

Acquisition of L2 Discourse Structure in CSL: A Study of Lexical Cohesion of CSL Learners

Siu-lun Lee

The Chinese University of Hong Kong, Hong Kong, China

This is a first study to look at English-speaking CSL (Chinese as a second language) learners' development of Cantonese discourse structure. A longitudinal study lasted for 18 months has been done to trace four CSL learners' acquisition of lexical cohesion devices. Learners' outcome in narrative story making and speech lessons are recorded and transcribed for the study. The data show that learners have steady growth in using lexical cohesion devices. SLR (Simple Lexical Repetition) is the most frequently used device. However, other types of lexical cohesion devices, such as Paraphrasing, CR (Co-reference), and E (Ellipsis) are less preferred devices as compared with native Cantonese speakers (in the control group). The research throws some lights to future research concerning acquisition of lexical cohesion devices as well as open up questions for language teachers and materials developers to consider.

Keywords: lexical cohesion, discourse structure, language acquisition, narrative storytelling, teaching CSL (Chinese as a second language), teaching Cantonese as a second language

Introduction

Language learners always concern about how a native level of competence can be reached. For native level of language competence, learners usually refer to whether the L2 (second language) pronunciation is native-like, whether the grammatical structure of the target language has been acquired. There is no exception in teaching CSL (Chinese as a second language) literature, lots of works has been done on the acquisition of the grammatical elements and phonological aspects. However, can native-like pronunciation and correct use of lexical items and grammatical structure alone satisfy the actual needs to use the target language? The answer is that pronunciation and correct use of grammatical rules alone cannot satisfy the actual needs to use the target language in real life situations (Nunan, 2004). Sociolinguistic competence (Sapir, 1921, 1929; Hymes, 1972; Widdowson, 1978) and pragmatic competence (Levinson, 1983) also pay a very important role. In addition to these theoretical framework, how learners develop discourse structure in terms of complexity when they are learning the target language is also worth-studying. Questions asking whether students have similar discourse structure, especially in connected utterances, to native language users, and how well learners can organize connected texts or speeches are very important for language teachers to look at (Halliday & Hasan, 1976; Hoey, 1991).

Some previous studies show that lexical cohesive devices are important in learners' oral discourses (WU, 2010; Kafes, 2012), in learners' reading proficiency (MacMillan, 2007), in learners' written discourse (CHEN,

2007), and in developing learners' metacognitive strategy (Wilawan, 2007). WU (2010) showed that lexical cohesion devices affects and correlate with the quality of oral English of EFL (English as a foreign language) learners and concluded that EFL students "should improve their use of cohesive devices to make coherent and tightly organized oral discourses" (p. 101). This paper is the first study of CSL learners' development of lexical cohesion devices. The paper shows a longitudinal study of the acquisition of lexical cohesion devices by learners of Cantonese as an L2. The study is based on concepts and categorization of lexical cohesion devices discussed by Hoey (1991).

Many linguists (Halliday & Hasan, 1976; Hasan, 1984; Winter, 1979) tried to study the cohesion of a text. Halliday and Hasan's study (1976) was one of the early studies about "cohesion". They said cohesion "result from the combination of semantic configurations of two kinds: those of register, and those of cohesion, which enables a text to function as a text" (Halliday & Hasan, 1976, p. 2). Hoey (1991) summarized most of the works and discussed various "repetition links". He constructed theories about lexical cohesion and repetition links. According to Hoey (1991), repetition links can be classified into "SLR (Simple Lexical Repetition)", "CLR (Complex Lexical Repetition)", "SP (Simple Paraphrase)", "CP (Complex Paraphrase)", "HR (Hyponymic Repetition)", and "CR (Co-reference)". Table 1 shows Hoey's definition of various repetition links.

Table 1

Repetition Links With English and Cantonese Examples—Adopted From Hoey (1991)

Types of repetition	Definitions	English examples	Cantonese examples
SLR	When a lexical item that has already occurred in a text is repeated with no greater alteration than is entirely explicable in terms of a closed grammatical paradigm.	Bear & Bear; Bears & Bear	报纸 boujǐ & 报纸 boujǐ (newspaper & newspaper); 佢地 kéuihdeih & 佢 kéuih (they & he/she)
CLR	When two lexical items share a lexical morpheme, but are not formally identical..., or when they are formally identical, but have different grammatical functions.	Drug & Drugging	食饭 sihkfaan & 食咗饭 sihkjófahhn & 食紧饭 sihkgán faahn (have meal & had meal & having meal)
SP	Whenever a lexical item may substitute for another in context without loss or gain in specificity and with no discernible change in meaning.	Volumn & Book; Produce & Cause (in certain context)	车 chē & 汽车 heichē (cars & cars); 因为 yānwai & 原因 yùhnyān (because & reason)
CP	When two lexical items are definable such that one of the items includes the other, although they share no lexical morpheme and include antonyms.	Sickness & Doctor; Happy & Unhappy (antonym); Writer & Writings; Tooth & Dental & Dentist	病人 behngyàhn & 医生 yīsāng & 医院 yīyún (sickness & doctor & hospital); 肥 fēih & 瘦 sau (fat & thin)
HR	When two lexical items are of hyponymic relationship.	Animals & Bears	动物 duhngmaht & 狗 gáu & 猫 māau (animals & dogs & cats); 点心 dīmsām & 虾饺 hāgāu (Chinese pastry & shrimp dumplings)
CR	Two items are interpreted as having identical referents... even though they are either unrelated as lexical items..., or related as superordinate to hyponym.	Augustus & the Emperor; Mrs. Thatcher & the Prime Minister some personal pronouns and demonstrative pronouns	梁振英 Lèuhng janyīng & 特首 dahksáu (CY Leung & HKSAR Chief Executive)

