

Corpus-Based Error Analysis of Chinese Learners' Use of High-Frequency Verb Take

Yanru Li¹

¹ School of Foreign languages, Taishan University, Taian, Shandong, China

Correspondence: Yanru Li, School of Foreign Languages, Taishan University, Taian, Shandong, 271000, China.

Received: January 7, 2022

Accepted: January 25, 2022

Online Published: January 26, 2022

doi: 10.5539/elt.v15n2p21

URL: <https://doi.org/10.5539/elt.v15n2p21>

Abstract

This study investigated the erroneous use of the high-frequency verb TAKE by the Chinese college learners of English as a foreign language (EFL), aiming to identify the similarities and differences between Chinese EFL learners, aimed at finding out more effective ways for the teaching and researching of the high-frequency verbs. Corpus-based Contrastive Interlanguage Analysis and Error Analysis were carried out in the present study, with the subcorpora ST4 and ST6 of CLEC (Chinese Learner English Corpus) as the learner corpora.

The analyses involved the misuse of the verb TAKE by the Chinese EFL learners. The error analysis of TAKE was based on the classification in the corpus CLEC. From the perspective of the overall frequency, the ST6 learners commit fewer errors than the ST4 learners. From the perspective of error types, the ST6 learners and the ST4 learners have much in common. That is, the error types of “wd” and “cc” take up an overwhelming part of all the errors in both corpora. These errors are caused by some interlingual and intralingual factors such as language transfer, overgeneralization, and communication strategy. In comparison, in the process of EFL learning, the non-English majors are interfered by their mother tongue to a larger extent than the English majors.

Keywords: corpus-based study, error analysis, high-frequency verb TAKE

1. Introduction

1.1 Background of the Study

Language has three basic elements: sound, word and grammar. As a combination of sound and meaning, the word is an integral part of the language system. There is no doubt that vocabulary has always been a focus of language teaching and researching. Traditionally, the vocabulary research investigated the meanings of words and synonyms. In more recent times, such investigations have been extended using corpus-based approaches to examine the ways that words are used, especially the use of a given word. Nation (2001) classifies vocabulary into four groups: high-frequency words, low-frequency words, academic terms and technical terms, among which the high-frequency words constitute the largest part. According to Altenberg and Granger (2001), the learners of English as a foreign language (EFL), even the advanced learners, have great difficulty in using the high-frequency verbs.

The verb TAKE is on the list of the fifteen high-frequency verbs (disregarding BE and modal auxiliaries) made by Svartvik and Ekedahl (1995). In the present study, among these fifteen high-frequency verbs, the verb TAKE ranks the ninth and the seventh respectively in the learner corpora—ST4 and ST6, the subcorpora of CLEC (Chinese Learner English Corpus). As for such a high-frequency verb, the author has only retrieved four essays from CNKI (Chinese National Knowledge Infrastructure). One of them makes an analysis of the verb TAKE in terms of the lexical patterns and the rest mainly investigate the collocational behavior of TAKE.

1.2 Research Questions and Objectives of the Study

The high-frequency verb TAKE is a polysemous word with 35 meanings listed in Longman Dictionary of Contemporary English and enters into a range of structures, such as delexical structures, phrasal verbs and many idiomatic phrases. Complex use of the verb TAKE causes great difficulty in the acquisition of this word.

By means of Contrastive Interlanguage Analysis (CIA), this study focused on the error analysis of Chinese EFL learners' use of the high-frequency verb TAKE, hoping to answer the following research question: What is the distribution of the errors of TAKE committed by ST4 and ST6 learners and what are the sources of these errors?

The Chinese EFL learners in the present study involve the non-English majors in ST4 and the English majors in ST6.

The corpus-based error analysis of the learners' use of this word may reveal some major problems in using it and provide some pedagogical implications.

2. Literature Review

2.1 Error Analysis

Errors are studied by means of Error Analysis (EA). Its heyday is in the 1960s and 1970s. Corder is the first advocator of EA in the modern sense. In 1967, Corder is inspired by the error analysis in the mother tongue acquisition and publishes the article *The Significance of Learner's Errors*, where the significance of Error Analysis is pointed out: (1) Regarding the teachers, EA provides them with information concerning what the L2 learners have acquired; (2) Regarding the researchers, EA can offer them evidence of how language is learnt; (3) Regarding the learners, EA can offer them the devices by which they can discover the rules of the target language.

