Quantitative Reconsideration of an L1 Effect on Asian Learners' L2 English Writing: A Study Based on the ICNALE

Shin'ichiro Ishikawa

iskwshin@gmail.com, Kobe University, Japan

ABSTRACT

A writer's L1 is generally believed to influence their L2 English writing significantly. However, the extent to which L1 backgrounds influence Asian learners' L2 English writing has not been wholly elucidated due to the lack of data covering various learners in Asia. Therefore, this study analysed more than one-million-word essays written by 2,318 Asian students with 18 regional backgrounds and more than 14 L1 backgrounds, which were taken from the International Corpus Network of Asian Learners of English (ICNALE) Written Essays and the ICNALE Written Essays Plus, currently under construction. The analytical focus was whether learners with the same or similar L1 backgrounds were agglomerated in a single subcluster despite the difference in the other parameters, such as essay topics. Hierarchical agglomerative cluster analyses that focused on the surface layer (words), the deeper layer (lexicogrammatical features), and the latent layer (textual factor scores) of student writing revealed that the degree of L1-based output similarities may be much lower than generally believed, which requires us to reconsider the traditional view that a writer's L1 is an absolute factor in determining the aspects of

their L2 use and establish a new analytical framework for discussion of Asian learners' L2 English.

Keywords: Learner Corpus Research, L1 effects, Asian learners of English, ICNALE

Introduction

Learner corpus research (LCR) has traditionally attributed the features observed in the target learner group to their L1s. Thus, L1 effects, L1 transfers, and crosslinguistic influences (CLI) have been widely discussed.

In the discussion of the merits of applying LCR to English language teaching, Gilquin (2023) writes as follows:

> Usually, learner corpora come with rich metadata about learners (e.g. their mother tongue or the number of years they have been learning the target language). This is valuable information for teaching, because it can lead to *customisation*, i.e., the adaptation of teaching or materials to the target learner population (e.g. Chinese-speaking learners, beginners). Gilquin and Granger (2021), thus, show that the phrase *as far as X BE concerned* is particularly common in corpus data produced by French-speaking learners, who often use it as a topic introducer at the beginning of the sentence…. Such a finding could lead to a pedagogical intervention specifically targeted at French-speaking learners…. (p. 282)

As Gilquin suggests, LCR may help textbook writers produce materials adjusted for a particular L1 speaker group. Such tailor-made material development seems to be a promising direction in the application of LCR. However, whether the tendency observed in a learner corpus, such as the overuse of a particular word or phrase, is directly caused by learners' L1 is not necessarily clear.

Gilquin and Granger (2021) used the data from the International Corpus of Learner English (ICLE) to compare the usages of passive verbs of four EFL groups (French-speaking Belgian, German, Korean, and Serbian students) and four ESL groups (Dutch, Hong Kong, Norwegian, and Tswana students). Subsequently, they found out that Belgian students used "concern" in a passive form much more than the other students. This is a noteworthy finding, but whether this tendency can really be linked to L1 French, in other words, whether this finding is applicable to French-speaking learners in general rather than to French-speaking learners in Belgium, remains unclear. Gilquin and Granger also conducted a hierarchical cluster analysis based on

the passive ratios for various verbs, which showed that eight groups were classified into three distinct clusters: Korean and Hong Kong students; Tswana and Serbian students; German, Dutch, Norwegian, and native English-speaking students. Based on this result, they insist that EFL and ESL may not be neatly divided and the fact that students with L1s belonging to the Germanic language family are clustered with native English speakers "could be due to the typological closeness of these languages to English". Although this might be possible, typologically unrelated groups (i.e. L1 Korean and L1 Chinese speakers and L1 Tswana [Setswana] and L1 Servian speakers) are also clustered, which may suggest that L1 is not a staple factor in determining one's choice of passive forms. This possibility is well recognised by the authors, who carefully add that the agglomeration of three Germanic L1 speakers might be due to "the language proficiency of these learner populations, or to other factors which cannot be further investigated here but which could outweigh the distinction between EFL and ESL."

The LCR literature often takes for granted that learners' L1s should have a decisive influence on their L2 English use. However, claiming an L1 effect reliably is extremely challenging for several reasons.

First, an L1 effect, if any, does not occur at all times; therefore, it is highly elusive. As Jarvis (2017) suggests, this happens only when learners' L2 is somewhat related to their L1; they use prototypical or unmarked language features frequently occurring in their L1, and they are already at the developmental stage of using them sufficiently.

Second, it is usually difficult to distinguish between the effects of learners' L1s and their nations or regions. "French learners of English", for instance, can be L1 French speakers worldwide. Simultaneously, they may be learners from various L1 backgrounds who live in France and receive English education there. Friginal (2018) reports that the ICLE (ver. 2) comprises eleven national sub-corpora, and they represent eleven native language backgrounds, but "there is no exact match between the [L1] backgrounds and the [national] sub-corpora". Thus, "learners with a Swedish language background, for instance, are represented in both the Finish and Swedish sub-corpora" and "the French sub-corpus consists of essays written in Belgium (by native speakers of French), and the German sub-corpus [consists] of essays written in Austria, Germany, and Switzerland", which he says may lead to a puzzle for corpus users (p. 86).

Third, other parameters may concern what seems to be the L1 effect. Jarvis (2017) says that the omission of an article by an L1 Swedish learner may be the result not of their L1 background but of their limited L2 proficiency, lack of control (namely, a kind of performance error), linguistic simplification strategy that many learners adopt, or just an individual trait. From a methodological viewpoint, to claim an L1 effect, one must confirm three

things: "intra-group homogeneity" as a within-group tendency, "inter-group heterogeneity" as a between-group tendency, and "cross-language congruity" as a linguistic tendency. Jarvis (2017) concludes as follows:

> …it is important to verify that the area of learners' performance that is suspected of being influenced by CLI is indeed characteristic of (via within-group similarities) and distinctive of (via between-group differences) learners representing a particular language background, and that their performance is furthermore motivated by the corresponding features or conventions of their background language(s). (p.17)

Fourth, discussing the L1 effect entails the risk of sample cherry-picking. L1 Japanese learners, for instance, sometimes omit the plural marker of a noun (e.g. three * apple), which is usually explained by the fact that the Japanese language does not distinguish between singular and plural nouns. Although this seems to be a possible explanation, an intentionally selected single example or episode does not guarantee the overall influence of L1 Japanese on L2 English use by L1 Japanese learners. Discussing the methodological features of corpus linguistics, McEnery and Hardie (2012) note that researchers "*must not* select a favourable subset of the data" to prove a hypothesis they want to support (p. 15). If they wilfully cherry-pick only the samples that suit their hypothesis and ignore all other inconvenient data, falsifying it would become theoretically impossible, which is clearly against falsifiability and total accountability as major principles of science (McEnery & Brezina, 2022, pp. 14-15).

As summarised above, due to many theoretical and methodological difficulties, the L1 effect on learners' L2 use has not been fully explained. Therefore, this study aims to reconsider the effect of L1 on L2 English use by Asian learners. Claiming an L1 effect for Asian learners of English might be more challenging than for their European counterparts, who have long been the main target of L1 effect studies (Tenfjord et al., 2017). This is because none of the Asian languages belong to the Germanic language family, and typological closeness is hardly observed between Asian languages and English.

When examining the L1 effect seen in Asian learners' L2 English use while paying due attention to intra-group homogeneity and inter-group heterogeneity, one would need a learner corpus that (i) collects data from learners with varied L1 and regional backgrounds in Asia, (ii) appropriately controls L2 output conditions, such as topic, length, and time, which are likely to influence L2 outputs, and (iii) includes data about learners' proficiency levels, which also significantly influence L2 outputs.

