

อิทธิพลของความถี่ของประเภทคำและผลกระทบของภาษาแม่  
ที่มีต่อการใช้โครงสร้างภาษาอังกฤษที่ไม่เป็นไปตามแบบแผน  
ในงานเขียนของนักวิชาการไทย  
The Influence of Word-Type Frequency and Mother Tongue  
Effects on Non-Conventional English-Structure Usage in Thai  
Academics' Writing

Michael Currie<sup>1</sup>

Kemtong Sinwongsawat<sup>2</sup>

Kathleen Nicoletti<sup>3</sup>

<sup>1,2,3</sup>Department of Languages and Linguistics,

<sup>1,2,3</sup>Faculty of Liberal Arts,

<sup>1,2,3</sup>Prince of Songkla University, Hatyai campus

<sup>1</sup>Corresponding author: miccurrie@hotmail.com

บทคัดย่อ

งานวิจัยชิ้นนี้รายงานผลจากการวิเคราะห์บทความจำนวน 26 ชิ้นของนักวิชาการชาวไทยเพื่อตีพิมพ์ในวารสารภาษาอังกฤษ โดยการวิเคราะห์ดังกล่าวอิงกับคำแนะนำซึ่งผู้ตรวจแก้ภาษาที่เป็นเจ้าของภาษามีให้แก่ผู้เขียนก่อนส่งบทความเพื่อรับการพิจารณาตีพิมพ์ในวารสาร ทั้งนี้ผลที่รายงานเกี่ยวกับการใช้ภาษาอังกฤษที่ไม่เป็นไปตามแบบแผนที่พบในบทความจากการวิเคราะห์นั้นเชิงปริมาณเป็นเพียงส่วนหนึ่งของผลการศึกษาที่ได้จากงานวิจัยที่ทำขึ้นเพื่อระบุขอบข่ายโครงสร้างภาษาอังกฤษที่มักเป็นปัญหาสำหรับผู้เขียนบทความชาวไทยโดยการใช้กระบวนการวิจัยแบบผสมผสานสำหรับในบทความนี้ผู้วิจัยได้ทบทวนการวิเคราะห์ข้อผิดพลาดในงานเขียนภาษาอังกฤษของคนไทยที่ได้มีการศึกษามาก่อนหน้า ซึ่งส่วนใหญ่เป็นการวิเคราะห์งานเขียน

ของนักเรียนระดับมัธยมศึกษาและนักศึกษาระดับปริญญาตรี ทั้งนี้เมื่อเปรียบเทียบผล  
การศึกษาจากงานวิจัยนี้กับผลที่ได้จากการศึกษาที่ผ่านมาพบว่า ปัญหาหลักที่พบ  
ในการเขียนซึ่งเกี่ยวข้องกับการใช้คำนำหน้านาม (articles) กริยา (verbs), บุพบท  
(prepositions) และ คำนาม (nouns) เป็นปัญหาที่ปรากฏในบทความจำนวน 26 ชิ้นที่  
ศึกษาเช่นกัน โดยพบปัญหาเหล่านี้มากกว่า 80% ของโครงสร้างการใช้ภาษาอังกฤษ  
ที่ไม่เป็นไปตามแบบแผนทั้งหมดที่วิเคราะห์ นอกจากนั้นผู้วิจัยยังได้วิเคราะห์จำนวนคำ  
แต่ละประเภทที่ใช้ในบทความพร้อมหาความสัมพันธ์ระหว่างข้อผิดพลาดที่พบใน  
การใช้คำแต่ละประเภท และความถี่ที่คำประเภทต่างๆ เหล่านี้ปรากฏในการเขียน  
บทความโดยได้ค่าความสัมพันธ์ที่ .86 ซึ่งการศึกษาก่อนหน้านี้ยังไม่มีการหาค่าความ  
สัมพันธ์ดังกล่าว ทำให้อาจเกิดความเข้าใจคลาดเคลื่อนเกี่ยวกับระดับความยากง่าย  
ของการใช้คำแต่ละประเภท ถึงกระนั้นก็ตามพบว่าประเด็นที่เป็นปัญหาในการเขียน  
ส่วนใหญ่ที่รายงานมีความสอดคล้องกันโดยเฉพาะประเด็นที่เกี่ยวข้องกับความแตกต่าง  
ระหว่างโครงสร้างภาษาไทยและภาษาอังกฤษ เมื่อเชื่อมโยงกับกรอบแนวคิดทฤษฎีการ  
เรียนรู้ ผลการศึกษาในครั้งนี้ชี้ให้เห็นว่าถึงแม้ความถี่ของการใช้คำประเภทต่างๆ  
ที่ปรากฏในบทความจะเป็นปัจจัยหนึ่งที่สัมพันธ์กับปัญหาที่พบในการใช้โครงสร้างภาษา  
ที่ไม่เป็นไปตามแบบแผน แต่ปัจจัยเรื่องภาษาแม่ยังคงเป็นปัจจัยหลักที่มีผลต่อการ  
เกิดข้อผิดพลาดในการใช้โครงสร้างภาษาอังกฤษที่ไม่เป็นไปตามแบบแผนที่ศึกษาอย่าง  
มีนัยสำคัญ ฉะนั้นจึงสรุปว่าการสอนภาษาอังกฤษในทุกระดับควรให้ความสำคัญ  
โดยเฉพาะอย่างยิ่งกับการสอนการใช้คำนำหน้านามเพื่อชี้เฉพาะ (definite articles)  
การใช้รูปกาลที่เป็นอดีตและปัจจุบัน (Past and Present Simple Tense) นามเอกพจน์  
และพหูพจน์ รวมทั้งคำบุพบท เพื่อลดปัญหาที่พบในการใช้ภาษาอังกฤษในการเขียน  
ระดับที่สูงขึ้น

**คำสำคัญ :** การเขียนเชิงวิชาการ, การวิเคราะห์ข้อผิดพลาด, โครงสร้างภาษา, ประเภทคำ,  
ความถี่ของคำ

## Abstract

The study reported analyzed 26 manuscripts written by Thai academics for publication in English language journals. The analysis was based on the suggestions for changes made by a native-English speaking editor prior to submission. The study aimed to establish which areas of English structure cause most problems to Thai writers at this level. The overall study employed a mixed-methods approach although the findings reported in this paper are derived from the quantitative analysis of the non-conventional usage (NCU) of English in the manuscripts. The paper reviews previous error analyses of English writing by Thais, mostly involving high school or undergraduate students. The paper compares the findings of those studies with those of the study reported, finding that four of the main problem areas identified in those studies (article, verb, preposition and noun usage) were also the structural aspects most frequently identified in the 26 manuscripts analyzed, accounting for more than 80% of the structure NCU. However, the study also analyzed the number of tokens of each word type in the manuscripts and identified a correlation of 0.86 between the errors in those word types and the frequency of occurrence of the word types in the manuscripts. Previous studies have not identified such a relationship thus potentially misrepresenting the level of difficulty presented by different word-type usage. Nevertheless the areas giving rise to most problems are also noted to coincide with areas where Thai language structure differs from English and based on an associative learning theoretical framework, the paper concludes that although the distribution of word types in the manuscripts is a significant influence on the pattern of NCU, the influence of the L1 is the major factor behind the main types of errors identified. The paper concludes that a relatively small number of areas should be targeted at all stages of the teaching of English, notably the use of the definite article, the past

and present simple tenses, singular and plural nouns and preposition use generally, in order that problems in these areas should not later become entrenched in higher level usage.