Background of the Subjects and Research Methodology

This section discusses the background of the research subjects and research methodology used. Four learners, whose mother tongue is English, are volunteered as subjects for this study. They learn Cantonese as an L2 because of job-related reasons. They all have strong motivation of learning the language and learn the language. The learners share similar characteristics, such as they grown up with same native language as well as in the same cultural background, and they have no prior background of Chinese. The Cantonese course they took is an Advanced Diploma program at university level in Hong Kong which lasted for two years. A longitudinal study is carried out and starts from their level one and monitors all the way through level five or six (there are totally six level in the curriculum) by looking at learners' outcome in narration stories (story marking/speeches tasks). Proficiency tests, adopted from OPI (Oral Proficiency Interview) (ACTFL (American Council on the Teaching of Foreign Languages), 2012a), are carried out every three months to check subjects' proficiency level. Level one is the most beginning level with zero knowledge of Cantonese, and level six is the most advanced level in the curriculum and the level of proficiency is equivalent to Advanced Low in ACTFL proficiency guidelines (ACTFL, 2012b). Table 2 shows the biographical data of the subjects. Table 3 shows different proficiency levels in different learning periods in the curriculum with equivalent proficiency levels in ACTFL proficiency guidelines.

Table 2

Biographical Data of the Subjects

Subjects	Age	Gender	Years of learning
Subject A	28	Female	2 years
Subject B	32	Female	2 years
Subject C	26	Female	1.5 years
Subject D	30	Female	1.5 years

Table 3

Equivalent to ACTFL Proficiency Guidelines at Different Stages of Studies

Learning period	ACTFL proficiency guidelines
3 months	Novice mid
6 months	Novice high
9 months	Intermediate low
12 months	Intermediate mid
15 months	Intermediate mid to intermediate High
18 months	Intermediate high to advanced low

The longitudinal data of learners' narration storytelling gathered from story making/speech tasks. Learners are required to give a five to 10 minutes speech over topics assigned by teachers every month in language classroom. The students prepared the speech or story over certain topics at home and present it in class. Teachers monitor and record the presentation. As in part of the class activities, teachers also give feedback or correct mistakes and errors in class after the oral presentations of the learners. The first five minutes of the speeches or stories are transcribed and analyzed in terms of the lexical cohesion devices. A total of 15 speeches of each subject were collected and transcribed into Romanization for analysis.

Two voluntary Cantonese native speakers, of the same gender and in the same age group, were selected as native speaker references. Native speakers were given the same 15 speech topics as the learners with same

conditions. Fifteen speeches were also recorded from the native speakers group. Data from the learners are transcribed, analyzed, and compared among learners themselves and with the native speakers group. Table 4 shows the topics of speeches and stories, which concerning life in Hong Kong and daily life as well as discussing on some cultural, socio-economic, and political issues.

Table 4

Topics of Speeches in the Data

Levels	Topics of the speeches
Level 1-2	Self-introduction, Introduce your family, Introduce your friends, Introduce your school/your ideal school, Describe a gathering.
Level 3-4	Describe an overseas trip, Describe a dream, Describe an accident, Describe a visit to some friends, Tell us your future plans.
Level 5-6	Present your view towards religion, Present your view towards fashion, Discuss environmental problems, Discuss and compare housing problems/education issues between Hong Kong and your country, Discuss some social issues in Hong Kong and in your country, etc.

The materials collected were transcribed and analyzed (without error correction). In this paper, Yale Romanization system is used to present the examples. Yale Romanization is a Cantonese transcription system well known transcription method, which shows the sounds and tones of the language. English translation is provided on the second line in italic. The transcribed texts were analyzed based on the Hoey's (1991) model of lexical cohesion.

Repetition Links in the Data

SLR

As Hoey (1991) suggested all repetition links allow the speakers to say something again and some new information can be added. One form of repetition is SLR.

SLR as defined in previous section, occurs when a lexical item has occurred in a text is repeated with no greater alternation than is entirely explicable in terms of a closed grammatical paradigm. In the data, SLR is a very common device for the learners to hook up the whole speech or passage (see Example 1).

Example (1) Yùhgwó góng ngóh gwok ge jūnggaau chihngyihng	line 1
<i>If we talk about religion in our country,</i>	
ngóh jauh gokdāk hóu nàahngóng.	2
<i>then I feel difficult to say,</i>	
Yānwaih ngóhge gwokgā ge jūnggaau tūhng kèihtā gwok ge jūnggaau hóu mtùhng.	3
<i>because the religions in my country is different from the religion in other countries.</i> (Mary, 2009b, B4.4: L1-4)	

The above is a speech on “My Opinion Towards Religion” (2009), we can see that “Jūnggaau” (religion) occurred and repeated every line. Thus “Jūnggaau” is the main theme of the speech, the repetition obviously wide-spread in the whole speech.

CLR

As defined above, CLR occurs when two lexical items share a lexical morpheme, but are not formally identical. Since Cantonese is a language does not have many verb inflections as compare to other European languages. Examples of this type are not very rich in the data. In the example below, the aspect suffix/marker “-gán” can be found (see Example 2).

