

## Chapter Three

### METHODOLOGY

This Chapter describes the research design, including materials and procedures, which are used to examine the results of acquisition sequence and error analysis.

#### 3.1. Material

Our materials have been taken from the CLEC, Chinese Learner English Corpus (CLEC), which consists of over 1,000,000 words, compiled and computerized by Professors Gui Shichun and Yang Huizhong in Mainland China. 50,000Fifty thousand words were collected from the writings for the test, guided writings and free writings. It is divided into five sections from ST2 to ST6, representing the high-school learners (ST2), freshman and sophomore college learners who are not English majors (ST3), junior and senior college non-English-major learners (ST4) and English majors of beginning (ST5) and advanced (ST6) proficiency respectively. The sizes of each CLEC component are summarized in Table 3.

**Table 3**

*Dispersion of CLEC Data (Gui and Yanget al., 2003; p3)*

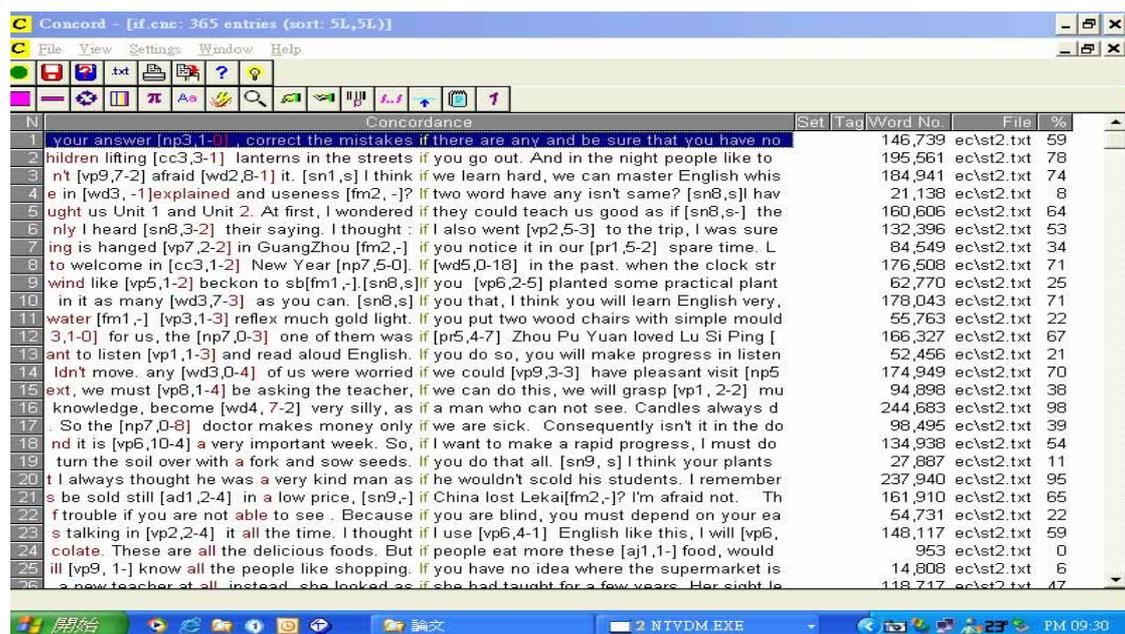
<b>Types</b>	<b>Lexis frequencies</b>
ST2	208088
ST3	209043
ST4	212855
ST5	214510
ST6	226106
Total	1070602

## 3.2. Procedure

### 3.2.1. Identifying If-conditionals

Wordsmith program (version 3.0) designed by Mike Scott (1998) was adopted. We used the CONCORD tool to extract the data needed.

First of all, because this study focuses mainly on “if”-conditionals, the initial step is to screen out the entire “if” sentences. A method for focusing the extraction on just these phrases is, first, to generate concordances for *if* in each sub-corpus, and, second, to have Wordsmith extract the recurring strings from that output and tally the frequency of each throughout the corpus (Figure 2).



