A Corpus-Based Study on the Most Frequently Used English Prepositions: A Conceptual Transfer Perspective

Tran Tin Nghi
Ho Chi Minh City University of Food Industry, Ho Chi Minh City, Vietnam

Tran Huu Phuc
University of Foreign Language Studies, The University of Da Nang, Da Nang City, Vietnam

Abstract—This corpus-based study aims to examine the frequencies of English preposition usages from the perspective of conceptual transfer. The authors have built up a corpus from third-year English major students’ Academic writing tasks (HSAC) as their Ending-course assignments from Ho Chi Minh City University of Food Industry for two consecutive semesters in the academic year 2019 – 2020. This corpus was used as the source of data in analyzing the frequencies of prepositions in their essays. Two hundred and fifty pieces of writing were examined. The concordance output was used to determine the frequency of English preposition usage and types of errors via Sketch Engine Software. The findings showed a negative relationship between prepositional senses and their collocations with certain Vietnamese linguistic features. The study also revealed that negative conceptual transfer was recurrent and systematic. Based upon the findings, a conceptual framework for teaching and learning functional words like prepositions was drawn for language teachers and learners in a foreign language context.

Index Terms—corpus-based approach, English prepositions, conceptual transfer, mother tongue, foreign language

I. INTRODUCTION

Language learning for non-English-speaking countries is challenging and takes a lot of time. English learners in these countries encounter many difficulties in their learning process, even those who are gifted at learning foreign languages. Burns and Richards (2018) added cognitive difficulties in acquiring a new language besides lexical and grammar difficulties. Many studies on errors in foreign language teaching have shown that there are cognitively complicated phenomena, especially negative conceptual transfer. Despite many arguments among foreign language researchers about the factors affecting second language acquisition, most of them agree that negative L1 transfer lacks state data that needs further studying and investigating (Dulay et al., 1982; Jarvis & Pavlenko, 2008).

Although Ministry of Education and Training started to introduce English to all levels of Vietnamese educational system, starting in the 1990s, students still encountered many problems so far. The importance of nouns, verbs, and adjectives is overestimated by learners because they assume these words are crucial. They devoted more time to learning nouns, verbs, and adjectives than other elements of speech in order to actively learn English vocabulary. They also feel that prepositions are simple terms, and as a consequence, they push for ‘the translation’ from their mother language to English. This transfer phenomenon is so common that it is possible that they will have a detrimental impact on fluent native-like command of the English language for years to come. As a result, numerous prepositional mistakes have been committed by Vietnamese speakers. There are numerous distinct viewpoints on prepositional errors, and each of these approaches offers a different set of recommendations for dealing with the learner’s mistakes.

Contrastive Analyses (CA) have been replaced by Error Analysis (EA) since the 1960s as a result of the CA hypothesis’s deficiencies in finding the language errors. However, both of these perspectives do not give a clear-cut picture of the causes of errors. In CA hypothesis, Lado (1957) stated that language elements, which are the same or nearly the same as the target language, are proved to have no difficulties for learners, and those with the differences cause a lot of difficulties for learners. In EA theory, researchers can distinguish between errors and mistakes. An error is systematic, while a mistake is not. Both theories attempt to explain a typology of errors that are named as omissive, additive or substitutes; word-order; or other errors that occur at different levels of language such as phonological, lexical, and syntactic levels. As a result, these typologies are problematic since they analyze linguistic data alone, resulting in less reliable causes of errors.

One more perspective on viewing errors is called the Conceptual transfer perspective. The first empirical Cognitive linguistics study that Cognitive linguists have recently experimented with to examine the semantic nature is English prepositions. Tyler and Evans (2001) claim that that ‘they [prepositions] are notoriously hard to learn and frequently subject to negative transfer.’ Depending on their context, prepositional meanings may take on a variety of different...
connotations, including those associated with time, location, manner, or even the abstract. The conventional approach to acquiring prepositions in a foreign language argues that rote memorization is the best method. Although it offers a number of benefits, this approach perceives prepositional meanings as unrelated to one another. The numerous applications of prepositions may be understood as connected in systematic ways, according to Cognitive Linguistics Perspective. That is also the most acceptable viewpoint to assist Vietnamese ESL students in comprehending the metaphorical uses of English prepositions in a clear and precise manner. In light of Cognitive Linguistics, this paper examines the most typical uses of English prepositions in the Vietnamese context from the perspective of conceptual transfer.