Example (2) Ngóh yáuh yātga <u>chē</u> .	line 1
<i>I had a <u>car</u>.</i>	
Ngóh hóu jūngyi <u>jāchē</u> , sóyíh yùhgwó yáuh géiyaht fongga gójahnsih, ngóh	2
<i>I loved to <u>drive</u>, so if there were a few days hoildays,</i>	
jihgéi <u>jāchē</u> fāan ngóhge hēunghá.	3
<i>I <u>drove</u> back to my hometown.</i>	
Yáuh yātyaht ngóh dōu <u>jāchē</u> fāan ngóhge hēunghá. Kèihsaht ngóh	4
<i>One day I also <u>drove</u> back to my hometown, In fact I</i>	
hóu guih. Yānwaih dōngsìh ngóh hóu mdākhaahn yātgoyuht yihseuhng	5
<i>was very tired. Because I was very busy, more than one month</i>	
móuh fongga, gēifùh léuhngyaht móuh fangaau. Ngóh jīdou gámge	6
<i>hadn't had vacation, almost two days hadn't slept. I knew</i>	
chihngyihng gójahnsih chīnkèih <u>mhóu jāchē</u> . Daahnhaih ngóh hóu	7
<i>I knew I <u>shouldn't drive</u> in that situation. But I loved to go home earlier</i>	
séung fāan ngūkkéi, yihché ngóh jaapgwaanjó <u>jāchē</u> fāan hēunghá.	8
<i>and I used to <u>drive</u> back to my hometown.</i>	
Ngóh hái gōuchūkgūnglauh <u>jāgānchē</u> , yíhgīng yehmáahn baatdím jūng la.	9
<i>when I <u>was driving</u> on the high way, that was already eight p.m.. (June, 2010, B5.4: L.4-12)</i>	

In Example (2), “jāchē” (*drive*), apart from occurring as SLR, was also found as CLR links with “chē” (*car*), “jāchē” (*drove*), and “jāgānchē” (*driving*) was found as repetition links. As discussed above, CLR is not as rich as and as many as SLR, because Cantonese is not an inflectional language (Matthews & Yip, 1994).

SP

SP occurs when a lexical item can substitute for another in context without loss or gain in specificity and with no discernible change in meaning. Most of them occur as synonymy.

In the scripts of “The Environmental Problems of Hong Kong” (2010), “chīnggit” (*clean*, literal meaning: clear & clean), and “gōnjehng” (*clean*, literal meaning: dry & clean) and be found.

CP

CP covers antonyms, and it also occurs when two lexical items are definable such that one of the items includes the other, although they share no lexical morpheme.

CP is also not rich in the data. In the scripts of “Hong Kong’s Environmental Problems” (2010), lexical items like, “gōnjehng” (*clean*), “chīnggit” (*clean*), “wūyihm” (*polluted*), and “wūjōu” (*dirty*) are good examples.

HR

HR occurs when two terms are of hyponymic relationship.

In the example of “My Opinion Towards Religion”, “jūnggaau” (*religion*) and “Fahtgaau” (*Buddism*),

“Tīnjyúgaau” (*Catholic*), “Gēidūkgaau” (*Christian*) can be found. In the speech about “Eating and Social Gathering” (2009), “sihkmaht” (*food*), “gāiyuhk” (*chicken meat*), “jyūyuhk” (*pork*), “choi” (*vegetables*), and “tōng” (*soup*) can be found. In the speech of “Hong Kong’s Social Problems” (2010), “gīngjai mahntàih” (*economic problems*), “gīngjai sēuiteui” (*economic recession*), “sātyihp” (*unemployment*), and “jāplāp” (*bankruptcy*) can be found.

CR

According to Hoey (1991), if two items are interpreted as having identical referents, they will be treated as repetitions—even though they are either unrelated as lexical items. There is an example of CR (see Example 3).

Example (3) Dáng ngóh gaasiuh, <u>kéuih</u> giujouh <u>John</u> . <u>Kéuih</u> ge bàhbāhah	line 1
<i>Let me introduce, he is called John. His father is</i>	
seuihsih yàhn. <u>John</u> ge màhmā haih Yahtbún chēutsai ge hòhnggwok	2
<i>Swiss. John’s mother is a Japan born Korean,</i>	
yàhn, <u>kéuih</u> haih ngóh ge pàhngyáuh. Yihgā <u>kéuihdeih</u> yātgā yàhn	3
<i>she is my friend. Now they (the whole family)</i>	
jyuh hái seuisih. Yihgā <u>kéuihdeih</u> hóu hōisām.	4
<i>live in Swiss. Now they are very happy.</i> (Mary, 2009a, A5.6:L1-4)	

In the above script kéuih (he/she) CRs with John and kéuih (his) in kéuihge bàhbā (his father) (line1), **Johnge màhmā** (John’s mother) CR with **kéuih** (she) (lines 2-3) and in lines 3 and 5, kéuihdeih (they) CRs with John’s family.

E (Ellipsis)

E means “omission”. There are a few lexical items can function E in Cantonese. The first one is “gám”. The meaning is equivalent to “so” in English as in “*I think so*” (ngóh dōu gám wah, literally means: I also so say), or “*in that way*” as in “*I don’t like to do that way*” (ngóh mséung gám jōuh, literally means: I don’t want so do). Another lexical item, “kèihtā” means “the other” or “another”. All these lexical items can function as E link. However, it is rarely found in the data.

A more complicated form of E is dropping the subjects which is referred to the Pro-drop parameter discussed in Chomskyan linguists (White, 1985, Licerias & Díaz, 1999). Cantonese is a “pro-drop” language, in which one of the properties is missing subjects and certain class of pronouns can be omitted. In the example, “it”, referring to the roasted chicken in the question is omitted in the first sentence. In the second sentence, “I”, the speakers, “you”, the listener, and “it”, the roasted chicken are omitted. This paper does not focus on the theoretical framework of acquisition of pro-drop parameter alone (since English, learners’ native language and Cantonese, are having different pro-drop parameters), however, concentrate on whether Cantonese L2 learners develop the use of this kind of E as one of the lexical cohesion devices in different stages of study (see Example 4).