Thus, this study analyses data from the International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa, 2023). The analytical focus will be on whether various L2 English essay samples written by college students (including some graduate students) from 18 regions in Asia are statistically clustered in terms of their L1 backgrounds rather than essay topics and L2 proficiency levels. Considering the possibility that L1 may influence different aspects of learner writing in different ways, this study conducts hierarchical agglomerative cluster analyses on three levels of learner writing: the surface layer (words), the deeper layer (lexicogrammatical features), and the latent layer (textual factor scores).

Method

Aim and RQs

This study re-examines the extent to which Asian learners' L1s influence the different layers of their L2 English writing. Thus, it compares the strength of the L1 effect, which is defined here as the tendency for all and only learners with the same or typologically related L1 backgrounds to produce lexically, lexicogrammatically, and textually similar L2 outputs. The analysis focuses on the surface (words), deeper (lexicogrammatical features), and latent (textual factor scores) layers of learners' L2 writing.

Unlike many previous studies discussing the localised relationship between particular errors and L1 background, this study adopts a total accountability approach and examines all learners' written outputs. An L1 effect could be reasonably claimed if learners with the same or typologically related L1 backgrounds are clustered despite differences in other parameters, essay topics, in particular. To achieve this goal, this study addresses three research questions.

RQ1: In terms of word frequency, do Asian learners' essays cluster according to their L1 background?

RQ2: In terms of lexicogrammatical feature frequency, do Asian learners' essays cluster according to their L1 background?

RQ3: In terms of textural factor scores, do Asian learners' essays cluster according to their L1 background?

Data

Choosing the appropriate data is of paramount importance when discussing the highly elusive L1 effect. Jarvis (2017) encourages researchers to ensure the following: (a) groups of participants are truly representative of

their identified populations; (b) the different groups are comparable in all possible ways except in relation to the independent variable(s) being investigated; (c) the different groups are given the same tasks in the same conditions (p. 17).

Therefore, this study chose data from the ICNALE Written Essays. The ICNALE, which currently comprises five data modules (essays, edited essays, monologues, dialogues, and assessments), collects data in a highly controlled manner (Ishikawa, 2023, p. 21). Regarding essays, conditions such as writing time (20–40 minutes), essay length (200–300 words), essay topics (a part-time job for college students [P] and nonsmoking at restaurants [S]), and reference use (spell checker use is required, while dictionary use is prohibited) were common to all participants.

The ICNALE Written Essays (WE) was released in 2012. The current version (V2.6) includes 4,800 essays written by 2,400 college students (including graduate students) from ten countries and regions in Asia: China (CHN), Hong Kong (HKG), Indonesia (IDN), Japan (JPN), Korea (KOR), Pakistan (PAK), the Philippines (PHL), Singapore (SIN), Thailand (THA), and Taiwan (TWN).

Since 2023, the ICNALE team has been developing an additional module, ICNALE Written Essays Plus (WEP), to expand regional coverage. Its current version (v 0.2) includes 1,200 essays by 600 college students (including graduate students) from eight Asian countries: Bangladesh (BGD), Brunei (BRU), India (IND), Cambodia (KHM), Laos (LAO), Malaysia (MYS), Myanmar (MMR), and Vietnam (VNM). Notably, unlike in the days when the original essay module was developed, automatic translators (e.g. Google Translator, DeepL) and generative AIs (e.g. Chat GPT, MS Copilot, Google's Gemini) have become considerably common among young people. The project team explicitly told the students not to use these writing assistance tools, required them to declare that they had not utilised any of those tools when writing, conducted a manual check, and excluded essays that seemed identical to other students' essays. Therefore, even if a few students partially utilised these tools, it would not have significantly influenced the analysis results.

When comparing learners, it is important to control for their L2 proficiency levels. Regarding this point, all the ICNALE participants were classified into four proficiency bands linked to the Common European Framework of Reference for Languages (CEFR)—A2, B1_1 (B1 Low), B1_2 (B1 Upper), and B2+—according to their scores in standard English proficiency tests, such as TOEFL, TOEIC, and IELTS, or their scores in the common vocabulary size test, which was administered to all the participants. As expected, proficiency distributions are not even among different countries and regions, but most of the students, especially those in English as a foreign

language (EFL) regions, belong to the B1 intermediate level. Therefore, to maximise the number of participants to be analysed and simultaneously control for possible proficiency effects, which are expected to influence learner outputs significantly, this study analysed the essays of learners only at the B1 Upper (U) and B1 Lower (L) levels, which enabled a closer look at the L1 effect.

Among 72 datasets (18 regions x 2 topics x 2 proficiency levels), the B1 Low data of the four regions (BGD, BRN, IND, and SIN) were excluded as the number of participants was smaller than ten. Thus, this study analysed 68 sets of essays classified by region, proficiency level, and essay topic. Each essay set was merged into a single text file and treated as a single "data item" for this analysis. The total number of participating students was 2,318, and that of words in their essays was 1,101,916.

Table 1 shows learners' regions, L1s, L1s' linguistic typologies (namely, language families), which are based on two language databases, Ethnologue 27 (Eberhard et al., 2024) and Glottolog 5.0 (Hammarström et al., 2024), learners' L2 proficiency levels (U or L), the number of participating students (N), the total number of words in part-time job essays (P), and non-smoking essays (S).

Table 1

Outline of the Corpus Datasets Analysed in the Current Study

Region	L1	Language family	L2 Prof.	N of students	N of words (P essays)	N of words (S essays)
BGD	Bengali	Indo-European	U	72	19,183	18,185
BRN	Malay	Austronesian	U	37	9,464	9,061
CHN	Chinese	Sino-Tibetan	L	232	56,949	53,420
			U	105	27,009	25,061
HKG	Chinese	Sino-Tibetan	L	30	7,481	7,006
			U	52	12,571	12,030
IDN	Indonesian	Austronesian	L	82	19,039	18,457
			U	83	20,069	19,475
IND	Hindi, etc.[1]	Indo-European	U	14	3,076	3,100
JPN	Japanese	Japonic [2]	L	179	40,491	39,176
			U	49	11,320	11,092
KHM	Khmer	Austro-Asiatic	L	24	6,445	6,238
			U	24	6,247	6,183
KOR	Korean	Koreanic [2]	L	61	13,642	13,286
			U	88	20,510	19,605
LAO	Lao	Kra-Dai [3]	L	14	3,231	3,061
			U	16	3,788	3,486

MMR	Burmese	Sino-Tibetan	L	54	13,305	12,611
			U	59	15,383	14,568
MYS	Malay	Austronesian	L	14	3,883	3,642
			U	74	19,640	18,410
PAK	Urdu	Indo-European	L	91	21,787	21,847
			U	88	20,777	20,807
PHL	Filipino	Austronesian	L	11	2,679	2,602
			U	176	44,503	42,748
SIN	Mixed	Mixed	U	134	33,316	32,490
THA	Thai	Kra-Dai [3]	L	179	40,921	39,874
			U	100	23,267	22,589
TWN	Chinese	Sino-Tibetan	L	87	20,663	19,362
			U	61	14,642	13,855
VNM	Vietnamese	Austro-Asiatic	L	11	3,980	3,661
			U	17	2,986	2,681
Total	---	---	---	2,318	562,247	539,669

Note. (1) The L1s of the participants from IND include Hindi, Urdu, and Punjab, all of which are mutually intelligible and belong to the Indo-European family; (2) Japanese and Korean languages have many features in common, and they are sometimes considered part of the Altaic languages, but most modern linguists do not admit it but see them as language isolates (Yurayong and Szeto, 2020); (3) Glottolog regards LAO/THA as part of the Tai-Kadai family.

