'Description of Errors' is difficult, especially when errors occur in the same sentence and it is not clear exactly which element is to be classed as being an error (Lennon, 2008). Burt and Kiparsky (1972, as cited in Hendrickson, 1978) state that errors at the sentence level cause greater problems in comprehension and they are referred to as 'global' errors (as opposed to 'local' errors which do not affect the comprehension of the overall text). Burt (1975, as cited in Hendrikson, 1978) recommends that teachers focus on global errors because once these errors have been corrected they can change a sentence from being incomprehensible to comprehensible. Then local errors can be dealt with later. Although the description of errors is certainly difficult at times, this is not always the case, and while it may be difficult to address an error if it not clear what the error is, it also may be the case that the overall category of error can be identified and addressed comprehensively. For example, in the sentence 'Dogs like running' it may be problematic at times to determine whether a student has correctly used the zero article, is avoiding using articles, or may be using articles inconsistently throughout an essay. Nonetheless, if inconsistent article use is noted it can highlight that article use is an issue to address.

The third point he makes is that the explanation of the 'Causes of Error' are speculative, as errors could be caused by a range of factors, for example: L1 interference, learning strategies and intrinsic difficulty of the structure being attempted (Lennon, 2008). As Lennon indicates, speculation is not completely reliable but I note that Corder does state that it may be difficult at times to account for causes of error (Corder, 1973). Finally, the 'Remediation of an Error' entails an evaluation of the gravity of the error, which is problematic because there is wide variation of judgement of the seriousness of errors, depending on the person making the judgement (Lennon, 2008). Studies in the UK and the USA support this as they show that there are differences in the way in which university lecturers assess the gravity of students' errors, depending

on factors such as age, discipline (or field of study) and the native speaker status of the lecturer (Janopoulos, 1992; Porte, 1999; Santos, 1988; Vann, Meyer & O'Lorenz, 1984).

Lennons' criticisms do not have significant problems for Corder's EA framework. Corder states that after errors have been identified, categorised and the causes suggested, there should be efforts to address them. If the assessor takes the decision that any particular error which has been identified is not serious enough to warrant further effort then that decision rests with the assessor and not Corder. If there is not a universal error gravity rating this does not prevent universities from creating standardised marking and does not mean that errors which are considered as requiring rectification, should not be addressed.

There are further criticisms. Dagneaux, Denness and Granger (1998) identify 5 main issues with the EA approach. Firstly, EA is always based on heterogeneous data. Heterogeneity of data can be difficult to deal with because there are natural differences that exist between people. I think it is the degree of disparity within a group which is important and that it also depends on what each researcher wants to do with the data. It is better to ensure that the differences in subjects, or data, are clearly noted and the importance of any differences are assessed. Corder addresses this issue and states that homogeneity is a 'more or less thing' (Corder, 1973, p. 264) and concludes that there can only be a reasonable degree of homogeneity.

The second issue which Dagneau et al. raise is that the EA categories are always fuzzy, that is, they are not well defined, they are subjective and that often hybrid categories are formed. Uniformity of categorization is appealing and there may be some validity in creating a single taxonomy of error category. I agree that it is problematic for consistency between studies, but it is necessary for the researcher to define their categories and any reasons for creating hybrid categories to suit the aim of their study. If the categories in a study are well defined, this does not significantly challenge the usefulness of the EA approach.

Thirdly, EA does not capture data on some linguistic phenomena, such as avoidance. Dagneaux et al. (1998) mention this as an issue which their computer

software can be used to solve. Their software does make it easier to deal with, however it is not impossible to address without software.

Fourthly, EA focuses on what the learner cannot do rather than focussing on what they can do. The fourth issue seems to have ignored one of the main assumptions which Corder makes explicitly about EA, which is that it does not focus on what a learner may have learnt because what is assumed to have been learnt may be just repeating set phrases or may be an acceptable utterance in itself but does not express what the student truly intended (Corder 1967, p168). Dagneaux et al. (1998) have perhaps included this criticism because their software is able to identify correct usage as well as highlight errors, however they do not recognize Corder's original disclaimer.

Finally Dagneaux et al. (1998) state that the results of EA give a static picture of the learner's ability. This is a valid point which is inherent in the type of tool that it is, but it does not mean that EA could not be used in longitudinal studies. This is demonstrated by Ellis (2003, as cited in Castillejos Lopez, 2011). A more valid limitation of EA is that when applied to a corpus it does not take into consideration individual learning. Although a particular group may have a certain commonality of errors, the distribution of errors across the data is not considered. If distribution is widespread, then some students may not have any problem with an error while others may have a more serious problem with it. This is explicitly addressed by Corder who points out that language classes most often occur in groups and that teaching is directed towards groups (Corder, 1973). Acknowledging these limitations means that Corder's 5 steps in EA are still useful in identifying errors in student writing.

This is supported by literature relevant to EA. After outlining some of the claimed shortcomings of EA, Ellis (2008) states that "Nevertheless, it has continued to figure in the study of L2 acquisition" (p. 45). A handful of studies published over the last two years (2014-2015) demonstrate this (cf Carrio-Pastor & Mestre-Mestre, 2014; Presada & Badea, 2014; Tizazu, 2014; Adjei, 2015; Phuket & Othman, 2015; Moqimpour & Shahrokhi, 2015). EA is considered to be useful less as a theory of linguistics but more as

a procedure or tool and, due in part to its resurgence in computer-based analyses of language, it helps to decide which L2 features to teach (Ellis 2008).

Lightbown and Spada (2006) point out that “...error analysis has the advantage of describing what learners actually do rather than what they might do...” (p. 82). Ellis and Barkhuizen (2005) put the argument that although EA is not the most favoured approach for researching L2 learning it is “still alive and well” (p.53). It is in its procedure for analysing errors that it is most useful and will be used in this thesis.

2.3 Contrastive Analysis Hypothesis

Lados’ (1957) CAH aimed to facilitate L2 learning by comparing the structures of the L1 and the L2 and identifying the difficulty or ease with which the L2 should be able to be learned based on the differences or similarities between them. The teacher is then able to target the difficult areas as appropriate to the situation. It focused on comparison between five areas: sound systems, grammatical structures, vocabulary, writing systems and culture (Lado, 1957). Lennon notes that vocabulary and culture were later dropped from detailed consideration over time (Lennon, 2008) but he does not propose a reason for this. Janda (2008) states that culture and language are treated as different fields of study in academia, and this might suggest part of the reason, however I think that Janda makes a very general statement and it is not supported by any statistics or examples.

The focus of this study is on written grammar errors and therefore I will not detail the comparison of sound systems or vocabulary. A comparison of writing systems is also not necessary because Lado uses Chinese and English as an example of a situation where the writing systems are so different as to their being almost no transfer from Chinese to English (Lado, 1957, Kramsch, 2007). In addition, Chinese students who have reached the level to be accepted into a foreign university will have also learnt the pinyin system of writing which means that the orthographic challenges of writing English are not major issues by the time the student becomes an undergraduate. In Mainland China

pinyin is taught to students either in late kindergarten or early primary school (Jing, Tindall, & Nisbet, 2006). Therefore a comparison of writing systems will not be of benefit in identifying written grammar errors.

Comparing two cultures is more complex. Lado provides the example of how Spanish has different words for 'legs' for animals and humans, (*patas* for animals and *piernas* for humans) while English uses the same word for both (Lado, 1957, p. 116). Lado however does not discuss specialised word usage such as 'hoof', 'paw' or 'trotter'. He states that the division in word usage provides an insight into how animals and humans occupy different places in the two cultures. It is not just language which reflects the differences between cultures, Lado illustrates the differences behind Spanish peoples' attitudes and Americans' attitudes to bullfighting to show how culture affects perception (Lado, 1957). As will be discussed in the findings of previous research, culture has a significant effect on writing.

The Comparative Analysis Hypothesis framework for comparison of grammatical structures is applicable in this study. Lado states that 'grammatical structure' refers to *form* linked to *meaning*. There are specific features which indicate meaning and relationships. He provides the example of adding a suffix, -s (the form) to a singular noun to make it a plural noun (the meaning) and that by omitting that suffix -s (form) it has the meaning of being singular (Lado, 1957, p. 53). Lado gives a number of further examples which are designed to demonstrate that it is not only the use of a grammatical feature which affects the meaning, but that its omission changes meaning. It is these features which together form a grammatical structure (Lado, 1957). Lado elaborates with the example of the English definite article 'the'. It does not change whether plural or singular, or referring to masculine or feminine. He contrasts this with the Spanish equivalent of the English definite article 'the' which changes for both gender and number, becoming *la* and *el* for feminine and masculine singular forms respectively, and *las* and *los* for feminine and masculine plural forms. Lado does mention the neuter form *lo* but does not discuss indefinite articles, where there is a similar distinction for gender and number, *un/una* for masculine/feminine singular and *unos/unas* for

masculine/feminine plural (Cazalaa, Cabot, & Palat, 1993). In each case the masculine plural is used to refer to mixed gender groups. These grammatical structures together form a complicated system and according to Lado it is impossible to calculate each one and its interactions with all the other grammatical structures while speaking or writing. He concludes that therefore they must become habitual for them to be used effectively (Lado, 1957).

As an example of the difference in grammar between the two languages, English grammar is inflectional and uses inflectional morphemes to indicate plurals, tense, case and person (in the sense of verb conjugation) (Plag, 2003). Chinese is not inflectional and uses word order and function words to indicate grammar (Jing, Tindall, & Nisbet, 2006). In the following example, the function word *le* can indicate simple past tense by attaching to the end of the verb. Unless otherwise indicated, all Chinese examples are invented by myself and checked with at least two native Chinese speakers.

- (1) *Wǒ mǎile yī běn shū*
 (I buy (past tense marker) one (classifier) book
 'I bought a book.'

Having outlined the basic assumptions of the grammatical structure, Lado lists six problems that he views are involved in learning a foreign grammatical structure:

1) *Transfer*. Sentence forms, number and other grammatical structures from the L1 are applied to the L2.

2) *Similarity/difference to determine ease (or difficulty) of learning the L2*.

The assumption is that the more similar two languages are to each other the easier it will be to master the L2, however in the case where the two grammatical structures are completely different, the new language structure will be difficult to learn.

3) *The difference between comprehension (recognition) and production.*

Language learners may be able to understand an utterance in the L2, showing comprehension of the grammatical structure, but not be able to apply the same grammatical structure when speaking or writing. An English language learner may well be able to understand the meaning of the simple present third person singular when it is read or heard, but they may be unable to produce it in writing or speech.

4) *Difference of form.*

A grammatical structure may have a similar meaning in both languages, such as the concept of a subject of a sentence, but have a different form. The form that marks the meaning can differ in two ways. If the word is changed by the same process, such as adding a function word for example, it will be a less difficult difference to master. It will be more difficult to master if the change is marked by inflection in one language and a function word in another, or if word order versus intonation marks the change. Here it is explicitly assumed that the meanings in each case are the same or similar. Lado provides the example of the English 'Who came?' and the Spanish equivalent *¿Quién vino?* To show that the function word occurs in the same position and functions in the same way, therefore the prediction is that this would be a relatively easy form to learn (Lado, 1957, p. 60).

5) *Difference of meaning.*

In the example above, the meanings are assumed to be similar. This is not always true because the meaning or concept in one language may be different in another. Lado's example is that for plurals in many languages, such as in English, number is divided into singular and plural, but in other languages the concept is expressed in singular, dual or plural (Lado, 1957, p. 65).

6) *Difference in distribution.*

This refers to the distribution of the form through a phrase or sentence. If a structure is similar in distribution it will be easier to learn, however if the distribution is different then it is more difficult to learn. Lado's example, again taken from Spanish, is that the plural marker must apply to the article, adjective and noun, whereas in English it only applies to the noun. For example, English is changed by adding the plural marker -s to the noun (in most cases), while in Spanish the definite article is changed to *los*, the noun receives the plural marker -s and the adjective receives the plural marker, -s (for nouns ending in a vowel) or -es (nouns ending in a consonant). This is shown in my invented example in Table 2.1 below.

Table 2.1

Distribution of plural markers for English and Spanish

	English	Spanish
Singular	The blue bird	El pajaros azul
Plural	The blue birds	Los pajaros azules

After identifying the six problems in learning a grammatical structure, Lado explains his General Procedure for comparing two grammatical structures and then provides more detail in the three-step Specific Procedures. The General Procedure is to analyse the L2 and compare it with the L1, structure by structure. For each structure in the L2, the procedure is to identify if there is a similar structure in the L1. In each case the assessor should determine whether the structure is indicated in the same manner, has the same meaning and similar distribution. The specific procedures state that the first step is to identify a structural description of the languages being compared covering *form, meaning and distribution*. The Second Step is to make a summary of the types of structure once they have been identified in Step One. An example of this would be to

list all the different ways of making a question and then incorporating them into a type of structure called 'Questions'. The Third Step is to make the actual comparisons between the two structures in the two languages. The examination of *form*, *meaning* and *distribution* allows the assessor to determine the difficulty the L1 speaker faces when learning the L2.

Lado states that where there are similar items in a grammatical structure these can be re-grouped into broader categories. To summarize the example given by Lado, he explains that the problems for a Spanish speaker learning English will include using the function word 'do' and the reversal of word order in question making because these are features which are not present in Spanish question making. This only considers writing and therefore does not include such aspects as intonation to form questions. The difference between making a statement into a question in English and Spanish is summarised in Table 2.2 in an invented example which was checked by consulting with native Spanish speakers from Spain and Latin America.

Table 2.2

Word order change in question formation in English and Spanish

	English	Spanish
Statement	You (do) like chocolate	<i>Te gusta el chocolate</i>
Question	Do you like chocolate?	<i>¿Te gusta el chocolate?</i>

This shows the word reversal required for making a question and inclusion of 'do'. In Spanish the statement does not have any changes to word order or insertion (Lado, 1957). For Spanish speakers the written English form requires the omission of the inverted question mark at the beginning of the sentence. Similarly word order reversal is used for making questions using 'can' in English and therefore reversal of word order for 'can and 'do' can be grouped together in Lado's system. This is a brief summary to show the idea behind grouping grammatical structures into larger patterns for evaluation of ease or difficulty of learning.

To briefly summarise other aspects of Lado's theory, Lado identifies issues with dialect differences and style issues. Lado discusses both written and spoken language and includes written genres such as poetry, plays and newspapers. Academic writing is a specialised style of writing and varies somewhat between disciplines and lecturers' preferences. Lado concludes that the initial analysis, or comparison between grammatical structures, should be checked with the students' actual ability because a problem may be more severe (or less apparent) than predicted by the Contrastive Analysis. Here Lado actually allows somewhat for individual variation. Differences between students will mean that the extent of the problem will be different from one student to another within the same language background, but generally the problems can be predicted (Lado, 1957).

The importance of culture in general terms of being able to understand each other across cultural differences includes the role that culture plays in language learning and teaching (Lado, 1957). His basic definition of culture is "a variety of ways to live" (p. 8), as being the same as the "ways of a people" (p. 110) and more formally as "structured systems of patterned behaviour" (p. 111). Lado's explanation uses examples such as different breakfasts from around the world to show how the meaning of what is considered breakfast can be different. He suggests analysing culture using the same pattern of *form* (here, a unit of behaviour), *meaning* (the connotation or nuance of significance) and *distribution* (the time and place of behaviour) which he utilised for analysing the other 4 aspects of language. Lado proposes guidelines for collecting cultural data, such as interviewing articulate and introspective members of a cultural group, systematic observation of activities, and testing perceptions of variation of behaviour to observe the result. Culture has an important role to play in academic writing. For Chinese students their thought pattern is affected by their culture which in turn has an effect on their L2 writing (Xu, 2012). If the difference in culture between China and the UK is large, it can be expected to be difficult for Chinese to write using English academic norms.