Example (4) A: sīgāi hóu hóusihk. bīngō sū ga?	line 1
Lit: Roost chicken is delicious, Who roost?	
<i>This roost chicken is delicious. Who roost it?</i>	
B: Mjīdou. jūng mjūngyi a?	2
Lit: Not know. like [QUESTION MARKER]?	
<i>I don’t know. Do you like it?</i>	

A Longitudinal Analysis of the Use of Lexical Cohesion Devices of the Four Learners

In this section, we look at the four subjects and see how many lexical cohesion links appeared in their speeches in different stages in the language course and also try to see the pattern of their development.

Subject A

Looking at Subject A's data, there are growths in each period (see Table 5 and Figure 1). There is a jump after 18 months. SLR is the device used relatively more frequently. There is a drastic change in the pattern of lexical cohesion devices adopted. CLR on the other hand is not frequent, as we have discussed above, verb inflection is not so rich in Cantonese therefore not much can be found. The learner started to use paraphrases (SP, CP) steadily after 15 months of study. The use of SP and CP links has a big jump in the 15th month (from zero to six and from zero to four respectively). The reason for that is because the learner has learnt enough vocabulary items after 15 months of study and she can handle them confidently.

However, hyponyms and CR are not used very much. No token of E is recorded. The lacking of HR may be because she still does not have the confidence to use in properly. It may also be the case that hyponyms are not well covered in the language course but this claim needs further research to confirm. The lacking of CR and E is interesting, because the lexical items for CR and E (such as personal pronouns, demonstrative, and modifiers) were learnt in the first three months. The data show that the main devices for Subject A to build up cohesion in her speeches are SLR and paraphrasing.

Table 5

Number of Lexical Cohesion Links in Subject A's Narration Stories

Length of learning the target language	SLR	CLR	SP	CP	HR	CR	E	Total No. of links
3 months	5	0	0	0	0	0	0	5
6 months	4	1	0	0	0	0	0	5
9 months	7	0	1	0	0	2	0	10
12 months	6	2	0	0	2	1	0	11
15 months	10	3	6	4	5	0	0	18
18 months	13	3	8	8	1	1	0	34

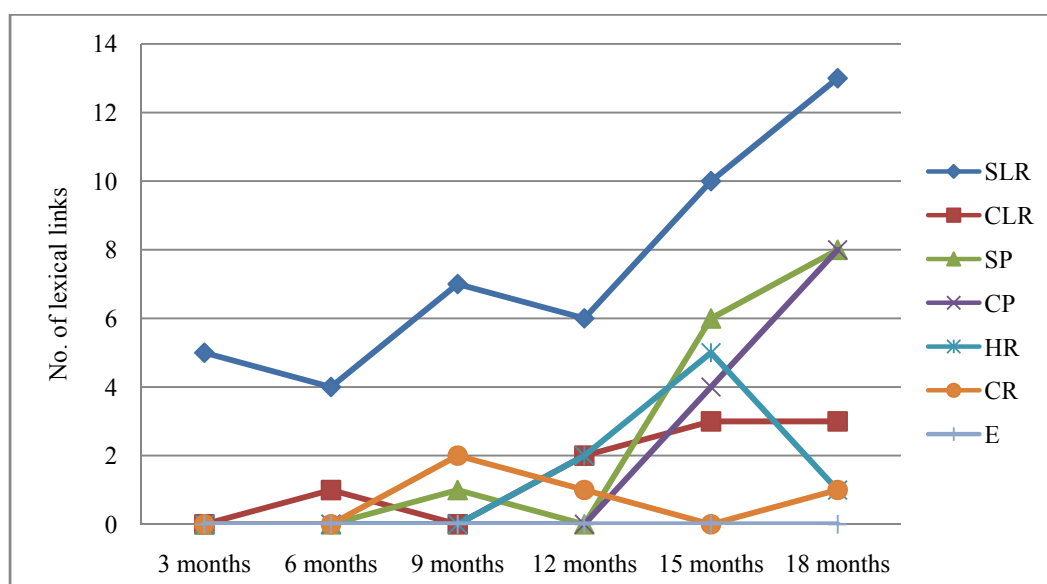


Figure 1. Number of lexical cohesion links in Subject A's narration stories.

Subject B

Subject B’s data show some similarities with Subject A’s. Both of them used SLR more frequently, but Subject B has a more evenly distribution of all the devices. There is a steady growth. The total use of cohesion devices has a big jump from the 12th month to the 18th month, from 17 to 32 tokens (see Table 6 and Figure 2).

Table 6

Number of Lexical Cohesion Links in Subject B’s Narration Stories

Length of learning the target language	SLR	CLR	SP	CP	HR	CR	E	Total No. of links
3 months	2	0	0	0	0	0	0	2
6 months	5	2	1	0	0	0	0	8
9 months	8	2	2	1	0	2	0	15
12 months	7	1	2	3	3	1	0	17
15 months	12	3	4	5	6	2	0	32
18 months	10	5	7	5	8	4	1	40

