▶ Second Language Acquisition of Chinese Conjunction Er ('And'): A Corpus-Based Study

漢語「而」的二語習得：一個基於語料庫的研究

SECOND LANGUAGE ACQUISITION OF CHINESE CONJUNCTION *ER* ('AND'): A CORPUS-BASED STUDY[*]

Yi-Chun Tseng[1] and Miao-Ling Hsieh[1, 2]
[1]Mandarin Training Center
[2]Department of English
National Taiwan Normal University

ABSTRACT

This study investigates L2 learners' acquisition of the conjunction *er* in Chinese and their general performance in conjoining clauses within a four-tier categorization of *er*. Our findings show that at the initial stage, learners tend to use *er* to connect clauses that involve a subordination relation. The preponderance of this use of *er* by intermediate learners is in sharp contrast with the finding that native speakers use more *er*'s to conjoin sentences/paragraphs instead. Equally remarkable is the progress that learners made from conjoining at an intra-sentence level to an inter-sentence level, showing a developmental pattern in interlanguage. Finally, non-target errors are found to be largely due to an under-differentiation of the syntactic and semantic requirements of different conjuncts.

Key words: Chinese conjunction, *er*, second language acquisition, corpus

---

1. INTRODUCTION

Er ('and') is generally taken to be a conjunction and is used frequently in modern Chinese. In the *Academia Sinica Balanced Corpus of Modern Chinese (Corpus-ASB)*, a corpus which contains about 8 million characters of spoken and written Chinese data across different registers and genres, *er* ranks the 17th highest (0.376%) in frequency. More remarkably, it is the only conjunction in the top 20. Despite its high frequency, the conjunction *er* has puzzled Chinese linguists for its simplicity in conjoining seemingly unrestricted clauses as well as its complexity in marking diverse semantic relations (cf. Wang 1989, Guo 2003, Chen 1994).

Without giving specific examples for the time being, we would like to point out that *er* shares similar linguistic features with the English coordinator *and*, though the two do not always behave in the same way. Semantically, interclausal semantic relations revealed by *er* can also be pure addition, contrast, cause-effect and so on, just like *and* in English (cf. Quirk et al. 1985, Liu et al. 1996, Lü 1999). Syntactically, *er* links adjective phrases, clauses and sentences/paragraphs, but not bare verbs or noun phrases, unlike *and*.

While *and* in English has been studied extensively (e.g. Ross 1967, Culicover & Jackendoff 1997, Progovac 1998, Schiffrin 1987, Schmidt 1980, Carston 1993, Dorgeloh 2004, etc.), Chinese *er*, by contrast, has received far less attention both in the analysis of its use and in second language acquisition (SLA). For its use, three analyses of *er* including Liu et al. (1996), Lü (1999) and Hong (2008) are reviewed in Section 2. As we will see, the classifications in these analyses fail to lay a clear basis for the analysis of *er* in a learner corpus. A finer-grained categorization is needed.

In SLA, there are two recent studies of *er*. In the first one, Li (2009) investigates various types of Chinese conjunctions used by L2 learners and their corresponding errors using a corpus-driven approach. One of the findings shows that *er* is misused quite frequently among coordinative conjunctions. In the second study, Zhang (2010) points out

that learners may overgeneralize *er* to conjoin grammatical categories that are not allowed by native speakers (i.e. bare verbs or noun phrase [1]). More specifically, both studies show that in some cases *er* should be replaced by *erqie* ('and', 'furthermore'), as in (1) (example from Li (2009: 42)):[2]

(1)   *我欣賞這個人物，也同情他沒有經歷過小時候的快樂，
      而歷經了風雨，吃了很多苦。

Wo	xinshang	zhe	ge	renwu,	ye	tongqing	ta
I	admire	this	CL	person	also	sympathize	he

meiyou	jingli-guo	xiaoshihou	de	kuaile,
not:have	experience-ASP	childhood	DE	happiness

er	lijing-le	fengyu,	chi-le	hen	duo	ku.
ER	experience-ASP	ordeal	eat-ASP	very	much	bitter

'I admire him for his character and also sympathize with him in that he has not experienced happiness in childhood but suffered a lot from his ordeal.'

These two SLA studies, however, do not give a systematic analysis of non-targeted errors, nor do they provide information about learners' use of *er* in general. Our study thus aims to give a finer-grained categorization of *er* and investigate the development of *er* in a learner corpus.

This paper is organized as follows. In Section 2, three different studies of Chinese *er* are reviewed, followed by our own categorization, which includes *er* as a subordinator, as a complementizer, as a coordinator and as a discourse marker, represented as er^S, er^{Comp}, er^C and er^D respectively. Following that, we address current research questions

---

[1] To conjoin two noun phrases, Chinese has the conjunction *he* 'and'. In fact, Chinese has a rich set of coordinative devices and some of them may be syntactically overlapped in taking the same word category (e.g. Chinese *he* can conjoin verb phrases as *erqie* does under certain circumstances), rendering the acquisition difficult for learners (cf. Wu 2005, Tsai 2006, Chen 2007).

[2] Abbreviations used in this paper are: ASP (aspect marker), BE (copular verb), CL (classifier or measure word), COMP (complement), DE (marker for genitive phrases or relative clauses), and PTC (particle).

Yi-Chun Tseng and Miao-Ling Hsieh

(Section 3) and methods applied in this study (Section 4). In the data analysis in Section 5, learners' correct uses of *er* are compared with native speakers' and their erroneous uses are sorted and exemplified. Finally, in Section 6, based on learners' data, we discuss issues pertaining to (1) learners' general usage of *er*, (2) difference in choosing *er* or *erqie* under a given semantic category (i.e. pure additive relation), (3) semantic requirements which are crucial in using er^S and er^{Comp}, and (4) related pedagogical implications for Chinese *er*. Finally, Section 7 concludes this paper.

2. ANALYSIS OF *ER*

Before the discussion of the conjunction *er*, it should be noted that we follow Chu's (2006) working definition of a Chinese clause/sentence: a Chinese clause contains at least one predicate, which is typically a verb phrase or an adjective phrase, though it may also be a noun phrase; a Chinese sentence comprises at least one clause or a block of clauses which are connected via explicit markers like conjunctions and adverbs, or by a topic chain (cf. Chu 2006: 270-290). For the study of our conjunction *er*, a clause is treated as the basic conjunct unit.

In this section, previous categorizations of Chinese *er* by Liu et al. (1996), Lü (1999) and Hong (2008) are reviewed and then followed by our categorization.

2.1 Previous Categorizations of *Er*

As mentioned above, *er* is used to conjoin adjective phrases, clauses and sentences/paragraphs, though not bare verbs (e.g. (2)) or noun phrases (e.g. (3)).

(2) *張三買而看了一本書。
     Zhangsan   mai *er* kan-le   yi   ben   shu.
     Zhangsan   buy ER read-ASP one CL   book
    'Zhangsan bought and read a book.'

128

(3) *張三買了一個蛋糕而一束花。

  Zhangsan   mai-le   yi   ge   dangao   ***er***
  Zhangsan   buy-ASP one CL   cake   ER
  yi-shu   hua.
  one-CL   flower
  'Zhangsan bought a cake and a bunch of flowers.'

In terms of the semantic relation between the two conjuncts of *er*, Liu et al. (1996) and Lü (1999) agree that *er* mainly marks **juxtaposition** (e.g. (4)), **continuation** (e.g. (5)) and an **adversative** relation (e.g. (6)), or it can pair with a preposition or conjunction (e.g. *yinwei* 'because') to indicate reason, cause, method, etc. (e.g. (7)), which will be called a 'correlative construction' henceforward. The following examples are taken from Lü (1999: 192-194)[3].

---

[3] In fact, Lü also lists two marked usages of *er*: (1) a conjunction of NPs (e.g. *you-**chun**-er-**xia** 'from spring to summer') and (2) a conjunction of NP and VP (e.g. ***Zuojia er bu xiezuo**, cheng-bu-shang zuojia* 'A writer who does not write cannot be called a writer.'), which is a non-typical usage of adversative. However, considering the fact that these usages are rarely introduced in textbooks for second language learners, we will not include them in our discussion.

(4) **Binglie (Juxtaposition)**
文筆簡練**而**生動。
Wenbi   jianlian   *er*   shengdong.
writing   concise   ER   lively
'(One's) writing is concise and lively.'

(5) **Dijin (Continuation)**
各組都取得了良好的成績，**而**以第三組的成績最為突出。
Ge zu   dou qude-le   lianghao de chengji,   *er*
each group   all get-ASP   excellent DE scores   ER
yi   di-san   zu   de chengji zui   weituchu.
by the:third group   DE score   most BE   prominent
'Each group has gained great scores, and the third group outperformed all others.'