Analytical Methods

This study investigates whether all and only learners with the same or typologically similar L1s are agglomerated in a single subcluster despite differences in other parameters, such as essay topics. These steps are necessary to confirm within-group similarities and between-group differences simultaneously (Jarvis, 2017).

This observation will focus mainly on whether the seven types of L1 typological groups (Table 2) and an additional pair of language isolates are represented in the dendrogram obtained from the cluster analyses. The first three groups share the same or practically the same L1s. Equating Lao with Thai is based on Enfield (2002), who concludes that "it is not possible to establish distinctions" between Lao and Thai, as well as Isan, a language spoken in northern Thailand, as "separate languages or dialects by appealing to objective criteria". The following four groups belong to any of the world's language families:

Table 2

L1 Typological Groups

#	Typological groups	N of the Regional Datasets Included in Each Cluster
1	L1 Chinese	12: CHN(4) /TWN(4)/HKG(4)
2	L1 Malay	6: BRN(2)/MYS(4)
3	L1 Thai	8: THA(4)/LAO(4)
4	Indo-European	8: BGD(2)/IND(2)/PAK(4)
5	Austronesian	14: BRN(2)/IDN(4)/MYS(4)/PHL(4)
6	Sino-Tibetan	16: CHN(4)/TWN(4)/HKG(4)/MMR(4)
7	Austro-Asiatic	8: KHM(4)/VNM(4)
x	Language isolates	8: JPN (4) [Japonic]/KOR (4) [Koreanic]

If the 68 items are statistically divided neatly into any of these typological groups, one can safely conclude that L1 is indeed a staple factor in determining aspects of Asian learners' L2 English use.

For objective data classification, this study adopts a hierarchical agglomerative cluster analysis, following Gilquin and Granger (2021) and many typology studies. In cluster analysis, individual data points are taken, and the closest ones are joined (agglomerated) in a step-by-step hierarchical procedure until a large cluster containing all data points is created (Brezina, 2018, p.154).

The initial distance is defined as the square root of (2-2r) when classifying variables (RQ1 and RQ2) and the Euclidean distance when classifying cases (RQ3). In both, Ward's agglomerative method was used.

Data classification was based on lexical, lexicogrammatical, and textual indices: the frequencies of 100 high-frequency words (RQ1) and 67 lexicogrammatical features (RQ2), as well as six textual factor scores (RQ3). To reduce the expected topic effect, the 100 words analysed in RQ1 did not include ordinary nouns or verbs (Table 3).

Table 3

A List of the 100 Words Analyses for RQ1

Words Used for the Analysis
the, be, to, a, and, in, of, they, it, for, I, can, not, we, their, to, that, will, should, you, but, with, that, if, also, at, because, or, from, who, many, on, so, our, all, some, good, by, as, this, there, when, other, important, for, t, bad, very, have, do, my, only, more, may, your, which, public, more, as, about, this, would, s, one, how, he, completely, even, what, an, such, so, like, must, well, than, no, just, these, however, all, around, while, that, up, why, his, most, especially, as, every, too, could, then, own, second, still, where, harmful, themselves

The analyses for RQ2 and RQ3 were based on the framework of a multi-dimensional analysis (MDA) proposed by Douglas Biber (Biber, 1988).

To classify various English genres, Biber examined the frequencies of 67 types of lexicogrammatical features (Table 4; Biber, 1988, p. 270) in 481 texts from 23 (17 written and six spoken) genres, most of which were taken from the LOB and London-Lund corpora. Subsequently, he conducted a factor analysis and extracted six factors that determined the textual feature of each genre (Table 5), which he called "dimensions" (Biber, 1988, pp. 127–155).

Table 4

A List of the 67 Kinds of Lexicogrammatical Feature Tags

Lexicogrammatical Features Analysed in Biber's (1988) MDA
1: past tense, 2: perfect aspect verbs, 3: present tense, 4: place adverbials, 5: time adverbials, 6: first person pronouns, 7: second person pronouns, 8: third person pronouns, 9: pronoun IT, 10: demonstrative pronouns, 11: indefinite pronouns, 12: DO as pro-verb, 13: WH questions, 14: nominalisations, 15: gerunds, 16: nouns, 17: agentless passives, 18: BY passives, 19: BE as main verb, 20: existential THERE, 21: THAT verb complements, 22: THAT adj complements, 23: WH clauses, 24: infinitives, 25: present participial clauses, 26: past participial clauses, 27: past participial WHIZ deletion, 28: present participial WHIZ deletion, 29: THAT relatives (subject position), 30: THAT relatives (object position), 31: WH relatives (subject position), 32: WH relatives (object position), 33: WH relatives (pied pipes), 34: sentence relatives, 35: adv. subordinator (cause), 36: adv. subordinator (concession), 37: adv. subordinator (condition), 38: adv. subordinator (others), 39: prepositions, 40: attributive adjectives, 41: predicative adjectives, 42: adverbs, 43: type/token ratio, 44: word length, 45: conjuncts, 46: downtoners, 47: hedges, 48: amplifiers, 49: emphatics, 50: discourse particles, 51: demonstratives, 52: possibility modals, 53: necessity modals, 54: predictive modals, 55: public verbs, 56: private verbs, 57: suasive verbs, 58: SEEM/APPEAR, 59: contractions, 60: THAT deletion, 61: stranded prepositions, 62: split infinitives, 63: split auxiliaries, 64: phrasal coordination, 65: nonphrasal coordinators, 66: synthetic negation, 67: analytic negation

Table 5

A List of the Six Textual Factors Obtained from the MDA

Factor	R^2	Features	Genres showing positive values	Genres showing negative values
Dim 1	84.3	Involved/ informational production	Telephone conversation	Official documents
Dim 2	60.8	Narrative/ non-narrative concerns	Romantic fiction	Hobbies Broadcasts
Dim 3	60.5	Explicit/Situation-dependent reference	Official documents	Broadcasts
Dim 4	16.9	Overt expression of persuasion	Professional letters Editorials	Broadcasts

| Dim 5 | 58.0 | Abstract/Non-abstract information | Academic prose | Telephone conversation |
| Dim 6 | 28.5 | On-line informational elaboration | Prepared speeches | General, science, mystery, and adventure fiction |

Note. R^2 represents to what extent each factor explains the variation in values.

This study used the multi-dimensional analysis tagger (MAT) v1.3.3 (Nini, 2019), a software tool reproduced from the original Biber tagger (Figure 1). It assigns 67 lexicogrammatical feature tags to the input texts and automatically calculates six types of factor scores.

As some of the dimension factors became negative, 10 points were equally added to each value to make all values positive, which was a necessary step for conducting cluster analysis (Tables 6 and 7).

Figure 1

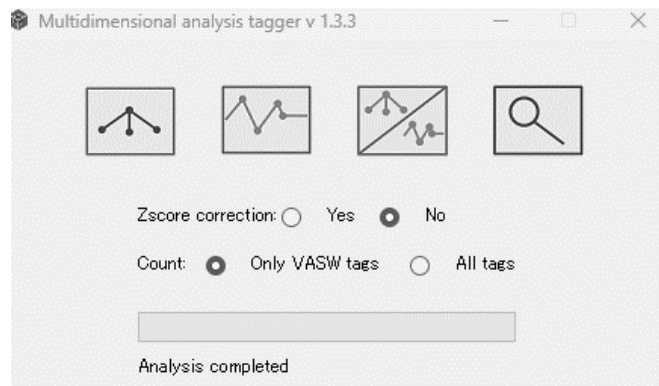*Interface of Nini's (2019) Multidimensional Analysis Tagger V1.3.3*

Table 6

Six Factor Scores of Sample Items (Before Correction)

Filename	D1	D2	D3	D4	D5	D6
BGD_PU	-3.47	-1.02	6.34	4.71	3.39	-0.96
BGD_SU	-3.06	-2.66	5.79	6.45	4.46	0.00
BRN_PU	4.56	0.44	6.29	4.91	4.34	-1.00
BRN_SU	5.65	-2.98	3.98	8.41	3.93	-0.38

Note. P/S in the filenames represents the type of topic (part-time job essays and nonsmoking essays), and U represents learners at the B1 upper level.