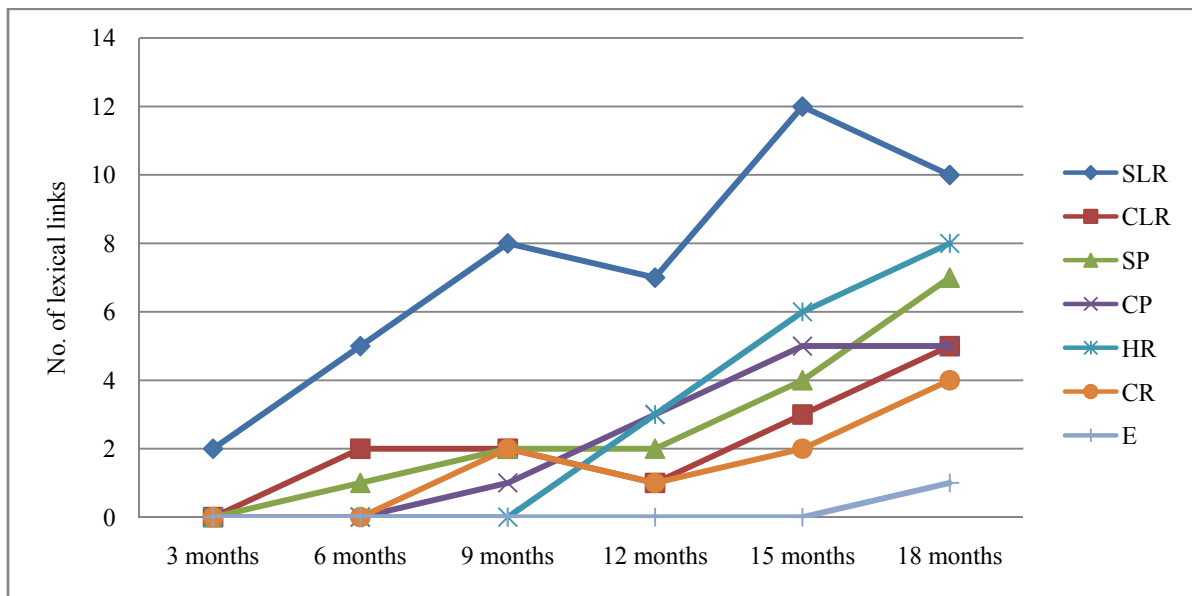


Figure 2. Number of lexical cohesion links in Subject B’s narration stories.

Subject C

Subject C’s data are quite similar to Subject A except that there are tokens of E in the data starting from the 15th month (see Table 7 and Figure 3). E is rarely found in Subjects A and B’s data.

Table 7

Number of Lexical Cohesion Links in Subject C’s Narration Stories

Length of learning the target language	SLR	CLR	SP	CP	HR	CR	E	Total No. of links
3 months	4	0	0	0	0	0	0	4
6 months	7	0	0	0	0	0	0	7
9 months	6	1	2	2	0	0	0	11
12 months	6	2	0	1	2	2	0	13
15 months	10	5	3	4	6	3	1	32
18 months	15	5	3	3	7	2	2	37

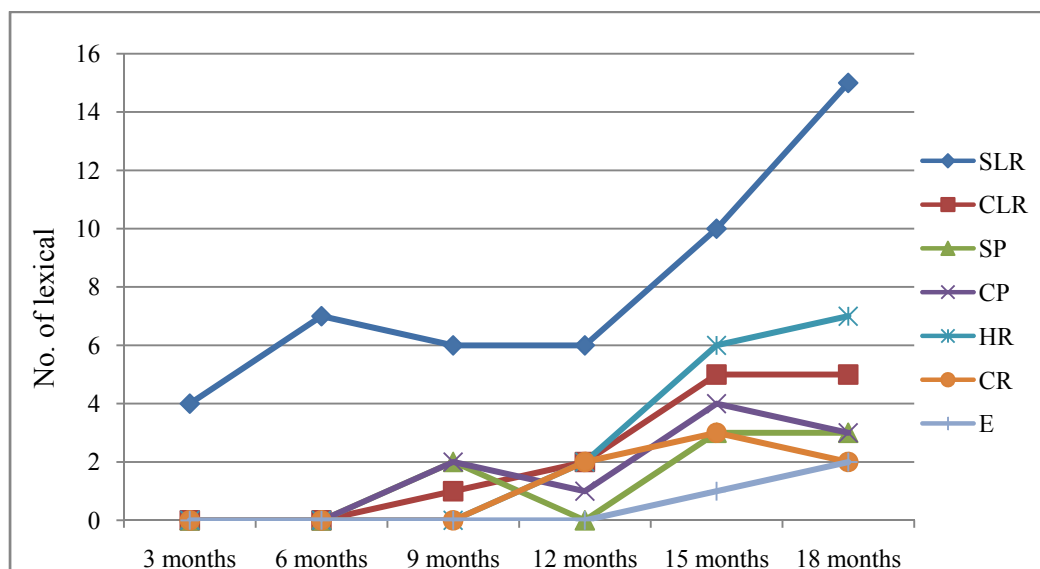


Figure 3. Number of lexical cohesion links in Subject C's narration stories.

Subject D

Subject D also has a steady growth. It started to have more than 20 links after nine months study of the target language (see Table 8 and Figure 4). Same as all the other subjects, SLR is the most frequently used device. It has a more evenly distributed used of all other devices except that E cannot be found in Subject D's data.

Table 8

Number of Lexical Cohesion Links in Subject D's Narration Stories

Length of learning the target language	SLR	CLR	SP	CP	HR	CR	E	Total No. of links
3 months	2	0	0	0	0	0	0	2
6 months	5	1	0	1	0	0	0	7
9 months	9	2	1	2	1	1	0	16
12 months	8	5	2	2	3	3	0	23
15 months	12	6	2	5	3	3	0	31
18 months	16	8	5	3	2	5	0	39