**Keywords:** Academic writing, error analysis, language structure, word type, word frequency

Academic writing falls within the broader genre of English for academic purposes (EAP) which Gillett (1996) noted, covers all aspects of the use of English in the academic field, although identifying writing as the most important, with accurate grammar and the formal language of the genre being crucial components. English is well established as an international language in academic publication and is widely accepted even in countries where English is not commonly spoken. Larsen and von Ins (2010), in reviewing a number of subject areas, noted a tendency for publication indexes to favor journals publishing articles in English and as an extreme example, Clarke et al. (2007) reviewed over 200,000 articles in the field of public health listed in the Science Citation Index and found that 96.5% were written in English. Further, as Vasconcelos, Soerenson and Leta (2009) noted, a lack of skill in English is a significant barrier for publication in international journals. There is therefore considerable pressure on academics to write in English but for those whose first language is not English, this can present a considerable challenge. This is the case in Thailand, which has no historical connection with the English speaking world either through colonial occupation or a common cultural heritage, and English skills are not widespread. In the latest EF-EPI survey (2018), Thailand ranked 64<sup>th</sup> of 88 non-English speaking countries for English proficiency.

The broader study from which this paper is drawn was based on the work of authors mostly based in the south of Thailand who had sought to have their work edited by a native speaker prior to submission for publication in English. The study arose from the editor's observation that a substantial proportion of the changes he recommended were based on the authors' difficulties in correctly using apparently simple aspects of English structure. This was surprising since the authors were generally able to control the

rhetoical aspects of the academic writing genre, yet had great problems in basic aspects of structure such as the use of articles, verb tenses and noun pluralization. No previous study in Thailand has investigated this phenomenon in writers at this level, who by their ability to succeed in writing articles in English for publication in academic journals, must be assumed to be among the top tier of English users in Thailand. This paper therefore offers a rare insight into the use of English by a professional group in Thailand for whom that use is dictated by the demands of their careers.

Although there have been previous findings of structure-usage problems in writing by Thais, past studies, of which more than 30 conducted since 2000 were traced and are detailed in Table 1, have focused almost exclusively on students during their education and on English writings either produced specifically for the purposes of the study or for purposes associated with their education. This study sought to establish whether the problems evidenced by the non-conventional usage (NCU) of English structure by these 26 Thai academics were similar to those which previous studies have identified among Thai students. This paper therefore presents the results of the analysis of the changes the editor recommended to the 26 papers and focuses on those changes recommended to remedy non-conventional language structure.

The following section reviews recent studies employing error analysis (EA) in Thailand and sets out the study's theoretical framework with respect to the identification of mother-tongue (or L1) influence on second language (L2) learning before setting out the research questions considered in this paper. The study's methodology and results are described and then discussed, and conclusions are offered in the final section.

## Literature review and research questions

### *Errors in second language learning*

The history of the treatment of errors in L2 learning can be traced back to the behaviorists who viewed language as a habit and employed the technique of contrastive analysis (CA) to identify structural differences between languages and to predict the errors learners would make in learning an L2 through L1 effects (Lado, 1957). But as the influence of behaviorism waned, the treatment of errors shifted to a perspective which viewed the learner as producing a transitional idiolect or interlanguage (IL) (Selinker, 1972). Under this approach, errors were viewed as indicative of learning and arose due to a variety of causes including external factors, such as the learner's L1 and the conditions under which the language was learned, as well as intra-linguistic and developmental factors, which were described by Richards (1970) as systematic errors common to learners with different L1s, including overgeneralization, ignorance or incomplete application of rules and the creation of false hypotheses.

Error analysis is a well-used technique of analyzing language production and the subject has its own published bibliography, which in its latest edition (Spillner, 2017) identifies more than 6000 studies. The genesis and conduct of EA is also set out at book-length in James (1998) and was usefully reviewed by Sompong (2014), who identified a number of recent studies conducted in Thailand. A more extensive review of Thai studies into aspects of EFL writing research in Thailand was carried out by Chuenchaichon (2014), who identified 48 studies conducted between 2004 and 2013 of which nine dealt with writing errors, although of those only four were studies which could strictly be said to constitute EA. Hinnon (2015) also reviewed EA studies in Thailand, citing nine, split into three groups, grammatical-lexical errors, L1 interference errors and errors in writing organization.

In fact the authors have traced 33 EA studies conducted in Thailand since 2000 which are detailed in Table 1. This list is not exhaustive and only includes generalized EA studies which included but were not necessarily restricted to grammatical errors. Thus, work which studied only particular areas are excluded, notably Pongpairaj's (2007) and Nopjirapong's (2011) studies of article errors. Also excluded are the only two studies traced, analyzing journal articles written by Thai academics (Amnuai and Wannaruk, 2012: Jaroongkhongdach, Watson-Todd, Keyuravong, and Hall, 2012) since they did not consider grammatical aspects of the authors' work. Of the 33 studies listed in Table 1, only Sereebenjapol (2003), Bootchuy (2008) and Runkati (2013) studied post-graduate students. Most of the studies analyzed work by undergraduate and high-school students (23 and 5 respectively), while only two (Chakorn, 2005; Hutyamanivudhi, 2001) considered writing in a non-educational context by analyzing business correspondence.

Table 1 shows the major areas identified in the studies as giving rise to errors (more than 5% of the total, or as identified in its abstract or in a citation in another work). Summing the areas identified shows that the largest number of references (25) were to errors in verbs, notably in tenses or forms, while prepositions (18) and articles and determiners (18) were the next largest categories. Other categories identified included fragments or the ellipsis of sentence elements (15), nouns including pluralization (14), subject/verb agreement (14), sentence structure/syntax (13) mechanical errors (13), lexical choice (12), word order (8) pronouns, adjectives and word form (4 each) conjunctions and subordinators (3) and adverbs (2). Whilst these figures are not intended as an accurate meta-analysis of the findings of the studies, they do broadly indicate the areas which the studies identified as causing most problems for their participants. Further, in considering the studies' methodologies,



it is notable that none considered the frequency of the occurrence of different word types in the participants' work in order to take that into consideration when assessing the level of difficulty faced by the users in using particular types of words in their work.

Finally, the cause of the errors identified were largely attributed either partly or wholly to L1 effects (24 studies) and/or based on Richards' (1970) classification referred to above (overgeneralization, etc. – 12 studies). Other causes referred to include carelessness, translating from Thai and poor language skills.

Table 1. Reports of generalized error analysis research in Thailand between 2000 and 2017

Study	Year	Writing from	Main problems detected	Attributed to
Thananart (cited in Sompong, 2014)	2000	U/g students	CS, VF, WC, Sp, transition signals	(Unknown)
Chownahe (cited in Bootchuy, 2008)	2000	H/s students	Adj, V, Trans, WO	Inter E, Intra E & Dev E
Likitratnaporn (cited in Bennui, 2008)	2001	3rd yr u/g students	SS & WO, art ellip	Trans
Hufyamanivudhi	2001	Company officers	V (22%), Prep (19%), Det (16%)	Inter E, overgen and lack of practice
Pongsriwet (cited in O'Donnell, 2016)	2001	1st yr u/g students	Sub/Vagr, VF & T, NPI, WF, frag, Art, Prep, Pro	Inter E
Stichai	2002	1st yr u/g students	Det, VT, Prep, Lex, Frag, WO	Inter E
Pongpairaj (cited in O'Donnell, 2016)	2002	1st yr u/g students	Syn & Lex E, Prep Coll, pl of n/c N, use of existential 'there'	Inter E, Ignorance of diff'ces in Thai & Eng GS
Lush	2002	3rd yr u/g students	Art, sing/pl Ns, interchanged & incorrect VT, Preps, Sub/V agr	(Descriptive study)
Aywaratana (cited in Bootchuy, 2008)	2002	4th yr u/g students	Punc, Sub/V agr, Frags & ROSs	Inter E
Khaourai	2002	U/g students	Preps, VT, Det, Contractions, SS, WCSS, WO, ME, Trans, overgen	IRA & IRR, Inter E & FCH
Serebenjapol	2003	Postgrad students	Art, V, NPI, Subordinators and conj	Carelessness, IRA, Inter E
Khamput	2004	Yr 11 h/s students	WO, Trans, ellip Sub	Inter E, Intra E & Dev E
Rurakwit	2004	U/g students	Frags, obj, V & Sub ellip, misplacement of advs & adjs	Inter E
Na-Ngam	2004	1st yr u/g students	Incomplete S, N, Sub/Vagr, Sp, VT, Art	Inter E, IRA, FCH, avoidance, carelessness
Chakorn	2005	Company officers	V (22%), Prep (19%), Det (16%)	Inter E

