CHAPTER FOUR: DATA ANALYSIS AND RESULTS

This chapter provides an analysis of the MWU errors in the corpus. In the first section, the overall results of the analysis will be presented. In the second section, a more detailed analysis of each category of the MWU errors will reveal the different types of deviation. At the end of this chapter, a summary of the findings will have answers to the research questions, posed at the beginning of this chapter.

4.1 MWU Errors in the Corpus

The corpus reveals four most prominent features of MWU errors which have been tagged as <MD> (modal structures), <IN> (infinitive structures), <JN> ('adjective + noun' collocations) and <CN> (connectors). The *WordSmith Tools* (WST) is used to generate the concordance lines for each MWU error which has been tagged. Table 4.1 below shows the erroneous MWU by categories, together with its respective frequency count, in numbers and percentage.

Table 4.1

Different categories of erroneous MWU structures

Types of MWU errors		Frequency Count	Percentage in the learner corpus
Grammatical Collocations	<MD> modal structures	190	46 %
	<IN> infinitive structures	61	15 %
Lexical Collocations	<JN> 'adj + n' collocations	90	22 %
	<CN> connectors	74	17 %
		Total = 415	Total = 100%

In this 40,000-word learner corpus, which was compiled from 90 essays, a total of 415 MWU errors were tagged. The errors are grouped into 'grammatical collocations' and 'lexical collocations'. In the 'grammatical collocations' category,

there are 190 <MD> errors and 61 <IN> errors. In the 'lexical collocations' category, there are 90 <JN> errors and 74 <CN> errors. The first column on the right in Table 4.1 also shows that 61% of the MWU errors in the corpus are ungrammatical collocations whilst the other 39% are erroneous lexical collocations.

These four categories of MWU errors are prevalent in this learner corpus as they occur consistently across the data. The concordance programme used, the WST, is able to generate the tagged MWU errors in concordance lines, in line with the source, which is the 'File' column. As shown below, the file numbers (e.g. *fs27*, *fs72* and so on); indicate that the MWU errors are not from one source. As exemplified below, it is evident that the MWU error 'beside that' occurs across the learner corpus. This evidence reveals the need to investigate further the occurrence of such MWU errors in this learner corpus.

N	Concordance	Set Tag Word No.	File
14	tolerance/to be tolerant@</IN> to each other. <CN>@Beside that/Besides that@</CN>, in a larg	319	1\fs27.txt
15	nguage. This is the benefic of earning english. <CN>@Beside that/Besides that@</CN>, learning	102	1\fs72.txt
16	> to know what information the lectural is giving out. <CN>@Beside that/Besides that@</CN> student	262	1\fs73.txt
17	on't need to be scared of getting lost or ask for help. <CN>@Beside that/Besides that@</CN>, English	77	1\fs67.txt
18	nicate with each other and get a better relationship. <CN>@Beside that/Besides that@</CN>, we can	208	1\fs79.txt
19	ng the family members if they are not tolerate. <CN>@Beside that/Besides that@</CN>, staying i	430	1\fs22.txt
20	success/to succeed@</IN> in the after days. <CN>@Beside that/Besides that@</CN>, a perso	212	1\fs16.txt
21	nt language and also at a good standard of English. <CN>@Beside those/In addition@</CN>, we can a	326	1\fs67.txt
22	ncourage to learn english when we were small. <CN>@Beside that/Besides that@</CN>, learning	145	1\fs68.txt
23	time, the family elationship will become worst. <CN>@Beside that/Besides that@</CN>, nowaday	404	1\fs52.txt
24	e problem that we face s not problem already. <CN>@Beside that/Besides that@</CN>, we will b	149	1\fs51.txt
25	ne-class" which provide by the lecturer in University.<CN>@Beside that/Besides that@</CN>, governm	371	1\fs20.txt

For a more systematic analysis, the following section will present the analysis of the errors according to the categories, as mentioned above. The concordance lines (hereafter CL), presented in the respective sub-sections, are examples used for

illustration purposes. For the complete output of CL for each category of MWU errors, refer to Appendix 5, 6, 7 and 8.

4.2   Analysis of Tagged MWU Errors

The analysis of the tagged MWU errors will describe the salient patterns identified in each of the MWU structure: <MD>, <IN>, <JN> and <CN>.

4.2.1   Modal structures <MD>

Modal auxiliaries are among the more difficult structures ESL/EFL teachers have to deal with because of the form of modals (Celce-Murcia & Larsen-Freeman, 1999). Some of the problems highlighted by them are: (1) the inflection of modals, and (2) the intervention of the infinitive 'to' between a modal and a verb. In problem (1), students have been told time and time again that present-tense verbs with third person singular subjects require an –s ending, however, they overgeneralise this rule to modals and produce structures such as [1]*'He <u>cans</u> play tennis'.

As for problem (2), students treat modals like other ordinary verb which requires an infinitive 'to' when another verb follows in sequence. For example, in this sentence, *'I want <u>to</u> go'*, the infinitive 'to' is used between the verb 'want' and 'go' and students who apply the same rules to the modal structures will produce an erroneous structure in the sentence – *'I can <u>to</u> go'*.