(6) **Zhuanzhe (Adversative)**
幼苗早管理，費力小**而**收效大。
Youmiao zao guanli,   feili   xiao *er* shouxiao da.
sprout   early take:care effort   smallER   yield   great
'Take care of the seeding as early as you can; efforts to be made are less but the yield is great.'

(7) **Introducing clauses or phrases preceding the main verb, marking reason, cause, method, etc. (Correlative construction)**
我們絕不能**因為**取得了一些成績**而**驕傲自滿起來。
Women   juebu   neng *yinwei*   qude-le   yixie
we   by:no:means   can because   get-ASP   a:little
chengji   *er*   jiaoaoziman-qilai.
score   ER   arrogant:conceited-ASP
'We should by no means be arrogant and conceited for the little progress we made.'

The two studies, however, diverge in the classification of a specific usage of *er*, that is, a conjoining of an affirmative and a negative phrase or clause in form (Liu et al. 1996: 176 and Lü 1999: 192), which will be labeled as 'antithesis' in this study. For example,

(8) **Antithesis**
我喜歡住在台北，而不喜歡住在高雄。
Wo xihuan   zhu zai Taibei   ***er*** bu xihuan
I   like   live in Taipei   ER   not like
zhu zai Gaoxiong.
live in Kaohsiung
'I like to live in Taipei, but don't like to live in Kaohsiung.'

For Lü, an example like (8) is adversative, while for Liu et al., it is juxtaposition. Yet, a similar example which involves an affirmative phrase and a negative one, i.e. *Zhe zhang hua de secai yan er bu su* 'The color of this painting is gay ER not gaudy,' is categorized as adversative in Liu et al. (2003: 196). Obviously, clearer criteria for different categories will be needed.

Now consider Hong's (2008) analysis. Unlike Liu et al. (1996) and Lü (1999), who characterize *er* from a semantic perspective, Hong (2008) first categorizes the use of *er* based on its syntactic function (**adverb** and **conjunction**) and then subcategorizes the conjunction category based on Halliday's (1994) semantic relations between English clauses (or clause complexes). Furthermore, Hong (2008) points out that *er* can be used as a discourse marker, conjoining several clauses. Returning to the examples in (4)-(8) above, *er* in (4)-(6) and (8) is a conjunction, while *er* in (7) is an adverb. For the use of *er* as a discourse marker, please see Appendix 1.

Under the conjunction category, Hong decomposes semantic relations between conjuncts of *er*'s into 'dissimilar additive', 'similar additive' and 'adversative'. See Table 1 below (examples and descriptions taken from Hong (2008: 60)).

Table 1. Subcategorizations of the conjunction *Er* in Hong (2008)

Subcategories	Descriptions and examples
Er₁: Dissimilar additive	Er₁-₁: *Liang ju duili* (antithesis) 老王是老師而不是學生。 Laowang shi laoshi *er* bu shi xuesheng. Laowang BE teacher ER not BE student 'Laowang is a teacher; he is not a student.' Er₁-₂: *Shijian jie yi* (two different events) 老王已有了妻子，而老張還是單身。 Laowang yi you qizi *er* Laozhang Laowang already have wife ER Laozhang hai shi danshen. still BE single 'Laowang already has a wife, while Laozhang is still single.'
Er₂: Adversative	*Yu yuqi xiang zuo* (counter-to-expectation) 老王在老家已有了妻子，而老張卻還一直幫他介紹對象。 Laowang zai laojia yi you qizi Laowang at hometown already have wife *er* Laozhang que hai yizhi bang ta ER Laozhang instead still always help he jieshao duixiang. introduce mate 'Laowang already has a wife back home, but Laozhang keeps introducing a mate to him.'
Er₃: Similar additive	*Liang shi xiangsi* (two similar events) 老王已有了妻子，而老張也結了婚。 Laowang yi you qizi *er* Laozhang ye Laowang already have wife ER Laozhang also jie-le-hun. marry-ASP-marry 'Laowang already has a wife, and Laozhang has married, too.'

Crucially, within the conjunction category, Hong takes the addition of an event to be *er*'s central function, which may introduce a similar event (similar additive), a different event (dissimilar additive) or an event that is counter-to-expectation (adversative). This approach gives a finer categorization, compared to two previous studies. However, the line between the adversative *er* and the dissimilar additive *er* may not be as clear as Hong assumes; context may play a dominant role. We will come back to the distinctions of the three subcategories in the next section.

In this paper, we will adopt Hong's classification, primarily in treating *er* as an adverb, a conjunction and a discourse marker, and propose a revised categorization of *er* with elaborated semantic subcategories that are more applicable in sorting learners' data.

2.2 Our Categorizations of *Er*

Building on Hong's three-tier system of *er*'s, we categorize *er* into four major groups: a subordinator (er^S), a complementizer (er^{Comp}), a coordinator (er^C) and a discourse marker (er^D). Furthermore, for a better understanding of the use of *er*, we extracted 500 randomized instances from a native speakers' corpus (i.e. *Corpus-ASB*).

2.2.1 Subcategories of er^S

In the er^S category, various semantic relations are brought into reality via the use of *er* as well as a preposition/conjunction, which will be called a **correlator**. Two clauses are in effect in a subordinate-main clause relation. In an unmarked structure, there is only one overt subject on the surface form, usually in the first clause.[4] Given the fact that

---

[4] In some cases, a different subject may be used:

(i)   他**因為**他太太太傷心**而**吃不下飯。
       Ta   ***yinwei***   ta   taitai tai   shangxin   ***er***   chibuxia   fan.
       he   because   he   wife too   sad   ER   can:not:eat   food
'Because his wife was very sad, he couldn't eat anything.'

previous studies fall short in giving a definite set of correlators or indicating their frequencies, we studied the 500 randomized instances extracted from *Corpus-ASB* and identified 106 correlators (See Appendix 2), which were further subcategorized into three major groups: **cause, reason** and **peripheral** (includes method/manner/contingency), as illustrated in (9) to (11) respectively:

(9) **Cause**

他**因為**太傷心**而**吃不下飯。

Ta	*yinwei*	tai	shangxin	*er*	chibuxia	fan.
he	because	too	sad	ER	can:not:eat	food

'Because he was very sad, he couldn't eat anything.'

(10) **Reason**

他**為了**養家**而**同時做了兩份工作。

Ta	*weile*	yangjia	*er*	tongshi	zuo-le
he	in:order:to	raise:family	ER	simultaneous	do-ASP

liang fen gongzuo.
two CL work

'In order to raise the family, he has to do two jobs simultaneously.'

(11) **Peripheral (e.g. method)**

他**經**別人的介紹**而**認識了李小姐。

Ta	*jin*	bieren	de	jieshao	*er*	renshi-le
he	**via**	other:people	DE	introduction	ER	know-ASP

Li xiaojie.
Li Ms.

'He got to know Ms. Li via other people's introduction.'

2.2.2 Subcategories of *er*^{Comp}

What is common in *er*^{S} shown above is that there exists an overt correlator (i.e. *yinwei, weile, jing*) in the first clause. However, *er* can also occur without an overt correlator and work as a complementizer. The second clause is usually headed by or can be supplemented with a light verb such as *lai* 'come' or *qu* 'go', or a causative verb such as *rang*

'make'. The semantic realization between the main clause and the complement is labeled here as **purpose** (e.g. (12)):

(12)　**Purpose**
他希望能贏得金牌**而**讓父母感到驕傲。

Ta	xiwang	neng	yingde	jinpai	*er*	rang
he	wish	can	win	gold:medal	ER	make

fumu	gandao	jiao'ao.
parents	feel	proud

'He wishes he can win a gold medal to make his parents feel proud.'

Similar to er^S, there is usually one overt subject dominating the predicate in each clause within a single complex event. The clause complex can be sequentially decomposed to a clause with background information, followed by another profiled as a foreground (cf. Hsueh 1991). For this reason, though the two clauses in (12) are seemingly two related events being added, the latter is actually more 'weighted' than the former. That is, his wish to win a gold medal is for the purpose of making his parents proud of him. Diachronically, the purpose *er* and er^S may have developed from *er* being a central coordinator for almost all syntactic categories to an adverbial marker or a complementizer undergoing 'Conjunctive Reduction' process (Mei 2002, Tsai 2008).