N	Concordance	Set	Tag	Word No.	File	%
1	your answer [np3,1-1], correct the mistakes if there are any and be sure that you have no			146,739	ec\st2.txt	59
2	children lifting [cc3,3-1] lanterns in the streets if you go out. And in the night people like to			195,561	ec\st2.txt	78
3	n't [vp9,7-2] afraid [wd2,8-1] it. [sn1,s] I think if we learn hard, we can master English whis			184,941	ec\st2.txt	74
4	e in [wd3,-1] explained and useness [fm2,-]? If two word have any isn't same? [sn8,s] hav			21,138	ec\st2.txt	8
5	ught us Unit 1 and Unit 2. At first, I wondered if they could teach us good as if [sn8,s-] the			160,606	ec\st2.txt	64
6	nly I heard [sn8,3-2] their saying. I thought : if I also went [vp2,5-3] to the trip, I was sure			132,396	ec\st2.txt	53
7	ing is hanged [vp7,2-2] in GuangZhou [fm2,-] if you notice it in our [pr1,5-2] spare time. L			84,549	ec\st2.txt	34
8	to welcome in [cc3,1-2] New Year [np7,5-0]. If [wd5,0-18] in the past, when the clock str			176,508	ec\st2.txt	71
9	wind like [vp5,1-2] beckon to sb [fm1,-]. [sn8,s] If you [vp6,2-5] planted some practical plant			62,770	ec\st2.txt	25
10	in it as many [wd3,7-3] as you can. [sn8,s] If you that, I think you will learn English very,			178,043	ec\st2.txt	71
11	water [fm1,-] [vp3,1-3] reflex much gold light. If you put two wood chairs with simple mould			55,763	ec\st2.txt	22
12	3,1-0] for us, the [np7,0-3] one of them was if [pr5,4-7] Zhou Pu Yuan loved Lu Si Ping [			166,327	ec\st2.txt	67
13	ant to listen [vp1,1-3] and read aloud English. If you do so, you will make progress in listen			52,456	ec\st2.txt	21
14	ldn't move. any [wd3,0-4] of us were worried if we could [vp9,3-3] have pleasant visit [np5			174,949	ec\st2.txt	70
15	ext, we must [vp8,1-4] be asking the teacher, if we can do this, we will grasp [vp1,2-2] mu			94,898	ec\st2.txt	38
16	knowledge, become [wd4,7-2] very silly, as if a man who can not see. Candles always d			244,683	ec\st2.txt	98
17	. So the [np7,0-8] doctor makes money only if we are sick. Consequently isn't it in the do			98,495	ec\st2.txt	39
18	nd it is [vp6,10-4] a very important week. So, if I want to make a rapid progress, I must do			134,938	ec\st2.txt	54
19	turn the soil over with a fork and sow seeds. If you do that all. [sn9,s] I think your plants			27,867	ec\st2.txt	11
20	t I always thought he was a very kind man as if he wouldn't scold his students. I remember			237,940	ec\st2.txt	95
21	be sold still [ad1,2-4] in a low price, [sn9,-] if China lost Lekai [fm2,-]? I'm afraid not. Th			161,910	ec\st2.txt	65
22	f trouble if you are not able to see. Because if you are blind, you must depend on your ea			54,731	ec\st2.txt	22
23	s talking in [vp2,2-4] it all the time. I thought if I use [vp6,4-1]. English like this, I will [vp6,			148,117	ec\st2.txt	59
24	colate. These are all the delicious foods. But if people eat more these [aj1,1-] food, would			953	ec\st2.txt	0
25	ill [vp9,1-] know all the people like shopping. If you have no idea where the supermarket is			14,808	ec\st2.txt	6
26	a new teacher at all, instead, she looked as if she had taught for a few years. Her sight le			118,717	ec\st2.txt	47