In this study, the analysis of the erroneous <MD> structures reveals that the learners have different problems with the modal auxiliaries. A common pattern can be

---

[1] This symbol (*) is used to mark incorrect sentences.

identified in the errors made by the learners and they are categorized into three groups: 'overgeneralisation of verb form', 'misformation of passive forms' and 'misformation of other modal structures' (refer to Table 4.2).

Table 4.2

Different categories of erroneous <MD> structures

Types of <MD> errors	Frequency Count	Frequency Count (%)	Percentage in the learner corpus (%)
Overgeneralisation of verb form	113	59	27
Misformation of passive forms	30	16	8
Misformation of other modal structures	47	25	11
Total	190	100%	45%

4.2.1.1 Overgeneralisation of verb form

Figure 4.1 shows the first category of <MD> errors – overgeneralization of verb form. Learners produce errors which deviate from the 'modal + base verb' structure by inflecting the verb with the –s inflection, past tense –ed or the –ing participle.

CL

1 for students so that they <MD>@can continues/can continue@</MD> their study. The minister

2   ant in a student's life. It <MD>@may affects/may affect@</MD> a student's future and determi

3   usion, a good education <MD>@will promotes/will promote@</MD> success. Being uneducated,

4 like each other. Quarrel <MD>@can caused/can cause@</MD> a lot of trouble. Moreover, when

5 ber of a large family, we <MD>@will learnt/will learn@</MD> how to tolerate how to cooperate

6   responsible teacher, he <MD>@must given/must give@</MD> his students a good education.

7 this is because when you <MD>@cannot done/cannot do@</MD> your work in time you can call

8 e in the house. Everyone <MD>@can helping/can help@</MD> you when you facing the proble

9   e very useful for us. We <MD>@must learning/must learn@</MD> English because the benefits

10 problem, the older ones <MD>@will giving/will give@</MD> you good opinions or good advices

Figure 4.1

Examples of <MD> errors: Overgeneralization of verb form

CL-1, CL-2 and CL-3 show how learners inflect the verb 'continue', 'affect' and 'promote' with the –s inflection while CL-8, CL-9 and CL-10 show how learners inflect the verb 'help', 'learn' and 'give' with the –ing participle. In CL-4, the use of past tense –ed of the verb 'cause' and the use of past tense in irregular verbs such as 'learnt' (CL-5) are also tagged as <MD> errors. In CL-6 and CL-7, the verbs used, 'given' and 'done', are in the past participle of the verb form. The tagged errors in this category of <MD> errors show that overgeneralising the verb form is a problem for learners in this corpus. Lack of awareness of the most basic 'modal + base verb' structure can be used as a reason to justify for the high frequency of errors in this category. We shall discuss the causes of such errors in greater detail in Chapter 5.

4.2.1.2 Misformation of passive forms

Another significant <MD> error is the misformation of passive forms, as exemplified in Figure 4.2.

CL

11 n cocuriculum after school. It <MD>@can be divide/can be divided@</MD> into three main ty

12 hat is correct and what things <MD>@cannot be do/cannot be done@</MD>. As we can see that

13   The children in a big family <MD>@would not be bully/would not be bullied@</MD> by othe

14 to succeed@</IN>. Education <MD>@can divided/can be divided@</MD> into two categories.

15   can get more knowledge that <MD>@can only found/can only be found@</MD> in english wri

16 arge groups. The atmosphere <MD>@will filled/will be filled@</MD with enthusiasm when a la

---

Figure 4.2

Examples of <MD> errors: Misformation of passive forms

The two most common types of deviation are the omission of the auxiliary 'be' and the wrong use of verb form. The various forms of deviation from the 'modal + be + verb (past participle)' are exemplified in the concordance lines above (refer to

Figure 4.2). Errors in CL-11, CL-12 and CL-13 show that learners do not use the verb in its past participle form. The verbs – 'divide', 'do' and 'bully' must be used as 'divided', 'done', and 'bullied' in order to form the passive <MD> structure. The tagged errors in CL-14, CL-15 and CL-16 show that learners have excluded the auxiliary 'be' and produce errors such as 'can divided', 'can only found' and 'will filled'.

The findings in this study further support what was found by Norman (1988: 165) who observed that active-passive sentence relationships cannot always be determined because grammatical voice is absent in Chinese (in Hinkel, 2002b). Therefore, as Hinkel (2004) suggests, Chinese learners need to learn the passive constructions of formal written discourse intensively. In the case of the Chinese learners in this study, they need to be taught overtly, the passive modal structures as: 'modal + be + verb (past participle)'.

4.2.1.3 Misformation of other modal structures

The third category of <MD> error is the misformation of other modal structures (refer to Figure 4.3). Throughout the process of analyzing the <MD> errors, it is found that the awareness of certain misformed structures are important and should be mentioned. The errors which have been tagged in this category can be grouped again into three different sub-categories according to the nature of the errors.

Errors in CL-17 to CL-21 can be termed as 'blend error'. 'Blends' is the fifth category of 'Target Modification Taxonomy', which was not considered by Dulay, Burt and Krashen (1982) in their second type of descriptive taxonomy – 'Surface Structure Taxonomy' (James, 1998). Both the taxonomies are specifications to

describe the errors. James (1998) added the 'blend' category to complement the Target Modification Taxonomy. When a learner is undecided about which of the two target languages he has in mind, the type of error that materializes is the blend error. Blending is exemplified in 'according to Erica's opinion', which arises when these two alternative grammatical forms – 'according to Erica' and 'in Erica's opinion' – are combined to produce an ungrammatical blend.