Study	Year	Writing from	Main problems detected	Attributed to
Banlomchon	2006	Yr 12 h/s students	DetS, WC, VF&T, Ns, SubVagr, Preps (57%), Lex (25%), ME (18%)	Inter E, overgen, IRR, IRA, FCH
Sattayatham and Honsa	2007	1st yr u/g students	Overgen, IRA, FCH	Inter E
Bennui	2008	3rd yr u/g students,	Voc, WO & SS, VT, SubVagr, inf, ellip G morph, Prep, Det, overgen and double marking, have/there conf	Inter E, Trans, L1 cultural knowledge and writing style
Sattayatham and Ratanapinyowong	2008	1st yr u/g students	No intro & conc, coh or org. Difficulty using Eng G and paras	Lack of writing ability at a paragraph level
Boothuay	2008	1st yr post-grad students	Ellip S elements (38%), compound & complex S (23%), WO (9%)	Inter E (47%), IRA (9%)
Jenwitheesuk	2009	3rd yr u/g students	Det, SubVagr, VT, Prep	Inter E, FCA ignorance of Eng SS and G
Ampornratana (cited in Phetdamuea and Ngongkum, 2016)	2009	Yr 12 h/s students	VF & T, Pro, Adj, subVagr, Prep, Conj, N, overuse of and ellip be, sub/obj rep	Inter E, Intra E and Dev E
Watcharapunyawong and Usaha	2012	2nd yr u/g students	VT&F, WC, SS, Art, Prep, mod &aux, NPI, Frag, Pro, ROSSs, SubVagr, inf/Ger, transitions	Inter E (assumed)
Runkati	2013	Postgrad students	Art (26%), N (18%), Prep (13%) Y (6%)	(Not stated)
Nonkukhetkhong	2013	1st yr u/g students	V, N, Poss, Art, Prep, Adj and Adv (47%), Syn (20%) substance & ME (19%) Lex (12%)	Inter E, overgen, IRR, IRA
Iamsiu	2014	2nd yr u/g students	WC (46%), SS (33%), SubVagr (13%), WO (11%)	Inter E
Andania	2015	Yr 7 h/s students	NPI, VT, SubVagr, WF, extraneous Sub, Punc	(Descriptive study)
Rattanadilok-Na-Phuket	2015	2nd & 3rd yr u/g students	Trans, VT, Prep, WC, VF, Sp	Inter E, Intra E, overgen, IRA, IRR
O'Donnell	2015	1st yr u/g students	VT, ME, Art & Det, Pro, NPI, SubVagr, Prep	Weak language skills, not noticing & Inter E

Study	Year	Writing from	Main problems detected	Attributed to
Pheiddannuea and Ngongkum	2016	2nd yr u/g students	Sub/Vagr, Det, ROS, Trans, VT, SS, NPI, Punc, VF, Prep, WC	Inter E, IRR, IRA, FCH
Promsupa, Varasarin and Brudhiprabha	2017	2nd yr u/g students	Morph E (82%) NPI (34%), Art (22%), Prep (5%), Syn (18%)	Inter E, Intra E & and Dev E
Suvarnamani	2017	1st yr u/g students	VT (Past/present simple) (72%), elip V be, Frag, Prep	Wrong pron, inconsistency, lack of care, Inter E, Trans
Sernsook, Liamninitir and Pochakorn	2017	2nd yr Eng major students	punc, Art, Sub/Vagr, V & VT, frag, capitalization, Prep, Sp	inter, lack of grammar knowledge, carelessness

**Key.**

h/s – high school; u/g - undergraduate; yr - year.

Problem types: Adj - adjective(s); Adv - adverb(s); Art - article(s); be/have - confusion of be and have; coh - cohesion; conc - conclusion; Conj - conjunction(s) Coll - collocation; Det - determiner(s); elip - elipsis; Frag - fragments; Ger - gerund; Inf - infinitive; G - grammar/grammatical; GS - grammatical structure; have/there conf - using have instead of there (is/are); intro - introduction; ME - mechanical error(s); morph - morpheme(s)/morphological; mods & auxs - modal and auxiliary verbs; morpheme(s)/ morphological; n/c - non-count; N - noun; NPI - noun plural(ization); Obj - object(s); org - organization; para - paragraph(s); Poss - possessive (s/case); Prep - preposition(s); Pro - pronoun(s); Punc - punctuation; rep – repetition; ROS - run-on sentences; S - sentence(s); SS - sentence structure; sing - singular; Sp- spelling; Sub - subject(s); Sub/Vagr - subject verb agreement; Syn - syntax/syntactical; V - verb(s); VF- verb form; VT - verb tense; Voc - vocabulary; WC - word choice; WF - word form; WO - word order.

Causes etc.: Dev - developmental; E - error(s) FCH - false concepts hypothesized: IRR - ignorance of rule restrictions; Inter - interlingual (including L1 transfer & mother tongue influence); Intra - intralingual; IRA - incomplete rule application; overgen - overgeneralization; pron - pronunciation: Trans -(direct) translation from Thai.

### *Mother tongue influence*

The preponderance of attributions to mother tongue influence in the EA studies of Thai students at various stages of their education noted above suggests that this is an aspect which would be of interest in the current study of authors, all of whom were at a more advanced stage in their learning and use of English. The effect of the L1 on the learning of an L2 has been a focus of research for many years dating back to the CA era referred to above, where areas likely to give rise to errors were predicted based on differences detected between languages. However as James (1998) and many other writers have noted, CA was to a large extent unsuccessful as a predictive methodology, although a form of post-dictive or diagnostic CA involving the comparison of languages to help explain errors in learners' ILs through cross-linguistic influence became an accepted part of EA.

While Odlin (2012) suggests that "... the problems related to cross-linguistic influence are so varied and so complex that there does not exist any really detailed theory of language transfer" he notes that there are two, widely recognized methods by which mother tongue influence is diagnosed. The first is the post-dictive comparison of languages mentioned above, which was first used by Selinker (1969) in his ground-breaking study of English word-order errors by Hebrew-speaking English learners, and both James (1998) and Odlin note that the most often used formula on which attributions of mother tongue interference are based are those involving a three-way comparison between the L2, the learners' L1 and their IL. The other method identified by Odlin, is the three-part framework suggested by Jarvis (2000) encompassing intra-L1 group and inter-L1 group comparisons and comparisons between the target groups' ILs and the L2. Jarvis's study used sample groups in Finland with Swedish and Finnish as their L1s and assessed the effect of those L1s on their learning of

English. However in the present study, no obvious comparable group of academics in Thailand existed who shared a similar socio-cultural context but had different L1s to enable inter-L1 group comparisons. Thus Jarvis's method was not practicable and that originally used by Selinker was therefore preferred.