2.2.3 Subcategories of er^C

Clauses conjoined by er^C are independent of each other in a loosely-tied relation, which coincides with the features of *accidental coordination* (cf. Dalrymple & Nikolaeva 2006). The clause complex in accidental coordination is assumed to be rephrasable by adding certain adverbs in the second clause such as *ye* 'similarly' or *que* 'however'. Crucially, these sentential adverbs may facilitate the clause complex to profile just one semantic relation when there are ambiguities; without them, er^C may have two readings at the same time (cf. Li 2012).

Basically, we assume the fundamental functions of er^C to be **additive** ($er^{C\text{-}A}$) and **contrastive** ($er^{C\text{-}C}$) (cf. Qiu 2005, Yan 2009). The former corresponds to Hong's (2008) similar additive, while the latter,

dissimilar additive and adversative. In the rest of this subsection, we give some typical instances to illustrate how we distinguish different semantic relations under $er^{C\text{-}A}$ and $er^{C\text{-}C}$ respectively.

2.2.3.1 Semantic relations in $er^{C\text{-}A}$

$Er^{C\text{-}A}$ adds one event to another for a **pure additive** relation such as 'buying a house' and 'buying a car' in (13); crucially, the additive adverb *ye* 'also' here is indispensable. On the other hand, the **continuation** *er*, which can conjoin two clauses without the gluing of an adverb, extends the given proposition via connecting two topics such as 'twelve months' and 'February' in (14) (topics here are understood semantically).

(13) **Pure additive**

張三買了房子，**而**也買了車子。

Zhangsan mai-le   fangzi   ***er***   ye   mai-le   chezi.

Zhangsan buy-ASP house   ER   also   buy-ASP   car

'Zhangsan bought a house and he also bought a car.'[5]

(14) **Continuation**

一年有十二個月，**而**二月有二十八天。

Yi   nian   you   shi'er   ge   yue,   ***er***

one   year   have   twelve   CL   month   ER

eryue   you   ershiba   tian

February   have 28   day

'There are twelve months in a year and 28 days in February.'

2.2.3.2 Semantic relations in $er^{C\text{-}C}$

    The second major subcategory in er^C introduces a contrastive relation. By 'contrastive', we follow Mann & Thompson (1992: 37) in defining it as a relation: (a) comprehended to be the same in many respects, (b) comprehended as differing in a few respects, and (c) compared with respect to one or more of these differences'. Finally, we include in our $er^{C\text{-}C}$ the semantic relations of **neutral contrast** (e.g. (15)), **antithesis** (e.g. (8), repeated here as (16)) and **adversative** (e.g. (17)).

---

[5] One of the reviewers noted that this sentence was unacceptable even with the occurrence of the adverb *ye*. We did, however, find some similar instances in *Corpus-ASB*. Here is one of them:

(ii)   各國在科學的基礎下，各有所長，**而**也互相交流…

      Ge   guo   zai   kexue   de   jichu xia,   ge   you

      each   country   at   science   DE   basis   under each   have

      suo   chang ***er***   **ye**   huxiang   jiaoliu…

      PTC   forte   ER   also   mutually   exchange

      'On basis of science, each country has its own forte and they also exchange (ideas) with each other…'

For the simplicity of demonstrating each subcategory, we will keep the original example.

(15)　**Neutral contrast**
台北冬天很冷，**而**高雄夏天很溫暖。
Taibei　　dongtian　hen　leng　***er***　Gaoxiong dongtian
Taipei　　winter　　very cold ER　Kaohsiungwinter
hen　wennuan
very　warm
'In winter, it is cold in Taipei and warm in Kaohsiung.'

(16)　**Antithesis**
我喜歡住在台北，**而**不喜歡住在高雄。
Wo　xihuan　　zhu zai Taibei　　***er***　bu　xihuan
I　　like　　　live in　Taipei　　ER　not　like
zhu　zai　Gaoxiong.
live　in　Kaohsiung
'I like to live in Taipei, but don't like to live in Kaohsiung.'

(17)　**Adversative**
張三很喜歡小孩，**而**他太太卻跟他相反。
Zhangsan hen　xihuan　　xiaohai　***er***　ta　　taitai
Zhangsan very like　　　children ER　he　　wife
que　　　gen　ta　xiangfan.
however　with he　opposite
'Zhangsan likes children very much but his wife is the opposite
of him.'

What *er* contrasts can be two events with two different subjects as 'Taipei' and 'Kaohsiung' in (15), and the reading is no more than a pure comparison; *er* can also contrast the actions done or states possessed by the same subject, usually by contrasting polarity items such as 'like' and 'dislike' in (16). Furthermore, *er*$^{C-C}$ can even express surprise or something unexpected. What is contrasted in (17) is 'Zhangsan's fondness of children' and 'Zhangsan's wife's dislike of children' with the baseline expectation that 'his wife likes children too'. Note that the

adverb *que* ('however') in (17) is especially crucial in yielding the adversative reading; alternatively, the context will play a dominant role, provided the information concerning the same shared expectation is explicitly given.[6]

One can see that there is an increasing degree of contrastiveness from (15) to (17), and the availability of the adversative reading depends either on the help of adverbs, or on the extent the interlocutors hold the same assumption in mind. In sorting our learners' data, we thereby turn to polarity items, adverbs and the context within this subcategory.

2.2.4 Subcategorization of *er*[D]

What has been addressed so far is conjoining at the clause level. Larger units can be combined by *er*, too. For *er*[D], a clause complex as a whole or a paragraph becomes the basic discourse unit. Unlike the less formal use of *and* in English academic writing (Dorgeloh 2004), *er*[D] is prevalent in formal Chinese writing or in formal speech. The discourse function of *er*[D] is akin to the Japanese continuative conjunction *ga*, which suggests inferences between two discourse units, assigning an interpretation that is relevant to a given context (Fukushima 2005)[7].

---

[6] Sentences without an adverb or a reasonably given context can be ambiguous in between adversative and neutral contrast reading. For example, if the adverb *zhi* 'only', which contributes to the adversative reading in this case, is removed, a neutral contrast of promotions between two people is also plausible.

(iii) 張三升了經理，**而**李四**只**升了組長。
    Zhangsan   sheng-le   jingli   ***er***   Lisi   **zhi**
    Zhangsan   promote-ASP   manager   ER   Lisi   only
    sheng-le   zuzhang.
    promote-ASP   section:chief
    'Zhangsan was promoted to the position of a manager but Lisi only to a section chief.'

In the actual counting in this study, when neither the adverb signaling adversative existed nor the context showed any sense of counter-to-expectation, the use of *er* was categorized as neutral contrast.

[7] The conjunction *er*, in accordance with most other languages, is derived from an older

Serving as a link to a previous discourse categories unit in an abstract manner, er^D can be subsumed under **contrastive** and **continuation**. Similarly, one can also find either the context or adverbs/adverbial phrases helpful in distinguishing these two subcategories. While *ling-yi-fangmian* 'on the other hand' signals a contrastive relation, *geng-jin-yi-bu* 'furthermore' signals a relation of continuation. Below is an example of er^D for a continuation reading; for a contrastive example, please see Appendix 1.

(18)   *er^D*: **continuation**

張三在北京碰到了暴風雪，所以回不了家。而他的家人因為聯絡不到他，他們都很擔心。

Zhangsan zai   Beijing   pengdao-le   baofengxue,
Zhangsan be   Beijing   encounter-ASP blizzard
suoyi   huibuliao   jia   Er ta de
so   can:not:return home   **ER** he DE
jiaren   yinwei   lianlubudao   ta,   tamen
family:member because   can:not:contact he,   they
dou hen danxin.
all very worried

'Zhangsan encountered a blizzard in Beijing and he couldn't go home **and** because Zhangsan's family couldn't contact him, they felt very worried.'

2.2.5 Summary

Our categorization of *er*'s for sorting learners' data is summarized in Table 2:

---

usage as a demonstrative pronoun (cf. Lan 1990, Ramat & Mauri 2011). When it functions as a discourse marker, it is arguably thought to be used to connect two discourse units by a referential linking (cf. Blühdorn 2008).

Table 2. Categorization of *er* for learner language analysis

Categorization	Semantic Relations		Examples
er^S		cause	(9)
		reason	(10)
		peripheral	(11)
er^Comp		purpose	(12)
er^C	*er*^C-A	pure additive	(13)
		continuation	(14)
	er^C-C	neutral contrast	(15)
		antithesis	(16)
		adversative	(17)
er^D		continuation	(18)
		contrastive	Appendix 1

3. RESEARCH QUESTIONS

Our study investigates how L2 learners combine different types of clauses in a grammatical and coherent manner via *er*. Binding two clauses together is undoubtedly more demanding due to the fact that the combination of clause types has to be legitimate for a given syntactic and semantic structure. Three research questions are addressed as follows:

(1) How do learners exploit *er* to conjoin conjuncts of different linguistic units? Do they show any preference to using *er* for specific semantic categories in comparison with native speakers?