Criticism of Contrastive Analysis

Jia and Tian (2012) argue that linguistic comparative analysis occurred in China long before Lado's work. They document examples of comparative analyses of language at a diachronic, synchronic, intralingual (changes within a language) and interlingual (changes between languages) level in China since 220 AD. Despite their strong evidence to support this, it is not until Lado that there is a theoretical framework which attempts to allow for prediction and provides step-by-step procedures and guidelines.

Wardhaugh, as summarised by Yang (1992), stated in 1970 that there is a strong version of the CAH which maintains the predictive power of the CAH and the weak version, which is useful to account for observed difficulties in language learning. These are called *a priori* and *a posteriori* versions of the CAH respectively. According to Yang there was a moderate version also posited in 1970 by Oller and Ziahosseiny which included the consideration that where concepts are more similar but still with minimal difference, they may in fact be more difficult to master due to confusion caused by that minimal difference (Yang, 1992). The strong version has been abandoned, mainly due to the lack of predictive power in practice, and the moderate version does not seem to have been taken up by researchers.

Yang (1992) published a comprehensive review of the studies supporting the value of the CAH (Broselow, 1984, Erdmann, 1973, Lehn and Slager, 1959, Rivers and Temperley, 1998, Wardhaugh, 1970) and those which were not supportive (Brown, 1987, Dulay and Burt, 1972, Hughes, 1980, Klein, 1986, Lance, 1969, Noblitt, 1972, Wardhaugh, 1970 – against the strong version, Whitman and Jackson, 1972, among others). The conclusion is that many of the claims against the CAH can be shown to have not fully understood the precepts or goals of the CAH or can be successfully counter-argued, chiefly by James (1985, as cited in Yang, 1992). Despite this, the CAH remains generally unpopular.

Kramsch reports that it is well recognised in the field of applied linguistics that the CAH was linked to behaviourism, which became discredited in linguistics.

Behaviourism and, by association, the CAH, became unpopular as a result of the move away from structural linguistics to cognitive linguistics and the accompanying thinking that language was not a finite list of rules but was infinitely complex. This movement was spearheaded by Chomsky's attack on Skinner's publication *Verbal Behaviour* (Kramsch, 2007). The criticism was refuted by MacCorquodale (1970) but despite this the CAH's popularity lessened. Sheen states that the CAH held the title of "*persona non grata*" (Sheen, 1996, p. 183) and Swan goes even further as to hint at deliberate bias among academics who were more interested in promoting their own views than accurately representing those of the CAH (Swan, 2007). Swan (2007) summarizes the main criticisms of Lado's CAH. Critics claimed that there was a degree of 'Overprediction' which Swan argues was not borne out by observation. He discusses the example of native English speakers having problems with learning the difference between Spanish *conocer* and *saber*, but Spanish speakers not having the similar problem when learning the English 'to know' (Odlin, 1989, as cited in Swan, 2007). Swan responds that Lado did not ever claim that there was bi-directionality in the CAH. The opposite claim was of 'Underprediction'. Here, the criticisms generally point to there being many mistakes made by language learners which are not predicted. Swan (2007) reports that Lado did not claim that all errors will come from L1 interference and so cannot necessarily be predicted. Yang also addresses this issue specifically by reporting that Lado did not specify that all errors would be predicted all the time (Yang, 1992).

The conclusion seems to be that Contrastive Analysis has seen some revival and is still considered a useful tool but less as a predictive tool and more so in its weaker version in a retrospective explanatory role. It has been given greater attention recently especially in conjunction with EA and there are suggestions that there is a future for the CAH. (Rustipa, 2011; Swan, 2007; Duskova, 1969, as cited in Yang, 1992).

According to Saville-Troike (2012) there has been a revival of contrastive analysis which has expanded in scope to include genre analysis and translation studies. This is supported by a number of publications based on the CAH in the 21st Century by: Granger, (2003), Degand (2004), Peterlin (2005), Laufer & Girsai (2008), Dervinyte

(2009), Wong & Dras (2009), Rahimpour & Dovaise (2011), Afraz & Ghaemi (2012), Gast (2013) and Roikiene (2014).

CAH remains useful in its post hoc or 'weak' version in two ways. The first is that it helps to understand the causes of the error. As Lightbown and Spada (2006) point out, EA provides a description of what errors learners make, but does not provide insights into why they do it. Yang (1992) summarizes the literature which states the usefulness of being able to gain insights into an error in the L2 by examining the L1 and from those insights being intuitive and obvious through observation. Secondly it is useful in teaching, as a way of increasing the metalinguistic awareness of the student and highlighting the difference in form, function and meaning and explicitly stating what it is that is to be learned (Yang, 1992; Tan 2007). CAH has a practical explanatory function (Rustipa, 2011).

2.4 English for Academic Purposes in China

EAP is relatively new in China and researchers are investigating what they should teach in EAP and how to teach it (Errey, & Li, 2005). The link between research and practice in EAP needs to be developed. An example is given by Li (2009), who reports the CNKI (China National Knowledge Infrastructure) statistic that there are over 150 papers by Chinese academics on the topic of fossilization, however Li claims that this topic is rarely given attention in teaching. This is due to the teachers being too busy teaching the syllabus or textbook materials to deal with latest research reports (Li, 2009). Teaching English in China has made some progress but still relies on old models of teaching which can have a negative effect on learning (Cai, & Zhang, 2013).

Few Chinese universities have established English Language Centres to provide any language consultancy, although in Hong Kong and Singapore, universities do increasingly provide such services (Cai, 2013). Only a minority of Chinese students will actually study abroad so for the majority of students a high level of proficiency in writing

is not required. Many of the English language classes are taught mostly in Chinese, due to the lack of the teachers' proficiency in English (Cai, 2013). Hyland (1997) reports that Hong Kong is monolingual and monocultural and therefore the students do not encounter English outside of the English classroom. This is true for Mainland China too, and even in the UK Chinese students complain that they do not have any opportunities to practice English due to the fact that they often share accommodation and mostly socialise with other Chinese students. The result of this is that many Chinese students enter the UK education system with little exposure to UK teaching methods and little preparedness for academic writing. The UK tertiary education system has responded to this need by implementing EAP programs designed to improve the students' English and to prepare them for academic life with study skills (Krishnamurthy, & Kosem, 2007).

There is some debate as to whether the focus of such programs should emphasize the academic skills, such as how to write an essay and conduct independent research, or target the language skills, such as grammar and vocabulary, but based on focus group discussions, Chinese students report that they dislike academic writing (Cai, 2013; Leedham, 2014). Academic writing courses in the UK are more concerned with literacy than grammar and although grammar errors are corrected, classroom time is used for discussing referencing and avoiding plagiarism. This approach does not meet Chinese students' demands for more grammar focus (Chuang, & Nesi, 2007). Hong Kong Chinese tertiary students also have greater trouble with language rather than with the academic requirements of writing, with the main problem area being vocabulary, but appropriate grammar use is also a concern (Evans, & Green, 2007). One recommendation is that EAP teachers should focus on language and content together, which will involve consultation between subject professionals and EAP teachers (Evans, & Green, 2007). They make the point that this does not mean that academic writing should be replaced with remedial English or General English, but that teaching materials should be content driven using task based approaches. Against this, it is argued that it is not the language issues which negatively affect international students but the academic requirements (Xing, Wang, & Spencer, 2008). Perhaps it is sometimes overlooked that

Chinese students do not always have an intrinsic interest in studying academic writing and are only studying English because it is required for their main area of study.

In China there are also psychological barriers which EAP students have to deal with. There is a lack of on-going interest in English because, in many cases, the students have studied English for a long time with no discernable use (Li, 2009). This may be compounded by their own lack of confidence in their ability. As they cannot make progress they doubt their learning ability (Li, 2009). There are other behaviour barriers which result in the students simply stopping putting any effort into learning (Li, 2009). The individual psychological barriers may be compounded by the requirement to learn about a different cultural system, which may be intrinsically uninteresting or simply difficult to learn. The importance of the cultural aspect of language learning however, is recognized by the Chinese Ministry of Education (2004, as cited in You, 2004) which recommends that English language teaching should include teaching of language learning strategies and cross-cultural skills.

2.5 Previous studies

Previous studies in the general area of Chinese students' academic writing at university level have focussed on Chinese students perceptions of aspects of academic writing (Cai, 2013; Cross 2006), reading strategies (Errey, & Li, 2005) applying cultural strengths in EAP (Jin, & Cortazzi, 2001), use of lexical chunking (Leedham, 2014), factors influencing confidence in EAP (Wang et al., 2013), Chinese immigrants in the USA (Lay, 1972), and academic writing in university in China (Zheng, & Park, 2013; Zhang, & Xie, 2014), Taiwan, (Tan, 2007), Singapore (Hu, 2007) and Hong Kong (Evans, & Green, 2007). Studies in academic writing courses in UK universities (Chuang, 2005; Chuang, & Nesi, 2006) have focussed on academic writing in EAP courses, not on discipline-specific texts.

Lay (1972) focussed on Chinese students at a North American university. The participants were children of Chinese immigrants, not students studying as international

students and therefore they had a family support network with many of them reporting having been born in the USA, which is a different situation from Chinese international students living away from their family and familiar environment. Some of the previous studies (Hu, 2007; Tan, 2007; Cai, 2013; Zheng, & Park, 2013; Zhang, & Xie, 2014) were conducted with students studying in universities in China, Taiwan, Singapore or Hong Kong, where the educational environment is different from the academic environment in the UK in terms of the level of independence of study and the type of performance expected from students academically (Cai, 2013).

Many of the studies were on essay topics that were not of an academic nature or did not focus on grammatical errors. The study by Tan (2007) included 96 participants who were non-English majors in a university in Taiwan. The essays in that study were students' responses to news reports posted by the researcher. These are not the same as the academic writing required by students in their coursework at university. Some studies did include students studying in UK universities but they were also not focussed on academic writing. The participants in the study by Chuang (2005) were all Mainland Chinese with IELTS 6.0 or equivalent and were in a Foundation programme at a UK university. All the topics were reported as being serious topics, however no examples were given of actual essay topics, except to mention generally that they deal with the European Monetary Union, genetic engineering, identity cards and restricting car usage. Although they were essays of 1,500 to 2,000 words (Chuang, & Nesi, 2006), by looking at IELTS websites it can be seen that these are typical of IELTS preparation topics rather than being academic topics. However, because the report did not include actual specific topics it is difficult to determine whether they are really academic essays or not. Green reports that international students themselves have voiced concerns that IELTS preparation courses do not prepare them for university academic writing (Green, 2007). The results of non-academic writing tasks cannot necessarily be applied to academic writing (Leedham, 2011).

Other studies have looked at discipline-specific texts. Leedham (2011) used data collected from students' academic writing at UK universities as part of their coursework.

However, these were analysed for features of writing such as lexical chunking, not grammatical errors. A two-year study in a Chinese university explored fossilization by starting with writing classes and then examining discipline-specific texts. However the researchers only looked at the first drafts of the those texts and not their final submitted drafts (Zhang & Xie, 2014) and as noted, the essays were not produced in a UK university environment. An associated aspect of examining discipline-specific texts is that the Chinese student studying in the UK has often spent three months in an EAP course and several months preparing their essay before submitting their first assessable essay. Previous research indicates that after a semester students have less cultural conflict and have adapted to some degree to the host culture and that this has an effect on their writing (Wang et al., 2013).

The previous studies mentioned have identified the main errors which Chinese students produce in different circumstances. They are summarised in Table 2.3 below.

Table 2.3

Summary of main errors identified in previous studies

Date	Author	Summary of Main Errors
1972	Lay	Pronoun Agreement
1984	Vann, Meyer & Lorenz	Spelling, Articles
1988	Santos	Articles
2005	Chuang	Missing Articles, Plural Errors
2006	Chuang & Nesi	Missing Articles, Plural Errors
2007	Tan	Word Choice, Verb Form, Missing Subject
2011	Leedham	Lexical chunking
2013	Zheng & Park	Omission of Articles, Omission of plural 's'
2014	Zhang & Xie	Article, Number, tense

As the above summary shows, there is no consistency between what the researchers have identified as being the main error, however there are some common errors which appear regularly in the top placing of rankings of errors, most notably article usage and number, (or errors in plural use).

Previous studies have also demonstrated the effect of culture on writing. Xing, Wang & Spencer, (2008) summarise 5 major differences in what Chinese would consider good writing style compared with English academic writing style, for example, the western 'Introduction-Body-Conclusion' structure contrasts with the Chinese 'Start-Sustain-Turn-Sum' approach, where the field or topic is established (Start), developed (Sustain), a different perspective is offered (Turn) and then the conclusion sums up the writers argument (Sum). Chinese writing has a rhetorical style which appeals to tradition, history and religion, major aspects of culture, and can therefore appear to lack structural coherence to a western reader who is looking for an analytical and logical pattern of thinking (Xu, 2012).

2.6 Conclusion

There is a gap in the previous literature which does not include studies which analyse grammatical errors made by Chinese undergraduate students at UK universities writing in English for their actual degree course. The primary aim of the present study was to determine answers to the following research questions:

1. What are the grammatical writing errors of Chinese undergraduate students in their discipline-specific texts in UK universities?
2. What are the linguistic causes of these errors?
3. How can EFL teachers address these issues?

Chapter 3: Methodology

In this section I will discuss the definition of the term 'Chinese students' and the identification of participants in this study. I will explain the source of the data and the parameters for inclusion, detail the development of the Error Correction Code, report on issues relating to the classification of errors, and give the rationale for marking by hand and not by a software program.

Chinese students

'Chinese students' is a term used to include students from China, Taiwan, Hong Kong, Malaysia and Singapore (Leedham, 2011). There are shared characteristics of the students from these countries which enable them to be considered as a single group. The first of these is literacy. The spoken language differs between people from these places but the ideographic writing system can be considered a common or shared characteristic as it allows for people to communicate in writing when verbal communication is difficult or impossible due to different dialects (Hu, 2001, as cited in Leedham, 2011; Milton, 2001) (whether it is truly an ideographic system is briefly discussed in the previous section on Chinese language). Secondly, the Grammar Translation approach, which focuses on grammatical analysis and translating between the L1 and L2, remains the chief approach to language learning. There are some inroads in the use of Communicative Language Teaching and Task Based Learning however these are still in the minority (Leedham, 2011). Finally many countries in this group share a Confucian cultural tradition. This includes shared beliefs about the value of education and the role of the teacher in society and the teacher's relationship with students (Leedham, 2011). Leedham (2011) also makes the point that students from these countries studying in the UK are likely to come from similar family backgrounds

and educational status. Given the high university fees for non–EU international students studying in the UK, it is likely that there is, for many students, a common level of socio-economic status. Therefore these students, although they may have different spoken varieties of Chinese, share a written commonality, education background commonality and a commonality of cultural tradition.

The written language can broadly be said to contribute to a degree of homogeneity as supported by Wang (1997, as cited in Wang, Tsai & Wang 2009).

Indeed, this uniformity of script and the homogeneity of the culture it promotes are the major reasons for calling Mandarin, Cantonese, and Taiwanese different dialects rather than different languages, even though their spoken forms are no less similar than the Romance languages. (p.411)

There is further support from Wang and Goodman (2014) who state that:

Because characters have the same meaning for speakers of all Chinese dialects, written Chinese is a great contributor to the unification and homogeneity of Chinese culture and played a major role in making China a single nation. (p. 627)

Norman (1988,) highlights that different spoken varieties (or dialects) share the same written form.