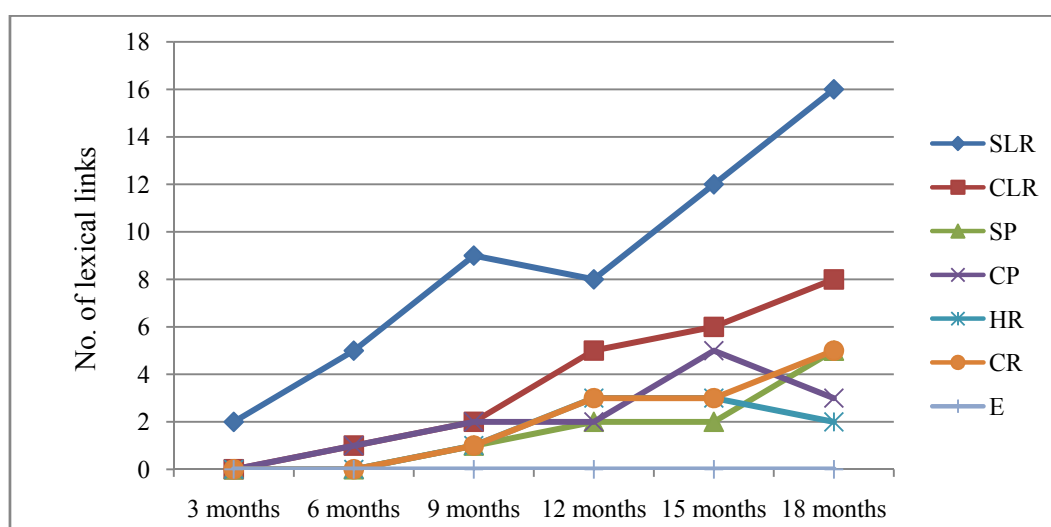


Figure 4. Number of lexical cohesion links in Subject D's narration stories.

Comparing the Four Subjects

To sum up, the four subjects show similar growth patterns, except Subject A is relatively slower in her growth. The use of lexical cohesion links grew during their study. The growth happened more drastically after the ninth month and the 15th month of study. The lexical cohesion links are difficult to be found before the sixth month of study. It is because during the first six months, most of the training is on lexical and sentence level. After six months, learners were given a text or a story to read and afterwards, there are some question and answer exercises. The exposure to an authentic text with systematic lexical cohesion devices may be helpful in developing their skills in forming a cohesive text and needs further research to verify. SLR is the most commonly used device (see Table 9 and Figure 5); it is not hard to understand, because SLP is the simplest one. Not until the end of the 12th and beginning of the 15th month, more vocabulary items, like antonyms, and hyponyms were learnt. This explains the drastic increase in started in the beginning of the 15th month. An interesting point is that although vocabulary items use for E and pronouns were learnt very early in the course, it was not used frequently by our subjects.

Table 9

Number of Lexical Cohesion Links of the Four Subjects After 18 Months

Subjects	SLR	CLR	SP	CP	HR	CR	E	Total No. of links
Subject A	13	3	8	8	1	1	0	34
Subject B	10	5	7	5	8	4	1	40
Subject C	15	5	3	3	7	2	2	37
Subject D	16	8	5	3	2	5	0	39
Native speakers	18	10	10	12	8	10	5	63

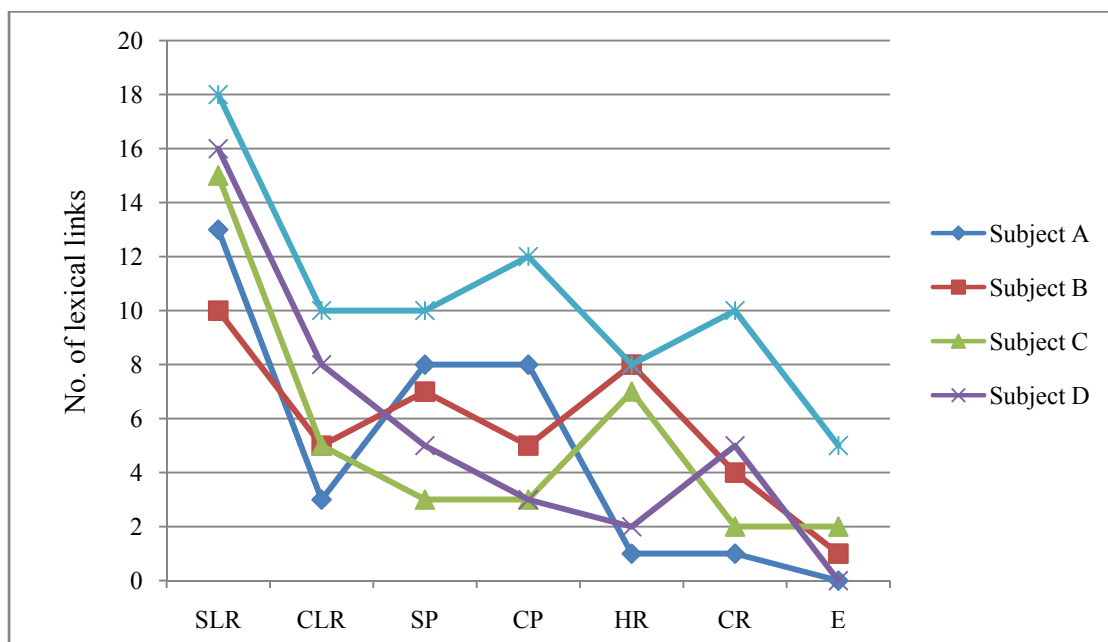


Figure 5. Number of lexical cohesion links of the four subjects after 18 months.

When comparing the data of the subjects after 18 months of study, all of them have quite similar numbers of links (30 links to 40 links) and similar patterns. It is not difficult to understand, because they all follow the same course and the same teaching material. Their trainings are almost homogenous in terms of teaching

environment, materials, and teaching methodology. If we compare the data with similar topics of speeches given by native speakers, the numbers of links of the subjects are far fewer than the native speakers (see Table 9 and Figure 5). The use of SLR is similar for both the learners and the native speakers; however, the occurrence of other links is limited when compared with the native speakers speeches. Especially the use of simple and CPs, hyponymic repetitions, CRs. E is not so frequently used in Cantonese even for native speakers as a device functioned in lexical cohesion.