Meanwhile, James (1998) suggested that the best evidence of language transfer is where non-standard usages from the learner's L1 are transferred to the L2. However such occurrences are unlikely to account for the majority of instances of interference, which are far more likely to be associated with the standard use of the L1, particularly in instances of low salience grammatical features of the L2 which are not shared by the L1 (Ellis, 2006). This is likely to lead to phenomena described as *perceptual blocking* or *overshadowing* which may make the perception and mastery of low-salience and redundant grammatical features, such as tense-related morphemes, articles and prepositions, more difficult, particularly for adult learners, if the equivalent categories do not exist in their L1.

Ellis (2006), drawing on associative learning theory, noted that L2 learners frequently fail to master aspects of language to which they are frequently and repetitively exposed because the cues which ought to promote learning are overshadowed by the greater salience of the existing L1. In divining meaning and regularities from L2 input, a learner thus begins by using the *cue weights* associated with their L1, but over time should tend towards the norms of the L2. However, the existing predominance of the L1 makes it highly unlikely that the learner will ever develop a native-like mastery and, as recently observed by Hartshorne, Tenenbaum and Pinker (2018), in a study of grammatical awareness involving almost 700,000 participants, due to *cue competition*, even bilinguals from birth rarely attain mono-lingual native-like performance in either of their languages. Ellis went on to suggest that the L2

learner may not only initially fail to recognize cues such as grammatical morphemes with low salience and high redundancy, but that exposure over time may actually make this tendency more marked as the L2 learner simply becomes habituated to the cues without forming regularized form-meaning mappings.

Thus, this study considered whether the NCU data suggested that the distribution of NCUs in the authors' written work showed an influence from their Thai L1 by comparing the main structural areas in which NCU occurred in their English with the equivalent areas in Thai, based on a summarized description appearing in Boutchuy (2008), and whether the types of NCU which gave rise to the need for changes in their manuscripts suggest the perceptual blocking and overshadowing noted by Ellis (2006) to be a feature of long-term L2 learners ILs.

Research questions dealt with in this paper

Therefore in this paper, the following questions are addressed:

1. What are the main areas of English language structure in which the participants' academic writing differs from conventional native speaker usage?
2. Does the distribution of non-conventional usage:
  - a) reflect the distribution of different word types in the participants' work?
  - b) suggest an effect from their Thai language backgrounds?

## Methodology

### *Source of data and selection of participants*

All the language data presented in this paper were derived from a corpus accumulated by the first author consisting of manuscripts edited by him while acting as a consultant for a publications clinic operated by the Graduate School's Research and Development Office (RDO) at a university in southern

Thailand. The editing of the papers was therefore not undertaken specifically for the study, but was conducted, in most cases, before the study commenced, and its primary purpose (which was to render the English in the paper acceptable for publication) was not directly related to the study. When data collection commenced in early 2016, the corpus consisted of 126 manuscripts but that figure had grown to over 200 by the time data collection was completed at the end of 2017. The sample of 26 manuscripts analyzed in the study was drawn from papers prepared both prior to and during the study period and altogether spanned the 8-year period 2010 – 2017.

The papers which formed the corpus related to a variety of academic disciplines ranging across the university's faculties which comprised four health science related faculties (nursing, dentistry, medical technology and Thai traditional medicine), seven science and technology faculties (science, engineering, computer science, agro industry, natural resources, technology and environment, environmental management), and five humanities and social science faculties, (H and SS, liberal arts, management sciences, economics and education.) They were generally edited without direct contact with the authors. Thus, when selecting papers for inclusion in the sample, the researchers normally had no information relating to the authors beyond their names and faculty affiliations and it was not possible to control for demographic factors.

Papers were selected purposively with the aim of including papers from all or most of the faculties represented in the corpus at the beginning of the study. In this, the sampling procedure was successful with papers being analyzed from 15 of the 16 different faculties indicated above, including at least one paper from each of the university's five campuses in Southern Thailand, and work also being included from writers from three other universities, who



had formerly worked or studied at that university. Nevertheless, the composition of the final sample was heavily influenced by the willingness of authors to participate in the study, since, as noted below, participation was voluntary and no papers were included in the sample without the specific agreement of the author. Further, the corpus included multiple works from some authors and the sample of papers was selected with no author being represented more than once, either as a principal or co-author.

Potential participating authors were approached individually by an e mail in which the study's purposes were detailed. The authors were assured of the anonymous use of their work and all the authors who agreed to participate later signed an informed consent form. On receipt of the authors' agreement, their work was analyzed as detailed below, following which they were interviewed although the content of those interviews is not dealt with in this paper. Of the 26 authors of the manuscripts in the sample, 17 were female and nine were male. All were Thai nationals with Thai as their L1 and all had received formal education up to bachelor's degree level in Thailand, with some completing higher degrees in Thailand, others studying abroad All held at least a master's degree with 21 also holding PhDs. At the time of editing their articles, all but one were employed as lecturers with one being a post-graduate student.

#### *Analysis of the manuscripts*

The manuscripts were analyzed in order to determine the reason that the editor (i.e. the first author) had recommended each change suggested at the time of editing the manuscript. As indicated above, the purpose of editing the manuscripts was not related to the purposes of this study, but was to render the manuscript in English expressing the author's meanings in a manner acceptable to the journal to which the paper was submitted for publication. The primary purposes therefore were to: 1) improve and correct structural and

grammatical usage and mechanical aspects of the writing; 2) improve the rhetorical style, cohesion and lexical use consistent with the usages of academic writing, and 3) suggest changes which would improve the reader's ability to understand the content including where necessary the information content.

In analyzing the changes suggested during the editing of the manuscripts, the first author developed his own coding nomenclature during a pilot study conducted in early 2016, the outcome of which formed the basis of [Author 1, Author 2 and Author 3 (pub. year) redacted] and covered four participants, all of whom are part of the sample included in the present study. The nomenclature was further extended during the remainder of the manuscript analyses but was rationalized at its conclusion to consist of 234 codes divided into six categories, *structure* (including grammatical usage, 192 codes), *lexical* issues (5 codes), *cohesion* (9 codes), *rhetorical style* (18 codes), *information* content (2 codes) and *miscellaneous* covering mechanical and non-language issues (8 codes). Further, the *structure* category was divided into 14 sub-categories: *word order*, *prepositions*, *verbs*, *articles and determiners*, *adverbs*, *nouns and compound nouns*, *word form*, *conjunctions*, *adjectives and modifiers*, *possessives*, *agreement*, *relative pronouns*, *pronouns*, and *miscellaneous structure*.

The coding was conducted manually with each change suggested being isolated and coded on a copy of the manuscript based on the researcher's own reason for suggesting the change. Codes for the *structure* sub-categories identified precisely what change was felt to be necessary; e.g. where a change of tense was necessary, the original and suggested tense were both identified in the code. For the non-structure categories the codes were less detailed since these were not the focus of the study. At the end of the coding, the individual

codes were recorded page-wise in the manuscript on an Excel spreadsheet so that they could be located to be checked or amended if necessary. The full list of tokens allotted to each code was then recorded on a cumulative spreadsheet.

The results obtained therefore represent an account of the editing of the manuscripts based on the editors' own, albeit, subjective reasons for recommending the changes he did. They are therefore an indication of the distance between the editor's native-speaker version of English and the writer's IL in the terms defined by Selinker (1972). Nevertheless in an effort to validate the approach adopted in analyzing the manuscripts, one of the aspects discussed with the participants during the interview was the extent to which the suggestions made had been adopted before the manuscript was submitted to the journal and whether, during the editorial and peer review process, it was necessary to further improve the English language (as opposed to the academic) content. In almost all cases the writers indicated having accepted all or most of the suggestions and only rarely was the writer asked to improve the English language content, although in a number of instances the papers were further reviewed by the editor in order to check changes to the academic content subsequent to peer review.