The Chinese language, especially in its written form, has always been one of the most powerful symbols of this cultural unity. The aptness of language as a symbol of cultural and even political unity was facilitated by the use of a script that for all practical purposes was independent of any particular phonetic manifestation of their language, allowing the Chinese to look upon the Chinese language as being more uniform and unchanging than it actually was. (p1)

Norman (1988) adds that Canton (as he calls Guangdong) and Peking (referring to Beijing) “have always used the same written language” (p. 2). He explains that this is not to say that other written forms of Chinese do not exist, but they have only been used for specific genres of regional literature (Norman 1988). Romsey (1987) states that “It is also true that when most Chinese think of a language that unites them as a people, the “common language” they have in mind is still fundamentally their written language” (p17). There is further supporting evidence from McNaughton (2005) who states that

The remarkable thing about the Chinese writing system...is that a literate native speaker of one dialect can communicate with another person of a different native dialect simply by writing down his thoughts. (p. xvii)

Leedham (2011) argued that the students in the BAWE database who have listed Chinese, Mandarin or Cantonese as their L1 can be considered to be homogenous from the point of view of their L1 written literacy. It could be argued that the group is not homogenous in terms of individuals’ sense of cultural identity, spoken language and political affiliation, however it is written academic English which is under consideration in this thesis and therefore these students can broadly be considered homogenous in respect of their L1 written literacy

Identification of database

An initial attempt to create a database based on networking with Chinese students did not result in a sufficient number of essays to create a small corpus. Of the students who responded, some did not provide the complete demographic information or requested essay grades. An internet search was conducted to identify alternate sources of students writings with the parameter that they must include students identifying themselves as having Chinese as their L1 and that the essays must be from

the students' actual majors. Databases which included essays written for IELTS exams, EAP courses and secondary school essays were specifically excluded. Essay banks, available commercially or through student associations, were not consulted because there are concerns as to their usefulness due to incomplete documentation and annotation (Nesi, 2008).

From an initial review of lists of linguistic corpora, many were discarded because: they did not include Chinese speakers, were still under development, focused on spoken language or multimedia or were country specific (e.g. Taiwan, Singapore, Hong Kong, USA). A list of 18 potentially relevant corpora was made for further investigation (see Appendix 2, which also details reasons for exclusion from the list). Of these, 17 were found to be either not discipline-specific texts (they were EAP, ESOL or TOEFL texts, business texts, such as job applications and resume writing), unable to be accessed due to a corrupted file or access being limited to academic staff of that university, essays collected from students studying in Mainland China, USA, Singapore, non university level, or a prohibitive cost. The BAWE (British Academic Written English) corpus was identified as being the only readily available source of students' essays.

3.1 The BAWE (British Academic Written English) Data Set

Permission to access the BAWE database was granted after submitting an application request to the database managers. The BAWE is a database of academic writing by both domestic and international university students, including undergraduate and Masters students studying in the UK. The database records information on a number of features such as the discipline, genre and word count, among others. The students indicated their L1 in the database but were not asked to indicate their country of birth, their nationality or status as overseas or domestic student (Heuboeck, Holmes, & Nesi, 2008). The students selected in this study all indicated their L1 as being 'Chinese' (that is, unspecified), 'Chinese (Cantonese)' or 'Chinese (Mandarin)'. Checking with two native

Chinese speakers from Mainland China (one an Associate Professor of English at a Chinese university and the other a Masters student in the UK) indicated that it can be assumed that when students stated their L1 as 'Chinese' and did not specify a 'dialect', it can be taken to be Mandarin. The reasons for including these different groups as a single group are discussed earlier in this section.

Students were paid by the database collectors in order to encourage participation from students studying in some fields which were under-represented. This means that the student contributions were all voluntary in the sense that, although they were paid, there was no obligation for them to submit their work. Essays graded at 60% or higher were included and varied in length from 500 to 10,000 words (Heuboek, Holmes, & Nesi, 2008). The BAWE was initially developed using data submitted from 2004 to 2007 as part of a research project on genres of writing in British higher education with contributions from students at the universities of Reading, Warwick and Oxford Brookes, including first year, second year, third year undergraduate students and Masters students, in 35 disciplines across 4 broad groups: Arts and Humanities, Life Sciences, Physical Sciences and Social Sciences (Heuboek, Holmes, & Nesi, 2008).

Each essay has a unique identification number to indicate which student submitted the essay. Some students submitted more than one essay and therefore a letter is added to the identification number to indicate which of the essays the student submitted. For example 0123a indicates that student 0123 submitted the essay and 'a' is to identify it as the first essay submitted. If the same student submitted a second essay it would be 0123b. This is regardless of year level. The information is contained in an Excel file which can be sorted by many features, including the student's year level and stated L1.

The essays are available in xml format (.xml) and text format (.txt). The xml format was difficult to read so a trial file was converted to Word document format. This file conversion included the entire markup (the computer coding system) which was distracting and difficult to read. The text file format was also difficult to read and so the text format files were converted to Word documents to enable ease of printing and

evaluation. This solved the markup problem however created a problem with the layout of the essays, mainly with headings, lists and spacing. Despite this, it was easier to read without the distractions of the markup notation and was therefore chosen as the form to use for the analysis of the essays.

Refinement of the Data Set

There were a number of students who provided more than 1 essay in their year level. To ensure there was no overrepresentation from any student in the corpus only one essay was included per student for the refined database. These essays were not previewed before being selected, that is, there was no 'cherry picking' of essays or selection based on the content on the essay. As mentioned above, the database administrators allocated a letter to each essay's identification number to identify it as a multiple entry. In the case where the student had provided more than one essay, in each case the essay with the letter 'a', or closest to 'a', was chosen. Four of the students whose essays were selected for inclusion in the study provided essays to more than one year level. Of these students, two students had their first year contributions retained but the third year contributions deleted. The other two students had the opposite, that is, their first year essays were not included but their third year essays were. There were 14 essays identified in the second year group but 10 of these had also submitted essays to either the first year or the third year group and therefore none of these essays were considered for inclusion because it only left 4 essays for the second year group.

A summary of information on the Chinese students whose essays were included is presented in table 3.1, including the average age of participants, stated L1 and gender. These are discussed in detail below.

Table 3.1

Summary of available student data on age, L1 and gender

	Age (average) Years Old	Number of Cantonese as L1	Number of Mandarin as L1 (incl 'Chinese')	Number of Males	Number of Females
First Year	19	7	11	6	12
Third Year	23	6	12	8	10
Total Number	21	13	23	14	22

Students' ages were not provided but students provided information about their date of birth and the date of submission of their essay. This information was used to calculate the age for each student at the time of submission of their essay. The average age for the first year group was 19 years old and for third year group it was naturally higher, 23 years old. One student in the third year group was 43 years old which would have affected the average, but only slightly. One student did not provide his date of birth and four students did not provide date of submission of essay (or it was not recorded). The average across both groups was 21 years old (all averages rounded to nearest whole number). There were 14 males and 22 females whose essays were selected for use in the study which makes a ratio of 37%: 58%. Percentages rounded up to nearest whole percent. This is approximately a 40/60 split.

Thirteen students stated their L1 was Cantonese, three stated it was Mandarin while 20 stated it was 'Chinese', that is, unspecified as to the variety of Chinese. As mentioned above, this is likely to indicate 'Mandarin'. This information is provided in detail for each student in Appendix 3.

The final data set consisted of 18 first year student essays and 18 third year student essays, covering the fields of Engineering, Biological Sciences, Economics, Business, Law, Politics, Chemistry, Mathematics, Food Sciences, HLTM (Hospitality, Leisure and Tourism Management) Agriculture, and Cybernetics and Electrical

Engineering. Generally many disciplines share the same characteristics of writing tasks and good writing skills are applicable across disciplines (Hu, 2007). The following are some examples of the essay titles for the first year essays followed by third year essay examples:

First year essays:

“Should the Dutch economy be seen as the first modern economy?”

“Humanoid Robotics in Artificial Intelligence”

“Business strategies demand discipline in the execution of long-term strategic plans and flexibility to address emergent changes. Discuss. Explain which one of two features is more critical in your view.”

Third year essays:

“Outline a version of the "first generation" or the "second generation" speculative attack model. Discuss the empirical evidence in support of the model you outlined. Briefly discuss the limitations and the extensions of the model.”

“The role of maternal effect genes in the development of the nematode *Caenorhabditis elegans*”

“Overcoming Seed Dormancy”

A complete list of essay titles is attached in Appendix 4.

The total word count was 64,565 words. Essay word counts ranged from 556 to 5693 (Appendix 5 includes the word count for each students’ essay). Aston (1997, as

cited in Krishnamurthy, & Kosem, 2007) defines a small corpus for written EAP texts as being 20,000 to 200,000. Therefore this sample falls within Aston's definition of being a small corpus.

3.2 Error Correction

The Error Correction Code

The Error Correction Code was developed by checking the essays and making a list of the errors as they were noted. Each error type was given an abbreviation. These abbreviations were adapted from common abbreviations used in error correction codes used in marking in EAP courses at UK universities, particularly from Reading University (Vicary, 2014). Although there are commonly used codes, there is no uniform agreement on error correction code abbreviations. From one perspective this is inconvenient for comparing studies, however from a practical point of view the codes need to be flexible to reflect the needs of the student and teacher using them. For example, a student may not need to be given exact information on what kind of error has been made in relation to article use, they just need to be informed that an error has been made, so in that case a code 'AR', signifying 'Article' may be sufficient. In this analysis however, it is useful to mark whether the student has made an error of omission ('MA' for 'Missing Article') or overuse of article ('RA' for 'Redundant Article') or incorrect use of definite article for indefinite article or vice versa ('IA' for 'Incorrect Article').

Based on the number of errors per type and the nature of the errors as they were identified, the final list of errors was shortened. Some error categories were deleted from further consideration and others were merged with other categories as appropriate. This is explained in the Results section. The Error Correction Code is included in Appendix 6, and includes error definitions.

Classification of Errors

Dulay et al. (1982) state that it is common for an error classification list, or error taxonomy, to focus on the linguistic item or feature being examined in a linguistic category. This is a useful approach for curriculum development and is also used in linguistic research as a mechanism for easily reporting findings. These categories include grammar (morphology and syntax), pronunciation (phonology), vocabulary (semantics and meaning), and style (discourse) (Dulay et al., 1982). The grammar linguistic category is the one appropriate for this study. As Dulay et al. (1982) point out, there may be differences in error categorisation decisions due to the variety of English being used as the linguistic norm. I decided to use my own norms of English as I have been employed as an English language teacher in various universities both in the UK and abroad, teaching students whose destinations were generally universities in the UK and the USA. Furthermore Dulay et al. (1982) state that there is a degree of subjectivity in this procedure. In some instances, it is ultimately simply a matter of judgement (Santos, 1988). In the process of correcting the errors, a total of 14 different categories were initially identified. These are listed below in Table 3.2

Table 3.2

Error Categories

Missing Articles	Plural Error	Word Order Reversed
Redundant Articles	Tense	Preposition
Incorrect Articles	Missing Word	Missing Possessive
Subject-Verb Agreement	Word Form	Redundant Possessive
Singular Error	Incomplete Sentence	

The Error Categories are for the most part self-explanatory and do not need to be discussed individually. A short description of each of the Error Categories is in Appendix 5.

Rationale for manually checking error identification

Error identification software has been developed and some software was considered for use, such as Markin, iSocrates and online software such as GrammarCheck. They were either no longer available, not compatible with the computer system used or required extensive time investment to set up. The errors could be manually marked in less time than that required to set the software packages up. Most importantly, the format of the essays made it too problematic to utilize automatic error identification software. There are limited independent reviews available for specific software to evaluate their accuracy and effectiveness. A review of online grammar services did not find any of the services to be better than manual checking (Nichol, n.d.). Furthermore, any program has to be manually checked to determine whether the error that has been identified is in fact an error.

The issues with automated software led to the decision to analyse the errors by hand. This raised other issues which were due mainly to the nature of correcting essays. Some errors are obvious, although in some instances there may be more than one way of analysing the error. Truscott (1996) discusses the difficulties that teachers, non teachers, native speakers and non native speakers may have in identifying errors and correctly categorizing them, even when such categorization is possible. This also applies to using software because human input is required in the software development stage. Computer software is based on human analysis to develop the tagging system required to build the program. This does not completely eradicate the issue of there being several possible correct forms which the analysers have to make a decision about regarding the most plausible correction (Dagneaux, et al., 1998). This corpus is relatively small and therefore it is not difficult to analyse. Manual analysis of data still remains relevant in EA (Tan, 2007) and the issues raised by Truscott mentioned above are borne in mind when analysing the errors.

3.3 Contrastive Analysis

Lado's (1957) work on contrastive analysis outlines the steps to compare two grammatical structures across languages. These have been outlined in the Background Literature section and will be applied to the findings described in the Results section. Due to motivation issues, teachers should focus on the main errors only (Al-khresheh, 2015) therefore the contrastive analysis will be conducted on the principal errors which result from the EA. This will compare the *form* that the grammatical structure takes in English and in Chinese, if applicable, the *meaning*, the *distribution* across sentences (or phrases) and will make an assessment on the expected ease or difficulty for a Chinese L1 speaker to learn.

This section has reported the source of the data and the issues in finalising the small corpus used. It has discussed the development of the Error Correction Code used in identifying errors in the students' essays and the rationale behind assessing the essays manually rather than by using software. In the next section the results of the analysis will be discussed and possible causes or reasons for the prevalence of these errors will be made using Lado's Comparative Analysis Hypothesis (1957). This will be developed into recommendations for approaches in teaching to redress these errors which adheres to the final step of Corder's EA (1967) of remediation.

Chapter 4. Results

The results section begins with reporting the results for the EA. Two of the error categories were conflated, that is, were merged into one category, because it was more useful to consider them as one category. Four categories were excluded from further consideration. The remaining categories were analysed to identify the main errors for inclusion in the second part of the analysis, the CAH framework. The results for the Contrastive Analysis are presented under each main grammatical category with an analysis on *form, meaning, distribution* and *ease or difficulty of learning*, following Lado's framework.

4.1 EA Results

The total number of errors for the first year essays was 501, and for the third year essays the total was 582 errors, making a total of 1,083 errors combined. This was from a total word count of 25,275 for the first year group, 39,290 for the third year group and a total word count of 64,565. The total errors for each error category will be analysed in more detail in following sections.

4.1.1 Categories not considered further

Some error categories were not considered further due to there being comparatively few errors, as are summarised in Table 4.1 below. Error in the use of Possessives was only 1 token for each of the Possessives categories identified, that is, Redundant Possessive and Missing Possessive. These both occurred in first year essays and were completely absent from third year essays. Word Order and Incomplete Sentences were also low, 8 and 9 tokens respectively. Although these categories were not considered for

further analysis, the number of errors is still included in the total number of errors identified because they are errors, but not main ones.

Table 4.1

Error categories with few tokens

Error Category	Number of total errors
Missing Possessive	1
Redundant Possessive	1
Word Order	8
Incomplete Sentences	9

This resulted in a total of 10 error categories: Missing Article, Redundant Article, Incorrect Article, Singular Error, Plural Error, Prepositional Error, Missing Word, Subject-Verb Agreement, Word Formation and Tense.