Table 7

Six Factor Scores of Sample Items (After Correction)

Filename	D1	D2	D3	D4	D5	D6
BGD_PU	6.53	8.98	16.34	14.71	13.39	9.04
BGD_SU	6.94	7.34	15.79	16.45	14.46	10.00
BRN_PU	14.56	10.44	16.29	14.91	14.34	9.00
BRN_SU	15.65	7.02	13.98	18.41	13.93	9.62

The result of a cluster analysis is usually presented as a dendrogram (also called a tree diagram), which graphically shows how items (i.e. variables or cases) are agglomerated step-by-step. Similar items are agglomerated first, dissimilar items are agglomerated later, and finally, all items are merged into a single large cluster. Interpretating item distributions, one usually chooses a cutting point and divides the items into several subclusters (SC). As each typological group included eight to ten datasets, this study determined the cut-off point such that each SC included approximately eight items.

When discussing the item distribution in the dendrograms obtained from the cluster analyses, this study pays special attention to the related item inclusion ratio (IR), which represents the percentage of typologically related items actually agglomerated in a single subcluster. The Discussion section also examines the region cover ratio (CR) and the unrelated item exclusion ratio (ER).
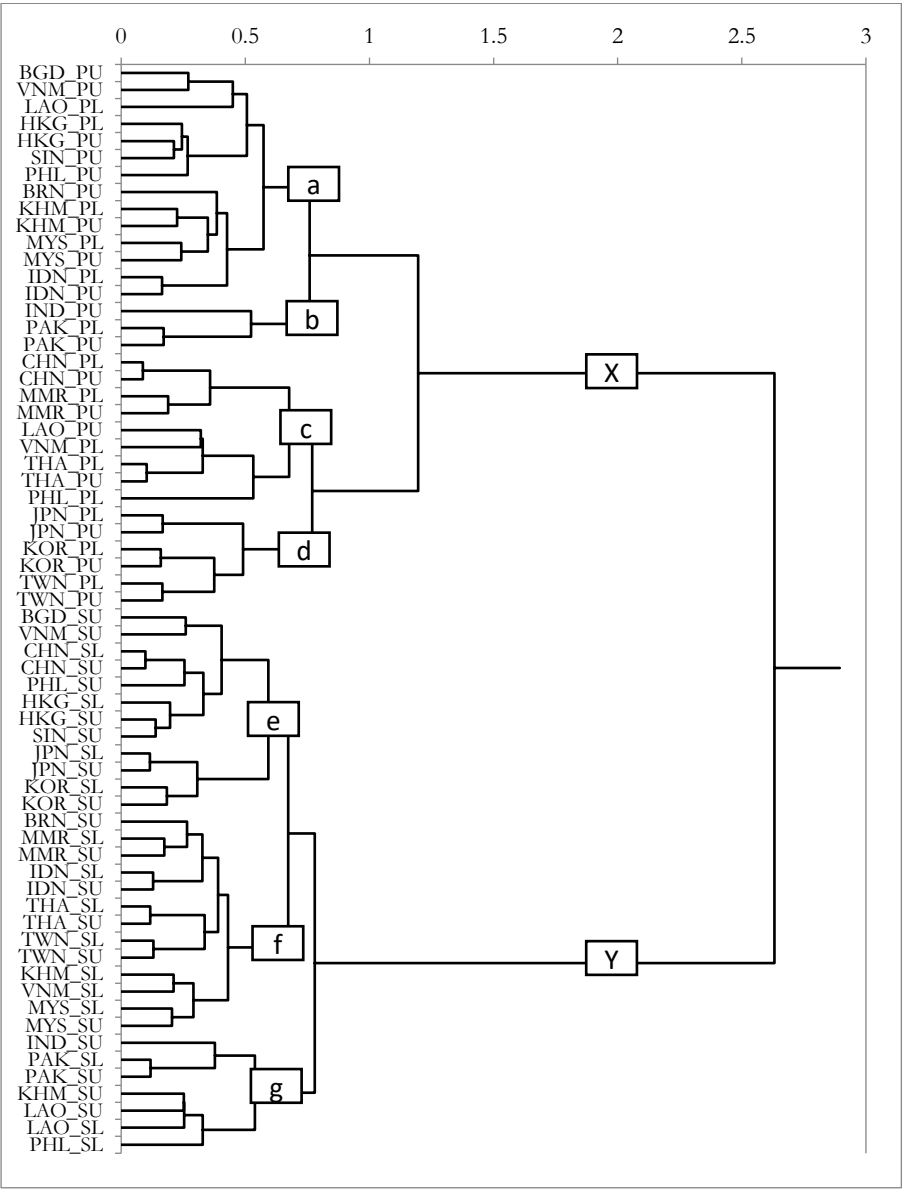
Results

RQ1 Word-based Classification

First, attention was paid to the strength of an L1 effect on the surface layer of L2 English writing by Asian learners. From the analysis of the frequencies of the 100 words, the dendrogram below was obtained, where P &S stand for essay topics, and U&L stand for proficiency levels (e.g. PU stands for part-time job essays written by B1 Upper-level students).

The items were divided into two large clusters that seemed irrelevant to the learners' L1 types. X comprised only part-time job essays, whereas Y comprised only non-smoking essays, suggesting that an essay topic is the most decisive factor in classifying Asian learners' L2 writing. X and Y were divided into four (a-d) and three (e-g) subclusters, respectively, although the latter was less distinct.

Figure 2

A Dendrogram from the Cluster Analysis of the Frequency of the 100 Words

Next, to examine the strength of the L1 effect, attention was paid to each of the seven typological groups (Table 2). If all the typologically related items are clustered and that cluster includes no typologically unrelated items—if the "all and only" criterion is fully met—one confirms an L1 effect

most satisfactorily. Therefore, the analysis focuses on the sub-clusters that contain the largest number of typologically related items and calculates the related item inclusion ratios (IR). For instance, if a typological group included ten data items in total and eight of them appeared in the same sub-cluster, the IR was calculated to be 80%.

First, the three same-L1 groups are discussed. Regarding the L1 Chinese group, four of the 12 items occur in SC_e (CHN:2, HKG: 2; 33.3%). A weak link is suggested between CHN and HKG, but a combination of all three regions (CHN, HKG, and TWN) is not observed in any of the subclusters. In the L1 Malay group, three of the six items occur in SC_a and SC_f (BRN:1, MYS:2; 100% combined). A somewhat stronger link is observed between the BRN and MYS, but an L1 effect comes after a topic effect, and both subclusters include many non-Malay items. In the L1 Thai group, three of the eight items occur in SC_c (LAO:1, THA:2, 37.5%). A weak link is suggested, but the other five items are scattered into three sub-clusters (SC_a, f, g).