Blended modal structures
CL
17   chemists and other scientists as we <MD>@could able/are able@</MD> to communicate with
18   @</MD> us in the future and we <MD>@may able/are able@</MD> to survive in any nation
19   te are written in english. Readers <MD>@must able/must be able@</MD> to understand
20   te our project when working. We <MD>@will able/will be able@</MD> to give more ideas
21   r grandparent by bluffing. These <MD>@will going to make/will make@</MD> other membe
Misformation of 'modal + adjective' structures
22   he family. In a huge family, some <MD>@may jelous/may be jealous@</MD> <IN>@to the
23   ment for their vacancy that is you <MD>@must good/must be good@</MD> in English. Good
24   together, work together, your life <MD>@will happy/will be happy@</MD everyday. It that
25   as before. Although more people <MD>@will better/will be better@</MD> but being a famil
Misordering of adverb
26   international walls. Besides, you <MD>@also can learn/can also learn@</MD> about the
27   country. The quality of a country <MD>@also will improve/will also improve@</MD> if a lot
28   without speaking Japanese, you <MD>@still can use/can still use@</MD> english to get the

Figure 4.3

Examples of <MD> errors: Misformation of other modal structures

The errors exemplified in CL-17 to CL-21 can be termed as 'blended modal structures'. The blending here involves a modal and phrasal modal. According to Celce-Murcia and Larsen-Freeman (1999: 139), multi-word forms ending in infinitive –to, which function semantically like true modals (in certain of their meanings), are called phrasal modals (they are also called periphrastic modals, pseudo modals, or

quasi-modals). Every modal seems to have at least one phrasal counterpart, and some modals have several:

Modal	Phrasal Modal
can, could	be able to
will, shall	be going to, be about to
must	have to, have got to
should, ought to	be to, be supposed to
would	used to
may, might	be allowed to, be permitted to

CL-17 shows that the modal 'could' has been blended with its phrasal modal counterpart to form 'could able'. Similarly, in CL-21, the modal 'will' with its own counterpart 'going to' forms a blend error of 'will going to make'. However, in CL-18, the modal 'may' is blended, not with its own counterpart, but a phrasal counterpart of another modal. This 'cross-blending error' is also exemplified in CL-19 and CL-20 where the modal 'must' and 'will' are blended with 'able'.

Referring to the suggested correction of error, it is interesting to find that 'able' can collocate with all the modals, but functioning as an adjective in this structure – 'must + be + able' (CL-19) or 'will + be + able' (CL-20). These two structures are the suggested corrections for the learner errors and they should be used based on its context to convey the intended meaning. In CL-19 and CL-20, the use of modal 'must' and 'will' in these sentences, express a strong degree of ability and thus, reinforces the voice and opinion of the writer.

Even though the frequency of blended <MD> errors is low, it is important to highlight the wrong combinations of modal and phrasal modal; and determine the possible combinations in standard English. Celce-Murcia and Larsen-Freeman (1999)

has identified this as a problem and as shown in the data, it remains a problem for the learners.

The second sub-category is misformation of 'modal + adjective' structures. Examples in CL-22 to CL-25 (refer to Figure 4.3) show that it is an error to use an adjective right after a modal. This error can be easily rectified by inserting the infinitive of the verb 'be' in between the modal and adjective. 'jealous', 'good', 'happy' and 'better' are all adjectives and they can form grammatical collocations with modals according to this frame of structure – 'modal + be + adjective'. Even though the data does not show high frequency of such errors, it is important to include this as a teaching point. This can help learners to reduce errors in their writings when they are aware of the various grammatical collocations for modals.

The third sub-category is misordering of adverb in <MD> structures. This is probably a typical problem for Chinese learners because in Mandarin, the use of adverb is before the modal and verb. CL-26, CL-27 and CL-28 exemplify how learners use the adverbs – 'also' and 'still' before the modal and verb, which is influenced by their first language. In Mandarin, '*ye ker yi xue*' (also can learn), '*ye yi ting hui jin bu*' (also will improve) and '*hai ker yi yung*' (still can use) are grammatically correct structures, however, the equivalent structures in English are considered grammatically wrong.

In standard English, the 'modal + adverb + verb' is the right grammatical collocation for this <MD> structure and a cross reference with the [2]LOB (Lancaster-Oslo-Bergen) corpus on the Web Concordance at

---

[2] The LOB Corpus has been mentioned in section 2.2.1. It is used as a reference corpus in this study.

confirms this. The examples below show the accurate use of the modal structures in the LOB corpus.

```
231   Granby Corner, Leicester. You can also buy it in  beechwood, and
798  es. A weekly  steamer service will also take you over to Turkey,
200  nd this, to a large extent, he can still do on the Solent  althou
610    Then the match went out. I could still see the small metal box
```

4.2.2   Infinitive Structures <IN>

The erroneous <IN> structures in this learner corpus can be categorized into: 'misselection of parts-of-speech (POS)', 'overgeneralisation of verb form' and 'misformation of infinitive structures' (refer to Table 4.3).