Learners Group Versus Native Speakers Group

Table 10 and Figure 6 show the average numbers of lexical cohesion links of the four subjects in the data. The result echoes with previous analysis that SLR is the most frequently used lexical cohesion devices and E is the least frequently used device.

Table 10

Average Number of Lexical Cohesion Links of the Four Subjects

Average	SLR	CLR	SP	CP	HR	CR	E	Total
Learnt for 3 months	3.25	0	0	0.25	0	0	0	3.5
Learnt for 6 months	5.25	1.5	0	0.25	0	0	0	7
Learnt for 9 months	7.5	1.25	1.5	1.25	0.25	1.25	0	13.5
Learnt for 12 months	6.75	2.5	1	1.5	2.25	1.75	0	15.75
Learnt for 15 months	11	4.25	3.25	4.5	5	2	0.25	30.25
Learnt for 18 months	13.5	3.25	5.75	4.75	4.5	2.75	0.75	35.25

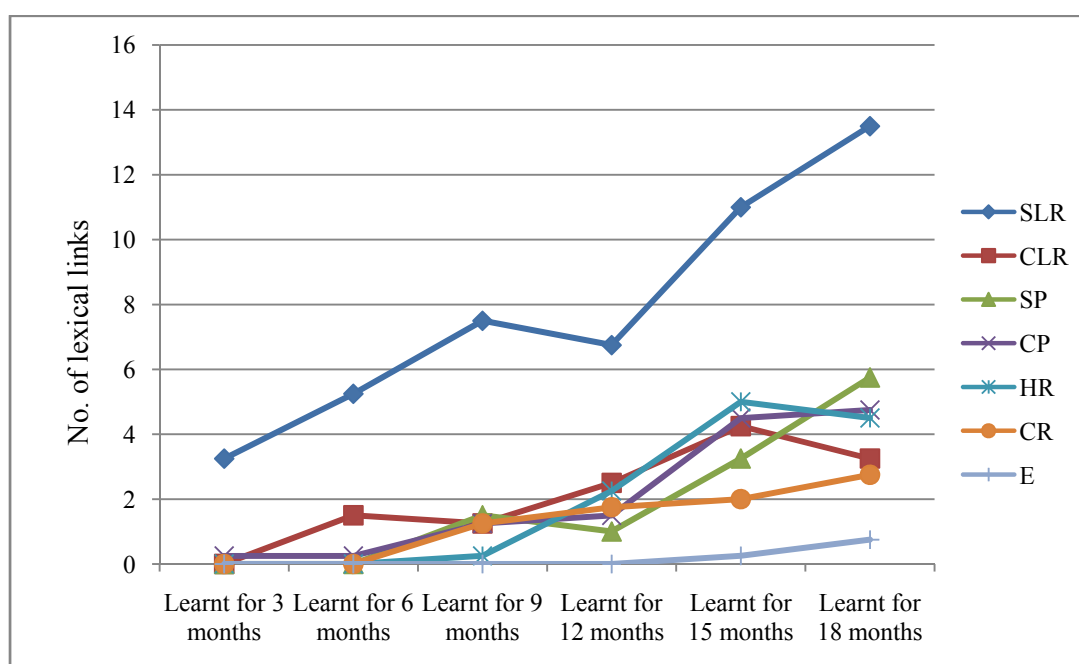


Figure 6. Average number of lexical cohesion links of the four subjects.

When comparing learners' outcomes at their 18th month of study with the native speaker-group (see Table 11 and Figure 7), the use of SLR is similar to that of the native speakers, however, the total number of lexical cohesion links used by learners is about half of that of native speakers (35.25 links/63 links), this indicates that the use of devices other than SLR is less than half of the usage found in native speakers group data, which indicates that the use of Paraphrasing, HR, CR, and E are far below the usage of native speakers.

Table 11

Comparison of the Use of Lexical Cohesion Links Between Learners Group and Native Speakers Group

Subjects	SLR	CLR	SP	CP	HR	CR	E	Total No. of links
Learner-group	13.5	3.25	5.75	4.75	4.5	2.75	0.75	35.25
Native speaker-group	18	10	10	12	8	10	5	63