Next, four similar-L1 groups are discussed. First, regarding the Indo-European group, three of the eight items occur in SC_b and SC_g (IND:1, PAK:2; 75% combined), but a combination of all three regions (BGD, IND, and PAK) is not seen in any of the subclusters. Regarding the Austronesian group, six of the 14 items occur in SC_a (BRN:1, IDN:2, MYS:2, PHL:1; 42.9%), but a combination of all four regions is seen only here. In the Sino-Tibetan group, four of the 16 items occur in SC_c (CHN:2, MMR:2), SC_e (CHN:2, HKG2), and SC_f (TWN:2, MMR:2) (75% combined); however, a link between all four regions (CHN, HKG, TWN, and MMR) is not observed in any of the subclusters. In the Austro-Asiatic group, three of the eight items occur in SC_a (KHM: 2, VNM:1; 37.5%), while the others are scattered across four subclusters (SC_c, e, f, g). Finally, regarding the two-language isolates, four of the eight items appear in SC_d and SC_e (JPN:2, KOR:2, 100% combined). A considerably strong link is suggested, but both sub-clusters include other items.
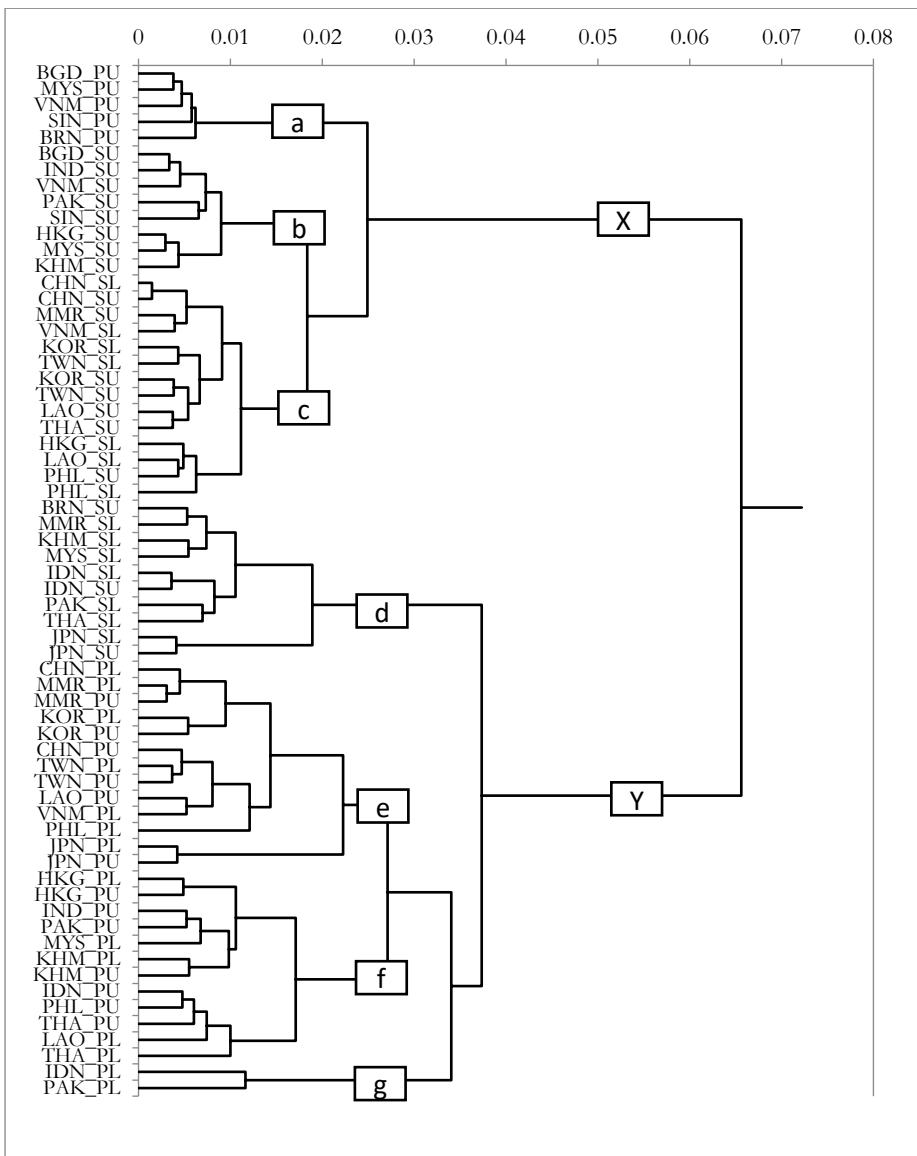
Thus, among the eight groups that include pairs of language isolates, full regional coverage is suggested for six groups, excluding the L1 Chinese and Indo-European groups. This seems to suggest a certain level of L1 effect, but the inclusion ratios are largely low, and the items often agglomerate with typologically unrelated items. For instance, regarding an L1 Malay group, SC_a includes BRN and MYS, but at the same time, various items from the unrelated L1 backgrounds (BGD, HKG, IDN, KHM, LAO, PHL, SIN, and VNM.) This exemplifies that the "all and only" criterion is hardly met.

RQ2 Lexicogrammar-based Classification

Next, attention is paid to the strength of an L1 effect on the deeper layer of L2 English writing by Asian learners. A dendrogram was obtained from the analysis of the frequencies of 67 lexicogrammatical features that were less likely to be influenced by an essay topic.

Figure 3

A Dendrogram from the Cluster Analysis of the Frequency of the 67 Lexicogrammatical Features

As in word-based classification, items are divided into two large clusters that are not related to learners' L1 types. X comprises five part-time job essays and 22 non-smoking essays, whereas Y comprises 27 part-time job essays and ten non-smoking essays. Although some mixes are observed, 81% of the X items are non-smoking essays, and 73% of the Y items are part-time job essays. Even when analysing lexicogrammatical features, a topic seems to be the primary factor influencing Asian learners' L2 writing. X and Y are divided into three (a-c) and four (d-g) subclusters, respectively, although the difference between SC_e and SC_f is subtle.
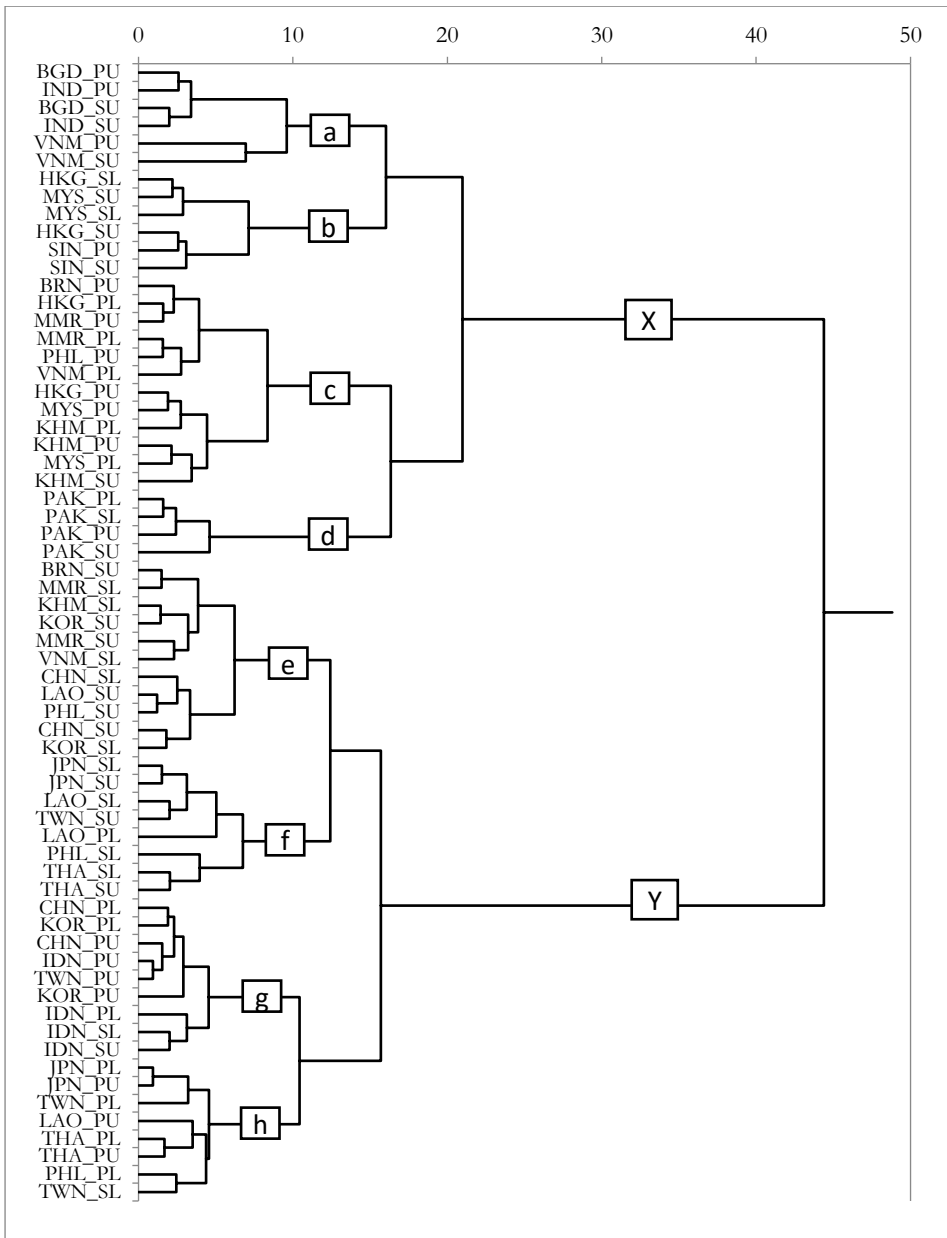
First, three same-L1 groups are examined. Regarding the L1 Chinese group, five of the 12 items occur in SC_c (CHN:2, TWN:2, HKG:1; 41.7%), but a combination of all three regions is seen only here. In the L1 Malay group, two of the six items occur in SC_a and SC_d (BRN:1, MYS:1; 66.6% combined), while the other items are widely scattered. Regarding the L1 Thai group, three of the eight items occur in SC_c (LAO:2, THA:1) and SC_f (LAO:1, THA:2) (75% combined), but both subclusters include many unrelated items.