In order to check the consistency of the coding procedure, an additional paper for each of two authors for whom more than one paper was included in the corpus was also analyzed, although the results of those analyses were not included in the main results. However the two papers analyzed for each author were then compared to establish if the allocation of codes was consistent, indicating both that the coding procedure was reliable and also that it was detecting consistent patterns of NCU and was thus a valid measure of the author's performance. The results of this comparison are shown in Table 2 and

as can be seen, at all levels (overall codes, structure codes only, structure sub-categories and overall categories) there were high and significant correlations detected confirming both the consistency of the coding procedure and also that its outcome was a valid reflection the authors' performance.

**Table 2.** Correlations between coding of two papers from two authors

		Author 1	Author 2
Paper included	NCU %w overall	9.37 %w	6.60 %w
in analysis	NCU %w structure only	5.56 %w	4.31 %w
Additional paper	NCU %w overall	3.19 %w	7.34 %w
analyzed	NCU %w structure only	1.76 %w	4.53 %w
Correlations:			
Overall codes (df, 232)		0.826***	0.960***
Structure codes only (df, 190)		0.856***	0.986***
Structure sub-categories (df, 12)		0.945***	0.993***
Main categories (df, 4)		0.962***	0.998***

Notes. NCU%w: non-conventional usages per 100 words in text

\*\*\* significant at  $p < .001$  (df as shown in parentheses in table)

### *Data analysis*

In order to compare the distribution of NCUs identified in the manuscripts with the distribution of word types, the original versions of the manuscripts prior to being edited were analyzed using the Wmatrix word-tagging application (Rayson, 2009) which produces a CLAWS (constituent likelihood automatic word-tagging system) v.7 tag-set. The word-type analysis derived from this tag-set employs 137 codes, the tokens from which were each allotted to one of the 11 sub-categories relating to particular word types out of the 14 *structure* sub-categories. Later, as described below, further adjustments were made to the coding alignment to render the NCU data and the CLAWS-derived data as closely comparable as possible and the two data sets were then compared using percentages to illustrate the distribution of NCUs to categories, sub-categories and individual codes, and Pearson product moment correlations to compare the distribution of word types and NCUs in the manuscripts. The assessment of the effect of the authors' L1 on their English writing was based on a comparison of the treatment of the main structural areas in which NCU was detected with the equivalent areas in Thai based on the description in Bootchuy (2008).

### **Results**

In this section, the study's results will be outlined based on the research questions identified in Section 2, the first of which asked:

What are the main areas of English language structure in which the participants' academic writing differs from conventional native speaker usage?

The 26 texts analyzed contained a total of 112923 words and 15993 changes were suggested by the editor of which approximately 56% (8931) were classified as relating to language structure, at an average rate of 7.91 NCUs

per 100 words. The NCUs categorized into the 14 *structure* sub-categories in the primary analysis which are the focus of this paper are illustrated in Figure 1 and detailed in Table 3 below.

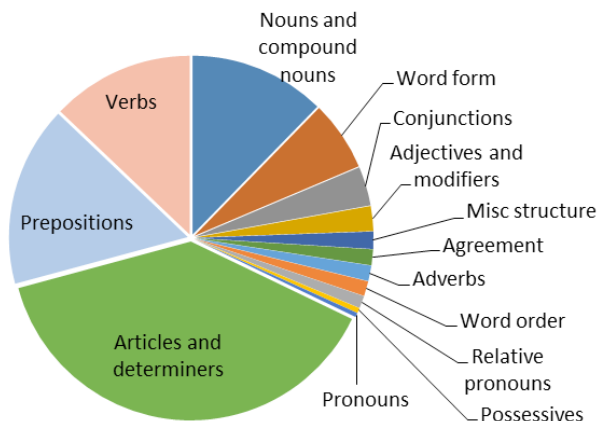


Figure 1. Breakdown of structure NCUs between 14 sub-categories

Table 3. Number of structure NCUs by sub-category

Structure sub category	Total NCUs
Word order	120
Prepositions	1456
Verbs	1152
Articles and determiners	3453
Adverbs	125
Nouns & compound nouns.	1099
Word form	565
Conjunctions	316
Adjectives and modifiers	201
Possessives	42
Agreement	128
Relative pronouns	100
Pronouns	34
Misc. structure	140
Structure NCU total (26 texts)	8931
Total words reviewed (26 Texts)	112923

Within the structure sub-categories, it can be seen that *articles and determiners* were overwhelmingly the largest source of NCUs accounting for almost 40% (3453 out of 8931). Of the remaining sub-categories, *prepositions*, *verbs* and *nouns and compound nouns* also produced substantial totals (1456, 1152 and 1099, respectively) and the four largest sub-categories together represented over 80% of the structure NCUs.

To complete consideration of research question 1, the NCU data were considered at the individual code level to try to identify the main types of NCUs within the individual sub-categories. As stated above, the largest single NCU type was the omission of the definite article *the*, and overall the absence of articles etc. in obligatory situations accounted for 2425 or 28% of the NCUs. Of the remaining tokens from the *articles and determiners* sub category, 763 involved the inappropriate use of *the*, with only 126 relating to the misuse of an indefinite article. Clearly there was considerable confusion among the authors about the use of articles in general, and in respect of the use of *the* in particular.

Among the other sub-categories, *prepositions* NCUs (1456) mostly concerned incorrect preposition choice (863) with the omission of a preposition or the use of a preposition where none was needed accounting for, respectively, 231 and 208 NCUs. From the *verbs* sub-category, 659 tokens related to inappropriate tense choice, of which 432 related to the over-use of the present simple tense, and of those, 367 concerned the use of the present simple tense instead of the past simple tense. Incorrect choice of the past simple tense itself accounted for 152 tokens, 93 in situations where the present simple was indicated and 50 where the present perfect was the appropriate choice. Of the remaining verb NCUs, 297 concerned the use of an incorrect verb form, while



87 related to incorrectly formulated or used passive voice constructions. Within the *nouns* sub-category, problems were overwhelmingly related to the incorrect use of plural and singular count nouns with 804 singular forms being incorrectly used instead of a plural, and inappropriately used plural nouns numbering 220.

Research question 2a asked:

Does the distribution of non-conventional usage reflect the distribution of different word types in the participants' work?

Therefore, in the next stage of the data analysis the CLAWS tag-sets were used to consider how closely aligned were the distributions of the NCUs and the tokens of the word types in their work.

**Table 4.** Word-type tokens per Wmatrix

Word class	Words
Prepositions	14547
Verbs	16217
Articles and determiners	17243
Adverbs	3301
Nouns and compound nouns	37876
Conjunctions	6649
Adjectives and modifiers	10460
Possessives	100
Relative pronouns	616
Pronouns	1333
Misc. structure	876
Not classified	3705
<b>Total</b>	112923
<b>Total classified</b>	109218

Table 4 presents the overall Wmatrix classification which was able to categorize around 97% of the words (109218 tokens) in the manuscripts into the 11 sub-categories used in the initial NCU analysis for which there were direct analogs, with the items not categorized being formulae, figures and letters in the text, foreign words and the infinitive marker *to* which was counted as part of a verb. However the 14 *structure* sub-categories on which the NCU analysis

had been conducted included *word order*, *word form* and *agreement* which had no analogs in Wmatrix, as well as some individual codes within the *adjectives and modifiers* and *miscellaneous* categories which related to clause or sentence level errors which could not be aligned with the Wmatrix distribution. Moreover from the individual codings within the *structure* sub-categories, it was found that the largest number of NCUs related to the omission of the definite article, *the* (1921 tokens) and that a further 504 tokens related to the omission of an article or determiner in an obligatory situation. Therefore, 70% of the overall tokens attributed to *articles and determiners* in the structure NCU analysis had no analog in the CLAWS tag-sets which do not record the use of the *no article* category within the English article system. As Swan (1996) notes, the use of no article before a plural or non-count noun signifies a general reference to the thing etc. denoted by that noun and the *no article* category signifies a distinct type of noun usage and represents an important part of the determiner system.