4.1.2 Conflated Categories

Articles

The use of incorrect articles and redundant articles, that is, the use of a definite article when an indefinite article was required (or vice versa) or the inclusion of an article when one was not required, was very low at only 48 total errors, across both groups, that is 4.4% of total errors. The Missing Articles categories had the highest number of errors for each group, 110 errors for first year essays and 197 errors for third year essays. These are presented in Table 4.2.

Table 4.2

Article Errors

Article Error	First Year	Third Year	Total
Missing	110	197	307
Redundant	29	9	38
Incorrect	5	5	10
Total	144	211	355

Redundant articles and incorrect article use are low in number but should not be excluded because they are a part of article usage and are linked to the missing articles categories. This is because missing articles may be due to avoidance of using articles and the redundant or incorrect article use may reflect attempts to apply rules of article use, although incorrectly. Therefore these errors can be conflated into one category and re-labelled ‘Articles’.

Number (singular/plural)

The initial results indicated that singular noun forms being used when plural noun forms were required was nearly 5 times higher in number than the plural form being used when the singular was required. That is the greater error was in missing plural nouns. This pattern was consistent across first year and third year errors. The error scores are displayed in Table 4.3. Singular noun forms are usually 4 times more frequent in English than plural nouns, although this is not consistent for all nouns (Taylor, 2012). The two forms were conflated together because the aim of the study is to identify which errors require further attention in teaching EAP and it is likely to be difficult to teach plural forms without discussing singular forms. The conflated category was labelled ‘Plurals’

Table 4.3

Number of errors for Singular and Plural Errors across First Year and Third Year essays

Error Type	First Year	Third Year	Total
Singular Errors	16	18	34
Plural Errors	76	89	165
Total	92	107	199

The resulting error categories were then ranked by order of error frequency and are presented in Table 4.4. The percentage of error for each category by year was calculated as a percentage of the total number of errors for that year, including those categories which were not analysed further due to low number of errors.

Table 4.4

Revised Error Category ranked by number and percentage for first year and third year essays*

Rank	Category	First Year Essays	Rank	Category	Third Year Essays	Total
1	Articles	144 (29%)	1	Articles	211 (36%)	355
2	Singular/Plural	92 (18%)	2	Singular/Plural	107 (18%)	199
3	Prepositions	70 (14%)	3	Word Form	65 (11%)	135
4	Missing Word	62 (12%)	4	Missing Word	57 (10%)	119
5	S-V Agreement	45 (09%)	5	Tense	51 (09%)	96
6	Word Form	39 (08%)	6	S-V Agreement	44 (08%)	83
7	Tense	38 (08%)	7	Preposition	39 (07%)	77
Total		490			574	1064

* percentages rounded to nearest figure and calculated as a percentage of the total errors for each year.

4.2 Major Error Categories

Hendrickson (1978) notes previous research which indicates that the main errors which impede communication should be corrected first, however he concedes that this is not always easy to identify due to teachers' familiarity with their students communication style and differences in tolerance to errors. Han (2002) concludes that corrective feedback needs to be focused to be effective and that teachers have to prioritize their efforts in correcting errors. This means that they should not address all errors. This is supported by Ellis (2009) who also makes a point that error correction should only be directed at features that are causing learners to have problems. Error correction for L2 learners needs to be restricted to the most important errors only (Al-kresheh, 2015). Hendrickson (1978) states that the high frequency errors should be the first to be corrected.

To consider the extreme case in this study, there was only one instance of a missing possessive out of a total word count of 64, 565 and a total number of errors of 1,083. That is 0.00015 (to nearest fifth decimal place) as a percentage of the total word count, and 0.09 (to the nearest second decimal place) as a percentage of total errors. Between such low levels of error counts and the more frequently occurring errors there has to be a cut off point due to space in the study. This is especially relevant because each of the errors is to be compared for form, meaning and distribution which is beyond the scope of the thesis. Therefore only the most frequently occurring errors will be included in the CAH procedure.

The top two error categories for the first year and third essays were the same, that is, Article Errors and Plurals for each group. Articles for the first year group were 29% of total errors and for the third year group they were 36%. For Plurals, the percent was the same for both groups, 18%.

For first year errors, the top two error categories combined made up for 47% of the total errors for that group and for the third year group the top two combined errors were 54% (figures rounded up to nearest whole per cent). Combining the two groups,

Articles make up 33% of total errors and Plurals make up 18%; together they make up 51% of total errors. The similarities between the rankings of errors and the similarities between the percentages that these errors, when combined, make out of the total of remaining error categories, mean that these two groups can be considered to be homogenous for the purposes of analysing the linguistic framework for the possible causes of these errors. The two groups may not be homogenous from other perspectives, such as the amount of time they have had in practice writing essays, the degree to which they have integrated with the local community and the amount of English language training they have had since starting study in the UK.

The next most common errors differ to various extents between the two groups. Prepositions are the next most frequent error for the first year group. Prepositions often occur as part of phrasal verbs, which cause difficulty for many people learning English. While some of the errors in prepositions occur with phrasal verbs they also occur in non-phrasal verb contexts. This indicates that preposition errors are not limited to their use in phrasal verbs. Subject-Verb agreement may be a distracting or annoying issue for some teachers but its rankings in this study shows it to be less frequent than other errors. This does not reflect the perception of S-V agreement as a serious (or otherwise) error by lecturers, but its frequency of occurrence. Missing Word as a category excludes missing words which are accounted for by other categories, such as articles and prepositions. Issues with sentence structure or vocabulary generally cause Missing Words. This is often related to students writing by translating word for word from their L1, a common issue with Chinese students (Li, 2007).

Tense, although ranked differently for the two groups, is similar as a percentage of errors for the two groups (8% and 9% respectively). Using the wrong tense is likely to be a problem depending on the context. In some instances the meaning will be clear regardless of the tense, in other contexts the meaning will be obscured by the incorrect tense. Word Form error refers to using an adjective when a noun was required, for example. This occurs as a low occurrence for the first year group but is the third most frequent error for the third year group at 11% of total errors for that group.

Some data anomalies

The average word count per essay for the first year group was 1,404 words and for third year essays it was 2,183 words. This is influenced by 2 essays in the third year group of over 5,000 words each. These 2 essays push the total word count for third year essays up by over 10,000 words, creating a large difference in the average word count between the first year essays and third year essays. If these two essays are removed from the data, the two subsets are more equal in number with the total word count for the third year group becoming 28,160 compared with the first year group total word count of 25,275 (although this does result in the first year group including 2 essays more than the third year group). One essay in the third year group contained no grammatical errors.

Overall the two groups can be taken together. Although the average word length is different and the percentage of errors is different, the types of errors and their rankings are the same between the two groups. Therefore the errors to be considered for the Contrastive Analysis are Articles and Plurals.

4.3 Contrastive Analysis Hypothesis Results

Following Lado's steps for the identified errors of Articles and Plurals I will examine whether the form exists in Chinese or not, and if it does, how it is represented, the meaning and distribution. Following this, the next step is to ascertain whether the structure can be designated as difficult or easy to learn for Chinese L1 speakers. Lado's Comparative Analysis Hypothesis (1957) calls for three specific steps, the first two of which (step one, identify the structural description of the languages being compared and step two, the summary of types of structure) have been documented by the Defense Language Institute (1974). Therefore I will focus on the third step, the comparison between the grammatical structures, with attention to the two structures identified as being the most frequent in this study. The role of articles and issues

surrounding plural forms in Chinese cannot be adequately discussed without considering the role of Classifiers and Measure Words in Chinese. Classifiers are an integral part of Chinese grammar and a brief summary is provided here.

Classifiers and Measure Words in Chinese

In Chinese, *liàngcí* refers generally to classifiers and measure words. Classifiers are associated with some shared feature of a group of nouns while measure words indicate the quantity of the noun (Tai & Wang, 1990, as cited in Her and Hsieh, 2010). *Gè* is the most general classifier, usually classifiers reflect some shared characteristic of the group of nouns they can be used with, such as the shape of the object. There is debate about the actual number of classifiers in Chinese and whether there is in fact any semantic difference between classifiers and measure words in Chinese. Her and Hsieh (2010) conclude that there is a semantic difference, being that classifiers do not have a meaning unless attached to a noun, whereas measure words can possess an independent meaning. This example is from Li, Huang and Hsiao, 2010:

(4) *Sān gè wǎn*
 three (classifier) bowl
 ‘three bowls’

(5) *Sān wǎn shuǐ*
 three bowl water
 ‘three bowls of water’

As classifiers do not have a meaning on their own, *gè* is only used to refer to the bowls in sentence 4. *Wǎn* is a measure word used to refer to the bowls of water in sentence 5. As a word by itself *wǎn* means ‘bowl’ as in sentence 4.

4.3.1 Articles

Form

Articles in English can be indefinite ('a/an'), definite ('the') or zero, that is, are not indicated in writing (but represented as \emptyset for discussion). There are four forms which an article can take, given that there are two forms for the indefinite article. The zero article has no written form and Berezowski (2009) argues that the zero form cannot actually be said to exist. He states that the use of the zero article masks other linguistic features and that it should not be counted as an article. In examining a corpus for errors in zero article use, it will only be apparent that an error has been made by actual use of an alternative article, because there is no way of knowing whether a student is using a zero article correctly or is avoiding using articles, because the form of both is identical. That is, it is the absence of an article which indicates both a zero article use, and also indicates avoidance of article use. Milton concludes that it is not possible to distinguish between errors in using the zero article for a definite or indefinite article or simply missing out an article (Milton, 2001).

There were a total of 38 errors classed as Redundant Articles in this study, that is, instances when an article was provided when the zero article was required. This indicates that students were using articles when a well-formed sentence would not have required an article, however, in comparison to the frequency of other errors this was rather low and was subsumed into the larger category of Article error. This is a practical issue because it is unlikely that only one article would be taught.

Chinese does not have articles (Defense Language Institute, 1974; Robertson, 2000; Milton, 2001; Chuang, 2005; Chuang & Nesi, 2006; Li & Yang, 2010; Zheng & Park, 2013) but it is noted that word order can be used to indicate the definite or indefinite noun (Robertson, 2000). Zhang (2004, cited in Zheng & Park, 2013, p. 1347) gives the following example of this:

- (2) *lái kèle*
 come guest (past tense marker)
 ‘a guest is coming’
- (3) *kè láile*
 guest comes (past tense marker)
 ‘the guest is coming’

In sentence (2) a guest has arrived who has not been identified as yet, that is, the person speaking may, or may not know, who the guest is, but they have not identified the guest to the hearer (an indefinite meaning). In sentence (3) the guest who has arrived is expected or known to both speaker and listener (a definite meaning).

Li and Yang (2010) and Milton (2001) point out that although there are no articles, the concepts of definiteness and indefiniteness can also be represented by other determiners ‘this’ (*zhè*) and ‘that’ (*nà*) and the plural forms ‘these’ (*zhèxiē*) and ‘those’ (*nàxiē*) for definite nouns. The English definite article ‘the’ is therefore either dropped altogether in Chinese or replaced by one of the demonstratives (Defense Language Institute, 1974). The indefinite ‘a/an’ is expressed by ‘one’ (*yī*) (Defense Language Institute, 1974). Chinese determinatives include demonstratives and numerals but also include classifiers, which are used to indicate definiteness (Milton, 2001).

Meaning

In English the definite article ‘the’ is used to refer to a particular noun. It is used when referring to a noun or noun phrase which is a known entity to both speaker and listener. It is used when referring backwards to a previously mentioned (indefinite) noun, referring forward, or when referring to categories, such as ‘The lion is a dangerous animal’ (Richards, & Schmidt, 2002, p. 32). The indefinite article ‘a’ (or ‘an’) refers to a

general noun or to an unspecified noun, such as a member of a category or an example of a category, such as 'A dog is a friendly animal' (Richards, & Schmidt, 2002, p. 32). The zero article is used for non-count nouns or plurals when they are first mentioned, such as 'Cats like sleeping' (Richards, & Schmidt, 2002, p. 32). This is except in special circumstances where a definite article is required. As well as the second mention, these include when used with a superlative or ordinal ('most', 'first'), specifiers ('same', 'only') shared knowledge, postmodifying 'Of' phrases ('the cost of...'), partitive 'of' phrases with plurals ('half of the people') and when a noun is modified by a proper noun being used as an adjective ('the Mandela effect') (although not when used in the possessive form) (Swales, & Feak, 2012, p 398-399).

Cheng and Sybesma (1999, 2005, cited in Wu, & Bodomo, 2009) claim that classifiers in Chinese are equivalent to articles in English. This is refuted by Wu and Bodomo (2009) who argue that classifiers contain a semantic meaning (but only when attached to a noun) whereas articles are functional words which only indicate definiteness or indefiniteness and therefore cannot be equated (Wu, & Bodomo, 2009).

Distribution

The comparison of distribution is therefore straightforward. In English articles appear before the noun. In Chinese, they are not present. The comparison is of a grammatical structure which exists in English with a non-existent form in Chinese. Therefore there is simply no distribution for articles in Chinese. The indicators of definiteness and indefiniteness (such as determiners 'this' (*zhè*) and 'that' (*nà*) for definite nouns) also occur before the noun. Although they are interspersed with a classifier, the classifier attaches to the noun.

Assessment of Ease or Difficulty to Learn

The article form does not exist in Chinese, the meaning is not represented equivalently and there is subsequently no distribution within Chinese. In English there are four forms of article (including the zero article) and the rules for applying them is not always clear to a second language learner and may be difficult to teach. For Chinese students, article use has often been reported in the previous literature as one of the most frequently occurring errors. English language teachers may provide only a basic and misleadingly incomplete explanation of using indefinite article for first mention and definite thereafter, with students being directed to other resources for further self study, or with articles errors being simply highlighted with the hope that repetition of correction will result in acquisition (Nickalls, 2011). In addition to the 'second mention' rule, Swales and Feak (2012) give 7 other rules for when the definite article must be used. Robertson (2000) notes that in academic writing Chinese students are not simply avoiding article use, they do include it as a feature of their interlanguage, they are just not using them at a native-speaker equivalent level. This makes article use one of the most difficult for Chinese students of English to master.

4.3.2 Plurals

Form

In English most plurals are formed by adding *-s* or *-es*. However there are a substantial number of types of exceptions (Jing, Tindall, & Nisbet, 2006) such as the *-f* and *-fe* endings, cases where *-ies* is added, where the plural form is the same as the singular form, where there is an internal vowel change, and the forms of some foreign loanwords. In addition, some nouns can only be plural (when used as nouns), such as

‘tweezers’. These basic rules are summarised with examples in Table 4.5 on the following page.

Table 4.5

Summary of English noun plural formation

	Plural	Example	Exceptions
Most Nouns	Add -s	Ant - ants	<i>Some</i> nouns ending in o – add es, e.g. potato – potatoes. This process is in transition. <i>Some</i> nouns ending in –f or –fe add –ves, e.g. elf – elves, knife – knives.
Nouns ending in sibilant consonant	Add -es	Bus - buses	
Nouns ending in y	a) Vowel + y – add -s b) Consonant + y – omit y and add -ies	Toy – toys Body - bodies	
Some nouns do not change		Sheep - sheep	
Maintained forms from older forms of English	Internal vowel change	Man – men	
Foreign nouns	Typically add –s or –es as per English nouns.	Tempo (singular) – tempi (plural in foreign form) Tempos (plural in English).	Some foreign plural forms are used e.g. stimulus – stimuli.