2.1.1 Interlanguage

EA provides a new method to investigate the learner language, a language system distinctive from either one's native language or the target language.

This language system is generally believed to be constructed by the learners based on the input to which one has been exposed. Nemser (1971) describes it as "approximate system" and Corder (1971) refers to it as "idiosyncratic dialect". The term "interlanguage" is used for the first time by Selinker (1969) in his essay *Language Transfer*; and his article *Interlanguage* published in 1972 provides for the position of interlanguage in SLA. Then the term "interlanguage" becomes popular in that it might be better understood to be seen as a continuum between the native language and the target language along which all learners traverse. The learner's language is systematic and dynamic which has its own characteristic system known as "built-in syllabus" (Corder, 1967), or "internal grammar" (Van Els et al., 1984), or "interlanguage competence" (Brown, 1994).

2.1.2 Procedures of EA

Corder (1974) proposes five procedures for the teachers and researchers to conduct the error analysis: collecting a sample of learner language, identifying errors, describing errors, explaining errors and evaluating errors. According to Ellis (1994), many studies do not include the fifth procedure which would not be elaborated here.

The first step of EA is to collect samples of the learner language. In this study, the data was collected from the subcorpora ST4 and ST6 of CLEC.

The second step of EA is to identify errors. The major task is to distinguish *error* from *mistake*. According to Corder (1981), *mistake* is a performance slip caused by fatigue, excitement, etc, which is quite random and can be readily self-corrected, whereas *error* is a systematic deviation made by learners for lack of the rules of the target language, which represents a lack of competence. Since EA should be restricted to the study of errors, it is necessary to eliminate mistakes from errors. In the present study, the learner corpus CLEC is error-tagged, which facilitates the retrieval of the errors of the verb TAKE.

The third step of EA is to describe errors. The analysis at this stage focuses on the classification of the errors. The descriptive taxonomy on the basis of linguistic categories is perhaps the simplest way of classifying errors. For example, in the research made by Burt and Kiparsky (1972), a number of linguistic categories are identified, such as the skeleton of English clauses, the auxiliary system, and sentential complements. In comparison, Politzer and Ramirez (1973) classify the errors into morphology, syntax and vocabulary. The errors in CLEC are first classified into 11 large categories ("fm", "vp", "np", "pr", "aj", "ad", "pp", "cj", "wd", "cc", and "sn") and subdivided into 61 types which would be elaborated in Chapter Three. The error analysis in the present study was based on this taxonomy.

The fourth step of EA is to explain errors. To explain the errors means accounting for why the error is made or exploring the sources of it.

According to Richards (1971a), some errors committed by L2 learners can be traced to L1 interference or negative transfer of their mother tongue, which is referred to as "interlingual errors". Besides, there are a large number of errors committed by the learners regardless of their mother tongue. Richards terms them "intra-lingual errors", which are subdivided into four categories: overgeneralization errors, ignorance of rule restrictions, incomplete application of rules, and false concepts hypothesized. Some unique errors are distinguished by researchers. Selinker (1972) labels communication-based errors which arise when the SL learners invoked

communicative strategies (e.g. avoidance and paraphrase) and Stenson (1974) identifies induced errors which occur when a teacher sequences or presents two linguistic items in a way which may cause confusion in the mind of the second language learners, or briefly speaking, these errors are the results of the instruction they receive.

2.1.3 Limitations of EA

EA comes into disfavor with the researchers after being used for a short period. Brown (1994) holds that it is dangerous to focus too much on the learners' errors. One charge is on the fact that EA can not provide a complete picture of learner language by only focusing on errors. Another charge is that EA cannot give reason for all the areas of L2 where learners have difficulty. Although errors do not arise in some structures, it does not necessarily mean that the learners have no difficulty in acquiring these structures. On the contrary, just because of the difficulty in using these structures, the learners tend to avoid using them deliberately. This is the so-called "avoidance" which cannot be explained by EA (Larsen-Freeman & Long, 1991). On the other hand, Levenston (1971) points out that there is another problem on the part of the learners—the overuse of some structures. This problem is opposite to avoidance and is labeled "over-indulgence". Only focusing on errors can not help the researchers account for such phenomena.