(2) What error types can be identified and categorized in our learner corpus? Can the *er/erqie* problem also be found? And to what extent can these two be differentiated by learners?

(3) How can learners' correct and erroneous uses of *er*s tell us about the restriction of its usage and related pedagogical implications?

4. METHODOLOGY

This study adopts a corpus-based approach (cf. Ellis & Markhuizen 2005, Gries 2008), comparing the data from a learner corpus and those from the *Academia Sinica Balanced Corpus of Modern Chinese* (*Corpus-ASB*). The learner corpus we used is *Learner Corpus - Test of Chinese as a Foreign Language (Learner Corpus - TOCFL)*, which was established by the Mandarin Training Center of National Taiwan Normal University in July 2010. By June 2011, 3,128 pieces of written texts with 1,002,710 Chinese characters had been collected mostly from the computer-based exam TOCFL (The Test of Chinese as a Foreign Language) (Zhang 2012). What deserves to be mentioned is that the stratification of language proficiency in *TOCFL* is based on the *Common European Framework of Reference (CEFR)*, including levels A2, B1, B2

and C1.[8] Each examinee may choose to take part in the exam of a specific level and only the data of those who passed the level by scoring 3 points or above on a 0-5 point scoring scale were taken into consideration in this study.

As for the makeup of the examinees' first languages, the examinees whose L1s are Japanese (25%), English (17%), Korean (13%), Indonesian (11%) and Vietnamese (7%) account for most of the data. Besides, the amount of data collected in each proficiency level is not evenly distributed, that is, B1 has the most data (53.64%), followed by B2 (20.24%), A2 (16.84%) and C1 (9.28%).

Given the fact that the conjunction *er* is taught at a relatively late stage as evidenced in the textbooks used at the Mandarin Training Center[9], we only concentrate on the B1, B2 and C1 data, without including those from A2. For comparison, 500 instances of *er* were extracted from *Corpus-ASB*. The functions of *er* along with their frequencies are identified and sorted; the general distribution of different uses by learners or native speakers can then be characterized.

In addition, learners' data are also compared and analyzed for errors or unconventional uses. To establish a criterion for errors, we follow Lennon's (1991) definition of an error, which is: '*…a linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speakers' native speaker counterparts*' (Lennon 1991: 182). On top of this, we took into consideration the comprehension end by rephrasing the definition as follows:

---

[8] The proficiency levels used in *TOCFL* are based on the classification of the *Common European Framework of Reference (CEFR)*, in which the beginner level corresponds to A2, basic to B1, intermediate to B2 and advanced to C1; those who want to pass the exam will have to score 3 points as the minimum on a 0-5 point scoring scale. *TOCFL* is conducted via computer under timed conditions. Typos pertaining to homophones are excluded in our analysis.

[9] *Er* first appears in Book 4 Lesson 3 (B4L3) in *Practical Audio-Visual Chinese 2nd edition* (2011), B2L14 in *Integrated Chinese* (2006) and B4L39 in *New Practical Chinese Reader* (2004).

A linguistic form or combination of forms which, under any condition, would not be processed smoothly for the subsequent discourse unit, and block the immediate comprehension by a native speaker.

Therefore, errors of *er's* in this study are identified in a more stringent way. In cases where learners' uses of *er* were at the borderline of acceptability, we further resorted to three native-speaker Chinese teachers for a re-examination according to the error definition indicated above.

5. RESULTS

There were, in total, 518 concordances of *er* in the learner corpus-TOCFL after 40 instances of typos and those used in idiomatic expressions (e.g. *zong-er-yan-zhi* 'generally speaking') were filtered out. Among them, 441 were judged to be correct uses, with the remaining 77 being somehow deemed non-target like or inappropriate. In this section, the frequency distribution of each subcategory of *er* in the two corpora is given and the error types are categorized and exemplified.

5.1 Frequency Distribution of *Er*'s in L1 and L2 Corpora

Frequencies of the four main categories of *er* used by non-native speakers (or NNS) and native speakers (or NS) are given in Table 3.[10] Interestingly, while er^S is found to be used as often as er^C by learners (i.e. 42.3% vs. 43.2%), the counterparts in NS data exhibit a predominant use of er^C over er^S (i.e. 50.4% vs. 21.2%). Moreover, native speakers strike a balance between er^S (21.2 %) and er^D (22.8%), and their use of er^D is nearly three times as much as that in NNS data (8.3%).

---

[10] The data here contain both the correct and the non-target like uses by the learners. When only the correct uses were counted, the result was: er^S: 186 (41.9%), er^{Comp}: 31 (7%), er^C: 184 (41.4%), and er^D: 43 (9.7%). Er^S and er^C still account for most of the data.

Table 3. *Er'*s used by non-native speakers and native speakers

	Non-native speakers		Native speakers	
	instances	percentages	instances	percentages
er^{S}	219	42.3%	106	21.2%
er^{Comp}	32	6.2%	28	5.6%
er^{C}	224	43.2%	252	50.4%
er^{D}	43	8.3%	114	22.8%
Total	518	**100%**	500	**100%**

More specifically, in Table 4 below, learners' data are arranged by our four major categories with their subcategories in comparison to the frequency of native speakers'.

Table 4. Frequency distribution of *er'*s used by NS and NNS

Categorization	Semantic relations	B1	B2	C1	Total	%	NS	%
er^S	cause	44	67	23	**134**	**25.9%**	27	**5.4%**
	reason	18	38	9	**65**	**12.5%**	22	**4.4%**
	peripheral	3	12	5	**20**	**3.9%**	57	**11.4%**
Subtotal		**219**		**(42.3%)**			106	**21.2%**
er^{Comp}	purpose	13	13	6	**32**	**6.2%**	28	**5.6%**
Subtotal		**32**		**(6.2%)**			28	**5.6%**
er^{C-A}	pure additive	43	51	15	**109**	**21%**	115	**23%**
	continuation	7	9	17	**33**	**6.4%**	44	**8.8%**
er^{C-C}	neutral contrast	3	3	3	**9**	**1.7%**	14	**2.8%**
	antithesis	11	27	4	**42**	**8.1%**	62	**12.4%**
	adversative	8	11	12	**31**	**6 %**	17	**3.4%**
Subtotal		**224**		**(43.2%)**			252	**50.4%**
er^D	continuation	9	7	11	**27**	**5.2%**	82	**16.4%**
	contrastive	4	6	6	**16**	**3.1%**	32	**6.4%**
subtotal		**43**		**(8.6%)**			114	**22.8%**
Sum		**163**	**244**	**111**	**518**	**100%**	**500**	**100%**

146

Patently, in learners' and native speakers' data, er^{C-A}, specifically the pure additive, takes the vast majority (21% vs. 23%). However, in terms of the discrepancy between two corpora, it is worth-noticing that while **the cause *er*** ranks the highest in learners' data (i.e. 25.9%), this is not the case in native speakers' (i.e. 5.4%); a similar case is found in **the reason *er*** (12.5% vs. 4.4%). On the other hand, what seems to characterize native speakers' data is a relatively greater use of **the peripheral *er*** in addition to **the discourse *er'*s** as just mentioned. In a nutshell, what distinguishes native speakers' use of *er* from the learners' is for one thing, that among the three subcategories of er^S, a variety of peripheral correlators for method (e.g. *yi*) or contingency (e.g. *sui*) are found to be more commonly used with *er* by L1 speakers, and for another, that a larger linguistic unit such as a sentence, instead of a dependent clause, is seen more often to be conjoined with *er*.