Therefore, in order to align the NCU analysis with the CLAWS tag-sets, a number of adjustments were made. Firstly, the *word form* and *agreement* tokens were allocated to their respective parts-of-speech sub-categories based on the part of speech which was incorrectly used in the manuscript, Secondly, for the *word order* sub-category, where the NCU related to the misplacement of a single word or to the ordering of adjectives before a noun, the tokens were moved into the appropriate word-type sub-category, with the remaining 88 more egregious word order NCUs being removed from the analysis. Also removed were 83 clause and sentence level NCUs from *adjectives and modifiers* and 110 NCUs from the *miscellaneous* category, with the excluded structure NCUs amounting to about 3% of the total, leaving 8650 NCUs to be compared with the CLAWS word-type data from the manuscripts.

For the *no article* NCUs it was decided to combine the *articles* and *determiners* and *possessives* sub-categories (since, as Swan (1996) notes, “ ‘s genitives” (p. 64) are used in place of articles, so effectively complement the determination system of English nouns and to compare the number of *articles* and *determiners* NCUs including the *no article* tokens, with the number of nouns recorded by the CLAWS tag-set rather than the numbers of articles, determiners and possessives, thus effectively counting the number of *no article* usages as well as those where articles were used correctly or incorrectly in the manuscripts. Adopting this measure increased the number of tokens classified in the texts from 109218 to 129751 by including the 20533 instances where nouns were used which were not preceded by an article, determiner or a possessive.

Table 5 below shows the final adjusted tokens of words and NCUs compared based on percentages, and Figure 2 a and b below show the distributions in chart form. A correlation coefficient was derived comparing the NCU's per word type with the number of tokens of that type (columns **B** and **A(i)** in Table 4) which produced a high correlation of  $r = 0.86$ , which was significant at  $p < 0.01$  (df, 12) Therefore the answer to research question 2a, was clearly affirmative, since based on the  $r^2$  of 0.745, almost three quarters of the variance in the distribution of the NCUs between the word types was related to the frequency of the use of those word types. This is not something which has been considered in studies in Thailand previously, nor does it seem to have been considered elsewhere. In fact such a high correlation would suggest that word frequency may be the main factor in the distribution of NCUs across word types. Therefore in considering whether there is evidence of an effect on NCU from the writers' Thai L1 it must be borne in mind that L1 effects cannot be the main factor in the way in which the NCUs are distributed between the different word types since this was largely related to the frequency of word types in the manuscripts.

Table 5. Word class tokens after adjustment, and corresponding NCUs

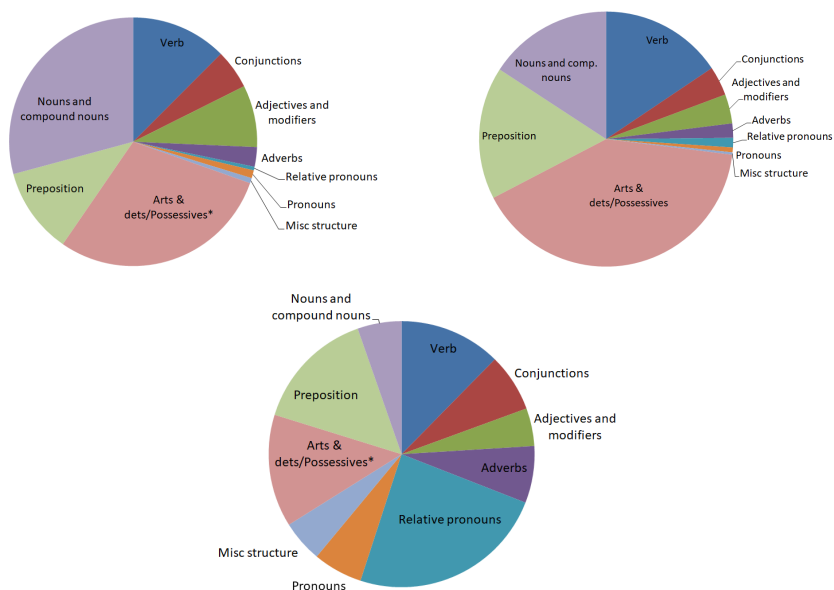
Word class	(i)	A	(ii)	B	C	D
	Words		% of total words	Structure NCUs	% of total NCUs	% of words in word class
<i>Prepositions</i>	14547		11.2%	1456	16.8%	10.0%
<i>Verbs</i>	16217		12.5%	1352	15.6%	8.3%
<i>Arts &amp; dets/Possessives</i>	37876*		29.2%	3495	40.4%	9.2%
<i>Adverbs</i>	3301		2.5%	154	1.8%	4.7%
<i>Nouns and comp. n's</i>	37876		29.2%	1369	15.8%	3.6%
<i>Conjunctions</i>	6649		5.1%	316	3.7%	4.8%
<i>Adjectives and modifiers</i>	10460		8.1%	324	3.7%	3.1%
<i>Relative pronouns</i>	616		0.5%	100	1.2%	16.2%
<i>Pronouns</i>	1333		1.0%	54	0.6%	4.1%
<i>Misc. structure</i>	876		0.7%	30	0.3%	3.4%
Total	129751			8650		
Excluded from analysis	3705			281		
Overall Total	133456			8931		

Notes:

\*Based on the number of nouns in the manuscripts (actual *Article and determiners* used in the manuscripts was 17243)

In column B, *Verbs* includes 200 tokens reclassified from *Word form* (99) and *Agreement* (101); *Articles & dets/Possessives* consists of 3453 tokens from *Articles and determiners* and 42 tokens from *Possessives*; *Nouns and comp nouns* includes 270 tokens reclassified from *Word form* (263) and *Agreement*(7); *Adverbs* includes 29 tokens reclassified from *Word form*; *Adjectives and modifiers* includes 206 tokens reclassified from *Word form* (174) and *Word*

*order* (32) and excludes 83 tokens at clause or sentence level; *Pronouns* includes 20 tokens reclassified from *Agreement*; *Misc. structure* excludes 110 tokens at clause or sentence level.



**Figure 2.** Adjusted distributions of:

- Word types in texts per Wmatrix (\*articles & determiners/possessives based on number of nouns) – column **A** in Table 5 above
- Structure NCUs (articles & determiners/possessives combined) (columns **B & C** in Table 5 above)
- Number of NCU's as a percentage of tokens of word type in the manuscripts (column **D** in Table 5 above, adjusted to 100%)

Research question 2b, asked:

Does the distribution of non-conventional usage suggest an effect from their Thai language backgrounds?

The main areas identified as giving rise to NCUs were: articles, verb tenses, noun pluralization and preposition use. All of these are areas which where English differs substantively from the Thai language, which, has no direct equivalent of the English article system, beyond the use of demonstrative adjectives crudely equivalent to the English *this*, *that* etc., nor does it indicate tense by inflecting verbs, favoring discrete time markers to indicate the time of an action where it is required, although these are often omitted where the time can be inferred from the context in which the verb occurs. Moreover, nouns are not marked for number which is added by way of a post-positioned classifier and although there are prepositions in Thai there is no one-to-one correspondence with English propositions or the situations in which they are used, nor are they used consistently to the same semantic effect, and are often omitted. (Bootchuy, 2008).