In short, the perspective of EA is still too narrow. With the learner corpora available, another important type of analysis—Contrastive Interlanguage Analysis becomes the new favorite of the SLA researchers.

2.2 Learner Corpora and Contrastive Interlanguage Analysis

Learner corpora can be constructed for the purpose of studying the learners' interlanguage much in the same way as the native speaker corpora are used for studying the language of the native speakers. The learner corpus is thought to be the meeting point of SLA research and corpus linguistic study. According to Granger (1998), since it is rooted in both corpus linguistic and second language acquisition studies, it can use the corpus approach to gain better understanding of the authentic learner language. She points out that "A learner corpus based on clear design criteria lends itself particularly well to a contrastive approach" (Granger, 1998:12) and refers to this new approach as Contrastive Interlanguage Analysis (henceforth CIA). CIA is different from the contrastive analysis in a traditional sense.

Two types of comparison are involved in CIA. One type is the comparison between the native language and interlanguage (NL vs. IL) and the other type is the comparison between different interlanguages (IL vs. IL).

As for NL/IL comparison—comparison between native language and interlanguage, its purpose is to unveil the features of learner language or interlanguage, such as "non-nativeness" and "linguistic strangeness". Before the learner corpora are available, interlanguage has been simply approached from the perspective of learners' errors. But now, by means of corpora, the SLA researchers can conduct quantitative analysis of interlanguage, investigating its quantitative features (i.e. overuse/underuse).

IL/IL comparison refers to the comparison between interlanguages of different languages. This type of comparison is mainly aimed at gaining a better understanding of the nature of interlanguage. By means of comparing learner corpora involving different varieties of English such as age, NL background, proficiency level, learning setting, etc, it enables the researchers to investigate the effect of the above variables on learner output (Granger, 1998).

By offering both qualitative and quantitative analysis of learner language, CIA can provide answers to some unsolved questions in the SLA research. This type of data analysis may help to design new pedagogical tools and classroom practices to more accurately target the need of the learners. By means of CIA, the present study attempts to explore the Chinese EFL learners' use of the verb TAKE, and provide some pedagogical implications.

2.3 Studies on High-Frequency Verb TAKE

Some researchers attempt to explore the Chinese EFL learners' use of the single verb TAKE, yet most of which is centered on the "verb + noun" collocation of TAKE. Some of the studies are made on the non-English majors while other studies are on the English majors. Zhou (2012) investigates the Chinese non-English majors' use of the high-frequency verb TAKE from the perspective of frequency and collocation. The study finds that as compared to the native speakers, the Chinese non-English majors tend to overuse the verb TAKE, but choose less various collocates and they could not use the collocates listed in English curriculum for universities well. Dong (2015) compares the essays in ST5 of CLEC written by the second year English majors with native-speaker learners. The study reveals that compared with the native speakers, the ST5 learners choose a large number of collocates with fewer collocation types. The researcher believes these problems are caused by their learning strategies and some inter- and intralingual factors. Li (2011) investigates the Chinese senior

English majors' feature in the use of the "verb + noun" collocations of TAKE. ST6 of CLEC is chosen as the learner corpus and LOCNESSS as the native speaker corpus. Besides the similar findings revealed in Dong's study, Li still finds that the collocates chosen by both ST6 and LOCNESS learners have different typicality across the two corpora and some collocates which are frequently used by the Chinese senior English majors are not found in the native speaker corpus.

Differently from the above researchers, Liu (2011) investigates the Chinese English majors' use of the verb TAKE in terms of lexical patterns. He categorizes the senses and patterns of the verb TAKE into 14 groups. The results show that compared with the native speakers, the distributions of the senses of the verb TAKE are quite different. Besides, the Chinese learners overuse the pattern of phrasal verbs and underuse the delexicalized construction of the verb TAKE.

Above all, most studies on Chinese learners' use of verb TAKE in China are either on English majors or on non-English majors. Few studies compare the use of TAKE by English and non-English majors. Besides, few researchers investigate the errors of the verb TAKE. This study aims to investigate the misuse of the verb TAKE by the non-English majors in ST4 and the English majors in ST6, hoping to shed some light on the teaching and research of the high-frequency verbs.