Next, four similar-L1 groups are discussed. First, in the Indo-European group, three of the eight items occur in SC_b (BGD:1, IND:1, PAK:1; 37.5%). Regarding the Austronesian group, four of the 14 items occur in SC_ d (BRN:1, MYS:1, IDN:2; 28.6%). In the Sino-Tibetan group, six of the 16 items occur in SC_c (CHN:2, TWN:2, HKG:1, MMR: 1) and SC_e (CHN:2, TWN:2, MMR:2) (75% combined). In the Austro-Asiatic group, two of the eight items occur in SC_b (KHM:1, VNM:1) and SC_f (KHM:2) (50% combined). Finally, regarding language isolates, four of the eight items occur in SC_e (JPN:2, KOR:2; 50%).

Full regional coverage is observed in seven of the eight groups, excluding the Austronesian group. Contrary to word-based classification, the L1 effect could be somewhat stronger in the deeper layers of learner writing, where the topic effect seems lessened. However, the related item inclusion ratios are not sufficiently high, and the "all and only" criterion is not clearly met.

RQ3 Textual Factor-based Classification

Finally, attention is paid to the strength of an L1 effect on the latent layer of L2 English writing by Asian learners. From the analysis of the six-factor scores, which were assumed to be the least likely to be influenced by a topic, a dendrogram was obtained.

Figure 4

A Dendrogram from the Cluster Analysis of the Six Textual Factor Scores

As in the previous classifications, items are divided into two large clusters that show no clear relationship with the learners' L1 types. X

comprises 17 part-time job and 11 non-smoking essays, whereas Y comprises 15 part-time job and 21 non-smoking essays. The two types of topic essays are sufficiently intermingled, and the topic effect, as expected, seems to be the least at the latent layer of learner writing. X and Y are divided into four subclusters, respectively (a-d; e-h).

First, attention is paid to the three same-L1 groups. Regarding the L1 Chinese group, three of the 12 items occur in SC_g (CHN:2, TWN:1; 25%), but a combination with HKG items is not observed. In the L1 Malay group, three of the six items occur in SC_c (BRN:1, MYS:2; 50%). In the L1 Thai group, four of the eight items occur in SC_f (THA:2, LAO:2; 50%).

Next, attention is paid to four similar-L1 groups. First, in the Indo-European group, four of the eight items occur in SC_a (BGD:2, IND:2, 50%), but a combination with PAK items is not observed. With regard to the Austronesian group, four of the 14 items occur in SC_c (BRN:1, MYS:2, PHL: 1; 29%), but a combination with IDN items does not occur at all. Regarding the Sino-Tibetan group, four of the 16 items occur in SC_c (HKG:2, MMR:2) and SC_e (CHN:2, MMR:2) (50% combined), but a combination of all four regions is not observed. In the Austro-Asiatic group, four of the eight items occur in SC_c (KHM:3, VNM:1; 50%). Finally, regarding language isolates, unlike in the previous analyses, the co-occurrence of JPN and KOR is not observed in any of the subclusters.

Thus, full regional coverage is observed in only three groups (L1 Malay, L1 Thai, and Austro-Asiatic). A topic effect is assumed to be the weakest at the latent layer of learner writing, which, however, does not lead directly to the increase in the strength of the L1 effect in item classification.

Discussion

This study conducted three types of cluster analyses while paying attention to three layers of learner writing: the surface layer (words), the deeper layer (lexicogrammatical features), and the latent layer (textual factor scores). The objective was to investigate whether L1 typological grouping and statistical clustering are in accordance—whether data items sharing the same or similar L1 backgrounds are statistically clustered in a single subcluster despite the difference in terms of other parameters, such as essay topics.

If a learner's L1 has a sole and absolute effect on learner writing, (1) all related regional varieties are covered in a single subcluster, (2) all data items from the same or typologically similar L1 backgrounds are included, and (3) no unrelated items are included. Based on the findings of RQ1–3, this section summarises the region cover ratio (CR), the related item inclusion ratio (IR), and the unrelated item exclusion ratio (ER). CR and IR indicate the degree of

intra-group homogeneity, and ER indicates the degree of inter-group heterogeneity.

For instance, in the word-based classification, the largest subcluster for the Austro-Asiatic group is SC_a. This typological group comprises eight items (four KHM + four VNM), and SC_a includes 14 items in total, of which only three are related to KHM or VNM. The three indices are calculated as listed in Table 8.

Table 8

A Sample of Calculation of the CR, IR, ER, IR/ER Mean Values

Variables	Values
a: Number of regions in a typological group	2
b: Number of regions included in the largest SC	2
c: Number of items in a typological group	8
d: Number of items in the largest subcluster	14
e: Number of related items in the largest subcluster	3
CR (b/a)	2/2= 100.0%
IR (e/c)	3/8= 37.5%
ER (e/d)	3/14= 21.4%
Mean of IR/ER	29.5%

When the target sub-clusters comprise items only from a single region, they are excluded from the analysis. When two or more sub-clusters contain the largest number of items, those subclusters are combined for the analysis.

Table 9 presents the values of (a) to (e), as well as the CR, IR, ER, and IR/ER means according to different typological groups and different layers of learner writing. The values of the language isolates group are excluded in the calculation of the "7 group" means.