Table 4.3

Different categories of erroneous <IN> structures

Types of <IN> errors	Frequency Count	Frequency Count (%)	Percentage in the learner corpus (%)
Misselection of POS	30	59	27
Overgeneralisation of verb form	21	16	8
Misformation of infinitive structures	10	25	11
Total	61	100%	15%

4.2.2.1 Misselection of parts-of-speech

Referring to section 3.3.2 on the infinitive structure, *the full infinitive* (Thomson and Martinet, 1986), which is also called *the simple infinitive* (Larsen-freeman, 2000), is described as '*to* + verb'. '*It would take months to travel to Mars'* is an example of how the infinitive is used in a sentence. The data shows a high frequency count of the 'to + success' structure, whereby 'to' is described as a preposition, when it is followed by a noun. Therefore, 'to success' is tagged as an

infinitive error. There are 30 occurrences of such erroneous structure and the errors are caused by the misselection of word class.

Figure 4.4 shows the <IN> structures which have been tagged as errors because a noun is used after 'to' when it is supposed to be an infinitive 'to + verb'. Examples in CL-1, 2, and 3, show that the noun 'success' is used instead of the verb 'succeed'. CL-4 shows that the misselection of the word 'emphasis' (noun) is considered as an erroneous infinitive structure used. The verb 'emphasise' is the right parts-of-speech (POS) to be used in the infinitive structure.

CL

1 n may get a stable job and easier <IN>@to success/to succeed@</IN> in life with stable job

2 ion and it is the key for a person <IN>@to success/to succeed@</IN> in life and have a bright

3 is very important for everyone <IN>@to success/to succeed@</IN> in the future. We start

4 hamuddin encourages students <IN>@to emphasis/to emphasise</IN> on learning of English.

Figure 4.4

Examples of <IN> errors: Misselection of POS

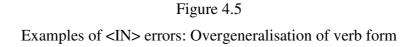*4.2.2.2 Overgeneralisation of verb form*

The second category of <IN> error is the use of wrong verb form, which is a deviation from the 'to + base verb' structure (Figure 4.5).

CL

5 things and a foresight which able <IN>@to helps/to help@</IN> them look into his future. This

6 ily. Everyone has their own plan <IN>@to built/to build@</IN> their family, is either large or s

7 ple, a chef also need to learn how <IN>@to cooking/to cook@</IN>. If compare a chef which ha

8 e lighter. Besides that, it is good <IN>@to having/to have@</IN> a large family. Well, the child

---

Figure 4.5

Examples of <IN> errors: Overgeneralisation of verb form

Instead of the base verb, verbs are used with an overinclusion of the –s inflection. Verbs are also used in the past tense form, as well as with the –ing participle; after the 'to' infinitive. These erroneous structures are exemplified in CL-5 (*to helps*), CL-6 (*to **built***), CL-7 (*to cook**ing***) and CL-8 (*to hav**ing***) (refer to Figure 4.5).

4.2.2.3 Misformation of infinitive structures

In the third category of 'misformation of infinitive structures', there are two structures which should be identified and the differences between them should be discussed. These two structures are: 1) the passive infinitive (e.g *The suggestion to be seen by a surgeon was never followed.*) and 2) the simple infinitive before an adjective (e.g. *As a baby, Einstein appeared to be retarded.*).

In Figure 4.6, CL-9 to CL-11 are examples of misformation of passive infinitive structures whereas CL-12 to CL-14 are examples of misformation of simple infinitive before an adjective.

Misformation of passive infinitive structures
CL
9   rent. Children are more likely <IN>@to be educate/to be educated@</IN> till form three, some
10   can bring many benefits to us <IN>@to be use/to be used@</IN> in different aspects. Student
11   ther if an agreement is unable <IN>@to reached/to be reached@</IN>. Besides that, members
Misformation of simple infinitive before adjective
12 t their child. Spend more time <IN>@to close/to be close@</IN> with them, teach them and
13   ple. In large family, you have <IN>@to tolerant/to be tolerant@</IN> with your siblings or
14 t's package, so you don't need <IN>@to scared/to be scared@</IN> of being cheated or get your

Figure 4.6

Examples of <IN> errors: Misformation of infinitive structures

Tagged errors in CL-9 to CL-11 (refer to Figure 4.6) show that learners have problems with the passive form of the <IN> structures. The passive infinitive is explained as a formation of 'to' + 'be' + 'past participle'. In using the passive infinitive structure, learners tend to forget that the lexical verb is in the past participle form, for example 'educated' (CL-9) and 'used' (CL-10). There is also a tendency that learners forget the 'be' (the infinitive of the verb 'BE') in the passive infinitive structure (CL-11).

CL-12 to CL-14 (refer to Figure 4.6) are examples of <IN> errors before an adjective. The construction of 'to' + 'be' which forms the simple infinitive, should be used before an adjective to form the right collocation. Even though there are only 10 occurrences of such errors in this category, it is pertinent to point out the differences of these two structures in order to teach the learners to form correct grammatical collocations in their writings. This <IN> structure should be differentiated from the passive form as students who do not know the POS of the word, may be confused with the different <IN> structures.