3. Methodology

3.1 Corpora Used in the Study

In this study, Chinese Learner English Corpus (CLEC) was employed as the learner corpus. It is a very authoritative Chinese learner corpus and has been widely used in the interlanguage study of Chinese English learners. Under the supervision of Gui Shichun and Yang Huizhong, this learner corpus collects 1,000,000 words of essays from the writing output of senior middle school students, college non-English majors and English majors. These students are of five different English proficiency, and accordingly, CLEC is divided into five subcorpora—ST2, ST3, ST4, ST5 and ST6. Table 1 can give a clear description of the five subcorpora.

Table 1. Brief Description of CLEC

Subcorpus	Level of learner
ST2	senior middle school students
ST3	non-English majors (Grade 1-2)
ST4	non-English majors (Grade 3-4)
ST5	English majors (Grade 1-2)
ST6	English majors (Grade 3-4)
Total	Chinese EFL learners

According to Yang (2002), the corpus is designed to facilitate the research on the development of the learners' interlanguage and on the difference analysis between Chinese EFL learners and the native speakers. Besides CLEC is an error-tagged learner corpus, which can help make the error analysis of the verb TAKE.

In the present study, the author only chose ST4 and ST6 for further investigation. Since the learners in ST4 and ST6 respectively represent the college non-English majors and English majors with higher English proficiency, they can best represent the Chinese college EFL learners.

3.2 Data Collection and Processing

WordSmith Tools 6.0 was adopted to identify and analyze the data in the corpora. It is one of the most common software used in corpus-based research. Since CLEC is error-tagged, it is quite easy to sort out all the errors concerning TAKE from ST4 and ST6.

In CLEC, the errors are first classified into 11 large categories (such as "vp", "wd", "fm", "cc", "sn", and "pr"). Then each category is subdivided into several groups marked with numbers (such as "vp1", "wd3", and "cc2"). Besides, the errors in CLEC are tagged with square brackets after the errors, as shown in the sentence "*We must take some means [cc3, 2-] to do with them*". [cc3] stands for the third error type of "cc"—"verb + noun" collocational error; "2-" refers to where the error appears ("-" shows the location of the error; "2" means there are two words before the error).

Altogether fifteen subcategories of errors involving TAKE were sorted out from the two corpora. They are "vp1", "vp2", "vp3", "vp4", "vp5", "vp6", "vp7", "vp9", "wd1", "wd2", "wd3", "wd4", "wd5", "wd7", and "cc3". It is worth noting that among these error types, such error types as "vp3" (agreement), "vp4" (finite/non-finite),

“vp6” (tense), “vp9” (modal/auxiliary) are not typical errors of TAKE committed by EFL learners but errors committed in the process of learning English, so they are not included in further analysis. To have a better understanding of the rest error types, explanations and examples of each error type were given below:

(1) vp1: error in transitivity

For example: I'll take [vp1, -2] active in the societism [fm2, -] things.

(2) vp2: error in form or use of phrasal verb and verbal phrase

For example: The social [wd2, -] will take the most advantage from [vp2, 4-] the death of the commodities.

(3) wd2: error in part of speech

For example: As a college student, I will take activity [wd2, -] part in the social active [wd2, 2].

(4) wd3: error in word choice

For example: My view on job-hopping is taking for [wd5, 2-2] one job.

(5) wd4: omission of a word

For example: Taking [wd4, 1-3] place of it is [vp6, s-] a political class.

(6) wd5: oversuppliance of a word

For example: Some people consider it is profitable to take up in [wd5, -2] one job in their lives.

(7) wd7: not clear word meaning

For example: Nowadays [fm1, -], fake commodities take up [wd7, 3-] in more and more field.

(8) cc3: improper verb and noun (phrase) combination

For example: We must take some means [cc3, 2-] to do with them.

After all the errors under examination were identified, the distributions of those errors in ST4 and ST6 were analyzed and then the most frequent errors committed by non-English majors and English majors were examined in detail to explore the possible sources for those errors under discussion.

In the last phase of the study, major findings were presented and some pedagogical suggestions were offered.

4. Results and Discussion

4.1 Distribution of Errors in ST4 and ST6

After sorting out all the errors of the verb TAKE, the frequencies and percentages of these errors in the two corpora were counted and calculated, as shown in Table 2.