II. THEORETICAL BACKGROUND

We applied Jarvis (2011) and Odlin (2005) frameworks to investigate the conceptual transfer hypothesis in SLA from their empirical evidence. Based on these frameworks, we develop the conceptual transfer framework in relation to the English prepositional usages of Vietnamese learners of English (See Figure. 1)

Figure 1. Conceptual transfer framework in Vietnamese context of learning English

A difference between linguistic and conceptual transfer is made in the transfer taxonomy. The former is mainly concerned with the mother tongue's and target language's linguistic forms and structures, while the latter is concerned with the many sorts of transfer that occur as a result of the mental conceptions that support those forms and structures. Linguistic transfer aims to explain 'transfer' by examining the similarities and differences between two languages' structural properties: the source and target languages — this kind of language transfer results in the linguistic representation of lexical level. Conceptual transfer, on the other hand, aims to understand another source of transfer by comparing the source and recipient languages' grammar and discourse levels and comparing conceptual categories.

When we discuss conceptual structures, we must apply our knowledge-based perspective to the language structures' concepts. In these circumstances, we must draw on our general knowledge of the world in order to investigate something (Keil, 1989; Murphy & Medin, 1985), since there is no one conceptual representation that corresponds to everything in the universe. Any concept must have an amalgam structure that satisfies three critical requirements: (i) it must possess knowledge of categories that pertain to a particular type of thing or group of things; (ii) it must possess expertise with borderline or fuzzy prototypes; and (iii) it must possess knowledge or belief regarding the internal structure. The conceptual structures have a strong resemblance to those of culture and cognition. Thought and language are two ways in which cognition manifests itself; certain conceptual representations are non-linguistic, as well (Jarvis & Pavlenko, 2008, p. 122). This point of view is also supported by Lakoff (1987), Ungerer and Schmid (1996), who asserted that cognition and culture are two sides of the same coin, respectively.

Inter-individual and inter-group differences are highlighted by cognitive representations whereas sociocultural and sociohistorical traits are emphasized by cultural elements. As a result, when we examine conceptual structures, we must consider "concepts" to be multimodal mental representations that include visual, auditory, perceptual, and kinaesthetic information, among other things (Barsalou, 2003). Consequently, social and political developments have a direct impact on our ability to communicate, and our conceptual knowledge is influenced by these developments as well. It's important to note that conceptual frameworks do not always have the same characteristics across members of the same community or speech group. For example, Vietnamese people employ the prepositional phrase "ngoài vườn" (viz., ‘outside of the garden’) to express the idea of "an item or animal in the garden," "Outside of the garden" or "Not in the garden" are translations of the English prepositional phrase if we use English – Vietnamese lexical equivalents.
Learners must understand a number of Vietnamese communities in order to comprehend a single Vietnamese sentence, each of which has its own culture and cognition of embodiment. "Ideas" and "conceptual categories" are two concepts that must be explored in depth because they are closely related. According to Murphy (2002), both words adhere completely to the tradition of concept research. "Concepts" are often defined as "mental representations of classes of things," whereas "categories" are defined as "the actual classes of things themselves" (Murphy, 2002, p. 5). In order to distinguish between "concepts" and "categories," we can use the terms "central and peripheral category members" when referring to the "whole concept" and "prototypes and mental imagery" when referencing the "mental representations of particular category members" (Jarvis & Pavlenko, 2008, p. 113). Both "language-independent concepts" and "concepts mediated by language," as defined by an experience-based approach, are part of conceptual language development. Despite acknowledging the existence of a language-independent mind, Whorf (1956) focused his attention on linguistic thinking, namely "thought insofar as it is linguistic (pp. 67-68)." Vygotsky (1978, 1986) makes use of the phrase "language mediation," however we did not incorporate his views into the research because Vygotsky himself was not interested in CLI. The importance of his influential notion of "language mediation" for later research based on "linguistic logical entities," "cognitive psychology," and other concepts must be acknowledged.