(Adapted from Burt, 2002; Greenbaum, & Nelson, 2002)

English nouns are made up of mass nouns as count nouns, however this is not as simple a categorisation as it appears because some mass nouns can be used as count nouns. For example, 'language' is a different concept when the word is used as a mass noun, that is, as the concept of language, or a count noun when referring to a language (Robins, 2000). Although some sources indicate that Chinese does not distinguish between count and mass nouns (Defense Language Institute, 1974; Chuang & Nesi, 2006) other research suggests that the concept of count nouns and mass nouns are essentially the same in Chinese and English (Jing, Tindall & Nisbet, 2006). Classifiers only attach to count nouns while measure words attach to both count nouns and mass nouns (Her, & Hsieh, 2010) however what is considered as a count noun and a mass noun differs between the two languages (Jing, Tindall & Nisbet, 2006). The choice of classifier is used to indicate whether a noun is a count or mass noun for e.g.

- (6) *Sān zhī xióng*
 3 (classifier) bear (s) – objects
 '3 bears'

- (7) *Sān zhǒng xióng*
 3 (classifier) bear (s) – species
 '3 bears'

In English '3 bears' could mean either three individual bears or three types of bear species (Krifka, 1995).

Context is one of the primary means of determining plurality in Chinese. Furthermore a numeral or a determiner can be used (Jing, Tindall, & Nisbet, 2006). In Chinese, nouns are not inflected for the plural form (Defense Language Institute, 1974). They depend on context, once the plurality has been established by use of a determiner then the determiner is no longer necessary. Jing, Tindall and Nisbet give the example of

(8) *Yi xiē xuéshēng zài jiàoshì lǐ, lǎoshī zhèng gěi xuéshēng jiǎng yīgè gùshi*

There are some student in the classroom. Teacher is telling a story to student.

‘There are some students in the classroom, the teacher is telling them a story.’

(Jing, Tindall & Nisbet, 2006, p. 131.)

Here the determiner *yī xiē* - ‘some’ is placed before the noun ‘student’ to indicate plurality. In the second part of the sentence it is not required and can be omitted because the context has already been established. If the sentence were to convey that there were some students in the classroom but the teacher was telling a story to only one student, the measure word *yī* to denote– ‘one’ and the classifier *gè* would have to be inserted in front of the noun ‘student’. So the plural and the singular form for *Xuéshēng* (student) are the same.

(8) *Nàgè xuéshēng shì cōngmíngde*

‘That student is intelligent’

(9) *Nàxiē xuéshēng shì cōngmíngde*

‘Those students are intelligent’

Plurals can also be formed by adding *-men* but only to certain nouns as discussed below. A noun can also be made plural by adding a number or a determiner (Jing, Tindall, & Nisbet, 2006).

(10) *Wǔ běn shū*

5 (classifier) book

‘5 books’

- (11) *Nàxiē shū*
Those book
'Those books'

Meaning

The concept of plurality is present in Chinese in relation to nouns, but it also exists in personal pronouns such as 'we' (*wǒmen*), 'them' (*tāmen*). In writing using characters, the difference between 'them' (masculine) and 'them' (feminine) (*tāmen*) is apparent but not in spoken Chinese or pinyin. The addition of the suffix *-men* can be used to make a plural, however it is not the equivalent of the plural suffix *-s* (or *-es*) because of its limited application. It only applies to words for people or groups of people (Jing, Tindall & Nisbet, 2006). It cannot be considered as an *-s* morpheme equivalent because the indication is one of collectiveness, or belonging to a group, not plurality (Iljic, 1994, as cited in Her & Hsieh, 2010).

Distribution

The plural form in English appears as an inflectional change on the noun. This is chiefly invariable in its placement (except for internal vowel changes), although its form is variable and inconsistent at times. In Chinese the plural marker is more diverse. It appears at the end of the noun when it is a person word such as *lǎoshīmen* ('teachers'), *péngyoumén* ('friends') however it appears before the noun when it is a classifier or measure word accompanied by a numerical. The classifier or measure word, although not a plural marker, appears prior to the noun and after the numerical. There is a consistent distribution for these plural markers.

Assessment of Ease or Difficulty to Learn

The plural marker in Chinese has a very different form from that in English, although there is possibly some degree of ease in replacing *-men* with the appropriate suffix in English, this is a limited case which only applies to groups of people. The formation of plurals by internal vowel change may cause difficulties for Chinese students because Chinese characters do not allow internal change (Ho, 1997, as cited in Jing, Tindall, & Nisbet, 2006) which means that students are expected to learn these plurals by rote. Because the plural marker in English is comparatively complicated, this can be expected to be a difficult structure to learn.

Chapter 5. Discussion

In this section I will discuss the main limitations of the study, the degree to which they affect the results and how they could have been addressed, if possible. I will discuss the results and assess the main solutions suggested by previous researchers in relation to their usefulness in redressing the errors raised in this study with the aim of answering the research questions raised in the first part of the thesis. This follows the final step inherent in Corder's EA procedure of 'Remediation' (Corder, 1974).

5.1 Research Questions

1. What are the grammatical writing errors of Chinese undergraduate students in their discipline-specific texts in UK universities?

Previous research in EAP in the UK and in non-UK universities indicates that the main errors include articles, plurals, prepositions, pronoun agreement, spelling, plural errors, word choice (vocabulary), verb form, missing subject and tense. The results of this study found that articles, plurals, prepositions, subject verb-agreement, possessives, word form, missing word, incomplete sentences and word order were all included as errors (Table 4.1). The two main errors were Article use and Plurals, which together accounted for over 50% of the written errors (Figure 4.1). This finding supports those of Santos, (1988), Chuang (2005), Chuang and Nesi (2006), Zheng and Park (2013), Zhang and Xie (2014) despite some of these previous results being from studies outside of the UK and none of them being discipline specific texts.

Articles.

The English definite article 'the' is the most frequent word in the British National Corpus (BNC), the indefinite 'a' is ranked 4th and 'an' is ranked 34th (Cook, 2008). The zero article was reported as being the 54th most used article, based on a corpus study by Masters (1987, as cited in Nickalls, 2011). Chuang and Nesi (2006) note that because the frequency of article occurrence in English is high, the incidence of error will likely be high as well. The high frequency of Article occurrence in English may account for the reason that it scores highly in EA results (Table 4.5). The misuse of articles is unlikely to cause significant problems in comprehension, however frequent article misuse may cause difficulty for a reader (Nickalls, 2011).

Article use is more highly tolerated than other errors according to Janopoulos (1992). However, that study used decontextualized errors rather than errors in a discourse. Albrechtson, Kendrickson and Faerch (1980, as cited in Vann et al. 1984) found that there is some evidence that it is not necessarily the gravity of the error but the frequency of the error which has a role to play in the level of irritation caused to the reader. Article errors are the most frequent of the errors found in this study.

Plurals

The plural formation of *-men* might be construed as being similar to English plural formation to some degree, because it signifies plural and is a suffix, however it only applies to certain specific plurals, that is one of collectiveness for people, and alternatives plural forms also exist. Singular nouns are 4 times as more likely to occur than plural nouns, this study found that there were 4 times as many plural noun errors than singular errors (Table 4.4). It is unlikely, therefore, that the high number of errors is due to the higher frequency of the word form. This has ramifications for assessing the cause of linguistic error.

2. What are the linguistic causes of these errors?

Touchie summarizes the main causes of error in language learning: L1 interference, Simplification (choosing simple forms instead of complex ones), Overgeneralisation (overextending a form to contexts where it does not apply), Hypercorrection (correcting a form which is already correct, thereby making it incorrect), Faulty Teaching (when the teacher has provided incorrect information), Fossilization (persistent errors resistant to correction efforts), Avoidance (choosing not use a difficult structure), Inadequate Learning (essentially a lack of knowledge), and False Concepts Hypothesized (attributing a feature as having a function that it does not have) (Touchie, 1986). As Lennon (2008) points out it is difficult to definitively determine the cause of error.

Odlin (1989) discusses the difficulty of considering the term interference and refers instead to negative language transfer and positive language transfer. The degree to which languages can be said to be distant affects transfer (or L1 interference) (Odlin, 1989). Typological factors can affect transfer of a structure; if a structure is present in one language it may transfer into the other language (Odlin, 1989). Consequently if the structure is not present, it cannot transfer.

Negative language transfer can occur through the morphemic, lexical, syntactical and discourse level (Li, 2007). Chinese is not an inflectional language whereas English is, therefore adding the suffix *-s* or *-es* to form plurals in English is an example of morphemic negative transfer. Chinese do not use articles and therefore experience difficulties when attempting to use articles in English, which is a result of negative syntactical transfer (Li, 2007).

Articles

The Contrastive Analysis between Chinese and English grammatical structures showed that there is no article in Chinese and that the definiteness or otherwise of a noun is

shown to be associated with the demonstrative used and word order. These are not proxies or equivalents to the English article. Furthermore the English article system is not simple and is usually learnt after plural formation (Cook 2008). As the students demonstrated difficulties with plural formation it is logical that there would be difficulties with article use. This is likely to be caused by L1 interference, however the role of imperfect or incomplete teaching may also play a role. There is possibly some degree of avoidance, but not completely, as students did display some use of articles. Whether they were avoiding using the articles due to lack of knowledge about when to use them or whether they were forgetting to use them due to L1 interference is difficult to assess completely without consulting the students themselves.

Plurals

Of the possible causes for error, language transfer is the most likely. It is unlikely to be a reflection of word frequency. The singular form occurs much more than the plural form which would lead to the conclusion that more errors would occur in the singular (most frequently occurring form) than the plurals (least frequently occurring form) if the errors reflected the relative word frequency occurrence. The results show that but there are more errors in the plural form than the singular form. Therefore it is not related to the greater frequency of singular forms and therefore has another cause. It is most likely language transfer.

The plural form in English is varied and is formed by inflection whereas in Chinese plurals are indicated by classifiers and measure words. It is possible to include poor teaching as one possible contributor because as Cai (2013) states, in China there is sometimes poor teaching because most teachers are Chinese native speakers and some may have imperfect language and teach much of the class in Chinese. It may also be the case that either students or teachers do not have much vested interest in the student passing the course. Another cause may be linked to motivation, or interest, as some

students may be interested in English only as a medium to study their chosen field abroad.

3. How can EFL teachers address these issues?

The findings of previous research on Chinese students' written grammatical errors can be grouped into three general areas: strategies, the role of culture, and the importance of motivation. The strategies tend to be general, and although some studies found that the main errors included article use and plurals, they did not specifically address them, with the exception of Lu (2010). Lu provides an actual exercise for teaching articles, although it is really only applicable to 'first mention' and 'second mention' use. Lu's main strategy for success relates to examples and suggestions for the role of the teacher in creating a comfortable and friendly learning environment for students.

Hu (2007) reports that some teachers have had some success with publishing students' work in the classroom to heighten students' awareness of issues such as writing for an audience, quality of writing and mutual learning. The particular EAP course Hu refers to reports that students were able to improve their IELTS scores by nearly two IELTS bands. However, I note that the course is a six-month course, whereas most UK based EAP courses are only 3 months or less, raising the question as to whether it is the techniques used or simply longer time exposure to teaching.

Tan (2007) lists the requirements for language learning improvement as being: learning vocabulary, receiving explicit grammar instruction on problem areas, guided reading for pleasure to improve input, selective error correction by class conference to avoid student and teacher de-motivation by having to correct every error, and raising awareness of the effect of L1 interference on L2 language learning. Tan does not provide any experimental evidence of the effect of any of these suggestions; they appear to be based on reflections of classroom teaching experience.

Other studies have reported on using technology in the classroom. Ean (2001) reports the lack of success in using IT to foster autonomous learning in Hong Kong students, although she does highlight that the study centred on Vocational Education students who may have been generally less motivated than university students might have been. Xing, Wang and Spencer (2008) found that an eLearning course was effective in academic writing, although the course focussed on highlighting cultural differences in academic writing. Chuang and Nesi (2007) found that the online learning program that they developed, GrammarTalk, raised awareness of grammar errors, especially in singular and plural use, and also fostered autonomous learning for Chinese students in a UK university. They state that it is still a work in progress as there is still some room for improvement. Li (2009) suggests that teachers can address grammatical issues using repetition drills to avoid L1 interference, and increase foreign language input. He does not provide detailed suggestions on how to do this. He also suggests that teachers increase output through English corners and speech contests, but does not provide any evidence to show that these have been effective.

Leedham (2014) recommends that one-to-one tutoring is required before academic writing skills can consistently improve, and she makes the point that this is necessary both for international students and home students. The suggestions given by Zhang and Xie (2014) are not very specific; they state that the requirements for academic writing improvement include attention to linguistic form, awareness, and training for self-monitoring.

Highly salient in previous research was the importance of culture in addressing grammar writing needs, specifically concerning Chinese and UK cultural differences in education. Flowerdew and Miller (1995) identify a framework for thinking about the concept of culture and how it applies to the cross-cultural classroom situation to improve teaching. However, they do not propose how to popularise or promulgate their framework. Although the Communicative Language Teaching precepts used in UK classrooms are at odds with the Chinese cultural teaching norms, some ideas are compatible, such as collaborative learning (Hu, 2002). Errey and Li (2005) reported that

the time required for cross-cultural adaption slowed down learning for Chinese students abroad. They suggest that overt discussion of different learning approaches and conscious recognition of the need to build cultural bridges will provide a solution to this problem. But this assumes that Chinese students want to integrate and that the students of host university cultures are amenable to this. Cross (2006) found that many Chinese students do in fact want to integrate and that language is a significant factor impacting on their success in social situations. He suggests that students can be encouraged to integrate with the local culture through homestays, peer mentoring, social interaction on campus and in class, sports and recreational activities and in-class groupwork. In the community outside of university he again suggests sporting clubs and, perhaps somewhat controversially, churches. He advocates for information provision at student orientation events, however it is not clear exactly what kind of information is provided, how it is provided or who it is provided by. Homestays are not always ideal, as the host family may be more concerned with receiving the financial benefit of hosting a student rather than exchanging cultures, and when there is a clash of cultures in these situations there may not always be a readily available or competent counsellor to help resolve misunderstandings, resulting in further alienation for the student.

Jian (2009) highlights the differences between Chinese and Western education learning styles and concludes that as China learns more from western pedagogy, it will not necessarily follow western techniques, but may adapt to meet its own needs, taking into consideration the Chinese cultural values. This provides a challenge to the notion that Chinese students are going to be getting their English language instruction from western universities. Xu (2012) concludes with a general claim that teaching EAP must focus on culture and teaching thinking strategies for Chinese students to be successful.

Finally, the role of motivation has been found to have a significant effect on students' second language learning. If attitude and confidence are positive, then Chinese students do better (Eerey and Li, 2005). Li (2009) suggests that teachers should be highly motivated themselves while Throssell & Zhao, (2011) advocate that teachers should motivate students, as motivated students perform better. This raises several

questions. Are there differences between what is meant by motivation between the two cultures? What is the best way to motivate students? Who should do it? Based on some internet discussion boards and personal conversations with Chinese people about their experiences of education in China, some teachers embarrass or shame their students as a means to make them study harder, while western educators may at times be overly liberal with their praise with the effect that the students do not feel that it is well earned. Oxford (1999) raises the question that if greater learner autonomy increases language proficiency, how much should, or can, teachers challenge students from countries such as China where there are generally lower levels of learner autonomy? This leads into the changing attitudes towards the role of internationalisation and what it means for universities.