Table 2. Overall Distribution of Errors across ST4 and ST6

Verb lemma TAKE	ST4		ST6	
	frequency	percentage	frequency	percentage
Errors of TAKE	47	8.6	14	3.3
Occurrences of TAKE	549	100	425	100

From Table 2, it can be seen that, in terms of the total sum of the errors of the verb TAKE, the ST4 learners outnumber the ST6 learners. The ST4 learners commit 47 errors, accounting for 8.6% of the entire TAKE occurrences, while the ST6 learners make 14 errors, constituting 3.3%. In some way, the comparatively higher error rate of the ST4 learners indicates that the non-English majors have more difficulty in using the verb TAKE. The distribution of each type of errors in ST4 and ST6 is shown in Table 3.

Table 3. Distribution of Each Error Type across ST4 and ST6

Error types	ST4		ST6	
	frequency	percentage	frequency	percentage
vp1	6	12.8	1	7.1
vp2	4	8.5	1	7.1
wd2	1	2.1	1	7.1
wd3	15	31.9	1	7.1
wd4	2	4.3	4	28.7
wd5	5	10.6	1	7.1
wd7	2	4.3	1	7.1
cc3	12	25.5	4	28.7

Note. In CLEC, the errors are divided into 11 groups (such as “vp”, “wd”, “cc”, “fm”, “pr”, “sn”, etc.) and subdivided into 61 types marked with numbers. The errors are tagged with square brackets after the error, as shown in the sentence “We must take some means [cc3, 2-] to do with them.” [cc3] stands for the third error type of “cc” (collocation) — “verb + noun” collocational error; “2-” refers to where the error appears (“-” shows the location of the error; “2” means there are two words before the error).

Table 3 shows that 8 error types of the verb TAKE are committed by the Chinese EFL learners. It can be seen that the error type “wd3” (the error in word choice) is the most frequent one in ST4, taking up 31.9% of all the instances of errors, and the most recurrent error type “wd4” (the omission of a word) in ST6 takes up 28.7%. In addition, the error type “cc3” (the error of improper V+N collocation) ranks the second in both ST4 and ST6, taking up 25.5% and 28.7% respectively. These three error types are the most typical and representative, and worthy of further analysis. The focus of the subsequent analyses is on the possible sources of these errors.

4.2 Possible Sources of Errors in ST4 and ST6

Since the sources of the error types “wd3”, “wd4” and “cc3” may have some overlaps, the author did not analyze these error types one by one but illustrated each source of the relevant errors with examples. Some interlingual errors (L1 transfer) and intralingual errors (overgeneralization, communication-based errors) were explored in the present study.

4.2.1 Language Transfer

According to behaviorist theories, the main obstacle of second language acquisition is the interference from the native language. Interference is the result of proactive inhibition, that is, what has been learnt or memorized before has influence on what is learnt later. If the native language and the target language share a meaning but use different ways to express it, it is likely that the learners may commit errors in the target language. The reason is that the learners tend to transfer the realization device from their NLs into the TLs (Ellis, 1986). In other words, the negative transfer will occur when the differences between the NLs and the TLs create learning difficulty which leads to errors. For example:

- (1) Taking [wd4, 1-3] place of it is a political class. (ST6)
- (2) For [wd5, s-] the third people take more care of their bodies. They take [cc3, -2] more exercises than before. (ST4)

The errors in these two examples are interfered by the differences between English and Chinese. In Example (1), the writer misused the phrase *Taking [wd4, 1-3] place of* (Correct form: *taking the place of*) at the result of omitting the definite article *the*. This may be due to the fact that in Chinese there are no definite articles. As for Example (2), given the context, one can infer that the writer wanted to use the collocation *take more exercises* to convey the meaning “做更多运动”. According to Longman Dictionary, the word *exercise* is an uncountable noun when it refers to physical activities which can be collocated with *do/take*. The correct form should be *take more exercise* rather than *take more exercises*. The error arose by reason that in Chinese there are no such marks to distinguish the countable and uncountable nouns, which causes great confusion to many Chinese EFL learners.