Furthermore, as pointed out earlier in section 4.2.1.2, the Chinese language does not explicitly distinguish between the active and passive constructions. In fact, as explained by Li and Thompson (1981), only the passive particle 'bei' 被 is predominantly used to indicate the few passive constructions in Chinese (in Hinkel, 2002b: 11). Therefore, learners need to learn the complicated functions of the 'be' verb not only in 'passive modal' structures, but also in 'passive infinitive' structures and 'simple infinitive before adjectives' as the 'be' constructions (in the TL) is much more diverse compared to the bei-constructions in Chinese.

4.2.3 'Adjective + Noun' Collocations <JN>

There are 90 occurrences of adjective-noun collocation which have been tagged as erroneous <JN> structure. An analysis of the tagged errors reveals some interesting findings of this lexical collocation in this learner corpus. First of all, all the <JN> errors can be categorized into different groups based on the type of errors. The <JN> errors can be broadly grouped into 'inappropriate collocations', 'misselection of adjectives', and 'misformation of compound adjectives'. Table 4.4 indicates the different categories of adjective-noun collocation errors and the frequency count for each group.

Table 4.4

Different categories of adjective-noun collocation errors

Types of <JN> errors	Frequency Count	Frequency Count (%)	Percentage in the learner corpus
Inappropriate collocations	41	46	10
Misselection of adjectives	34	38	8
Misformation of compound adjectives	15	16	4
Total	90	100%	22%

A certain pattern can be detected among the tagged <JN> errors and the errors can be grouped according to the different types of error production. Each category will be explained further, in turns.

4.2.3.1 Inappropriate collocations

There are 41 occurrences of inappropriate <JN> collocations and some of these tagged errors are exemplified in Figure 4.7. These erroneous <JN> structure does not provide an accurate meaning to the context, even though it is still possible to

guess the intended meaning of the attempted collocation. The tagged errors displayed below are considered as collocation errors because such structures are not used in standard written English.

CL

1 Learning of english become a <JN>@fixed subject/compulsory subject@</JN> for him if he wis

2 coming examination. With a <JN>@good control/good command@</JN> of English, we do not

3 ll get to higher scholarship. A <JN>@high scholarship/notable scholarship@</JN> will gurantee

4 ember of a large family. The <JN>@maiden advantage/main advantage@</JN> is where you a

5 ave learned English since the <JN>@primary study/primary education@</JN> because learnin

6 th the parents did not have a <JN>@stable profit/stable income@</JN>, the family members ar

7 ng will also gone if we have a <JN>@well education/good education@</JN>. Besides that, havin

---

Figure 4.7

Examples of <JN> errors: Inappropriate adjective-noun collocations

For example, 'compulsory subject' (CL-1) is the more accurate <JN> collocation instead of 'fixed subject' when describing a subject in school which students have to learn. It is interesting that Lu (2002) in her research also finds that her research subjects use the collocation 'fixed subject' to mean 'compulsory subject'. L1 influence is obvious here because '固定*gu ding* 科目*ker mu*' literally means '<u>fixed subject</u>' in English. Similarly, 'good command' of a certain language is a more accurate collocation to use, compared to 'good control' (CL-2). Other examples of such collocation errors are shown in CL-3 to CL-7. These <JN> collocations – 'high scholarship', 'maiden advantage', 'primary study', 'stable profit' and 'well education', are considered as L2 collocation errors.

4.2.3.2 Misselection of adjectives

The second category of <JN> collocation errors is 'Misselection of adjectives'. There is a total of 34 <JN> errors grouped as 'Misselection of adjectives'. A few examples of such collocation errors are shown in Figure 4.8. The tagged errors in Figure 4.8 exemplify the misselection of the lexical item which is of a different word class. The examples in CL-8 to CL-17 show that learners who are not aware of the different POS of the word, will tend to misselect the lexical item of a different word class instead of an adjective.

> **CL**
>
> **8**   the level is not high, he might lose his <JN>@desire job/desired job@</JN>. In addition, know
>
> **9**   nglish also used in a business between <JN>@difference country/different country@</JN> that
>
> **10**   rists come visits, our country will face <JN>@economy problems/economic problems@</JN>.
>
> **11**   ssful@</IN>. In this view english is an <JN>@importance tool/important tool@</JN> <IN>@to
>
> **12**   need some peace. When you need a <JN>@peace place/peaceful place@</JN>. Suddenly
>
> **13**   arder/may be harder@</MD> to get a <JN>@silence moment/silent moment@</JN>, a peace
>
> **14**   eadership skill, speaking skill or even <JN>@socialize skills/social skills@</JN> in their school
>
> **15**   yone in this world want to become a <JN>@success people/successful person@</JN>. But
>
> **16**   at having a good education will have a <JN>@brightly future/bright future@</JN>. <CN>@
>
> **17** h good education, a person may earn a <JN>@highly income/high income@</JN>. Therefore, a

Figure 4.8

Examples of <JN> errors: Misselection of adjectives

Table 4.5 identifies the words with morphological similarities which cause confusion to learners. These words are found to be problematic for the learners in this corpus. They tend to cause confusion to learners because of their similarities, in terms of morphology. For example, in CL-8 (refer to Figure 4.8), 'desire job' is a wrong <JN> collocation. As shown in Table 4.5, 'desired' is the adjective which collocates with the noun 'job'.