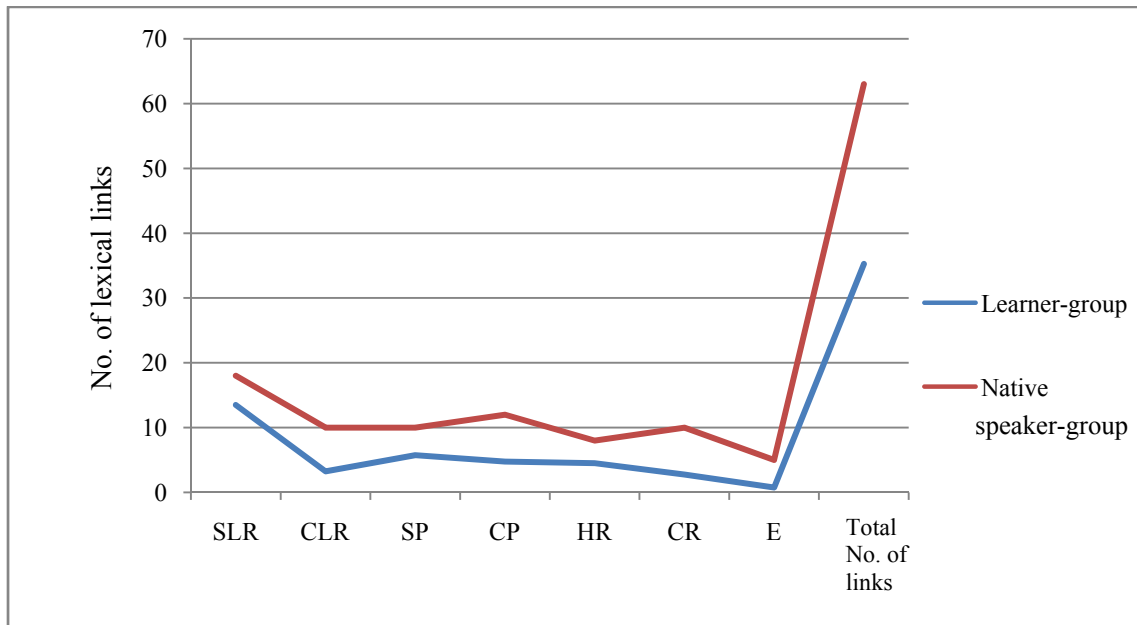


Figure 7. Comparison of the use of lexical cohesion links between learners group and native speakers group.

Discussion and Conclusions

This is a preliminary study concerning acquisition of discourse structure with special reference to lexical cohesion by CSL learners. The research data show that CSL learners have developed in lexical cohesive devices when their oral proficiency level increases during their studies. CSL Learners' use of SLR is similar to that of native speakers, however, the use of other cohesion links are far below the complexity of that of native speakers. The findings echo with some previous studies about EFL learners (WU, 2010; Kafes, 2012) that the main cohesive device use is repetition, while other devices are rarely employed in learners' language outputs. Kafes (2012) studied Turkish EFL learners ability in composing cohesive texts and compare the similarities between lexical cohesive devices learners employ in composing cohesive texts in English. Kafes' (2012) study shows that repetition of the same lexical items was the most frequently used type of lexical cohesion device. Kafes' (2012) study shows that the use of SLR is the most frequently used lexical cohesion devices as comparing to other devices, like paraphrasing, HR, CR, and E in learners' native language, and it is not difficult to understand the frequent occurrence of SLR since SLR is the easiest device to use in composing lexical cohesive text. WU (2010) suggested that it may be due to the "limitation of vocabulary and the fear to make mistakes" (p. 101) and that leads to and overuse of repetition, however, the "employment of synonymy, antonym and superordinate is too far from satisfactory" (p. 101).

The data in this study show that there is not enough growth in Simple Paraphrasing SP, CP, and HR after 18 months of study. The slow development could be because of lack of lexical items learnt and/or lack of

confidence of using the newly learnt vocabulary. Further research on how the textbooks and language inputs affects the development of these links would be helpful to explain the phenomenon. The limited use of CR and E may be because of interfering from the learners' native language and/or the fear to avoid mistakes. Detailed explanations and suggestions to CSL teachers and materials developers required further studies to find out.

References

- ACTFL. (2012a). *Proficiency testing*. Retrieved from <http://www.actfl.org/professional-development/certified-proficiency-testing-program/testing-proficiency?pageid=3348>
- ACTFL. (2012b). *ACTFL proficiency guidelines 2012*. Retrieved from http://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf
- CHEN, X. F. (2007). Lexical cohesion in Chinese college EFL writing. *CELEA Journal*, 30(5), 46-57.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman Group Ltd..
- Hasan, R. (1984). Coherence and cohesive harmony. In J. Flood (Ed.), *Understanding reading comprehension: Cognition, language and the structure of prose*. Newark, Delaware: International Reading Association.
- Hoey, M. (1991). *Patterns of lexis in text*. Oxford: Oxford University Press.
- Hymes, D. (1972). On communicative competence. In J. B. Pride, & J. Holmes (Eds.), *Sociolinguistics*. Harmondsworth: Penguin.
- June. (2010). Describe an accident. *Language data collected in The Chinese University of Hong Kong*.
- Kafes, H. (2012). Lexical cohesion: An issue only in the foreign language?. *English Language Teaching*, 5(3), 83-94.
- Levinson, C. S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Liceras, J. M., & Díaz, L. (1999). Topic-drop versus pro-drop: Null subjects and pronominal subjects in the Spanish L2 of Chinese, English, French, German and Japanese speakers. *Second Language Research*, 15(1), 1-40.
- MacMillan, F. (2007). The role of lexical cohesion in the assessment of EFL reading proficiency. *Arizona Working Papers in SLA & Teaching*, 14, 75-93.
- Mary. (2009a). Introducing my friend. *Language data collected in The Chinese University of Hong Kong*.
- Mary. (2009b). My opinion towards religion. *Language data collected in The Chinese University of Hong Kong*.
- Matthews, S., & Yip, V. (1994). *Cantonese: A comprehensive grammar*. London: Routledge.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.
- Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt, Brace & World.
- Sapir, E. (1929). The status of linguistics as a science. In E. Sapir (1958), *Culture, language and personality* (D. G. Mandelbaum Ed.). Berkeley: University of California Press.
- White, L. (1985). The "pro-drop" parameter in adult second language acquisition. *Language Learning*, 35(1), 47-62.
- Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.
- Wilawan, S. (2007). Lexical cohesion and metacognitive strategy training: An integrated approach to main idea comprehension. *Leeds Working Papers in Linguistics and Phonetics*, 12, 102-125.
- Winter, E. O. (1979). Replacement as a fundamental function of the sentence in context. *Forum Linguisticum*, 4(2), 95-133.
- WU, S. X. (2010). Lexical cohesion in oral English. *Journal of Language Teaching and Research*, 1(1), 97-101.