Figure 2. A screenshot of the sample of if-conditionals

We began to survey the occurrence of if-sentences. Some disqualified materials were eliminated from the corpora. For example, “only if”, “as if”, “what if”, “if necessary”, “if any” and “if possible” ...etc. and “even if”--a concessive marker, were excluded because they are not the conditional targets we aimed. Additionally, some

pro-forms are used to substitute a whole sentence, for instance: *If so, please let me know* (Halliday & Hasan, 1976). “if so” and others, such as “if not”...etc., were also singled out for their absence of subject and verb (though they can be inferred, the risk of misinterpretation may occur), since our purpose lies in the verb change in the sentences. Besides, the pattern “V+if “ was also ignored as well as the pattern “imperative + if-clause”. The former is equivalent with “whether”, while the latter performs a function, which is discourse-oriented rather than conditional. Moreover, those with incomplete structure were not considered, including the sentence introduced by *if* but not being a conditional (see example [1]) and the sentence without its main clause (example [2]).

[ 1 ] *If [fm1,-] affects not only the child himself but also the family and...~[ST3, 205905, 84%]*

[ 2 ]... *[sn9,-] for example food clothes house. and [fm3,-] if [wd5,-] you even can [wd1,-] keep a younge [fm1,-] face. ~[ST3, 205532, 84%]*

### 3.2.2. Identifying Errors

The entire CLEC was error-tagged on the basis of an error-coding scheme (see Table4 below). The error-tagging scheme was devised by Gui et al. and Yang (2003) to depict the errors produced by the Chinese English learners. The error codes consist of 11 categories, each includes several sub-categories. Because we only concern with the structure of conditionals, other categories are excluded except Verb Phrase. Among the subcategories of VP errors, only Vp6 (tense), Vp8 (mood), and Vp9 (modal and auxiliary) concern primarily the verb change in the if-C and MC in conditionals<sup>9</sup>, they

---

<sup>9</sup> Vp7 was excluded because it describes voice errors, that is, error in the use of voice: active for

are listed in the following Table.

**Table 4**

*Error Tagging Taxonomy (Adapted from Gui and Yang, et al. 2003, p5-6)*

Type No.	Error code	Category	Description
6	Vp6	Tense	Error in tense use within a sentence; the sequence of tenses between sentences
8	Vp8	Mood	Error in the use of mood: imperative, subjunctive; improper structure of conditional sentences
9	Vp9	Modal/auxiliary	Misuse of modal/auxiliary verbs; wrong form of modal verb (or auxiliary verb) and verb combination (e.g. tense form, voice form, etc.)

Before delving into the verbal errors, it is necessary to deal with the special category--VP9 errors first. As mentioned before, Chou's (2000) study yielded a problem that she considered "the existence of modals" as one of the grammatical categories which generated errors on accounts of its contribution to syntactic complexity. However, this factor involves more with the speaker's original intention to express her subjective-ness and less with the conditional construction itself. Because modal is a means to convey the degrees of tentativeness and possibility, thus it plays a more flexible role in the conditionals. So the inaccurate selection of modals should not be included, it is the "forms of tense" needed to be considered in examining conditional constructions.

The VP9 errors are regarded as the misuses of modal or auxiliary verbs in Gui et

---

passive or vice versa.

al.’s tagging scheme, so we separated those which are “form-change errors” involved with modals, for instance: a past-form modal error in the present real conditionals, and included them with the main verb errors. Hence, the errors in these categories all pertain to the form changes of conditionals.<sup>10</sup> We will speak directly to the issue of the verb and modal misuses concerning the learners’ ability of using conditionals.

To make the definitions clear and precise, we have defined tense errors as the misuses of wrong verb (or modal) forms to indicate a particular time frame, while mood errors as the misuses of wrong verb (or modal) forms to denote the degrees of unreality. However, the distinction is not a clear-cut area. Although Gui *et al.* distinguished tense and mood errors, but in fact, some errors are unable to be categorized easily in the conditionals due to their ambiguity. Therefore, when analyzing the data, we decided to combine the VP6 and VP8 errors together to represent the deviated uses of verb and modal forms, instead of dividing the two types of errors by inferring the author’s intention.