Table 9

CR, IR, ER, and IR/ER Mean Values for All the Typological Groups

Word-based Classification										
Groups	SC	a	b	c	d	e	CR	IR	ER	IR/ER
Chinese	e	3	2	12	12	4	66.7	33.3	33.3	44.4
Malay	a+f	2	2	6	27	6	100.0	100.0	22.2	74.1
Thai	c	2	2	8	9	3	100.0	37.5	33.3	56.9
Indo-Erp	b+g	3	2	8	10	6	66.7	75.0	60.0	67.2
Aust	a	4	4	14	14	6	100.0	42.9	42.9	61.9
Sino-T	c+e+f	4	4	16	34	12	100.0	75.0	35.3	70.1
Austro-A	a	2	2	8	14	3	100.0	37.5	21.4	53.0

Lang Isol	d+e	2	2	8	18	8	100.0	100.0	44.4	81.5
7 group	-	-	-	-	-	-	90.5	57.3	35.5	61.1

Lexicogrammar-based Classification

Chinese	c+e	3	2	12	27	8	66.7	66.7	29.6	54.3
Malay	a+d	2	2	6	15	4	100.0	66.7	26.7	64.5
Thai	c+f	2	2	8	26	6	100.0	75.0	23.1	66.0
Indo-Erp	b	3	3	8	8	3	100.0	37.5	37.5	58.3
Aust	d	4	3	14	10	4	75.0	28.6	40.0	47.9
Sino-T	c+e	4	4	16	27	12	100.0	75.0	44.4	73.1
Austro-A	b	2	2	8	8	2	100.0	25.0	25.0	50.0
Lang Isol	e	2	2	8	13	4	100.0	50.0	30.8	60.3
7 group	-	-	-	-	-	-	91.7	53.5	32.3	59.2

Textual Factor-based Classification

Chinese	g	3	2	12	9	3	66.7	25.0	33.3	41.7
Malay	c	2	2	6	12	3	100.0	50.0	25.0	58.3
Thai	f	2	2	8	8	4	100.0	50.0	50.0	66.7
Indo-Erp	a	3	2	8	6	4	66.7	50.0	66.7	61.1
Aust	c	4	3	14	12	4	75.0	28.6	33.3	45.6
Sino-T	c+e	4	3	16	23	8	75.0	50.0	34.8	53.3
Austro-A	c	2	2	8	12	4	100.0	50.0	33.3	61.1
Lang Isol	-	2	-	8	-	-	-	-	-	-
7 Group	-	-	-	-	-	-	83.3	43.4	39.5	55.4

If L1 is the sole and stable source of the effect on Asian learners' L2 English writing, CR, IR, ER, and IR/ER values are all expected to be close to 100%, whereas if L1 is not related at all to L2 use, those values would be close to 0%.

Table 9 shows that the ranges of the CR, IR, ER, and IR/ER values at three layers of learner writing are 66–100%, 25–100%, 21–67%, and 41–74% (or to 82% if including the language isolate group), respectively. These facts seem to suggest that an L1 influences Asian learners' L2 English writing to some degree; however, its practical effect is less decisive and stable than is usually believed.

Regarding the practical effect of an L1, one may assume that it is rather a region effect, which, however, is not necessarily supported by the current data. In word-based classification, for instance, four KUM items belong to three different subclusters (SC_a, f, g), and the same number of PHL items belong to as many as four different subclusters (SC_a, c, e, g). These findings suggest that an L1 effect exists separately from a region effect.

Among the seven L1-based typological groups, the IR/ER mean values in textual factor-based classification, which seem to reflect the practical L1 effect most clearly in terms of both intra-group homogeneity and inter-group heterogeneity, are relatively higher for the L1 Thai group (66.7%) and the Indo-European and Austro-Asiatic groups (61.1%), whereas they are the lowest for the L1 Chinese group (41.7%).

The exceptionally weak link between the three regions in the L1 Chinese group (CHN, TWN, and HKG) should attract attention. The fact that the L2 usage of learners in mainland China, Taiwan, and Hong Kong does not sufficiently overlap suggests that Asian learners' L2 English use might be influenced not solely by their L1 but by a much greater variety of factors, as suggested in many of the previous studies. Regarding the design of learner corpora, Gilquin (2015) notes the following:

> Unsurprisingly, many of the variables that affect the nature of interlanguage concern the learners themselves. Some of these variables are general, being applicable to any speaker/writer, native or not, e.g. age, gender, country/area, mother tongue. Other variables are more specifically relevant to learners, like the parents' native languages, the language(s) spoken at home, the learner's proficiency level, exposure to the target language inside the classroom (e.g. number of years spent learning the target language, pedagogical materials used) and outside the classroom (e.g. contact with the target language in everyday life, stays in target-language countries), or knowledge of other foreign languages…. The PAROLE Corpus is an example of a learner corpus that offers a particularly wide variety of measures, including motivation, listening comprehension skills, grammatical and lexical competence, aptitude for grammatical analysis and phonological memory… (p. 17).

In the Asian context, variables affecting interlanguage or learners' L2 use might also include the status of the English language in society, people's belief in English, English speakers, and acquisition of English, the weight of English in an entrance examination, and the major teaching method (e.g. grammar-oriented or communication-oriented, direct, or indirect methods). Considering such a limitless variety of factors, taking the L1 effect for granted and seeing it as an absolute source of the differences observed in the L2 English use of various Asian learners seems to be least appropriate.

Finally, the analysis sheds light on the relationship between learners with L1 Japanese and Korean backgrounds. As mentioned above, the two languages have many common features in morphology and syntax: both have case particles and multilayered honorific systems, and their sentence structures follow the same SOV pattern. According to Tsunoda (2009), who compares 19 kinds of linguistic features, including an SOV order, the relationship between nouns and adpositions, and an interrogatory marker, among others, of more than 100 languages worldwide, the two languages show the same patterns in terms of 18 of the 19 features (p. 282). However, in this analysis, pairing between the two languages is not observed in the textual factor-based classification, which may support the modern linguists'

view that they should be considered typologically different types of languages, although at the surface lexis level, two learner groups show high similarity.

The findings summarised above exemplify that, at least for intermediate learners in Asia, an L1 effect exists, but they are much less stable and decisive in comparison to the case of European learners, who have long been the target of previous L1 effect studies. In the Asian context, taking an L1 effect for granted and discussing the lexical or linguistic features of "L1 Chinese learners" or "learners with Indo-European L1 background", for instance, might not have a sufficient empirical foundation.

Conclusion

To re-examine to what extent Asian learners' L1s influence the surface layer (words), the deeper layer (lexicogrammatical features), and the latent layer (textual factor scores) of their L2 English writing, this corpus-based study analysed more than one-million-word topic-controlled essays written by 2,318 students from 18 regions and with more than 14 L1 backgrounds to investigate three research questions.

First, regarding RQ1 (word-based classification), when averaging the values of the seven typological groups, the CR, IR, ER, and IR/ER mean values were 90.5%, 57.3%, 35.5%, and 61.1%, respectively. The values were higher for the L1 Malay and Sino-Tibetan groups and lower for the L1 Chinese and Austro-Asiatic groups.

Regarding RQ2 (lexicogrammar-based classification), the mean values of CR, IR, ER, and IR/ER were 91.7%, 53.5%, 32.3%, and 59.2%, respectively. The values were relatively higher for the Sino-Tibetan, L1 Thai, and L1 Malay groups and lower for the Austronesian group.

Regarding RQ3 (factor score-based classification), the four values were 83.3%, 43.4%, 39.5%, and 55.4%, respectively. These values were relatively higher for the L1 Thai, Indo-European, and Sino-Tibetan groups and lower for the L1 Chinese and Austronesian groups.

From a macro viewpoint, intra-group homogeneity tends to reach approximately 50% in many cases, which suggests that an L1 effect on Asian learners' writing can be confirmed to some extent, although it comes after an essay topic as the primary source of effect. Meanwhile, inter-group heterogeneity tends to remain at approximately 35%—different typological groups are not mutually exclusive, and a considerable overlap exists. If one adopts the most rigid framework in the conceptualisation of an L1 effect (Jarvis, 2017) and admits it only when the "all and only" criterion is fully met, one would need to be sufficiently careful in discussing an L1 in the analysis of Asian learners' L2 English use.