Ellis (2006, p. 175) notes that theoretical approaches to L1 interference have tended to concentrate on the transfer of features from the learner's L1 to the L2, whereas perceptual blocking and overshadowing from areas of the L1 in respect of structures that are of low salience and high redundancy are a more potent source of mother tongue interference. The four areas giving rise to over 80% of the NCUs detected in the writers' manuscripts are all such structural features of English of low salience and high redundancy. On that basis it would be difficult to conclude otherwise than that cross-linguistic influences and perceptual blocking caused by differences between the L2 and L1 are a major causative factor of the NCUs recorded in those four areas, and that these effects resulted in the L1 shaping the writers' performance in the L2.

## Discussion

The results presented above clearly show that the occurrence of NCUs in the writings analyzed was heavily weighted towards four areas, *articles and determiners*, *verbs*, *prepositions* and *nouns*, accounting together for over 80% of the overall total of 8931 *structure* NCUs, and for almost 89% of the word-level errors.

Further, within those four areas, the distribution was also concentrated among a small number of NCU types. For *articles and determiners*, the omission of the definite article, *the*, was overwhelmingly the largest issue and combined with the use of *the* in situations where no article was required accounted for more than 80% of the *articles and determiners* NCUs (2875 out of 3495), and overall for 32% of all the *structure* NCUs recorded. This agrees with the findings of Nopjirapong (2011), in which, respectively, 42% and 20% of the article errors recorded in her study of 2<sup>nd</sup> year English major students at Srinakarinwirot University were due to the omission or unnecessary use, respectively, of *the*. In the present study it was apparent that with no direct analog in their L1 these Thai writers clearly found great difficulty in correctly using articles or indeed, in the case of the definite article, in using it at all, and as Ellis (2006) notes, this feature of English is of both low salience and high complexity (p.167). Moreover, the results are in accord with those of Pongpairroj (2007), who found that Thai L1 learners of English made significantly more article errors in her study than did French L1 learners concluding that this was because French L1 learners "...have this functional category in their grammars..." (p. 116) whereas Thai L1 learners do not.

For the *verbs* sub-category, the largest area of NCU was in confusion between the past and present simple tenses which together accounted for 460 out of 1152 tokens of *verbs* NCU, or 34%. This finding is broadly in agreement



with that of Suvarnamani (2017), who also found that transposing these two tenses was the major source of verb-tense error, in her study of 1<sup>st</sup> year undergraduates, accounting for over 73% such errors. For Thai learners of English, the lack of an equivalent auxiliary and verb-inflectional system in Thai and the fact that Thai speakers generally divine the time of an action from time markers or context, would render verb form of low salience in understanding the time of an action and as Ellis (2006) notes: "When two cues jointly predict an outcome, the more salient one may be learned and the less salient may not...". (p. 179). Similarly for *nouns*, the main source of errors was in respect of the misuse of singular or plural forms, accounting for over 90% of the 1099 tokens. Ellis (p. 167) highlights the plural *s* ending as being of low salience in view of the number of different uses to which the *s* ending is put in English, and a Thai speaker would be very likely to overlook the noun ending in favor of other more salient indicators of quantity such as numbers or quantifiers. Finally for preposition use, for which almost 90% of the 1456 tokens were due to incorrect choice, omission or unnecessary use, their low salience, the lack of form-meaning similarities between prepositions in Thai and English, and the fact that English prepositions often having multiple and not necessarily contiguous usage (see, for instance, Ellis, p.167, relating to the various uses of *in*) clearly created problems for these authors in their writing of English.

In comparing the findings of the study with those of the 33 EA studies traced in Thailand since 2000, although the participants in the present study are not directly comparable with those in the 33 studies listed in Table 1, it is notable that the three areas most often identified as producing errors in those studies are also the three areas in this study that produced the highest levels of NCUs (*articles and determiners*, *verbs* and *prepositions*) with the other major sub-category in the present study, *nouns* and in particular noun

pluralization also being identified frequently in those studies. In addition, most of the 33 studies cited mother tongue effects as one of the causes of the errors identified. The participants in those studies were in the main composed of undergraduate students with the next largest group being high school students so that the consistency between earlier results and those in the present study suggests that as Thai learners progress from the intermediate phases of learning to more advanced stages at which they have a real need to be able to actively and accurately use the language, the areas of NCU remain largely similar, suggesting an underlying and persistent influence from their shared Thai L1.

Therefore while the present study found that the distribution of NCUs was heavily influenced by the distribution of word types in the manuscripts, the major structural problems which these writers experience in using English appear to have their source in the blocking and overshadowing effect of their Thai L1. This study therefore offers compelling evidence of the likelihood that the areas identified as producing the largest numbers of NCUs in this study are those where learners will experience the greatest problems in the production of conventional English structure, and those areas therefore need to be targeted during all stages of their learning of English.

### **Conclusion: implications, future directions and limitations**

The study reported investigated non-conventional English language structure in manuscripts written by Thai academics with the aim of establishing to what extent the distribution of NCUs was related to the frequency of occurrence of different word types in the manuscripts, and also whether there was evidence of an effect from the writers' Thai L1. The results clearly show that these two factors are heavily implicated in the pattern of NCU with word

frequency accounting for approximately three quarters of the variance in the occurrence of NCUs and mother tongue effects being the most likely cause of NCUs in the four major areas in which the writers experienced difficulty in producing structurally correct English.

#### *Pedagogical implications*

The implication of these results is that those particular areas identified as giving rise to the greatest number of NCUs bear a disproportionate importance for Thai academics who aspire to publish their work in English, firstly because the word types identified as problematic are those which constitute the main types of words employed in writing academic articles and also because, based on the congruence between the findings of previous studies of Thai learners of English at earlier stages of development and the academics in the present study, those are the areas which give rise to the largest numbers of errors in written work by Thai users of English at all levels from high school upwards. Pedagogically, this means that a relatively small number of areas, particularly the use of the definite article *the*, the use of verb tenses and in particular the present and past simple tenses, the use of plural and singular nouns, and preposition use and collocations generally need to be targeted to prevent the possibility of errors in these areas becoming entrenched at later stages of language development.

#### *Future directions*

This paper presented only overall data from the sample, in line with how previous studies analyzing errors in Thai L1 learners of English have presented their findings. A future paper drawn from the findings of the broader study will consider inter-individual and intra-individual variation in the NCU data. In addition, a further paper will attempt to show how the individual experiences of the participants in learning and using English influenced the level of NCU in

their writing and speech as well as their meta-linguistic awareness of English, based on the data derived from interviewing the authors of the 26 papers analyzed.

#### *Limitations and suggestions for future research*

The main limitations of this study are the relatively small size of the sample and the difficulty of ensuring that this was representative of Thai academics who publish articles in English. Moreover, the restriction of the articles analyzed to those edited by only one native English speaking editor might introduce an element of personal language 'prejudices' into the outcome of the analysis. Future studies with English users at this level might usefully be based on work from more than one source university/region of Thailand and include edited work from more than one editor.

#### References

- Amnuai, W., & Wannaruk, A. (2012). Investigating move structure of English applied linguistics research article discussions published in international and Thai journals. *English Language Teaching*, 6(2), 1-13.
- Andania, R. A. (2015). *The grammatical accuracy, cohesion and coherence of Thai students' English Writing in Darawithaya School, Narathiwat*. Unpublished master's thesis, Sunan Ampel State Islamic University, Surabaya, Malaysia.
- Banlomchon, J. (2006). *An error analysis of free compositions in written English by Thai high school students*. Unpublished master's thesis. Christian University, Bangkok, Thailand.
- Bennui, P. (2016). A study of L1 interference in the writing of Thai EFL students. *Malaysian Journal of ELT Research*, 4(1), 72-99.