Table 4.5

Examples of words with morphological similarities

Noun	Adjective	Verb	Adverb
desire	desired		desirably
difference	different		
economy	economic		
importance	important		
peace	peaceful		
silence	silent		
	sociable (person) social	socialise	socially
success	successful	succeed	successfully
	bright		brightly
	high		highly

An interesting finding from the tagged errors is that learners tend to misselect the lexical item in its noun word class – 'difference country' (CL-9), 'economy problems' (CL-10), 'importance tool' (CL-11), 'peace place' (CL-12), 'silence moment' (CL-13), and 'success people' (CL-15). In CL-14, 'socialise' (a verb) has been misselected instead of the adjective 'social'. In CL-16 and CL-17, the adjectives 'bright' and 'high' have been wrongly substituted with the adverbs 'brightly' and 'highly'.

These erroneous structures occur because other parts-of-speech (of the same word) have been wrongly used to form the intended 'adjective + noun' collocation. The wrong choice of adjective is caused by POS (parts-of-speech) confusion. From

the selected concordance lines in Figure 4.8, it is evident that learners have problems identifying the word in its right word class, which is an adjective, to form the right collocation with the noun.

Lexical items and their word class variation can cause confusion to learners who are unaware of the right collocation of words. Therefore, it is important to teach collocations in vocabulary lessons to expose learners to the right collocations of words.

4.2.3.3 Misformation of compound adjectives

Another interesting finding from the analysis of the <JN> structure is the inability to form compound adjectives, as shown in the tagged errors in Figure 4.9.

CL

18 key to success. A <JN>@good educated people/well-educated person@</JN> will also have more

19 an everything. A <JN>@good education people/well-educated person@</JN> will always get firs

20 This is because a <JN>@high educated person/highly-educated person@</JN> knows what is th

21 They deserves a <JN>@high-paid job/highly-paid job@</JN> or even go to the professional. Th

22 h. Actually it is a <JN>@short sighness/short-sighted view@</JN>. Learning English is very imp

---

Figure 4.9

Examples of <JN> errors: Misformation of compound adjectives

Referring to Figure 4.9, the examples in CL-18 and CL-19 show that learners have difficulties in describing a person who has a good education using the 'adjective + noun' structure. It is understood by learners that 'good' is an adjective, but it does not collocate with 'education' and 'people/person' to form the 'adjective + noun' structure. Instead, 'well', which is an adverb, is used to form a compound adjective 'well-educated' to describe the noun 'people/person'.

65

Other such errors in this category can be detected in CL-20, CL-21 and CL-22. In CL-20, 'highly' (adverb) and 'educated' (adjective), forms a compound adjective 'highly-educated' to collocate with the noun 'person'. Similarly, 'highly' and 'paid' (verb) forms 'highly-paid', a compound adjective to collocate with the noun 'job'. In CL-22, the learner most probably intended to write 'short-sighted', but what was produced is an error – 'short sighness'. In this case, the learner not only failed to form a compound adjective but he was also unaware that 'short-sighted' pairing with 'view' is a common collocation in this context. In total, there are 15 occurrences of such errors.

4.2.4   Connectors <CN>

A great deal of attention began to be paid to the need for connectors as a means of achieving cohesion in ESL writing, partially as a result of the influence of Halliday and Hasan's work on cohesion in written text (Celce-Murcia & Larsen-Freeman, 1999). In the ESL composition literature, some of the problems related to the use of connectors are: (1) problems with forms, (2) problems in choosing between connectors with similar meanings, (3) punctuation problems, and (4) overuse of certain connectors (Celce-Murcia & Larsen-Freeman, 1999).

The analysis of the tagged <CN> structures in this corpus revealed that the learners in this study also have similar problems as those mentioned above but the analysis of the learner errors in this corpus reveals other errors which are unique in this learner corpus. A total of 74 <CN> structures have been tagged as errors. The errors are classified into: 'distortion in the <CN> structure', 'misselection of <CN> structure', and 'misformation of <CN> structure' (refer to Table 4.6 below).

Table 4.6

Classification of <CN> errors

Types of <CN> errors	Frequency Count	Frequency Count (%)	Percentage in the learner corpus (%)
Distortion in the <CN> structure	35	46	8
Misselection of <CN> structure	26	35	6
Misformation of <CN> structure	13	18	3
Total	74	100%	22%

In classifying lexical errors, James (1998: 145-151) divided formal errors of lexis into these categories: formal misselection, misformations and distortions. His scope of discussion goes deeper into these three broad categories, but the explanation is limited to word(s) level only. The discussion on the classification of <CN> errors in this study will be based on the main categories suggested by James (1998) but based on MWU errors instead of word(s). The tagged <CN> errors in this learner corpus are divided into these three groups: 'distortion of structure', 'inappropriate use of connectors' and 'misformation of structure'. Each will be discussed further below.