Despite this, spoken communicative problems are also an issue for Chinese students studying in the UK. Students increasingly have to participate in group discussions to complete projects and to participate in group presentations. In addition to addressing the academic style issues and grammatical writing problems, students' spoken communicative skills can create barriers which may lead to them being unable to participate fully. A common complaint I hear from Chinese students relates to group discussions. By the time they have understood the discussion, decided what their contribution should be and then formulated their thoughts into English, the discussion has moved on and they have not participated. This may be further complicated by cultural misunderstandings of expected roles and group norms. The scope of the problem includes the situation when a student approaches a teacher and is unable to clearly state the reason they are seeking help, or explain their ideas, and due to cultural issues may be unable to interact with their lecturer or tutor, for example, by disagreeing, or correcting their tutor if the tutor has misunderstood them. Therefore the relevance of English language skills should not be ignored due to the fact that some lecturers may be tolerant of written errors.

The role of Internationalisation of English

The Internationalisation of English has a weak version, which refers to the teaching of English as a Second Language to enable overseas students to study in English language medium universities. There is also a strong version which includes the university adopting a policy of inclusion of international students as full members of the university community in a mindset of incorporation of cultural diversity. This leads to an awareness of the different varieties of English spoken by students at university (Marlina, 2013). The strong version of internationalisation goes further than broadening acceptance of English varieties. Kirkpatrick (n.d., as cited in You, 2004) states that the academic writing style of Chinese students should be allowed to incorporate Chinese rhetorical style, and Cao (n.d., as cited in You, 2004) suggests that this style of writing should be appreciated by western university staff because it is an indicator of Chinese cultural identity. Associated with the concept of internationalisation of English is ELF (English as a Lingua Franca), that is, the varieties of English spoken by people who use English to communicate with each other but none of whom have English as a native language (although it does not necessarily preclude native English speakers from participating). The ramifications are that features of language which may not be native-speaker like but do not significantly impede communication, are best not focused on in language training, freeing up time and effort on more significant aspects of language and communication. This also provides a framework for avoiding significant focus on aspects of language which may not be learnable until after there has been significant exposure to the language (Seidlhofer, 2005).

The practice of teaching English across the Curriculum (EAC) has gained renewed support in recent years. At university level in the UK this has included calls for teaching academic literacy, that is, not only language but also the academic conventions and style required for each discipline. It is suggested that academic literacy teaching is carried out by subject specific teachers in collaboration with academic literacy experts (Wingate, 2015). Dudley and Evans (2001, cited in Cargill and O'Connor, 2012) outline 3

levels of partnership which they recommend for second language learning. The first is co-operation, whereby language specialists provide language training with input from subject specialists who provide information on the students subject specific language requirements and the type of tasks they will be required to do. The second is collaboration, whereby they work together outside of the class to design classroom activities. The third type is team teaching, in the same classroom.

EFL researchers are suggesting that subject lecturers take on even more duties than they already have. This will entail not only further resources and investment in staff development (Wingate, 2015) but requires the subject teacher's commitment and willingness to be involved. EFL researchers are also calling for there to be some form of training on culture, or cross-cultural adaption and cross-cultural communication. Again there are problems as to who will teach these skills, in what environment and the content of what they will teach. At a recent seminar, a Chinese language teacher strongly objected to having to teach Chinese culture because his interest and expertise are in teaching Chinese language (Personal communication at the 'Perspectives on intercultural communication and intercultural communicative competence seminar, Southampton University, November 11, 2015). EAP tutors have difficulty in adequately teaching all of the material on the syllabus of a typical pre-session course, without having to consider teaching cultural issues as well. A result of this is that EAP tutors do not have a great deal of time to focus on any specific grammatical structure for any lengthy period of time.

Limitations of this study

One of the chief limitations of this study comes from the limitations of the corpus that the data came from. The data in the corpus, although paid for, was voluntary in that the students did not have to contribute if they did not want to. Furthermore the number of essays collected from students who state that their L1 is Chinese results in a small

corpus. The actual number of students who submitted essays is even smaller because many of the students submitted more than one essay. There are attempts in the UK to collect a greater number of essays from international students, with some universities collecting all international students essays into a corpus without paying the students and requesting that the students actively opt out, or refuse to have their work used. There are ethical issues with this approach because students may not feel confident in refusing to assist the university which they hope will grant them a qualification.

A further limitation is that the BAWE data set does not collect information on Country of Birth. It is possible that some of the students who have identified Chinese as being their L1 may have done so out of a sense of cultural identity, while they were perhaps born in the UK or spent considerable years of schooling in an English speaking country and education system.

There is an issue with the method which relates to the subjectivity of the assessor of errors and the associated difficulty in determining what is an overt error and a covert error. This could have been overcome by having a second or third marker with differences in error evaluation being resolved between the assessors, or with an average score being calculated. Realistically this was not possible. There are further issues with the demographic characteristics of the assessor as discussed by (Janopoulos, 1992; Porte, 1999; Santos, 1988; Vann, Meyer & O'Lorenz, 1984) related to age and native speaker status. Older teachers and native speakers tend to be more error tolerant.

It is a further limitation of the study that there was only one assessor and the errors were not analysed by other assessors. Given that these issues have been highlighted, the assessor is an experienced English language teacher and that the results are to some degree in concordance with the findings of studies in EAP courses and non-UK universities, the results can be taken with some tentative level of acceptance.

5.2 Recommendations

The field of EAP in UK universities is continuing to develop and depending on the success that the internationalisation movement has in encouraging subject tutors to participate in academic literacy teaching and moving towards a greater tolerance of non-native varieties of English, the attention given to article use and plurals by Chinese students may vary. If successful, these grammatical structures may be seen as, although frequently occurring, not significantly impeding comprehension.

On the other hand, universities may respond to calls for international students to achieve higher standards of English before being admitted to university courses. Funding will play a crucial role here. If the international students continue to be significant funding sources, the English language requirements may be lowered or abolished altogether. If the government (or other funding streams) supports universities, then universities will be in a position to demand higher levels of English from students, allowing subject teachers to spend more time on research and teaching their subject than teaching English or study skills. It is a complicated argument as to how universities should be funded, who accesses the education provided by universities and the role of universities as education and research institutions. Further arguments as to the aspects of the politics of access to university being controlled as a method of perpetuation of class power include, in an international context, issues of racism and colonialism and this has ramifications for English language teaching and the issue of native-speakerism. These broader issues are beyond the scope of this thesis. Anecdotal evidence seems to support claims that due to international students having unsatisfactory levels of English, the course content is simplified. As a result, degrees and the universities associated with those particular degrees may lose their value in the long term. As the debate continues, there is a further aspect which is that not only are lecturers pressured to spend more time understanding different Englishes and different cultures, but the expectations on their pastoral duties are increasing too. If salaries are not increased accordingly, the risk is that lecturers either cut corners academically, (its

easier to give a pass to a borderline essay than to justify failing it), perform poorly or self select out of the system and move to either more lucrative positions in government or industry or less demanding positions. In my experience there are some students who do not want to be speaking a form of incorrect but acceptable English, but want to know that their English language skills are, if not native-speaker like, at least do not require specialised language listening skills for the listener. Their voice needs to be heard in the debate too.

In any case, the requirements for international students to have some cultural knowledge and cross-cultural communication skills will still need to be addressed either by the universities themselves or by the students prior to attending university.

1) I recommend that a strategy for implementation is developed for there to be an open, inclusive and productive discussion about the role and development of English as a Lingua Franca and its role in education, the role of English Across the Curriculum and the Internationalisation of education and English under different funding structures. As Seidlhofer (2005, p. 339) notes, the claims of ELF and its ramifications are accepted by some and 'deplored' by others. The combative language and attitudes encountered in seminars and private discussions do not foster rational and logical consideration of the phenomena of ELF but engenders emotionally negative responses and delays and frustrates the design of suitable programmes to meet the needs of students.

2) If UK universities decide that article use and plurals are to be taught, there needs to be sufficient time in the syllabus for these to be taught, with realistic expectations of the level which can be reached given the complex nature of the structures and the competing demands of other skills which are required for academic success. The strategies should be developed taking into consideration the results of previous studies which include highlighting differences in language and the potential of using software programs such as GrammarTalk and the sequence of acquisition research which indicates that plural forms should be mastered before articles are taught (Cook, 2008).

This has ramifications for the time spent teaching pre-sessional courses and in view of the exportation of UK education abroad, the place.

Students' attitudes and requirements should also be taken into consideration. Further research into the potentially diverse needs of Chinese and other international students would reveal the extent to which they wish to have their ELF skills evaluated against being taught to speak a native speaker variety of English.

3) The importance of cross-cultural adaption in second language learning has been well established. Further research is required to determine how to teach it to Chinese students intending to study abroad, what to teach, who is to teach it, when it should be taught, with consideration of the funding structures required. I recommend that a comprehensive program be developed which includes an action plan for establishing cultural competency training as an inherent part of providing services to students, and not as an optional adjunct which can easily be ignored or deleted due to funding and time constraints. This should take into consideration the attitudes of students and EAP teachers who may have diverse opinions about their role in the cross-cultural competency training.

References

- Adjei, A. A. (2015). Analysis of subordination errors in students' writings: A study of selected teacher training colleges in Ghana. *Journal of Education and Practice*, 6(8), 62-77. Retrieved from http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwie6lr-kq7LAhUH1RQKHUWCogQFgggMAA&url=http%3A%2F%2Fwww.iiste.org%2FJournals%2Findex.php%2FJEP%2Farticle%2Fdownload%2F21022%2F21278&usg=AFQjCNF0wleSJRzGPbPIp yyXixs5_kNByg
- Afraz, S., & Ghaemi, H. (2012). The effect of focused written corrective feedback of contrastive analysis on efl learners' acquisition of verb tenses. *Journal of Educational and Instructional Studies in the World*, 2(4), 48-61. Retrieved from <http://www.wjeis.org/FileUpload/ds217232/File/06.afraz.pdf>
- Al-khresheh, M. H. (2015). A review study of interlanguage theory. *International Journal of Applied Linguistics & English Literature*, 4(3), 123-131. doi:10.7575/aiac.ijalel.v.4n.3p.123
- Arthur, N. (2015). Counseling international students in the context of cross-cultural transitions. In P. B. Pedersen, W. J. Lonner, J. G. Draguns, J. E. Trimble & M. R. Scharron-del Rio, (Eds.), *Counseling across cultures* (7th ed., pp. 301 – 324). Thousand Oaks, CA: Sage.
- Bassetti, B. (2007). *Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language*. London, UK: Birkbeck ePrints. Available from: <http://eprints.bbk.ac.uk/528>
- Berezowski, L. (2009). *The myth of the zero article*. London, UK: Continuum.
- Brooks, R., Waters, J., & Pimlott-Wilson, H. (2012). International education and the employability of UK students. *British Educational Research Journal*, 38, 281–298. doi:10.1080/01411926.2010.544710
- Burt, A. (2002). *The A to Z of correct English*. (2nd ed.). Oxford, UK: How To Books.

- Cai, L. J. (2013). Students' perceptions of academic writing: A needs analysis of EAP in China. *Language Education in Asia*, 4(1), 5 – 22.
doi:org/10.5746/LEiA/13/V4/I1/A2/Cai
- Cai, Q., & Zhang, J. P. (2013). Corpus resources and their use in English teaching. *Proceedings of the second international conference on systems engineering and modeling*. Beijing, China: Paris: Atlantis Press. Retrieved from
www.atlantis -press.com/php/download_paper.php?id=5520
- Cargill, M., & O'Connor, P. (2012). Identifying and addressing challenges to international publication success for EFL science researchers: Implementing an integrated training package in China. In R. Tang (Ed.), *Academic writing in a second or foreign language* (pp. 21-44). London, UK: Continuum.
- Carrio-Pastor, M. L., & Mestre Mestre, E. M. (2014). Lexical errors in second language scientific writing. *International Journal of English Studies*, 14(1), 97-108.
Retrieved from <http://files.eric.ed.gov/fulltext/EJ1042860.pdf>
- Castillejos Lopez, W. (2011). Error analysis in a learner corpus. What are the learner's strategies? A survey on corpus based research. *Spanish Association for Corpus Linguistics*. Universidad de Murcia. Retrieved from
<http://www.um.es/lacell/aelinco/contenido/pdf/45.pdf>
- Cazalaa, B., Cabot, P., & Palat, A. J. (1993). *Spanish dictionary*. Paris: France. Larousse.
- Chuang, F. Y. (2005). Article misuse: a neglected problem in Chinese EAP student writing. doi:10.1.1.121.6756
- Chuang, F. Y., & Nesi, H. (2006). An analysis of formal errors in a corpus of L2 English produced by Chinese students. *Corpora*, 1(2), 251-271.
<http://dx.doi.org/10.3366/cor.2006.1.2.251>
- Chuang, F. Y., & Nesi, H. (2007). GrammarTalk: International students responses to an online grammar resource. In M. Edwardes (Ed.), *Proceedings of the BAAL Conference 2007*, Edinburgh, UK: British Association for Applied Linguistics.
Retrieved from: http://www.baal.org.uk/proc07/04_fei_yu_chuang.pdf
- Conlon, G., Litchfield, A., & Sadlier, G. (2011). Estimating the value to the UK of

- education exports. (BIS Research paper number 46). Retrieved from <https://www.gov.uk/government/publications/education-exports-estimating-their-value-to-the-uk>
- Corder, S. P. (1967). The significance of learner's errors. *IRAL - International Review of Applied Linguistics in Language Teaching*, 5(1-4), 161–170.
doi:[10.1515/iral.1967.5.1-4.161](https://doi.org/10.1515/iral.1967.5.1-4.161)
- Corder, S. P. (1973). *Introducing applied linguistics*. Middlesex, UK: Penguin.
- Corder, S. P. (1974). Error analysis. In J. L. P. Allen, & S. P. Corder, (Eds.), *Techniques in applied linguistics: Vol 3* (pp. 122-154). Oxford, UK: Oxford University Press.
- Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford, UK: Oxford University Press.
- Cook, V. (2008). *Second language learning and language teaching*. (4th ed.). London, UK: Hodder.
- Cross, J. (2006). Engaging Diversity At An Australian University: Chinese International Students' Perceptions Of Their Social Experience While Studying In Australia. *Proceedings of the EDU-COM 2006 International Conference. Engagement and Empowerment: New Opportunities for Growth in Higher Education*, Edith Cowan University, Perth Western Australia, 22-24 November 2006. Retrieved from <http://ro.ecu.edu.au/ceducom/72>
- Crystal, D. (2010). *The Cambridge encyclopedia of language*, (3rd Ed). Cambridge, UK: Cambridge University Press.
- Dagneaux, E., Denness, S., & Granger, S. (1998). Computer aided error analysis. *System*, 26, 163-174. Retrieved from:
http://www.researchgate.net/profile/Sylviane_Granger/publication/222308230_Computer-aided_error_analysis/links/09e4151332c3838a13000000.pdf
- Defense Language Inst., Monterey, CA. & Defense Language Inst., Monterey, CA. (1974). *A Contrastive Study of English and Mandarin Chinese*. [Washington, D.C.] : Distributed by ERIC Clearinghouse, <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?acc>

no=ED105774

- Degand, L. (2004). Contrastive analyses, translation and speaker involvement: the case of Puisque and Aangezien. In M. Achard & S. Kemmer (Eds.), *Language, culture and mind*. 251-270. Retrieved from https://www.researchgate.net/publication/265074247_Contrastive_Analyses_Translation_and_Speaker_Involvement_the_Case_of_Puisque_and_Aangezien
- Dervinyte, I. (2009). Conceptual emigration and immigration metaphors in the language of the press: A contrastive analysis. *Studies about Language*, 14, 49-55. Retrieved from: http://www.kalbos.lt/zurnalai/14_numeris/09.pdf
- Dulay, H. C., Burt, M.K., & Krashen, S. D. (1982). *Language two*. Oxford, UK: Oxford University Press.
- Ean, A. N. P. (2001). IT as a mediator for learning in a self-access Hong Kong VET context: Gender difference and motivation. International Vocational Education and Training Association, Conference 2001, Jamaica. Retrieved from <http://www.voced.edu.au/content/ngv%3A24815>
- Ellis, R. (1997). *Second language acquisition*. Oxford, UK: Oxford University Press.
- Ellis, R. (2008). *The study of second language acquisition*. (2nd ed.). Oxford, UK: Oxford University Press.
- Ellis, R. (2009). Corrective Feedback and Teacher Development. *L2 Journal*, 1(1), 3-18. Retrieved from <http://escholarship.org/uc/item/2504d6w3>
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford, UK: Oxford University Press.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368. doi:10+10170S0272263106060141
- Errey, L., & Li, H. J. (2005). Shift in Chinese EAP learners' perceptions of reading strategies. *The East Asian Learner*, 2(1), Retrieved from <http://owwww.brookes.ac.uk/schools/education/eal/eal-2-1/vol2-no1-errey-li.html>

- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6, 3–17.
doi:10.1016/j.jeap.2006.11.005
- Flowerdew, J., & Miller, L. (1995). On the notion of culture in second language lectures. *TESOL Quarterly*, 29(2), 345-373.
- Gast, V. (2013). *Contrastive analysis*. In M. Byram & A. Hu (Eds.), *Routledge encyclopedia of language teaching and learning*. (2nd Ed.). 153-158. London, UK: Routledge.
Retrieved from: <http://www.personal.uni-jena.de/~mu65qev/papdf/CA.pdf>
- Granger, S. (2003). The corpus approach: a common way forward for contrastive linguistics and translation studies? In S. Granger, J. Lerot & S. Petch-Tyson (Eds.), *Corpus-based approaches to contrastive linguistics and translation studies*. 17-30. Retrieved from http://sites-test.uclouvain.be/cecl/archives/Contr_Ling_&_TranslationRodopi1.pdf
- Green, A. (2007). *IELTS washback in context: preparation for academic writing in higher education*. Cambridge, UK: Cambridge University Press. Retrieved from <https://books.google.nl/books?id=iRSILmh1MdMC&pg=PA358&lpg=PA358&dq>
- Greenbaum, S., & Nelson, G. (2002). *An introduction to English grammar*. (2nd ed.). London, UK: Pearson.
- Han, Z. (2002). Rethinking the role of corrective feedback in communicative language teaching. *RELJ Journal*, 33(1), 1-34. Retrieved from <http://rel.sagepub.com/content/33/1/1.short>
- Hannas, W.C. (1997). *Asia's orthographic dilemma*. Honolulu, USA: Hawai'i University Press.
- Henderson, M. M. T. (1985). The interlanguage notion. *Journal of Modern Language Learning*, 21, 23-27. Retrieved from <http://people.ku.edu/~mmth/Interscam.pdf>
- Hendrickson, J. M. (1978). Error correction in foreign language teaching: recent theory, research and practice. *The Modern Language Journal*, 62, 387 – 398.
doi:10.2307/326176
- Her, O.S., & Hsieh, C. T. (2010). On the semantic distinction between classifiers and

- measure words in Chinese. *Language and Linguistics*. 11(3), 527-551. Retrieved from http://www3.nccu.edu.tw/~osh/C-M_Distinction.pdf
- Heuboeck, A., Holmes, J., & Nesi, H. (2008). *The BAWE corpus manual*. Warwick: Warwick University.
- Hu, G. W. (2002). Potential cultural resistance to pedagogical imports: the case of Communicative Language Teaching in China. *Language, Culture and Curriculum*, 15(2), 93-105, doi:10.1080/07908310208666636
- Hu, G. W. (2007). Developing an EAP writing course for Chinese ESL students. *The RELC Journal*, 38(1), 67-86. doi:10.1177/0033688206076160
- Hyland, K. (1997). Is EAP necessary? A survey of Hong Kong students. *Asian Journal of English Language Teaching*, 7, 77-99. Retrieved from: <http://www.cuhk.edu.hk/ajelt/vol7/art5.htm>
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Abingdon, UK: Routledge.
- Janda, L.A. (2008). From cognitive linguistics to cultural linguistics. *Slovo a smysl/Word and Sense*, 8, 48-68. Retrieved from <http://slovoasmysl.ff.cuni.cz/node/222>
- Janopoulos, M. (1992). University faculty tolerance of NS and NNS writing errors: a comparison. *Journal of Second Language Writing*, 1 (2), 109-121. Retrieved from <http://www.sciencedirect.com/science/article/pii/106037439290011D>
- Jia, H. W., & Tian, J. F. (2012). Contrastive analysis in China: today and yesterday. *Theory and Practice in Language Studies*, 2(11), 2269-2276. doi:10.4304/tpls.2.11.2269-2276.
- Jian, H. (2009). A contrastive study of cultural diversity of learning styles between China and the United States. *International Education Studies*. 2(1), 163-166. Retrieved from <http://www.ccsenet.org/journal/index.php/ies/article/view/410/383>
- Jin, L. X., & Cortazzi, M. (2001). *Cultural synergy: Using Chinese strengths in EAP*. Retrieved from <https://www.baleap.org/baleap/conference-events/pims/pim-reports/focus-chinese-learners/opening-plenary-session>
- Jing, L., Tindall, E., & Nisbet, D. (2006). Chinese learners and English plural forms. *The*

- Linguistics Journal*, 1(3), 127-147.
<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=3A9EA1F6F667D54A4B4F7AB55C81953A?doi=10.1.1.418.4475&rep=rep1&type=pdf>
- Kramsch, C. (2007). Re-reading Robert Lado, 1957, *Linguistics across cultures*. *Applied linguistics for language teachers* [Review of the book *Linguistics across cultures. Applied linguistics for language teachers*, by R. Lado]. *International Journal of Applied Linguistics*, 17(2), 241-247. doi:10.1111/j.1473-4192.2007.00149.x
- Krifka, M. (1995). Common nouns: A contrastive analysis of Chinese and English. In G. Carlson, & F. J. Lelletier, (Eds.), *The generic book* (pp. 398-411). USA: The University of Chicago Press. Retrieved from <https://books.google.nl/books?hl=en&lr=&id=WUih45y9jcUC&oi>
- Krishnamurthy, R., & Kosem, I. (2007). Issues in creating a corpus for EAP pedagogy and research. *Journal of English for Academic Purposes*, 6, 356–373. doi:10.1016/j.jeap.2007.09.003
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Michigan, USA: University of Michigan Press.
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694–716. doi:10.1093/applin/amn0
- Law of the People's Republic of China on the Standard Spoken and Written Chinese Language. Retrieved from http://www.npc.gov.cn/englishnpc/Law/2007-12/11/content_1383540.htm
- Lay, N. D. S. (1972). *Chinese language interference in written English*. Retrieved from <http://wac.colostate.edu/jbw/v1n1/lay.pdf>
- Leedham, M. (2011). *A corpus-driven study of features of Chinese students' undergraduate writing in UK universities*. (Doctoral thesis). Available from Open Research Online. Retrieved from <http://oro.open.ac.uk/29228/>
- Leedham, M. (2014). 'Enjoyable', 'okay', or 'like drawing teeth'? Chinese and British students' views on writing assignments in UK universities. *Journal of Academic*

- Writing*, 4(1), 1-11. <http://dx.doi.org/10.18552/joaw.v4i1.128>
- Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly*, 28(1), 81-101.
- Lennon, P. (2008). Contrastive analysis, error analysis, interlanguage. In S. Gramely & V. Gramely (Eds.), *Bielefeld introduction to applied Linguistics* (pp. 51-60). Bielefeld, Germany: Aisthesis Verlag. Retrieved from <http://www.homes.uni-bielefeld.de/sgramley/CA-ErrorAnalysis-Interlang-Lennon.pdf>
- Li, D. (2009). Activating strategies to fossilization for English learners in China. *English Language Teaching*, 2(4), 75-77. doi:10.5539/elt.v2n4p75
- Li, H. Y., & Yang L. R. (2010). An investigation of English articles' acquisition by Chinese learners of English. *Chinese Journal of Applied Linguistics*, 33(3), 15-31. Retrieved from www.celea.org.cn/teic/91/10091702.pdf
- Li, J. (2007). Negative language transfer in college English writing by Chinese students: Problems and strategies. *Teaching English in China Journal*, 30(2), 88- 94. Retrieved from <http://www.celea.org.cn/teic/72/72-88.pdf>
- Li, P., Huang, B., & Hsiao, Y. (2010). Learning that classifiers count: Mandarin-speaking children's acquisition of sortal and mensural classifiers. *Journal of East Asian Linguistics*, 19(3), 207-230. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3606901/pdf/nihms422825.pdf>
- Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford, UK: Oxford University Press.
- Lu, T. H. (2010). Correcting the errors in the writing of university students' in the comfortable atmosphere. *International Education Studies*, 3(3), 74-78. Retrieved from <http://www.ccsenet.org/journal/index.php/ies/article/viewFile/6775/5314>
- MacCorquodale, K. (1970). On Chomsky's review of Skinner's Verbal Behaviour. *Journal of the Experimental Analysis of Behavior*. 1(13). 83-99. doi:10.1901/jeab.1970.13-83
- Mair, V. H. (1991). What is a Chinese "dialect/topolect?" Reflections on some key Sino-English linguistic terms. *Sino-Platonic Papers*, 29, 2-31. Retrieved from

- http://www.sino-platonic.org/complete/spp029_chinese_dialect.pdf
- Marlina, R. (2013). Globalisation, internationalisation, and language education: an academic program for global citizens. *Multilingual Education*, 3(5), 1-21.
doi:10.1186/2191-5059-3-5
- McNaughton, W. (2005). *Reading & writing Chinese* (3rd ed.). Boston, USA: Tuttle.
- Milton, J. (2001). Elements of a written interlanguage: a computational and corpus based study of institutional influences on the acquisition of English by Hong Kong Chinese students. In G. James, (Ed.), *Research reports volume 2* (pp. 1-125).
Retrieved from <http://repository.ust.hk/ir/bitstream/1783.1-1055/1/john.pdf>
- Moqimipour K., & Mohsen Shahrokhi, M. (2015). The impact of text genre on Iranian intermediate EFL students' writing errors: An error analysis perspective.
International Education Studies, 8(3). doi:10.5539/ies.v8n3p122
- Nesi, H. (2008) Corpora and English for Academic Purposes. Retrieved from
http://www2.warwick.ac.uk/fac/soc/al/research/collections/bawe/papers/corpora_and_eap.pdf
- Nichol, M. (n.d). Grammar checking software is soft on errors. *Daily Writing Tips*.
Retrieved from <http://www.dailywritingtips.com/grammar-checking-software-is-soft-on-grammar-errors/>
- Nickalls, R. (2011). How definite are we about articles in English? A study of L2 learners' English article interlanguage during a university preessional English course.
Proceedings from the 2011 Corpus Linguistics Conference, Retrieved from
<http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/Paper-92.pdf>
- Norman, J. (1988). *Chinese*. Cambridge UK: Cambridge University Press.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge, UK: Cambridge University Press.
- Oxford, R. L. (1999). Relationships between second language learning strategies and language proficiency in the context of learner autonomy and self regulation.
Revista Canaria de Estudios Ingleses, 38, 109-126. Retrieved from

- [http://publica.webs.ull.es/upload/REV%20RECEI/38%20%201999/08%20\(Rebecca%20L.%20Oxford\).pdf](http://publica.webs.ull.es/upload/REV%20RECEI/38%20%201999/08%20(Rebecca%20L.%20Oxford).pdf)
- Parr, C. (2015, February 5). Best University Workplace Survey: unease at inadequate language skills. *Times Higher Education*. Retrieved from: <http://www.timeshighereducation.co.uk/news/best-university-workplace...rvey-unease-at-inadequate-language-skills/2018345.article?sm=2018345>
- Peterlin, A. P. (2005). Text-organising metatext in research articles: An English–Slovene contrastive analysis. *English for Specific Purposes*, 24, 307–319. doi:10.1016/j.esp.2004.11.001
- Phuket, P.P.R.N., & Othman, N.B. (2015). Understanding EFL students' errors in writing. *Journal of Education and Practice*, 6(32), 99-106. Retrieved from: <http://files.eric.ed.gov/fulltext/EJ1083531.pdf>
- Plag, I. (2003). *Word-formation in English*. Cambridge, UK: Cambridge University Press.
- Porte, G. (1999). Where to draw the red line: error toleration of native and non-native EFL faculty. *Foreign Language Annals*, 32(4), 426-434. <http://onlinelibrary.wiley.com/doi/10.1111/j.1944-9720.1999.tb00873.x/abstract>
- Presada, D., & Badea, M. (2014). The effectiveness of error analysis in translation classes. A pilot study. *Porta Linguarum*, 22, 49-59. Retrieved from: http://www.ugr.es/~portalin/articulos/PL_numero22/4%20%20DIANA%20PRESADA.pdf
- Rahimpour, M. & Dovaise, M. S. (2011). A phonological contrastive analysis of Kurdish and English. *International Journal of English Linguistics*, 1(2), 73-82. doi:10.5539/ijel.v1n2p73
- Ramsey, S. R. (1987). *The languages of China*. Princeton USA: Princeton University Press.
- Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. (3rd ed.). London, UK: Pearson.
- Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, 16(2), 135–172.

doi:10.1191/026765800672262975

- Robins, D. (2000). Mass nouns and count nouns in classical Chinese. *Early China*, 25, 147-184. doi:10.2307/23354276
- Roikiene, D. (2014). Contrastive analysis of incorporation in English and Lithuanian. *Res Humanitariae*, 16, 235–248. Retrieved from <http://journals.ku.lt/index.php/RH/article/viewFile/1023/1206>
- Rustipa, K. (2011). Contrastive Analysis, Error Analysis, Interlanguage and the implication to language teaching. *Ragam Jurnal Pengembangan Humaniora*, 11(1), 16-22. Retrieved from http://www.polines.ac.id/ragam/index_files/jurnalragam/paper_3%20apr_2011.pdf
- Sanders, R. M. (1987). The four languages of ‘Mandarin’. *Sino Platonic Papers*, 4. Retrieved from http://www.sinoplatonic.org/complete/spp004_mandarin_chinese.pdf
- Santos, T. (1988). Professors' reactions to the academic writing of nonnative-speaking students. *TESOL Quarterly*, 22(1), 69-90.
- Saville-Troike, M. (2012). *Introducing second language acquisition* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, 59(4), 339-341. doi:10.1093/elt/cci064
- Selinker, L. (1972). Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching*, 10,(1-4), 209–232. doi:[10.1515/iral.1972.10.1-4.209](https://doi.org/10.1515/iral.1972.10.1-4.209)
- Sheen, R. (1996). The advantage of exploiting contrastive analysis in teaching and learning a foreign language. *IRAL - International Review of Applied Linguistics in Language Teaching*, 34(3), 83-197. doi:[10.1515/iral.1996.34.3.183](https://doi.org/10.1515/iral.1996.34.3.183)
- Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students*. (3rd ed.). USA: University of Michigan Press.
- Swan, M. (2007). Follow-up to Claire Kramsch’s ‘classic book’ review of Lado, 1957, linguistics across cultures: History is not what happened: the case of contrastive