An experienced process is used to develop language-independent presentations (or mental conceptions), rather than the use of verbal representation to convey information. In other words, language-mediated ideas are produced as a result of the evolution of language socialization, during which the learning of words and the acquisition of their categories have an impact on one another over an extended period of time. Murphy (2002) has synthesized his research on "word meaning development" in the process, where 'word learning progresses creates changes in conceptual structure’ (p. 402).

There have been numerous investigations into how different aspects of linguistic fields affect language transfer. The distinction between linguistic and conceptual transfer is one of the cross-language influence taxonomy dimensions (Jarvis & Pavlenko, 2008). Linguistic transfer is concerned with the relationships that exist between language forms and structures in two different languages (or languages and cultures). According to Nghi (2020), conceptual transfer, on the other hand, is the influence of a language’s language-mediated conceptual categories on the language usage of a second or additional language. Phonetical transfer (Bleses et al., 2011; Burt & Dulay, 1978; Cambier-Langeveld & Turk, 1999; Chávez-Péon et al., 2012; Gordon et al., 2002), morphological transfer (Lidner & Johnston, 1992; Sagarra & Ellis, 2013), syntactical transfer (Biria & Ameri-Golestan, 2010; Paradis & Genesee, 1996; Yip & Matthews, 2000), semantic and pragmatic transfer (Allami & Naeimi, 2011; Bu, 2011; Golestani et al., 2009), and sociolinguistic transfer (Bayley, 2007; Behrens & Neeman, 2004; Meyerhoff, 2009) are some of the subfields of linguistic transfer research that have been studied.

These findings reinforced Oldin's hypothesis that “transfer can occur in all linguistic subsystems.” A number of recent studies on conceptual transfer examined the similarities and differences in conceptual representations that corresponded to linguistic subcategories of the source and recipient languages, such as English vocabulary usage (Huiping & Yongbing, 2014; Jakovleva, 2012; Phúc et al., 2018; Zhang & Luo, 2017), number cases (Charters et al., 2012), concepts (Oldin, 2005), L2 writing (Omar, 2018), and language proficiency and age (Bagherian, 2012). This means that conceptual transfer is considered a hypothesis that L1 transfer in a person’s use of specific target language items originates from their ‘conceptual knowledge and thought patterns’ (Jarvis, 2007, p. 44).

### III. Methodology

The corpus HUFI Student Academic Corpus (HSAC) data were built up from writing tasks of HUFI English majored students in their End-course examination during two consecutive semesters in the academic year 2019 – 2020. Two hundred and fifty writings involving English prepositions in the corpus were examined to interpret the most common English prepositions in their writing articles submitted to the Research Methodology Subject (see Table 1).

| Table 1. HUFI STUDENT ACADEMIC CORPUS INFORMATION |
|---------------------------------------------------|
| COUNTS INFO                                      |
| Tokens                                           | 2,185,327 |
| Words                                            | 1,792,679 |
| Sentences                                       | 98,774   |
| Documents                                       | 250      |

We choose Cambridge Academic English Corpus (CAEC) to be the comparative data in our study. The CAEC is an Academic English corpus comprised of a sample of texts gathered from undergraduate and postgraduate students at a variety of US and UK universities. Academic English corpus materials include university lectures, seminars, presentations, essays, research articles, and university coursebooks. The topics are nearly the same or sub-types of the...
topics in HSAC because these topics are very common with language teaching and English literature. It would be appropriate to discuss the features of Vietnamese learners’ (of English) conceptual transfer compared to the data of American learners. CAE has a total of 100,190,245 tokens, 85,683,246 words, 4,578,508 sentences, 1,951,384 paragraphs, and 154,457 documents.