Many researchers such as Farghal and Obiedat (1995) find that many EFL learners have a tendency to use literal translation. For example:

- (1) If I change a new job, I must take [wd3, s-] a long time to adapt [vp1, -s] to a new surrounding. (ST4)
- (2) From [wd7, -1] me, I already [ad1, -1] had taken [cc3, -3] the home teacher, and the second work [wd3, 2-] before this term. (ST4)

In Example (1), the writer intended to convey the meaning “如果我换了新的工作，我需要很长的时间来适应新的环境”。The verb TAKE has the meaning *need/require* which was illustrated in Chapter Three. The writer just literally translated the Chinese into English, ignoring the syntactic restrictions of the verb TAKE in terms of the meaning *need/require* in the English language. The correct form should be “It must take me a long time to adapt to a new surrounding”. The error in this example results from a syntactic transfer from the learner’s mother tongue. As for Example (2), given the context, “I already [ad1, -1] had taken [cc3, -3] the home teacher” can be translated as “我已经做过家庭教师了”，which is also a literal translation. In Chinese, “做过家庭教师” actually means “做过家庭教师的工作”。With the word “工作” omitted from their utterance, it is acceptable. However, it is not the case in English. The correct form should be “I have already taken the job of home teacher”.

4.2.2 Overgeneralization

According to Richard (1971b), overgeneralization refers to errors which occur when the learners output deviant structures based on other structures in the TLs. The learners usually extend the use of a linguistic item or a grammatical rule beyond its acceptable use. The analysis in this study revealed that the Chinese EFL learners tend to create a structure based on the one that they are familiar. For example:

- (1) And the developed contries [fm1, -] took hold of [cc3, 2-3] the world economic [fm1, -], that made the [wd5, 2-3] most developing countries’s [np4, -] people [wd4, s-] in deep trouble.
- (2) The [wd5, 1-1] factories shall take effect [wd3, -] to protect themselves. 采取措施

In Example (1), the writer misused the collocation *took hold of* (in Chinese “抓住”). The collocation *take hold of the world economy* can be translated as “抓住世界经济” which is unacceptable. According to Longman Dictionary, the word *hold* has the meaning *control*, and the whole sentence should be translated as “发达国家控制世界经济给发展中国家的人民带来很大麻烦”。When the word *hold* means *control*, it is usually used in the collocations *get/keep/have a hold of*. It can be assumed that the writer’s output of *take hold of* was actually a deviant structure of the collocation *take control of*. Similarly in Example (2), the writer intended to convey the meaning “采取措施做某事”。The structure *take effect to* is the duplicate of the phases *take measure/step/action to*.

Allerton (1984) points out that the choice of the delexical verbs is quite arbitrary. The choice of the verb *take* in the combination *take a stroll* instead of *make* is not motivated semantically and could be largely language-specific, which creates great problem in SLA. It is unavoidable that many deviant structures of such kind exist in ST4 and ST6. For example:

- (1) The appearance of fake commodities not only influence the trade but also take [cc3, -2] harm to the people. (Correct form: do harm to)
- (2) Doing a job consistantly [fm1, -] help [vp3, 2-] you take further progress [cc3, 2-] in this field. (Correct form: make further progress)
- (3) Something mus [fm1, -] be taken [wd3, s-] to protect and heed the nature [cc3, 2-]. (ST6)

(Correct form: Something must be done to...)

Since the verb TAKE can be used in DVC and collocated with a range of nouns, ignoring the arbitrariness of the word choice or the idiom principle of English language, the writers of these three sentences overgeneralized and analogized the use of DVC, and reproduced the inappropriate collocations.

4.2.3 Communication-Based Errors

Selinker (1972) introduces the term “communication-based errors”. This kind of error arises when the second language learners invoke communicative strategies. He perceives communication strategy as a by-product of the learner’s output with limited target language knowledge. Communication strategies consist of attempts to deal with problems of communication that have arisen in interaction (for example, avoidance, paraphrase, conscious transfer, appeal for assistance and mime—use of nonverbal device to refer to an object or event). According to Gui (2007), when the learners could not find the exact verb in certain collocation, they resort to some polysemantic words, and thus delexical verbs (such as *make, give, take, get, have, and do*) become their first choice. For lack of lexical knowledge of the target language, they tend to use these delexical verbs to substitute the exact verbs and errors arise. For example:

- (1) He should not be haste which [pr1, s-] may take [wd3, -1] waste in opposition [wd3, 1-] (ST4)
- (2) It is also good for the young men themselves to take military service [cc3, 2-]. (ST6)
- (3) Take [wd3, -s] our best to do them well. (ST4) (Correct form: try our best)
- (4) They take [wd3, -3] attention to the relation between nutrition and health. (ST4) (Correct form: pay attention to)
- (5) If the man only takes [wd3, -1] forward to reach the goal, but not do it step by step. (ST4) (Correct form: try our best)