The results of this study may require us to reconsider the traditional view that an L1 always plays a decisive role in L2 acquisition and use. LCR, especially contrastive interlanguage analysis (Granger, 1996; Granger, 2015) as its standard analytical method, which encourages one to compare an interlanguage with a native language ("native language" is reconceptualised by "reference varieties" in a revised model) and also compare interlanguages of learners with different L1 backgrounds, owes a lot to the theoretical approaches introduced in applied linguistics in the early days.

For example, contrastive analysis (Lado, 1957), which marks the beginning of modern applied linguistics, aimed to identify the difficulties in L2 acquisition for a particular L1 speaker group by comparing the syntactic structures and cultures of the two languages and revealing the mechanism of interference or negative L1 transfer. Subsequently, error analysis (Corder, 1967) aimed to collect, identify, describe, explain, and evaluate learner errors in L2 outputs because systematic errors, considered the result of the collision between L1 and L2, could represent a learner's "built-in syllabus" or a natural sequence of L2 acquisition. Further, interlanguage analysis (Selinker, 1972) redefined learners' L2 outputs as "interlanguage", which was regarded as a stable, patterned, and independent system existing between one's L1 and L2. Like their predecessors, LCR tends to pay special attention to learners' L1. Granger (1996) mentions that contrastive analysis can be connected to contrastive interlanguage analysis with predictive and diagnostic links, and Gilquin (2008) proposes combining contrastive and interlanguage analyses in the framework of LCR to detect, explain, and evaluate L1 transfer.

Contrary to these traditional views, this study presents a considerably different picture of the relationship between learners' L1 backgrounds and their L2 writing. Of course, one must be careful about the gap between previous studies and this study. The former usually adopt a narrowly localised approach, focus on a particular error, and consider whether it can be explained by a particular L1, which may be called a "weak version" approach (Wardhaugh, 1970) because its staple aim lies in diagnosis of the reasons for learners' (erroneous) L2 use. Conversely, this study chose a global analytical approach, examined the whole of learners' outputs that cover words, lexicogrammatical features, and textual factors, and considered whether L1 similarities lead to stable and quantitatively confirmable similarities between L2 output data, which might be called a kind of "strong version" approach because it tries to predict learners' L2 patterns from their L1 backgrounds.

This study has illuminated several noteworthy facts, as summarised above, but it includes several methodological limitations. First, the data size is not necessarily enough, especially with the students from the regions newly covered in the Written Essays Plus (WEP) module, such as Laos and Vietnam. As WEP is still an ongoing project, the author plans to conduct a

verification study after collecting additional data. Second, students' detailed L1 backgrounds, including "the parents' native languages" and "the language(s) spoken at home" (Gilquin, 2015, p. 17), have not been surveyed in the ICNALE. The project team required all the collaborators to collect data from the local students having the same L1 backgrounds and the team also asked each of the participants to report their mother tongue. The data from overseas students and exchange students were all excluded. However, some students may speak two or more languages as de facto mother tongues. How bilingual or multilingual students' L2 outputs differ from monolingual students' ones could be a new research topic. Third, this study analysed the essays of the students only at the B1 Upper and the B1 Lower levels. Limiting the proficiency to the B1, to which most of the Asian students belong, helped lessen the proficiency effect and focus on the L1 effect. The relationship between an L1 effect and a proficiency effect of a wider range should be a topic for further research. Finally, this study depended on statistical analyses, and it did not discuss learner outputs in sufficient detail from a qualitative viewpoint. How statistical classification is reflected in learner texts and what type of essays characterise each of the typological groups need to be further explored.

Although it is necessary to pay attention to these limitations, the results of this study warn us not to attribute any tendency observed in the output of a particular learner group easily and solely to their L1 backgrounds and not to apply the findings from the studies on European learners directly to Asian contexts without careful data-based consideration. In comparison to European learners, linguistic similarity between English and the mother tongue is not necessarily a prerequisite for Asian learners. Although the LCR in the early period was closely linked to the European context, it is now spreading widely and globally. Considering this shift, one naturally needs to pay more attention to multilingual and multicultural perspectives and global voices from learners around the world to create a new globalised version of LCR.

Acknowledgements

About the Author

Shin'ichiro Ishikawa: A professor of Applied Linguistics at the School of Languages & Communication, Kobe University, Japan. His research interests include corpus linguistics, statistical linguistics, TESOL and SLA. He has published numerous academic papers and books in applied linguistics. He was a leader of the ICNALE Learner Corpus Project.

References

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.

Corder, P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching, 5*, 161–170.

Eberhard, D. M., Gary, F. S., & Fennig, C. D. (Eds.). (2024). *Ethnologue: Languages of the world*. (27th ed.). SIL International. Online version: http://www.ethnologue.com.

Enfield, N. J. (2002). How to define 'Lao,' 'Thai,' and 'Isan' language? A view from linguistic science. *Tai Culture, 7*(1), 62–67.

Friginal, E. (2018). *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*. Routledge.

Gilquin, G. (2008). Combining contrastive and interlanguage analysis to apprehend transfer: Detection, explanation, evaluation. In G. Gilquin, S. Papp, & M. B. Díez-Bedmar, *Linking up contrastive and learner corpus research* (pp. 1–33). Brill.

Gilquin, G. (2015). From design to collection of learner corpora. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 9–34). Cambridge University Press.

Gilquin, G. (2023). Written learner corpora to inform teaching. In R. R. Jablonkai & E. Csomay (Eds.). *The Routledge handbook of corpora and English language teaching and learning* (pp. 281–295). Routledge.

Gilquin, G., & Granger, S. (2021). The passive and the lexis-grammar interface: An inter-varietal perspective. In S. Granger (Ed.), *Perspectives on the L2 phrasicon: The view from learner corpora* (pp. 72–98). Multilingual Matters.

Granger, S. (1996). From CA to CIA and back: An integrated contrastive approach to computerised bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast: Papers from a symposium on text-based cross-linguistic studies, Lund 4–5 March 1994* (pp. 37–51). Lund University Press.

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, *1*(1), 7–24.

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2024). Glottolog 5.0. Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.5281/zenodo.14006617 Available online at http://glottolog.org

Ishikawa, S. (2023). *The ICNALE guide: An introduction to a learner corpus study on Asian learners' L2 English*. Routledge.

Jarvis, S. (2017). Transfer: An overview with an expanded scope. In A. Golden, S. Jarvis, & K. Tenfjord (Eds.), *Crosslinguistic influence and distinctive patterns of language learning: Findings and insights from a learner corpus* (pp. 12–28). Multilingual Matters.

Lado, R. (1957). *Linguistics across cultures*. University of Michigan Press.

McEnery, T., & Brezina, V. (2022). *Fundamental principles of corpus linguistics*. Cambridge University Press.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Nini, A. (2019). The multi-dimensional analysis tagger. In T. B. Sardinha & M. V. Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 67–94). Bloomsbury Academic.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209–232.

Tenfjord, K., Jarvis, S., & Golden, A. (2017). Introduction. In A. Golden, S. Jarvis, & K. Tenfjord (Eds.), *Crosslinguistic influence and distinctive patterns of language learning: Findings and insights from a learner corpus* (pp. 1–11). Multilingual Matters.

Tsunoda, T. (2009). *Sekai no gengo to nihongo* [Japanese and the world languages]. Kurosio Publishers.

Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly*, *4*, 123–130.

Yurayong, C., & Szeto, P. U. (2020). (De-)Altaicisation as convergence and divergence between Japonic and Koreanic languages. *Japanese/Korean Linguistics*, *27*, 1–16.