![Figure 2. CAEC topics based on the subjects in the sub-corpora](image)

IV. RESULTS AND DISCUSSION

We reported the frequency of occurrence of specific terms in conceptual transfer categories. With the help of Sketch Word Software, we calculated the normalized frequencies per 10,000 words of the English prepositions in both corpora. Table 2 presented the most common English prepositions written by Vietnamese students and native students.
From data in Table 2, among the 40 English prepositions commonly used by Vietnamese students and European students and lecturers, only 4 English prepositions have the same normalized frequencies rate. They are of with 6.29 ~ 7.77, in with 4.24 ~ 6.86, to with 1.67 ~ 3.69, and for with 1.63 ~ 2.77. Even though Vietnamese learners have mostly used these four English prepositions like the native students, they have a difference in normalized frequencies. The HSAC’s normalized frequencies are only half of CAE’s number. In fact, when we look further in the two language systems, Vietnamese and English, we realized that the numbers of English prepositions are much larger than Vietnamese prepositions. The Vietnamese prepositions are only approximately up to 22% in comparison with those in English language.

![Figure 3. The comparison between Vietnamese and English prepositions](image-url)
The difference was so enormous that Vietnamese learners of English get lost in their search for appropriate English prepositions when they need to use them in their English production. The results revealed that Vietnamese students use four English prepositions: ‘OF’, ‘IN’, ‘TO’, and ‘FOR’ with nearly half of the frequencies compared to native students.

For the extended list of English prepositions in two corpora, Vietnamese students utilize just 85 of the most frequently used English prepositions, which is less than a quarter of the list of European students. Some of English prepositions are rarely used by Vietnamese students such as alongside, albeit (functions like Preposition), unto (same as “to” - old use), underneath, notwithstanding, amid, cum (used to join two nouns, like bedroom-cum-study), qua, vis-à-vis, lest, outwith, circa, sans, ere, chez, betwixt, par, considering, concerning, astride, downstream, north, forth, afore, etc.

When we set up the concordance with the VPS formula: [verb] + [10 most common prepositions], we calculated the distribution of observed prepositions in two corpora. Below are some of the results from our observations (see Table 3).

**TABLE 3**

| Order | Observed Prepositions | HUFI Student Academic Corpus | Cambridge Academic Corpus |
|-------|------------------------|-------------------------------|---------------------------|
|       |                        | VPS Normalized frequencies   | Numbers of distributions  | Percentages |
| 1     | of                     | 5503                          | 62912                     | 8.75        | 5454 | 12645 | 43.13 |
| 2     | in                     | 4956                          | 42521                     | 11.66       | 3037 | 77119 | 3.94  |
| 3     | to                     | 37895                         | 55143                     | 68.72       | 51237| 89413 | 57.30 |
| 4     | for                    | 1855                          | 16321                     | 11.37       | 1974 | 27691 | 7.13  |
| 5     | on                     | 1052                          | 11058                     | 9.51        | 509  | 17528 | 2.90  |
| 6     | with                   | 426                           | 10577                     | 4.03        | 548  | 22419 | 2.44  |
| 7     | as                     | 1437                          | 13117                     | 10.96       | 3794 | 32987 | 11.50 |
| 8     | from                   | 274                           | 7059                      | 3.88        | 554  | 13870 | 3.99  |
| 9     | by                     | 1694                          | 6945                      | 24.39       | 3459 | 19546 | 17.70 |
| 10    | at                     | 190                           | 6433                      | 2.95        | 430  | 10183 | 4.22  |

From the data in Table 3, only prepositions ‘to’, ‘for’, ‘as’ and ‘from’ remained at the same rates with each other in both corpora. It meant that conceptual presentations for the meaning of these four prepositions had reached the level of native speakers, and there was a very little L1 conceptual transfer for kinds of prepositions.

Certain verbs are more often used by Vietnamese learners, such as assist, use, take, study, feel, and teach, meet, speak, learn, make and do; whereas western students prefer to use the verbs such as work, spend, write, earn, thank, and come, explore, come, understand, dream, etc. From these data, we may hypothesize that the discrepancies in high frequency verbs across learner groups are due to differences in L1 conceptual category systems (see Figure 4 and 5).