In Example (1), the writer intended to express the meaning “欲速则不达”, which is usually translated as “haste makes waste”. Here the word *make* is not a delexical verb; it has the meaning *produce*. Since the verb *take* does not have this meaning, it cannot be used in this collocation and should be replaced by the verb *make*. In Example (2), the writer used the collocation “take military service” to convey the meaning “服兵役、参军”. Even though the verb *take* is polysemantic, it does not have the meaning “服务、参加” and should be substituted by the word *join*. As for Examples (3), (4) and (5), the writers could not find the exact verbs in the target collocations and simply chose the delexical and polysemantic verb *take* to sustain the communication.

In this part, the main possible causes of the errors committed by the ST4 and ST6 learners are investigated. Some inter- and intralingual errors are identified in ST4 and ST6. The frequency and percentage of each error source in ST4 and ST6 is counted and calculated as shown in Table 4.16.

Table 4. Distribution of Each Error Source in ST4 and ST6

Error sources	ST4		ST6	
	frequency	percentage	frequency	percentage
Interlingual errors (L1 transfer)	10	34.5	2	22.2
Intralingual errors	19	65.5	7	77.8
overgeneralization	11	38	5	55.6
communication strategy	8	27.5	2	22.2

From Table 4, it can be seen that the intralingual errors takes up 65.5% in ST4 and 77.8% in ST6, the interlingual errors only account for 34.5% and 22.2% respectively in ST4 and ST6. That is, the Chinese advanced EFL learners commit much more intralingual errors than interlingual errors, which is in accordance with the findings by Taylor (1975) that the elementary learners tend to produce more transfer errors than the intermediate or advanced learners while conversely, the intermediate or advanced learners tend to produce more intralingual errors than the elementary learners. In addition, comparatively speaking, in the process of SLA the non-English majors are interfered by L1 to a larger extend than the English majors.

It is worth noting that errors can have more than one source. When one researcher identifies certain error source as interlingual, another researcher may identify it as intralingual. This study simply investigated the sources of each error from one perspective, and further studies could explore the sources of the TAKE errors from some other perspectives.

5. Conclusions

5.1 Summary of Major Findings

In terms of the total sum of the errors of the verb TAKE, the non-English majors in ST4 commit much more errors than the English majors in ST6, which indicates that the non-English majors have more difficulty in using this verb.

The errors are analyzed based on the taxonomy of Gui and Yang (2003). Three large categories of errors of the verb TAKE (“vp”, “wd” and “cc”) arise in ST4 and ST6, with most errors are distributed in the level of word and collocation. These errors are subdivided into 8 types (“vp1”, “vp2”, “wd2”, “wd3”, “wd4”, “wd5”, “wd7” and “cc3”). In ST4 the top two recurrent error types are “wd3” (substitution) and “cc3” (“verb + noun” collocation) while in ST6, “wd4” (omission) and “cc3” are the most frequent.

Possible sources of “wd3”, “wd4” and “cc3” errors in ST4 and ST6 are explored from the perspectives of interlingual errors and intralingual errors. The interlingual errors involve errors resulting from the L1 interference while the intralingual errors in this study are mainly caused by overgeneralization and the adoption

of communication strategy (such as avoidance). Besides, the findings show that both the English and non-English majors commit much more intralingual errors than interlingual errors.

5.2 Pedagogical Implications

Based on the above analysis, the author suggests that the three large categories of errors of the verb TAKE, “vp”, “wd” and “cc”, should be paid enough attention when teaching the word TAKE. Besides, Altenberg and Granger (2001) proposes that the concordance-based exercises extracted from native corpora are a useful resource for raising advanced learners’ consciousness of the collocational and structural complexity of the high-frequency verbs. By means of concordancing, the teachers can provide large numbers of authentic examples of a given verb to the students. In the course of their autonomic learning, the students can also refer to the corpus by themselves to get a comprehensive understanding of the language use. With genuine input, the students can produce more pure English expressions.