- analysis [Review of the book *Linguistics across cultures. Applied linguistics for language teachers*, by R. Lado]. *International Journal of Applied Linguistics*, 17(3), 414-419. doi:10.1111/j.1473-4192.2007.00163.x
- Tan, H. M. (2007). A study of EFL learners' writing errors and instructional strategies. *Journal of Kun Shan University*, 4, 113-122. Retrieved from http://ir.lib.ksu.edu.tw/bitstream/987654321/12216/2/10.p113~p122_D12-10英語學習者之寫作錯誤分析及教學策略之探討.pdf
- Taylor, J. R. (2012). *The mental corpus; How language is represented in the mind*. Oxford, UK: Oxford University Press.
- Throssell, P. & Zhao, Y.Y. (2011). Developing motivation in ESL learning in the Chinese context. *Journal of the World Universities Forum*, 4. Retrieved from <http://wuj.cgpublisher.com/product/pub.173/prod.303>
- Tizazu, Y. (2014). A linguistic analysis of errors in the compositions of Arba Minch University students. *Advances in Language and Literary Studies*, 5(3), 191-205. doi:10.7575/aiac.all.v.5n.3p.195
- Touchie, H. Y. (1986). Second language learning errors their types, causes, and treatment. *Japan Association for Language Teaching Journal*, 8(1), 75-80. Retrieved from <http://jalt-publications.org/jj/articles/1571-second-language-learning-errors-their-types-causes-and-treatment>
- Truscott (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327-369. Retrieved from http://lecture.ecc.utokyo.ac.jp/~cwpgally/references/2009S_TEW_Truscott_original_article.pdf
- UK Council for International Student Affairs. (2015). *International student statistics: UK higher education*. Retrieved from <http://www.ukcisa.org.uk/Info-for-universities-colleges--schools/Policy-research--statistics/Research--statistics/International-students-in-UK-HE/>
- Universities UK. (2013). *Parliamentary brief. The UK's relationship with China: Universities*. Retrieved from

- <http://www.universitiesuk.ac.uk/highereducation/Documents/2013/UKandChina.pdf>
- Universities UK. (2014). *International students in higher education: The UK and its competition*. Retrieved from <http://www.universitiesuk.ac.uk/highereducation/Documents/2014/InternationalStudentsInHigherEducation.pdf>
- Vann, R. J., Meyer, D. E., & Lorenz, F. O. (1984). Error gravity: a study of faculty opinion of ESL errors. *TESOL Quarterly*, 18, (3), 427-440.
- Vicary, A. (2014). *Grammar for writing*. Reading, UK: Garnet.
- Wang, S., & Goodman, K. (2014). Making sense of Chinese reading: Yi and Xing. *Open Journal of Modern Linguistics*, 4, 621-640.
<http://dx.doi.org/10.4236/ojml.2014.45054>
- Wang, E., Matsuo, V., Matsuo, H., & Pascoe-Chavez, D. (2013). Factors associated with the perceived English competency of Chinese international students in ESL/EAP classrooms: the influence of motivation, social networks and beliefs on authority. *Proceedings from The Asian Conference on Cultural Studies 2013 Official Conference Proceedings 2013*. Osaka, Japan. Retrieved from http://iafor.org/archives/offprints/accs2013-offprints/ACCS2013_0341.pdf
- Wang, F., Tsai, Y., & Wang, W. S. F. (2009). Chinese literacy. In D. Olson & N. Torrance. (Eds.), *The Cambridge handbook of literacy* (pp. 386-417). Cambridge University Press: New York USA. Retrieved from https://books.google.com/books?id=8TyOC9nqEokC&pg=PA411&lpg=PA411&dq=homogeneity+of+written+cantonese+and+mandarin&source=bl&ots=iUAwxXlGOe&sig=bm0SzKv5NfzBzfFCAzGsZdlqo4&hl=en&sa=X&ved=0ahUKEwjsucTv5XLAhWGTBQKHXP_DQQ6AEIQzAF#v=onepage&q=homogeneity%20of%20written%20cantonese%20and%20mandarin&f=false
- Wingate, (U). (2015). Academic literacy across the curriculum. Towards a collaborative instructional approach. Retrieved from http://eacconference.elc.polyu.edu.hk/docs/eac2015/Ursula_abstract.pdf

- Wong, S. M. J., & Dras, M. (2009). Contrastive analysis and native language identification. In L. A. Pizzato & R. Schwitter (Eds.), *Australasian language technology association workshop 2009 proceedings of the workshop*. 53-61. Retrieved from: <https://aclweb.org/anthology/U/U09/U09-1.pdf>
- Wu, A. (1991). Thinking about prof. John deFrancis. *Sino-Platonic Papers*, 27, 229-233. Retrieved from http://sino-platonic.org/complete/spp027_john_defrancis.pdf
- Wu, Y. C., & Bodomo, A. (2009). Classifiers ≠ Determiners. *Linguistic Inquiry*, 40, (3), 487–503. doi:10.1162/ling.2009.40.3.487
- Xing, M. J., Wang, J. H., & Spencer, K. (2008). Raising students' awareness of cross-cultural contrastive rhetoric in English writing via an e-learning course. *Language Learning & Technology*, 12,(2), 71–93. Retrieved from <http://www.hull.ac.uk/php/edskas/articles/LLT2008.pdf>
- Xu, X. Y. (2012). Cultural factors in EAP teaching — influences of thought pattern on English academic writing. *Cross-Cultural Communication*, 8(4), 53-57. <http://dx.doi.org/10.3968/j.ccc.1923670020120804.2307>
- Yang, B. G. (1992). A review of the contrastive analysis hypothesis. *Ren wen, she hui ke xue pian*, 19, 133-149. Retrieved from <http://fonetiks.info/bgyang/db/92cah.pdf>.
- You, X.Y. (2004). New directions in EFL writing: a report from China. *Journal of Second Language Writing*, 13, 253–256. doi:10.1016/j.jslw.2004.09.002
- Zhang, H. W., & Xie, J. (2014). Interlanguage fossilization in Chinese EFL writing —an empirical research of 20 English major students. *Sino-US English Teaching*, 11(4), 248-258. Retrieved from <http://www.davidpublishing.com/davidpublishing/Upfile/4/21/2014/2014042104386414.pdf>
- Zheng, C., & Park, T.J. (2013). An analysis of errors in English writing made by Chinese and Korean university students. *Theory and Practice in Language Studies*, 3(8), 1,342-1,351. doi:10.4304/tpls.3.8.1342-1351

Appendices

Appendix 1. Tone markers in pinyin

A detailed description of the phonetic system in Chinese is not required for this analysis because the focus is on writing, however for completeness the appropriate diacritic markers will be included with the pinyin that is, the system of Romanised letters to represent the phonetic system in Chinese. Standard Chinese has 4 tones, and what is sometimes referred to as a neutral tone. In Chinese characters, or *hànzì*, these are not marked. In the pinyin system the four tones are marked as:

First Tone – flat tone, represented by (1) or (ˉ)

Second Tone – rising tone, represented by (2) or acute accent (´)

Third Tone – falling and then rising tone, represented by (3) or (˘) (often joined together)

Fourth Tone – falling tone represented by (4) or the grave accent (`)

The neutral tone is sometimes referred to as the Fifth Tone, but is not marked.

Sometimes the required tones are not marked in pinyin because although the tone is important for identifying the meaning of a spoken word in isolation, the meaning of a word can be derived from the context and therefore the diacritic marking is not always required when writing a basic sentence or phrase. Tone sandhi, the altering or omission of a tone, occurs in certain circumstances in spoken Chinese but these are not reflected in the diacritic markings in pinyin. Cantonese pinyin nowadays uses numbers to represent tones and not tone markers (as it did in the Yale system).

Appendix 2. List of potential corpora of Chinese students' academic writing.

- 1) BAWE (British Academic Written English) Students studying in UK with Chinese as their L1.
- 2) Lancaster Corpus of Academic Written English (LANCAWE) (EAP not discipline-specific texts)
- 3) The Bilingual Corpus of Chinese English Learners (BICCEL) Chinese (file corrupted according to website <http://www.corpus4u.org/threads/2288/>).
- 4) The Cambridge Learner Corpus (CLC) (ESOL exams – not discipline-specific texts)
- 5) The Chinese Academic Written English (CAWE) Corpus Chinese Hong Kong (relates to students writing in mainland China only)
- 6) The Chinese Learner English Corpus (CLEC)– not accessible to non University staff
- 7) The ETS Corpus of Non-Native Written English (TOEFL English not discipline-specific).
- 8) The Hong Kong University of Science & Technology (HKUST) learner corpus not university level
- 9) The Indianapolis Business Learner Corpus (IBLC) (job applications etc, not academic texts).
- 10) The International Corpus Network of Asian Learners of English (ICNALE) various(short essays, 300 words, Chinese student contributors are writing in Chinese colleges)
- 11) The International Corpus of Crosslinguistic Interlanguage (ICCI) (focus is on young learners)
- 12) The International Corpus of Learner English (ICLE) prohibitive cost
- 13) The Learner Corpus of Essays and Reports (available to Department of English of Hong Kong PolyU staff only).
- 14) The Longman Learners' Corpus (only commercially available)
- 15) The Montclair Electronic Language Database (MELD) Students in USA.
- 16) The NUS Corpus of Learner English Students in Singapore
- 17) The Spoken and Written English Corpus of Chinese Learners (SWECCCL) In China
- 18) The TELEC Secondary Learner Corpus (TSLC) Secondary School in Hong Kong

Appendix 3

List of students by Student ID number including essay identification letter and Date of Birth (DOB), Gender (g), Date of Submission of essay (DOS) and Age (at submission of essay) and L1. (M=Mandarin, C= Cantonese, U=unspecified).

First Year	DOB(g) DOS AGE L1	Third Year	DOB DOS AGE L1	
0008a	1985m 2004 19 M	0018c	1982m 2005 23 C	
0041a	1985f 2005 20 M	0155a	1983m 2006 23 U	
0071a	1987f 2005 18 C	0197a	1983f unknown U	
0080a	1984f 2004 20 U	0234j	1983f 2005 22 U	
0100a	1986f 2006 20 U	0254h	1982m 2006 24 U	
0123a	1985m 2005 20 C	0287a	1984f 2003 19 U	
0125a	1986m 2005 19 M	0357b	1985f unknown C	
0132a	1986f 2005 19 C	0378b	1986f 2007 21 C	
0142a	1985f 2006 21 C	0396a	1986m 2007 21 U	
0271a	1987f 2006 19 U	0410a	1984m unknown U	
0351a	1986m unknown C	0434a	1985f 2007 22 U	
3018a	1983f 2002 19 U	3010a	1982f 2004 22 C	
3047b	1984f 2004 20 U	3039a	1980m 2005 25 C	
3085a	1983m 2004 21 C	6008a	1984f 2005 21 C	
3093a	0000m 2006 U	6102a	1983m 2006 23 U	
6025b	1983f 2004 21 U	6107a	1986f 2006 20 U	
6081a	1984f 2005 21 U	6150a	1984f 2006 22 U	
6082a	1986f 2005 19 C	6215a	1963m 2006 43 U	
Total = 18	316	Total = 18	351	667
	Average Age = 19		Average Age =23	Average Age= 21

Appendix 4. Essay Topics

A number of the essays were untitled.

First Year

'Economics and the Structure of Industry Marketing/ Industry Exercise'

'Experiment 14: Growth curve of *Serratia marcescens*'

'Should the Dutch economy be seen as the first modern economy?'

'Theoretical models of consumer behaviour often stress the importance of perceived personal wealth, but models based on empirical data still rely heavily on the influence of current personal income. Why?'

'Experiment 11: Properties of enzymes and the kinetics of enzyme action'

'Business strategies demand discipline in the execution of long-term strategic plans and flexibility to address emergent changes. Discuss. Explain which one of the two features is more critical in your view.'

'Report'

'Analysis of Pepper vs Hart'

'Explain the importance of the case of *DONOGHUE V STEVENSON* 1932 AC 562 in the development of the law of negligence.'

'Business strategies demand discipline in the execution of long-term strategic plans and flexibility to address emergent changes. Discuss. Explain which one of two features is more critical in your view.'

'Ethics First term assessment'

'Humanoid Robotics in Artificial Intelligence'

'Determination of kinetic parameters (K_m and V_{max}) for the fumarase- catalyzed reaction'

'SURFACE TENSION'

'Features and Procedures - Improving Resolution in Microscopy'

Third Year

'Racing Engines'

'Assignment 1'

'Outline a version of the "first generation" or the "second generation" speculative attack model. Discuss the empirical evidence in support of the model you outlined. Briefly discuss the limitations and the extensions of the model.'

'What, if anything, has membership of a political community in common with membership of a family? Can this tell us anything about our obligations?'

'Heat Exchanger Design Exercise'

'Monopoly And Resource Allocation'

'The scale plan of what you intended the robot to draw'

'Extraction and Characterization of Natural Products'

'ST323 Multivariate Statistics Assignment One'

'There is no room for the introduction of "contributory negligence" on the part of claimants into the realm of liability of trustees for breach of their duty of care'. Discuss.

'The role of maternal effect genes in the development of the nematode *Caenorhabditis elegans*'

'Analysis on Bards Hall Hotel'

'Microbiology Lab Report (Expt 1)'

'Report for Artificial Intelligence'

'DSP Laboratory Session'

'Literature Review of Project - Whey Protein Concentrate in Ice Cream'

'OVERCOMING SEED DORMANCY'

Appendix 5. List of students by Student ID number including essay identification letter and word count.

First Year	Word Count	Third Year	Word Count	
0008a	1,706	0018c	5,437	
0041a	1,157	0155a	3,469	
0071a	1,766	0197a	2,761	
0080a	987	0234j	2,787	
0100a	973	0254h	1,049	
0123a	1,435	0287a	2,096	
0125a	869	0357b	905	
0132a	1,483	0378b	2,739	
0142a	1,624	0396a	1,335	
0271a	1,320	0410a	5,693	
0351a	558	0434a	3,017	
3018a	2,208	3010a	658	
3047b	2,892	3039a	2,428	
3085a	3,108	6008a	1,181	
3093a	1,420	6102a	852	
6025b	561	6107a	938	
6081a	715	6150a	1,147	
6082a	556	6215a	798	
Total = 18	Total = 25, 275	Total = 18	Total = 39,290	Total = 64,565
	Average Word Count per essay = 1,404		Average Word Count per essay = 2,183	Average Word Count Total =1,793

Appendix 6. Error Correction Codes and Category Definitions

MA = Missing Article. Assessed that an article should have been used but was not.

RA = Redundant Article. Assessed that an article should not have been used, but one was provided.

IA = Incorrect Article. An article was provided however it was the incorrect article, either a definite article used for an indefinite article or vice versa, or where 'a' was used instead of 'an', or vice versa.

S-V = Subject verb agreement incorrect. Assessed as the verb form did not match the subject.

S'/P = Should have been singular. Assessed as requiring a singular form but the plural was provided.

S/P' = Should have been plural. Assessed as requiring a plural form but a singular form was provided.

T = Incorrect tense. Assessed as being inconsistent with the rest of the paragraph.

MW = Missing Word. Assessed as requiring another word for completion, not included in other categories.

WO = Word Order. Assessed as requiring the word order to be reversed to be correct.

Pr = Preposition incorrect or missing. A preposition is included but is not the correct preposition or a preposition is required but is missing.

MP = Missing Possessive. The possessive is required but has not been provided.

RP = Redundant Possessive. No possessive is required but one has been provided.

WF = Word Form. The form of the word is incorrect, for example, an adverb is required but an alternative form such as an adjective is provided.

S = Incomplete sentence. The sentence is constructed in such a way that it is not complete on its own and requires either further information to be complete or should be linked to a subsequent sentence.