On the other hand, in a cross-sectional view of learners' language development, the data of *er'*s across three proficiency levels (i.e. B1, B2 and C1) with reference to L1's (e.g. the right most bar without patterns) are shown in Figure 1 (note: the categories are arranged in a descending order according to native speakers' use) :

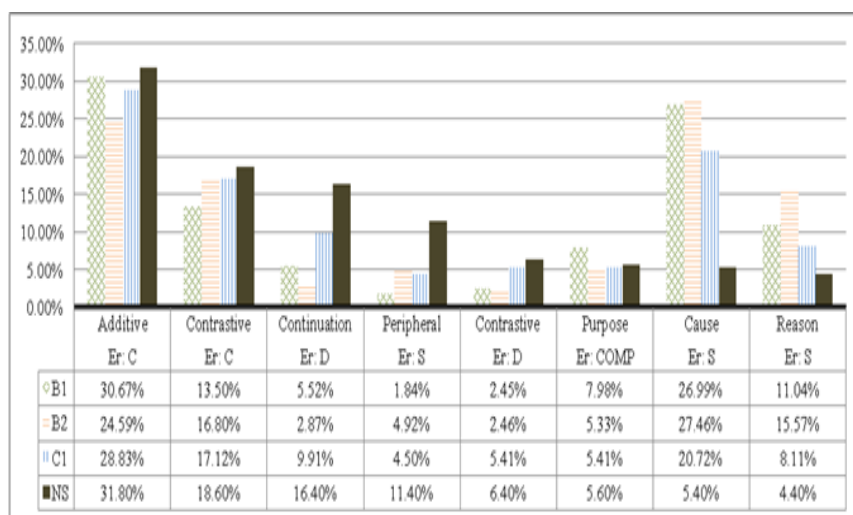|	Additive Er: C	Contrastive Er: C	Continuation Er: D	Peripheral Er: S	Contrastive Er: D	Purpose Er: COMP	Cause Er: S	Reason Er: S
B1	30.67%	13.50%	5.52%	1.84%	2.45%	7.98%	26.99%	11.04%
B2	24.59%	16.80%	2.87%	4.92%	2.46%	5.33%	27.46%	15.57%
C1	28.83%	17.12%	9.91%	4.50%	5.41%	5.41%	20.72%	8.11%
NS	31.80%	18.60%	16.40%	11.40%	6.40%	5.60%	5.40%	4.40%

Figure 1. Frequencies of learners' use of *er*s by proficiency levels

147

As shown in Figure 1, while learners' use of the subcategories of additive and contrastive in er^C exhibits a similar distribution to native speakers', what seems to rise according to the proficiency level is that of **peripheral** (from 1.84% in B1 to 4.5% in C1) in er^S, and those of continuation (from 5.52% to 9.91%) and contrastive (from 2.45% to 5.41%) in er^D. But as will be discussed below, while learners across the three levels expressed lots of the additive semantic relation in er^C, or more precisely, the pure additive, the error rate in B1 is as high as 48.8% (21 out of the total 43 pure additive uses).

5.2 Our Categorizations of *Er*

Among the total 518 instances of *er*, 74 (i.e. 14.3%) are assumed to be non-target like or inappropriate. Most of them fall into the main categories of er^C and er^S, within which **pure additive** (6.8%) and **cause** (4.8%) prevail (See Table 5). No errors were found in the use of er^D. However, learners' being free of errors in using er^D does not necessarily mean they acquire its usage perfectly; instead, it may be under disguise of avoidance or underuse, which is evident especially if we compare the use of er^D by L1 speakers.

Table 5. Non-target like occurrences and percentage of *er*'s

Categories	Semantic relations	B1	B2	C1	Total	% in all uses
Er^{C-A}	Pure additive	21	11	3	35	6.8%
Er^S	Cause	17	7	1	25	4.8%
Er^S	Reason	5	0	1	6	1.2%
Er^{C-C}	Antithesis	1	4	0	5	1.0%
Er^S	Peripheral	0	2	0	2	0.4%
Er^{Comp}	Purpose	1	0	0	1	0.2%
Er^{C-A}	Continuation	0	0	0	0	0
Er^{C-C}	Neutral contrast	0	0	0	0	0
Er^{C-C}	Adversative	0	0	0	0	0
Er^D	Continuation	0	0	0	0	0
Er^D	Contrastive	0	0	0	0	0
Sum		45	24	5	74	14.3%

For clause conjoining, pure additive in er^{C-A} (6.8%) and cause in er^S (4.8%) rank the two highest categories among all errors. In non-target like uses of the pure additive *er*, it involves a missing associative component, which explicitly marks the relation between two clauses, or in other cases it should be replaced by *erqie* in the construction *budan…er*(qie)* 'not only…but also'. As in cases of the cause *er*, the errors are in relation to the ungrammatical presence of a second subject in the second clause (Clause 2) in contrast with that in the first one (Clause 1). See Table 6 below (note: a target form (TF) reflects a desirable substitute form for learners' use of *er*; a target domain (TD) is the specific semantic relation involved):

Table 6: Error types of *er*

Error types	Descriptions	
Misuse	misuse of *er* for another conjunction due to similar additive function	
	Instances (%)	TF/TD
	32 **(43.2 %)**	**TF:** *erqie* ('furthermore') accounts for the most (others are like *bingqie* ('and' for coordinative VPs), *he* ('and' for NPs) or *huo* ('or')) **TD:** er^C (pure additive, continuation)
	misuse of *er* for *erqie* in the correlative pair construction	
	Instances (%)	TF/TD
	7 **(9.5 %)**	**TF:** *budan…erqie* 'not only…but also' **TD:** er^C (pure additive)
Missing	missing a correlator in Cause 1 or missing a causative/ light verb, a mirrored component or an additive adverb in Clause 2	
	Instances (%)	TF/TD
	25 **(33.8 %)**	**In Clause 1:** **(a) a correlator-** **TF:** *yinwei* 'because', *weile* 'for', *cong* 'from' **TD:** er^S (cause, reason, peripheral)

		In Clause 2:
		(a) a causative verb or a light verb
		TF: *rang* 'make', *lai* 'come'
		TD: er^{Comp} (purpose)
		(b) a mirrored component
		TF: *neng* 'can', *yinggai* 'should', *shi* 'be'
		TD: er^C (antithesis)
		(c) an additive adverb
		TF: *ye* 'also', *hai* 'still', 'even'
		TD: er^C (additive)
Structural mismatch	a different subject in Clause 2 for er^S	
	Instances (%)	TF/TD
	10 **(13.5 %)**	**TF:** *yinwei…suoyi* 'because…so' **TD:** er^S (cause)
Total	**74** **(100%)**	

Learners' errors are further categorized into three types and are exemplified as follows.

Type A Misuse: *Er* was used for another conjunction characteristic of a similar additive function (e.g. *he, erqie,* and *bing*), which accounted for the most errors (i.e. 43.2%). For example, *er* was overgeneralized to conjoin two NPs as in (19):

(19) *世界上除了這件事以外，很多人或事情讓你覺得難過，還有他們會試試看你的意志、信、**而**心情。

       Shijie   shang   chule   zhe jian shi   yiwai,
       world   on   beside   this CL thing   exclusive,
       hen duo   ren   huo shiqing   rang   ni
       very many   people   or thing   make   you
       juedenanguo,   haiyou   tamen   hui shishikan
       feel broken:hearted also   they   will test:a:little
       ni   de   yizhi,   xin,   ***er***   xinqing.
       you   DE   will,   faith ER   mood

       'In this world, beside this thing, there are many people or other things that make you feel broken-hearted; in addition, they will also test your determination, faith and mood.' (Korean, B2, 3 points)

Besides, *er* was also used as a reduced form for *erqie* in the correlative pair construction ***budan***…***erqie*** 'not only…but also', as in (20):

(20) *我們外國人來說唱歌是最好的國語表演。**不但**不需要什麼材料，**而**大家都可以參加。

       Dui women   waiguoren laishuo,   change   shi   zuihao
       for we   foreigner to:speak   sing:song BE the:best
       de   guoyu   biaoyan.   ***Budan***   bu   xuyao
       DE Chinese   performance.   not:only   not   need
       whatever cailiao,   ***er***   dajia   dou keyi canjia.
       sheme   material ER everyone all   can   participate

       'For foreigners like us, singing is the best performance in Chinese. Not only does it not need any props, but it is easy for everyone to participate.' (Japanese, B1, 3 points)

In fact, most of the errors that occurred in pure additive in $er^{\text{C-A}}$ were what previously were assumed to be misused with *erqie*, as in (21), which we will argue to involve more than just a substitution in Section 6.2.1.

(21) *我得到他的簽名,我跟他握手,**而**我和他一起拍照。

      Wo dedao   ta   de   qianming, wo gen ta
      I   get   he   DE   signature I   with him
      woshou,   ***er***   wo he ta   yiqi   paizhao.
      shake:hand   ER   I   with him together take:picture

      'I got his signature, I shook hands with him and I took a picture with him.' (Japanese, B1, 5 points)

Type B Missing: **a correlator** in Clause 1, or **a verbal element** or an **associative component** in Clause 2 is missing, degrading the sentence acceptability (33.8%). The errors centered on the TD of er^{S} in general, **antithesis** in $er^{\text{C-C}}$ as well as **pure additive** in $er^{\text{C-A}}$. In the errors found in er^{S}, it seems that learners simply grafted one clause onto another as in (22), leading to an obscure reading and decoding difficulties. Note that it could be saved by adding an explicit correlator such as *yin* 'because of' to enhance the logical relation between the two clauses.