---

Figure 4. An example of the use of English preposition “OF” in HSAC
V. CONCLUSION

Conceptual categories for Vietnamese prepositions are frequently transferred in the interlanguage of English-majored students. L1 conceptual representations mainly occur at lexical and grammatical levels. From the comparative results between the two corpora, conceptual transfer errors are evident and systematic with the different rates of usage, specifically in the following:

(1) Vietnamese English learners have less difficulty using four common English prepositions: "of," "in," "to," and "for";

(2) Vietnamese learners of English made some inappropriate uses of English prepositions due to their lack of equivalent terms in their mother tongue;

(3) Based on a comparison of two corpora, the frequencies of errors made by Vietnamese English-majored students are rather high and systematic, and L1 conceptual transfer counts for such errors.

REFERENCES

[1] Allami, H., & Naeimi, A. (2011). A cross-linguistic study of refusals: An analysis of pragmatic competence development in Iranian EFL learners. Journal of Pragmatics, 43(1), 385-406.

[2] Bagherian, A. (2012). EFL Learners’ L1 Conceptual Transfer and Its Relation to Their Language Proficiency and Age. International Journal of Applied Linguistics and English Literature, I(4), 152-160.

[3] Barsalou, L. W. (2003). Situated simulation in the human conceptual system. Language and Cognitive Processes, 18(5-6), 513-562. doi: https://doi.org/10.1080/01690960344000026

[4] Bayley, R. (2007). Sociolinguistic Variation: Second language acquisition: a variationist perspective. In R. Bayley, & C. Lucas, Sociolinguistic Variation: Theories, Methods, and Applications (pp. 133-144). Cambridge: Cambridge University Press.

[5] Behrens, S. J., & Neeman, A. R. (2004). Focus on Accent: A Sociolinguistic Perspective of Diversity in the Classroom. Research in the Teaching of Developmental Education, XX(2), 40-47.

[6] Biria, R., & Ameri-Golestain, A. (2010). The Impact of Syntactic Priming on English Language Learners’ Production: A Transfer Study of Indirect Questions/Requests. Journal of Language Teaching and Research, I(5), 583-590.

[7] Bleses, D., Basboll, H., Lum, J. A., & Vach, W. (2011). Phonology and lexicon in a cross-linguistic perspective: the importance of phonetics – a commentary on Stoel-Gammon’s ‘Relationships between lexical and phonological development in young children.’ Journal of Child Language, XXXVIII(1), 61-68.

[8] Bu, J. (2011). A Study of Relationships between L1 Pragmatic Transfer and L2 Proficiency. English Language Teaching, V(1), 32-43.

[9] Burns, A., & Richards, J. C. (2018). The Cambridge Guide to Learning English as a Second Language. Cambridge: CUP.

[10] Burt, M., & Dulay, H. (1978). Some Guidelines for the Assessment of Oral Language Proficiency and Dominance. TESOL Quarterly, XII(2), 177-192. doi: https://www.doi.org/10.2307/3585608

[11] Cambier-Langeveld, T., & Turk, A. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. Journal of Phonetics, XXVII(3), 255-280.

[12] Charters, H., Dao, L., & Jansen, L. (2012). Think of a number: conceptual transfer in the second language acquisition of English plural-marking. CognTextes, 1-12, doi: https://doi.org/10.4000/cognitextes.611

[13] Chávez-Péon, M. E., Bernhardt, B. M., Adler-Bock, M., Ávila, C., Carballo, G., Fresneda, D., Stemberger, J. P. (2012). A Spanish pilot investigation for a crosslinguistic study in protracted phonological development. Clinical Linguistics & Phonetics, 26(3), 255-272.

[14] Dulay, H. C., Burt, M. K., & Krashen, S. D. (1982). Language two. New York: Oxford University Press.

[15] Golestani, N., Rosen, S., & Scott, S. K. (2009). Native-language benefit for understanding speech-in-noise: The contribution of semantics. Bilingualism: Language and Cognition, 12(3), 385-392.