4.2.4.1 Distortion in the <CN> structure

The first category is 'distortion in the structure'. 'Distortions', as explained by James (1998: 150), are "the intralingual errors of form created without recourse to L1 resources", "forms non-existent in the TL(target language)", and the errors are "the result from the misapplication of one or more of the processing operations: omission, overinclusion, misselection, misordering, and blending (of the Target Modification Taxonomy)". Even though this is one of the categories used by James to classify lexical errors, 'distortions' can also be used to categorise this group of tagged <CN>

errors. These <CN> errors have the "characteristics of distortions", as described by James. The <CN> errors are intralingual errors – 'altered' structures – recognizable as a structure in the TL (intended to be used as such in the TL), yet, they are non-existent structures in the TL. Most importantly, the errors can be analysed and explained using one of the processing operations in the Target Modification Taxonomy.

Figure 4.10 shows that there are four different types of distortion in these <CN> errors.

Omission of the –s inflection
CL   **1 face is not problem already. <CN>@Beside that/Besides that@</CN>, we will be able to feel**   **2   as a high english education. <CN>@In other word/In other words@</CN>, learning English is**
Overinclusion of the –s inflection
3 is an international language. <CN>@For examples/For example@</CN>, it is used during politici   **4 ng with their sibling and etc. <CN>@On the other hands/On the other hand@</CN>, they maybe**
Overinclusion of articles
5   try also try to learn English. <CN>@For an example/For example@</CN>, China, is the populou   **6   cial from the other country. <CN>@For the example/For example@</CN>, you may enjoy the E**   **7   emerge to get the attention. <CN>@In a conclusion/In conclusion@</CN>, being a member in a**
Misuse of prepositions/articles
8 future/bright future@</JN>. <CN>@An addition/In addition@</CN>, having a good education   **9   can have a good application. <CN>@As conclusion/In conclusion@</CN>, learning english is**   **10 s life/successful life@</JN>. <CN>@As example/For example@</CN>, many of the child in the**   **11   ts to take care each of them. <CN>@In the other hand/On the other hand@</CN>, the children**

Figure 4.10

Examples of <CN> errors: Distortion in the structure

The first one is the omission of the –s inflection in these connectors: 'beside**s** that' and 'in other word**s**' (refer to CL-1 and CL-2). It is interesting to find that the

omission of the –s inflection in 'besides that' is very common in the learners' writing, even though it is a very basic addition connector.

The second type of distortion is the overinclusion of the –s inflection in the <CN> structure. CL-3 and CL-4 exemplifies how the overinclusion of –s inflection produce errors such as these: 'for example**s**' and 'on the other hand**s**'. In the third category of distortion, the examples in CL-5, CL-6 and CL-7 show how articles are being overincluded into the linking devices. Inserting 'an' and 'the' into these structures: 'For example' and 'In conclusion', causes distortions such as 'For **an** example' (CL-5), 'For **the** example' (CL-6) and 'In **a** conclusion' (CL-7).

The fourth and final type of distortion is not listed in the Target Modification Taxonomy, but one which I have created based on a common pattern identified in the errors – 'misuse of prepositions/articles'. From the examples in CL-8, CL-9, CL-10, and CL-11, the tagged errors reveal how the misuse of prepositions/articles distorts the structure of the connectors: 'an' is used instead of 'in' for the addition connector – '*in addition*' (CL-8), 'as' instead of 'in' for the conjunctive '*in conclusion*' (CL-9), 'as' instead of 'for' in '*for example*' (CL-10) and 'in' instead of 'on' in '*on the other hand*' (CL-11).

These examples of tagged <CN> errors are solid pieces of evidence which prove that second language learners do have form problems when using connectors in their writing, as mentioned by Celce-Murcia and Larsen-Freeman (1999). A more significant finding from the analysis is how and why learners distort the structures, and the impact of this finding on the pedagogical implications, which will be discussed in Chapter 5.

4.2.4.2 Misselection of <CN> structures

In the second category, there are 26 instances of 'misselection of <CN> structures' in the learners' writing. 'Formal misselection' is the first classification of lexical errors used by James (1998) to categorise errors of the malapropism type – pairs (or triples) of words that look and sound similar (e.g. parricide/patricide; accessory, accessary). Figure 4.11 shows that the tagged <CN> errors are also of malapropism type. The erroneous <CN> structures may or may not be similar-looking or similar-sounding as the TL, but the intended meaning and function are similar to the TL.

For example, in CL-12, the learner probably intended to use a clarifying connector ('in other words') to signal to the reader that the following sentence will clarify what has been said in the previous sentence, instead of 'in one word', which is not a connector in the TL. Errors such as these occur when learners think they know the accurate structure or assume that they are using the right structure.

---

12 mother, father, brother and sister. <CN>@In one word/In other words@</CN>, older family

13 help friends which are in trouble. <CN>@In opposite ways/On the other hand@</CN>, frien

14   things that happened around us? <CN>@In additional that/In addition@</CN>, nowadays

15   communicate with the properly. <CN>@If like this/As a result@</CN>, they will leave and

16   important and useful for our life. <CN>@In the other way of saying/In other words@</CN>,

17   family member of a large family. <CN>@Like example/For example@</CN>, we cannot

18 children will be the ones suffering. <CN>@Long story short/In a nutshell@</CN>, being a

19well-educated professions are born. <CN>@Same as the world/In the same way@</CN>, the

---

Figure 4.11

Examples of <CN> errors: Misselection of structure

Other examples in this category are 'in opposite ways' (CL-13) and 'in additional that' (CL-14). These 'misselected structures' are probably substitutes for 'on the other hand' and 'in addition'. This is probably also a 'communication strategy' – where learners avoid the concept or alternatively resorting to paraphrase or language switch when they do not know the right structures in the TL (James, 1998).