(22) ?我經常在街上看到許多流浪的狗**而**感到難過。

      Wo jingchang zai jie   shangkandao xuduo
       I   often   at   street   on   see many
      liulanggou   ***er***   gandao   nanguo.
      stray:dog   ER   feel   sorrow
      'I often see there are many stray dogs on the street and feel sorrowful (because of their situations).' (Korean, C1, 3 points)

Similarly, the use of a parallel structure can help the acceptability of a sentence. A case in point is given in (23). The learner in this example failed to provide a corresponding copular verb *shi* in Clause 2 to contrast with *bu shi* in Clause 1. Such a parallel structure is required for expressing the **antithesis** relation; gapping between two clauses (i.e.

'She is not that kind of person…' vs. 'She *(is) the kind of person …') is unacceptable in the formation of an antithesis structure.

(23) *我保證我所經(註：見)過的美英**不是**那樣的人，**而**___喜歡
幫助別人的人…。

Wo	baozheng	wo	suo	jing-guo	de	Meiying	bu	shi
I	promise	I	PTC	pass-ASP	DE	Meiying	not	BE

nayang	de	ren,	*er*	(shi)	xihuan	bangzhu
that:kind	DE	person	ER	BE	like	help

bieren	de	ren…
other:people	DE	person

'I guarantee that Meiying whom I passed (note: knew) is not that kind of person, but is a person who likes to help other people.'
(Korean, B2, 3 points)

Finally, the missing of an associative element concerns the pure additive relation and it is entangled with another conjunction *erqie*, which we will discuss in Section 6.2.1.

Type C Mismatch: a clause complex occurred with a second different subject in Clause 2 in the TD of cause or reason, rendering a **mismatch** (13.5%). Sentential comprehensibility in (24) was largely degraded due to the incongruous second subject 'I'. Such a use of *er* aroused competing readings for the fact that the correlator (i.e. *yinwei*) marked a cause relation, yet the use of the two different subjects was reminiscent of a contrastive one.

(24) *大家好！**因為**你們下個月來台灣**而**我已經做一個旅行計畫。

Dajiahao,	*yinwei*	nimen	xia	ge	yue
Everyone	hello because	you	next	CL	month

lai	Taiwan	*er*	wo	yijing	zuo	yi	ge
come	Taiwan	ER	I	already	make	one	CL

lüxing	jihua.
travel	plan

'Hello, everyone, because **you** will come to Taiwan next month, **I** have already made a travel plan.'
(Korean, B1, 3 points)

6. DISCUSSION

L2 learners' uses of *er*s have been examined under our categorization and their data have exhibited divergence from that of L1s. In this section, how learners exploit different functions of *er*s for meeting Chinese coherence or **texuality** (cf. Chu 2006) is first discussed. In addition, error types, especially those in additive and cause domains, are further analyzed. Finally, pedagogical implications are discussed.

6.1 The Use of *Er* for Structuring Chinese Texts: NS vs. NNS

As can be observed from the frequency distribution in the NS corpus, *er* is about equally distributed between er^S (21.2%) and er^D (22.8%), with er^C (50.4%) being the most dominant. In contrast, learners tend to use er^S (42.3%) to conjoin clauses with a subordinate-main connection at an intra-sentence level, especially in marking a cause relation (25.9%); conversely, they are less likely to use *er* to conjoin sentences/paragraphs, yielding only 8.6%. Consequently, learners develop along the spectrum from er^S to er^C and then to er^D and finally become more like native speakers in using er^D in organizing sentential units.

Let us now focus on the semantic relations that are equally distributed in the two corpora. That is, **additive** (er^{C-A}) and **contrastive** (er^{C-C}), which comprise nearly half of the entire data (e.g. 43.2% in learners' data and 50.4% in L1's). Functionally speaking, er^{C-A} conjoins events that are not as closely related as those in er^S. It profiles the sameness of two events, with the first event either added (i.e. pure additive) or further elaborated (i.e. continuation) by the second event. On the other hand, er^{C-C} contains two events but profiles their differences. Furthermore, it involves speakers' evaluation and its sentential semantic focus may be argued to fall on the second clause (i.e. by saying 'I want to drink coffee but I do not want to drink tea' as in the antithesis, the speaker may imply that tea is not on his wish list; otherwise, there is no need in telling the speaker what he/she doesn't want). Conceivably, the coordinator *er* can be deemed as the prototypical use in structuring Chinese texts.

In short, Chinese texuality can be functionally realized as various semantic relations by the connective device *er*. However, since *er* does not bear much semantic load in itself, it may be inclined to recruit more linguistic devices for delivering a specific interclausal semantic relation by narrowing down many a potential reading. Thus, as we shall see in the following subsections, the fact that different functions of *er*s may be divergent in their preference for specific linguistic representations may indicate a more demanding task for learners to use it correctly.

6.2 Learners' Errors

In identifying learners' errors of *er* (i.e. misuse, missing, and mismatch), errors such as that NPs should not be conjoined by *er* are less controversial (i.e. misuse), whereas errors such as a replacement by *erqie* cannot be easily characterized. In the following section, the dilemma of choosing *er/erqie* is first discussed, followed by the learners' ignorance of the unique semantic structure required by er^S and finally the related pedagogical suggestion.

6.2.1 Choice of *er* and *erqie* in denoting pure additive semantic relation

Previous studies have been keen on pointing out that in the non-target like instances, *er* should be replaced by *erqie* without, however, pinpointing to what extent these two conjunctions are overlapped. Our analysis shows that this type of learner error exclusively belongs to the pure additive domain. For example, in the previous instance (21) made by a learner (re-indexed here as (25)):

(25)　*我得到他的簽名，我跟他握手，**而**我和他一起拍照。
　　　Wo dedao　　ta　de　qianming, wo　gen　ta　　woshou,
　　　I　 get　　　he　DE　signature　I　　with　him　shake:hand
　　　er　_①_ wo　_②_ he　ta　yiqi　　paizhao.
　　　ER　　　I　　　　with him　together　take:picture
　　　'I got his signature, I shook hands with him and I took a picture with him.' (B1)

Besides the replacement of *er* by *erqie* (e.g. position ①), the sentence can still be saved by adding an additive adverb such as *hai* 'furthermore' (e.g. position ②) in the last clause. What the additive adverb (or a conjunctive adverb as called by Liu (1996)) does here is to mark the relation explicitly through association. Consequently, the lack of such an associative element may be the cause rather than an ostensible confusion with *erqie*. Since both *er* and *erqie* can convey an additive semantic relation, simply correcting learners' infelicitous use of *er* with another conjunction does not help much in explaining how it works in conjoining clauses.

To see how the use of the additive adverb *ye* interacts with different conjunctions (e.g. zero coordinator, *erqie* and *er*), consider the examples in (26) to (28).

(26) 張三買了房子，也買了車子。
Zhangsan mai-le   fangzi ,   ***ye***   mai-le   chezi.
Zhangsan buy-ASP house   ALSO   buy-ASP   car
'Zhangsan bought a house; he also bought a car.'
(a zero coordinator, an additive adverb ye)

(27) 張三買了房子，而且(也)買了車子。
Zhangsan mai-le   fangzi   ***erqie***   ***(ye)***   mai-le
Zhangsan buy-ASP house   ERQIE   ALSO   buy-ASP
chezi.
car
'Zhangsan bought a house and he also bought a car.'
(conjunction *erqie*)

(28) 張三買了房子，而也買了車子。
Zhangsan mai-le   fangzi   ***er***   ***ye***   mai-le   chezi.
Zhangsan buy-ASP house   ER   ALSO   buy-ASP   car
'Zhangsan bought a house and he also bought a car.'
(*er*^{C}, an additive adverb *ye*)

157

In (26), the clauses are conjoined by the zero coordinator, which is assumed to be the default linkage form in Chinese (cf. Chao 2002). The additive adverb (e.g. *ye* 'also') in this case helps to signal the association between two events. In (27), *erqie* is used to conjoin clauses. *Ye* is optional, since the conjunction manifests itself clearly enough in expressing its additive function. Finally, in (28), *ye* is a must when *er* is used.

Tsai's (2006: 47) analysis of *erqie* may help illustrate the same idea. In her analysis, *erqie* presupposes **informativeness** by narrowing down or by controlling the set of possible worlds; simply put, either an adequate propositional strength or an associative adverb, such as *ye* ('also'), is vital for speakers to manage the possible worlds grammatically. In a similar way, *er* requires a contextualization cue. In this case, an additive adverb is preferred in helping to fulfill the pure additive relation. When such cues are nonexistent, it leads to an infelicitous use of pure additive *er*, which native speakers might be inclined to replace with *erqie*, for the sake of enhancing the proposition strength, given that additive adverbs are optional in clauses conjoined by *erqie*.