5.3 Limitations of the Study and Suggestions for Further Study

As for the misuse of the verb TAKE, the analyses in the present study do not cover all the error types and only focus on three error types “wd3”, “wd4” and “cc3” which have high frequencies in both ST4 and ST6. A further analysis of other error types may dig out more interlingual and intralingual error sources, such as lexical transfer and ignorance of rule restrictions.

References

- Allerton, D. J. (1984). Three (or four) levels of co-occurrence restriction. *Lingua*, 63(1), 17-40. [https://doi.org/10.1016/0024-3841\(84\)90030-5](https://doi.org/10.1016/0024-3841(84)90030-5)
- Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22(2), 173-194. <https://doi.org/10.1093/applin/22.2.173>
- Brown, H. D. (1994). *Principles of language learning and teaching*. New Jersey: Prentice Hall Regents.
- Burt, M., & Kiparsky, C. (1972). *The gooficon: A repair manual for English*. Rowley, Mass.: Newbury House.
- Corder, S. P. (1967). The significance of learner’s errors. *International Review of Applied Linguistics in Language Teaching*, 5(1-4), 161-170. <https://doi.org/10.1515/iral.1967.5.1-4.161>
- Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *International Review of Applied Linguistics in Language Teaching*, 9(2), 147-160. <https://doi.org/10.1515/iral.1971.9.2.147>
- Ellis, R. (1986). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Review of Applied Linguistics in Language Teaching*, 33(4), 315-332. <https://doi.org/10.1515/iral.1995.33.4.315>
- Granger, S. (1998). *Learner English on computer*. London: Longman.
- Guobing Liu. (2011). A study on Chinese English learners’ use of the lexical patterns of the delexical verb TAKE. *Contemporary Foreign Language Studies*, (10), 14-19.
- Hao Zhou. (2012). A corpus-based comparative study on the high-frequency verb TAKE used by non-English majors in China. *Journal of Mudanjiang Education Institute*, (1), 41-42. <https://doi.org/10.3969/j.issn.1009-2323.2012.01.022>
- Huizhong Yang. (2002). *Introduction to corpus linguistics*. Shanghai: Shanghai Foreign Language Education Press.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. London: Longman.
- Levenston, E. (1971). Over-indulgence and under-representation: aspects of mother tongue interference. In Nickel (Ed.), *Papers in Contrastive Analysis* (pp. 115-121). Cambridge, Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nemser, W. (1971). Approximative systems of foreign language learners. *International Review of Applied Linguistics*, 9(2), 115-124. <https://doi.org/10.1515/iral.1971.9.2.115>
- Politzer, R., & Ramirez, A. (1973). An error analysis of the spoken English of Mexican-American pupils in a bilingual school and a monolingual school. *Language Learning*, 23(1), 39-61. <https://doi.org/10.1111/j.1467-1770.1973.tb00096.x>

- Richards, J. (1971a). Error analysis and second language strategies. *Language Sciences*, 17, 12-22.
- Selinker, L. (1969). Language transfer. *General Linguistics*, 9, 67-92.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(1-4), 209-231. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Shichun Gui, & Huizhong Yang. (2003). *Chinese Learner English Corpus*. Shanghai: Shanghai Foreign Language Education Press.
- Shichun Gui. (2007). Uncertain judgment and the acquisition of the delexical verbs by Chinese EFL Learners. *Foreign Language Teaching and Research*, 39(1), 3-12. <https://doi.org/10.3969/j.issn.1000-0429.2007.01.001>
- Svartvik, J., & Ekedahl, O. (1995). Verbs in private and public speaking. In B. Aarts & C. Meyer (Eds.), *The Verb in Contemporary English: Theory and Description* (pp. 273-289). Cambridge: Cambridge University Press.
- Taylor, B. (1975). The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL. *Language Learning*, 25(1), 73-107. <https://doi.org/10.1111/j.1467-1770.1975.tb00110.x>
- Yinyan Dong. (2015). A corpus-based study on the use of non-lexicalized verbs by Chinese English learners: A case study of the high-frequency word TAKE. *Journal of Changchun Institute of Education*, 31(20), 109-110. <https://doi.org/10.3969/j.issn.1671-6531.2015.20.047>
- Yue Li. (2011). The use of TAKE's verb-noun collocations by Chinese learners of English. *Education Science & Culture Magazine*, (20), 145-147. <https://doi.org/10.3969/j.issn.1672-7894.2011.20.089>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).