The tagged errors in CL-15 to CL-19 show that learners do not know the exact structures in the TL, and resort to translating from their L1. However, in Mandarin, linking devices are not fixed phrases, unlike the English language. In fact, sentence connectors in Mandarin do not exist as MWUs. Therefore, it is interesting to see how learners form the substitute expressions which are typically used in L1 speech. For example, 'If like this' in CL-15, is used as a substitute for 'As a result', a connector to express effects/results. In Mandarin, '*ru guo zhe yang*' (if like this) is typically used in informal speech only, even though it does resemble the meaning and function in the TL.

Other examples of such 'misselected <CN> structures' are shown in CL-16 to CL-19. In CL-16, 'in the other way of saying' is a direct translation from the L1 – '*ling yi ju hua suo*' to substitute 'in other words' and 'like example' – '*bi ru*' (in Mandarin) – is used as a substitute for the exemplifying connector, 'for example' (CL-17). 'long story short' has been directly translated from '*chang hua duan suo*' (CL-18) and 'same as the world' is probably translated from '*tong zai shi jie*' to mean 'in the same way' in the TL (CL-19).

The tagged errors in this learner corpus exemplify this finding by Gilquin et al. (2007: 328) – "Particularly striking is learners' tendency to use expressions, which are

more typical of speech than of writing". As a result, the learner language produced is what Granger refers to as 'foreign-sounding' (in James, 1998: 125) or 'unidiomatic expressions' (Gilquin et al., 2007).

4.2.4.3 Misformation of the <CN> structure

The third category of classification is 'misformation of the <CN> structure' (Figure 4.12). Even though there are only 13 occurrences of such errors, it is pertinent to mention how the learners misformed the <CN> structure. James (1998) defines such errors as 'words' that are non-existent in the TL, but originate in the L1 or created by the learner from the resources of the TL itself. The misformed structures in this learner corpus are slightly different from what has been defined by James. The <CN> errors can be words which are existent or non-existent in the TL, and they are re-created by the learner from the TL; but they are considered as errors because they are non-existent sentence connectors in the TL.

20   in this world are English. <CN>@Futher more/Furthermore@</CN>, you will know more
21   member of a large family. <CN>@None the less/Nonetheless@</CN>, they need to share all
22 language and it is prevelent. <CN>@Now a day/Nowadays@</CN>, many country used the
23 is classified by most country. <CN>@Now days/Nowadays@</CN>, english has become an im
24   <CN>@Now a days/Nowadays@</CN>, there are many large fa
25 its of learning it is unlimited. <CN>@Nows a day/Nowadays@</CN>, about all the university
26   ins, the more ideas we have. <CN>@Example/For example@</CN>, I as a student, facing so
27 monthly usage of the family. <CN>@Example/For example@</CN> like the electrical and

Figure 4.12

Examples of tagged <CN> errors: Misformation of structure

Even though this study is on MWUs, 'furthermore', 'nonetheless' and 'nowadays' are being discussed here, to highlight the misformation of these single

words, which have been re-created as multi-words. Examples in CL-20 and CL-21 show how these single words – 'furthermore' and 'nonetheless' are misformed as multi-word units – 'further more' and 'none the less'.

A more interesting finding in this learner corpus is how learners form a few variations of multi-word units for this single word – 'nowadays'. It is indeed a surprising finding that 'now a day' (CL-22), 'now days' (CL-23), 'now a days' (CL-24), and 'nows a day' (CL-25) can be reformed and misformed as a MWU for 'nowadays', which is a single unit. These errors were tagged in various learners' essays instead of one.

The tagged errors in this category highlighted an important issue – to use connectors as 'units' of language – single or multi-words. Learners need to be taught explicitly that 'furthermore', 'nonetheless', and 'nowadays' are connectors which can only exist as single units, in order to avoid misformation of such <CN> errors. Similarly, in CL-26 and CL-27, 'For example' also exists as a multi-word unit and reforming this connector into a single unit – 'example', is considered as an error. Therefore, learning connectors as 'chunk expressions' – single units and multi-word units help learners to be aware and reduce occurrences of such misformation of errors.

4.3   Summary of findings

From the detailed analysis of the tagged errors, the findings reveal that the Chinese-educated Malaysian learners of English have difficulties with these MWUs: the modal structures <MD>, the infinitive form <IN>, the adjective-noun collocations <JN>, and connectors <CN>. The <MD> structure is the category which has the highest frequency count of errors. The analysis shows that the errors can be categorized into: 'overgeneralization', 'misformation', 'misselection', 'distortion', and 'inappropriate use of collocations'. Now that there are answers to the first two research questions, we shall look at the factors that cause such difficulties in MWUs and the pedagogical implications of these findings in the next chapter.