Generally speaking, while *er* may seem to be equivalent to *erqie* in the pure additive domain, the former requires the help of an additive adverb to control the relation between two conjuncts, which should not be surprising given the fact that *er* can also be used for other semantic relations.

6.2.2 Semantic structure for er^{S} and er^{Comp}

Clausal conjuncts must be tailored in order to fit into specific syntactic and semantic relations required by different categories of *er*. This is especially true for er^{S} and er^{Comp} since there is a strong tendency for only an overt subject to be present. When learners failed to unite two clauses together followed by such a constraint as in the example (24) (repeated as (29) below), the superficial sentence form (i.e. neutral contrast) contradicts its intended meaning (i.e. cause), rendering a competition between two possible readings.

(29) *大家好！因為你們下個月來台灣而我已經做一個旅行計畫。
　　　Dajiahao, ***yinwei***　nimen　 xia　 ge　 yue
　　　Everyone hello because　 you　　　 next CL　 month
　　　lai　　 Taiwan　 ***er***　 wo yijing　　 zuo　　 yi　 ge
　　　come　 Taiwan　 ER　 I　 already　　 make　　 one CL
　　　lüxing　　 jihua.
　　　travel　　 plan
　　　'Hello, everyone, because **you** will come to Taiwan next month, **I**
　　　have already made a travel plan.'
　　　(Korean, B1, 3 points)

The sentence can be remedied by either moving the subject of Clause 2 (i.e. 'I') to the beginning of the clause complex or by using another functionally similar construction *yinwei…suoyi…*(literally 'because…so'), which is capable of accommodating two subjects within its scope (i.e. *yinwei* Subject1+V1, *suoyi* Subject2+V2), though a change of a causal construction may more or less influence the compactness of the entire sentence.

In Baker & Stewart's (2002) analysis of Consequential Serial Verb Construction (CSVC), each of two verbs introduces a distinct subevent in the overall composite event, which is similar to the function of purpose *er*. To some extent, Clause 2 in all of the subcategories in *er*[S] and in *er*[Comp] may be deemed as indicating a result or a consequence, only differing in whether a correlator is taken. As a consequence, two clauses work together to indicate a single composite event within which the very same subject is either the agent or the patient of two VPs. Given that Chinese does not mark the identity of the subject via morphosyntactic devices (cf. Lehmann 1988), the constraint for presence or absence of the subject is relatively subtle for learners to be aware of.

All in all, different uses of *er* are subject to syntactic constraints or inclined to have certain semantic/discourse preferences. Crucially, each use of *er* may be argued to have an optimal representation in terms of a given semantic structure, and the more the structural demand of a use is satisfied (i.e. by means of contextualization cues), the more it approaches native speakers' prototypical usage.

6.2.3 Pedagogical implications for *er*

From the empirical data discussed above, a picture emerges of learners' general performance in using Chinese *er*. At the beginning stage, learners across three language proficiency levels show a remarkable preference in using *er*^{S}, especially in marking cause and reason, to conjoin a subordinate clause with a main clause. It may be argued that a correlator together with *er* form a formulaic, which may relatively facilitate learners' acquisition (cf. Ellis & Barkhuizen 2005). With an anchor-like correlator which marks a specific semantic relation, all learners have to do is fit into the construction a subordinate-main clause combination. As learners' language proficiency advances, they endeavor to exploit *er*'s discourse function as well as the peripheral *er*. In other words, they have exhibited both the complexity in conjoining different clausal types with larger linguistic units and the diversity in specifying certain a peripheral semantic relation via a contextually appropriate correlator.

To investigate L2 learners' acquisition of Chinese *er* is by no means focusing on the conjunction per se, but rather, it is more of probing how they map specific linguistic forms or structures onto the conceptual domain (cf. Clark 2004). For a function word without much semantic loading for itself as *er*, the process of clause conjoining may be done in a more implicit way. Moreover, since most adult L2 learners are well-equipped or generatively entrenched, using Pienemann's (2008) term, with their L1s, they have more or less accustomed to the way their L1s encode events in a specific way. In order to acquire different uses of *er*, learners have to re-entrench themselves in the way events are encoded in Chinese.

Therefore, the error analysis helps us understand learners' inner process in mapping linguistic forms onto conceptually eventualized experiences. But given the fact that errors can be resulted from a variety of different sources such as inter/intra-lingual transfer or context of learning, there is no guarantee that one can be unambiguously identified. For example, the instances of using *er* to conjoin two NPs may either come from an interlingual transfer (e.g. an equivalent of the coordinator *and* for English L1 speakers) or an intralingual transfer (e.g.

overgeneralize *er* or misuse it with the Chinese conjunction *he*). In fact, a more plausible account for most of the errors in using *er* may be the context of learning. Hong (2008) compared four textbooks and concluded that they either scarcely provide a systematic view of *er'*s general functions or simply annotate it with other conjunctions (e.g. *erque* 'and' and *danshi* 'but') without much explanation.

Apparently, since *er* is present in nearly all the major semantic relations in Chinese, learners may assume that any clauses can be conjoined if they can be computed to denote a complex event. Overall, the errors of *er*s are, by and large, covert and global; they do not only erode interclausal coherence but also smear sentential comprehensibility (cf. Burt & Kiparsky 1974, Corder 1971), and that is also why only when a larger linguistic context is taken into consideration that learners' infelicitous uses become obvious.

As to the instructional order of *er*, Hong (2008: 87-99) attributes the learning difficulties of *er'*s to structure (e.g. syntactic categories of conjuncts), semantics (e.g. monosemantic or polysemantic) and cross-language comparison (e.g. Chinese-English), suggesting the overall instructional sequence to be: dissimilar additive> adversative> similar additive> adverb> discourse. Mapped to our categorization, the order should be: $er^C > er^S > er^D$.

However, when learners' performance is taken into consideration as in Table 7, er^C is the category to pay special attention to because it is marked by ambiguities between subcategories (i.e. neutral contrast vs. adversative) and by differential needs of context-dependent adverbs. Er^S, on the other hand, is less ambiguous, given the correlators ahead can set different tones in sentences; moreover, errors pertaining to structural mismatch (i.e. subject position/numbers), which were shown mostly in er^S, are relatively overt and easy to point out.

Table 7. Error rates in different categories of *er*

Categorization		Errors	Instances	Error rates
er^{S}		33	219	15.1%
er^{Comp}		1	32	3.1%
er^{C}	er^{C-A}	35	142	24.6%
	er^{C-C}	5	82	6.1%
er^{D}		0	43	0

Therefore, in light of learners' data, we suggest a pedagogical order: (er^{S}, er^{Comp})> (er^{C-A}, er^{C-C})> er^{D}, which corresponds to the compositional stages from intra-sentence to inter-sentence. To begin with, we should focus on the correctness of a single sentence, directing learners' attention to subject issues between constructions with or without *er* (e.g. *yinwei…suoyi* vs. *yinwei…er*). Second, as for er^{C-A} and er^{C-C}, which stand on a vital position from clausal to sentential combination, synonymous uses with each of their syntactic and contextual requirements should be compared systematically; not only can this facilitate learners' differentiation towards similar conjunctions (e.g. *er* vs. *he*), but also modulate the way they encode events that optimally match the sentence structure of different *er*'s subcategories in Chinese. Finally, through learning the adverb and the conjunction *er*, learners should have accumulated enough knowledge about semantic relations revealed by different sentential units, and it will be natural for them to continue or contrast two paragraphs when exploiting er^{D} to structure Chinese texts.

7. CONCLUSION

This study investigates learners' acquisition of the conjunction *er* and their general performance in conjoining clauses within a four-tier categorization. Our findings show that at the initial stage, learners tend to use *er*[S] to connect clauses with a subordination relation. The study of their development further shows that learners have progressed from conjoining clauses at an intra-sentence level to an inter-sentence level.

As for learners' errors, we find linguistic devices such as correlators, associative elements (i.e. additive adverbs) and/or the requirement of a parallel structure are crucial in legitimizing learners' infelicitous use of *er*. In fact, each subcategory of *er* has its own syntactic and semantic requirement. To provide a detailed description and analysis for each subcategory of *er* is beyond the scope of our study. Yet it is hoped that with our study, it is clear that a systematic introduction of different uses of *er* with their constraints, and a finer differentiation with other conjunctions will be beneficial for L2 learners.