N	Concordance	Set	Tag	Word No.	File	%
1	bit afraid. If we [wd4,1-4] not careful we can [vp6,s-] fall. Finally we reached the top of th			2,044	英Nf.cnc	27
2	ppy. I think [vp6,25-5] if I was a child , I will [vp6,6-3] get these present[fm1,-]. how hap			6,530	英Nf.cnc	46
3	e class. If we could do it really [ad1,6-] , it is [vp6, 2-6] not too difficulty [wd2,1-] to stud			9,966	英Nf.cnc	59
4	d4,1-3] again and again. if [fm3,-]you listen [vp6,2-1] clear [cc5,1-]. only [fm3,-]put on t			16,797	英Nf.cnc	94
5	ething [wd3,21-0]. If all the people who came [vp6,-] there do [vp5,9-1] this. the air will be			17,339	英Nf.cnc	88
6	himself? If he is [vp6,s-] an old man, I'll think [vp6, 2-2] he can't work. he is [vp6,s-] a pool			18,259	英Nf.cnc	92
7	-] are very happy to answer you. If you was [vp6,s-] not satisfied [fm1,-] with the super			17,777	英Nf.cnc	90
8	e [vp5,1-2] backon to sb[fm1,-].[sn6,s-]if you [vp6,2-5] planted some practical plants[cc4,			384	英Nf.cnc	23
9	tests. If I was more carefully [wd2,4-0]. I will [vp6,6-8] be [vp2, 2-5] more better [ad3,1-]i			10,293	英Nf.cnc	60
10	ught if I use [vp6,4-1] English like this. I will [vp6,s-s] learn it more quckly [fm1,-] than			1,667	英Nf.cnc	26
11	ay "Hello" or mod [fm1,-] if someone knew [vp6, 2-1] each [wd4,5-0] , they'll say congra			6,308	英Nf.cnc	47
12	an [vp3,1-] when you listen once if you didn't [vp6,s-] listen clear [wd2,s-] or want to listen			15,547	英Nf.cnc	97
13	he can't [vp6,s-1] got money himself? If he is [vp6,s-] an old man, I'll think [vp6, 2-2] he ca					
14						
15						
16						
17						
18						
19						
20						

Figure 3. A sample of error tags

Following Moffie’s (2000) suggestion to focus on verb forms, we will investigate the deviations in both the if-Cs and the MCs. Accordingly, the errors have to be searched and listed by its error code (Figure3). By doing so, it is possible to count the frequency of each error type, so that the acquisition sequence and error

<sup>10</sup> The errors which contain: modal + to, modal + finite verb, no verb follows the modal, misuse of the meanings of modals, and combination of two modals, are left out.

taxonomy can be determined. The results will be displayed in Chapter 4.

After all the target if-sentences and the errors were found out, the next step was to determine the conditional type of each sentence. Since the conditional type has not been identified yet in CLEC, we needed to annotate this particular information according to the classification in Table1.

### **3.2.3. Identifying Conditional Types**

When identifying the type of the conditional sentence, we had to refer to the original context where the target construction was embedded. This was also a semi-automatic procedure done through the function of “View Text”, which enables the necessary context to be displayed. Only by considering the whole picture could the conditional type be accurately determined.

Although the four conditional types listed in Table1 capture the majority of conditionals in CLEC, there are still many instances, which can not be fit into any of the categories. Hence, those alternative conditional structures were separated from our targets<sup>11</sup>. Besides, there are some ambiguous cases which are vague in type assignment. To reduce the effect of uncertainty, these kinds of cases were also grouped together and termed as “others” category.

### **3.3. Coding Errors**

The tagging system partially developed for this study is best explained with reference to an example, in which the coding was enclosed within square brackets: [vp6, 1-2, t3]. The symbols “Vpx” stands for error type, the numbers represent word sequences, and “tx” means conditional type.

---

<sup>11</sup> For example, the pattern “if+ modal, modal” does not belong to any of our four conditional types.