- Bootchuy, T. (2008). *An analysis of errors in English writing by a group of first year Thai graduates majoring in English*. Unpublished master's thesis, Kasetsart University, Bangkok.
- Chakorn, O. O. (2005). An analysis of lexico-grammatical errors of Thai businesspeople in their English business correspondence. *NIDA Language and Communication Journal*, 10, 70-94.
- Chuenchaichon, Y. (2014). A review of EFL writing research studies in Thailand in the past 10 years. *Journal of Humanities*, 11(1), 13-30.
- Clarke, A., Gatineau, M., Grimaud, O., Royer-Devaux, S., Wyn-Roberts, N., Le Bis, I., & Lewison, G. (2007). Bibliometric overview of public health research in Europe. *European Journal of Public Health*, 17 (Supplement 1): 43-49. [Author 1, Author 2, & Author 3 Reference redacted]
- EF-English Proficiency Index. (2018). Retrieved from <https://www.ef.co.th/epi>.
- Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27/2: 164-194.
- Gillet, A. (1996). What is EAP? *IATEFL ESP SIG Newsletter*, 6, 17-23.
- Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, 177, 263-277.
- Hinnon, A. (2015). Common errors in English writing and suggested solutions of Thai university students. *Humanities and Social Sciences*, 31(2), 165-180. Retrieved from file:///C:/Users/Acer/Downloads/32281-Article%20Text-71843-1-10-20150320%20(2).pdf.
- Hutyamanivudhi, N. (2001). *An error analysis of business correspondences concerning construction project at Italian-Thai Development Public Company Limited*. Unpublished master's Thesis, King Mongkut's Institute of Technology, North Bangkok.

- Iamsiu, C. (2014). *An analysis of grammatical errors in Srinakarinwirot University students' writing*. Unpublished master's thesis, Srinakarinwirot University, Bangkok.
- James, C. (1998). *Errors in language learning and use: Exploring error analysis*. London: Longman.
- Jaroongkhongdach, W., Watson-Todd, R., Keyuravong, S., & Hall, D. (2012). Differences in quality between Thai and international research articles in ELT. *Journal of English for Academic Purposes*, 11, 194-209.
- Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in them interlanguage lexicon. *Language Learning* 50(2): 245–309.
- Jenwitheesuk, T. (2009). *A study of persisted syntactic errors in writing of the 3rd year students of English for international communication program*. Proceedings of the International Conference on the Role of University in Hands-On Education, Rajamangla University of Technology Lanna. Chiang-Mai, Thailand (pp. 982-986).
- Khamput, P. (2004). *First language interference on second language writing of mattayomsuksa 5 students*. Unpublished master's thesis, Khon Kaen University. Thailand.
- Khaoural, P. (2002). *A study of error analysis in English composition: case study of English major students of Rajabhat Institute Nakhon Pathom*. Unpublished master's thesis, Mahidol University, Bangkok. Abstract retrieved from: <http://mulinet11.li.mahidol.ac.th/thesis/scan/4337020.pdf>.
- Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.

- Larsen, P. O., & von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Sci-entometrics*, 84(3), 575–603.
- Lush, B. (2002). Writing errors: A case study of Thai students' writing errors. *Thai TESOL Bulletin*, 15(1). 75-82.
- Na-Ngam, S. (2004). *Common grammatical errors in foundation English 1 written assignments of Prince of Songkla University student with high and low English entrance examination scores*. Unpublished master's thesis, Prince of Songkla University Hatyai, Thailand.
- Nonkukhetkhong, K. (2013). *Grammatical error analysis of the first year English major students*, Udon Thani Rajabhat University. Proceedings of the Asian Conference on Language Learning, 2013 (pp. 117-126). Retrieved from [http://iafor.info/archives/offprints/acll2013-offprints/ACLL2013\\_0068.pdf](http://iafor.info/archives/offprints/acll2013-offprints/ACLL2013_0068.pdf).
- Nopjirapong, R. (2011). *An analysis of article errors in Thai university students' composition*. Unpublished master's thesis, Srinakarinwirot University, Bangkok.
- Odlin, T. (2012). Cross-linguistic influence in second language acquisition. In C.A. Chappelle (ed.) *The encyclopedia of applied linguistics*. London, Wiley-Blackmore (pp. 436 – 486).
- O'Donnell, L. (2015). An error analysis of Thai EFL learner's controlled dialogue writing. *International Journal of Social Science and Economic Research*, 1(8), 1169-1171.
- Phetdannuea, F. & Ngonkum, S. (2016). An analysis of interlingual errors and intralingual errors in Thai EFL students' writing at Khon Kaen University. *Khon Kaen University Research Journal HS (GS)* 4(2), 35-51.

- Pongpairoj, N. (2007). Asymmetric patterns of English article omissions in L2A. In C. Gabrielatos, R. Slessor, & J. W. Unger, (eds.), *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching*, Vol. 1. (pp. 103-119).
- Promsupa, P., Varasarin, P., & Brudhiprabha, P. (2017). An analysis of grammatical errors in English writing of Thai university students. *HRD Journal*, 8(1), 93-104.
- Rayson, P. (2009). *Wmatrix: a web-based corpus processing environment*. Computing Department, Lancaster University, UK. Retrieved from <http://ucrel.lancs.ac.uk/wmatrix/>.
- Rattanadilok-Na-Phuket, P., & Othman, N. B. (2015). Understanding EFL students' errors in writing. *Journal of Education and Practice*, 6(32), 99-106.
- Richards, J. C. (1970). *A non-contrastive approach to error analysis*. Paper presented at the TESOL convention, San Francisco. Retrieved from <https://files.eric.ed.gov/fulltext/ED037721.pdf>.
- Rurakwit, M. (2004). *Negative transfer, communication strategies, and learning strategies in the English writing of Thammasart University*. Unpublished master's thesis. Thammasart University, Bangkok.
- Sattayatham, A., & Honsa, S. (2007). Medical students' most frequent errors at Mahidol University, Thailand. *Asian EFL Journal*, 9(2), 170-194.
- Sattayatham, A., & Ratanapinyowong, P. (2008). Analysis of errors in paragraph writing in English by first year medical students from the four medical schools at Mahidol University. *Silpakorn University International Journal*, 8, 17-38.
- Selinker, L. (1969). Language transfer. *General linguistics*, 9(2), 67.



- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10( 3).
- Sereebenjapol, P. (2003). *An analysis of the errors in English which graduate science students make in the discussion section of their thesis*. Unpublished master's thesis, Mahidol University, Bangkok. Abstract retrieved from <http://mulinet11.li.mahidol.ac.th/thesis/2546/4236311.pdf>.
- Sompong, M. (2014). Error analysis. *Thammasat review*, 16(2), 109-127.
- Spillner, B. (2017). *Error Analysis in the world. A bibliography (Vol. 37)*. Berlin, Frank & Timme GmbH.
- Srichai, C. (2002). *Analysis of errors in written work by first year business administration students at Prince of Songkla University, Hat Yai*. Unpublished master's thesis, Prince of Songkla University, Hatyai.
- Suvarnamani, S. (2017). A study of grammatical and lexical errors in descriptive writing of first year arts students at Silpakorn University. *Silpakorn University Journal of Social Sciences, Humanities, and Arts* 17(2), 239-264.
- Swan, M. (1996). *Practical English usage*. Oxford, Oxford Univ. Press.
- Vasconcelos, M. R., Sorenson, M. M., & Leta, S. (2009). A new input indicator for the assessment of science and technology research. *Scientometrics*, 80, 217-230.
- Watcharapunyawong, S., & Usaha, S. (2012). Thai EFL students' writing errors in different text types: The interference of the first language. *English Language Teaching*, 6(1), 67