REFERENCES

Baker, Mark C., and Osamuyimen T. Stewart. 2002. A serial verb construction without constructions.Unpublished manuscript at: http://ling.rutgers.edu/people/faculty/.

Blühdorn, Hardarik. 2008. Subordination and coordination in syntax, semantics and discourse: Evidence from the study of connectives. *'Subordination' versus 'Coordination' in Sentence and Text*, ed. by Cathrine Fabricius-Hansen and Wiebke Ramm, 59-85. Amsterdam: Benjamins Publishing.

Burt, Maria., and Carol Kiparsky. 1974. Global and local mistakes. *New Frontiers in Second Language Learning*, ed. by John Schumann and Nancy Stenson, 71-80. Rowley, Mass: Newbury House.

Carston, Robyn. 1993. Conjunction, explanation and relevance. *Lingua* 90: 27-48.

Chao, Yuen-ren. 1980. *A Grammar of Spoken Chinese*. Translated by Panghsin Ting (2002) (Revised edition). Hong Kong: The Chinese University of Hong Kong.

Chen, Bao-qing . 1994. Xian-Qin lianci *er* yufa yuyi kaocha [Semantic and syntactic investigation of the conjunction *er* in the pre-Qin period]. *Research in Ancient Chinese Language* 1: 89-96.

Chen, Jyun-gwang. 2007. *Contrastive Analysis and Its Applications in Language Pedagogy*. Taipei: Crane Publishing.

Chu, Cheng-hsi. 2006. *A Discourse Grammar of Mandarin Chinese.* Translated by Wen-guo Pan (2006). Beijing: Beijing Language and Culture University Press.

Clark, Eve V. 2004. How language acquisition builds on cognitive development. *Trends in Cognitive Sciences* 8.10: 472-78.

Corder, S. P. 1971. Idiosyncratic dialects and error analysis. *International Review of Applied Linguistics in Language Teaching* 9.2: 147-160.

Culicover, Peter W., and Ray Jackendoff. 1997. Semantic subordination despite syntactic coordination. *Linguistic Inquiry* 28.2: 195-217.

Dalrymple, Mary, and Irina Nikolaeva. 2006. Syntax of natural and accidental coordination: evidence from agreement. *Language* 82: 824-849.

Dorgeloh, Heidrun. 2004. Conjunction in sentence and discourse: sentence-initial *and* and discourse structure. *Journal of Pragmatics* 36: 1761–1779.

Ellis, Rod, and G. Barkhuizen. 2005. *Analyzing learner language*. Oxford: Oxford University Press.

Fukushima, Tatsuya. 2005. Japanese continuative conjunction *ga* as a semantic boundary marker. *Language and Communication* 25: 81-106.

Gries, Th. Stefan. 2008. Corpus-based methods in analyses of second language acquisition data. *Handbook of Cognitive Linguistics and Second Language Acquisition*, ed. by Peter Robinson and Nick Ellis, 406-431. New York: Routledge, Taylor & Francis Group.

Guo, Xi-liang. 2003. Gu-Hanyu xuci yanjiu pingyi [Reflections on the studies of functional words in old Chinese]. *Linguistic Sciences* 1: 87-98.

Halliday, M. A. K. 1994. *An Introduction to Functional Grammar*, 2nd ed. London: Arnold.

Hong, Shen-hsing. 2008. *Xiandai Hauyu fuju fenlei yu guanlian ciyu jiaoxue- yi er wei li [The semantic classification of compound sentences in modern Chinese and the pedagogical application of correlative connectors- Chinese er as an Example].* Master thesis. Taipei: National Taiwan Normal University.

Hsueh, Feng-sheng. 1991. Shilun lianci er de yuyi yu yufa gongneng [On the semantic and syntactic functions of the conjunction *er*]. *Studies in Language and Linguistics* 1: 55-62.

Lan, Ying. 1990. Cong shaoshu minzu yuyan kan *er* de xuhua yanbian *[On functional derivation of er based on Chinese ethnic minority]. Research in Ancient Chinese Language* 1: 64-70.

Lehmann, Christian. 1988. Towards a typology of clause linkage. *Clause Combining in Grammar and Discourse*, ed. by John Haiman and Sandra A. Thompson, 181-225. Amsterdam & Philadelphia: J. Benjamins.

Lennon, Paul. 1991. Error: some problems of definition, identification, and distinction. *Applied Linguistics* 12.2: 180-196.

Li, Chao-yu. 2009. *Zhong-gao-ji jieduan Hauyu xuexi-zhe lianci de shiyong qingkuang fenxi ji jiaoxue duice yanjiu [Intermediate and advanced L2 learners' usage of Chinese conjunctions analysis and corresponding teaching strategies study].* Master thesis. Shanghai: East China Normal University.

Li, Lin. 2012. Xiandai Hanyu shumanyu zhong lianci 'er' de kaocha (Investigation of conjunction *er* in modern written Chinese). *Xiandai Hanyu Xuci Yanjiu yu Duiwai Hanyu Jiaoxue*, ed. by Hu-yang Qi. Shanghai: Xuelin Press.

Liu, Le-ning. 1996. *The grammaticalization of Chinese conjunctive adverbs*. Doctoral dissertation. Florida: University of Florida.

Liu, Yue-hua, Wen-yu Pan, and Wei Gu. 1996. *Shiyong xiandai Hanyu yufa (Modern Chinese Grammar)*. Taipei: Shida Shuyuan.

Lü, Shu-xiang. 1999. *Xiandai Hanyu Babai Ci (Eight Hundred Words of Modern Chinese)* (Revised edition). Beijing: Shangwu Publishing.

Mann, William C., and Sandra A. Thompson. 1992. Relational discourse structure: a comparison of approaches to structuring texts by 'contrast'. *Language in context: Essays for Robert E. Longacre*, ed. by Shin Ja J. Hwang and William R. Merrifield, 19-45. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

McEnery, Tony and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice.* Cambridge: Cambridge University Press.

Mei, Guang. 2002. Yingjie yi-ge kaozheng-xue he yuyan-xue jiehe de Hanyu yufa-shi yanjiu xin jumian [A new approach for research of Chinese grammar via integration of textual criticism and linguistics]. *Gujin Tongse: Hanyu de Lishi yu Fazhan,* ed. by Da-an He, 23-47.Taipei: Institute of Linguistics, Academia Sinica.

Pienemann, Manfred. 2008. A brief introduction to Processability theory. *Processability Approaches to Second Language Development and Second Language Learning*, ed. by Jorg-U. Kesier, 9-29. Cambridge: Cambridge Scholars Publishing.

Progovac, Ljiljana. 1998. Structure for coordination. *Glot International*, 3.7: 3-6 (Part I) and 3.8: 3-9 (Part II).

Qiu, Xie-jun. 2005. Lianci *er* yufa gongneng shixi (On the grammar function of the conjunction *er*). *Journal of Guangxi Teachers College* 26.3: 96-103.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comparative Grammar of the English Language*. New York: Longman. Inc.

Ramat, Anna G., and Caterina Mauri. 2011. The grammaticalization of coordinating interclausal connectives. *The Oxford Handbook of Grammaticalizatio*n, ed. by Heiko Narrogk and Bernd Heine, 656-667. New York: Oxford University Press.

Ross, John R. 1967. Constraints on variables in syntax. Doctoral dissertation. MIT.

Schmidt, Maxine. 1980. Coordinate structures and language universals in interlanguage. *Language Learning* 30.2: 397-416.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Tsai, Hui-chin. 2006. *Additive adverbs and coordinators in Chinese: a case study of ye 'also', erqie 'and' and he 'and'*. Master thesis. Hsichu: National Chiao Tung University.

Tsai, Wei-tien. 2008. Tan gu-Hanyu he Taiwan Nandaoyu zhong de binglie-shuaijian xianxiang (On conjunctive reduction in ancient Chinese and Formosan languages). *Oriental Linguistics* 3: 5-14.

Wang, Li. 1989. *Hanyu Yufa-shi (The History of Chinese Grammar)*. Beijing: Shangwu Publishing.

Wu, Yun-fang. 2005. He, yu, bing, er lianjie weici-xing chengfen-shi de qubie (Distinction of Chinese conjunctions *he*, *yu,*b*ing* and *er* in conjoining predicates). *Linguistic Researches* 1: 13-16.

Yan, Li-ming. 2009. Biaoshi duibi de lianci *er* (An analysis on *er* as a conjunction of indicating contrast). *Journal of College of Chinese Language and Culture of Jinan University* 1: 89-94.

Zhang, Dan. 2010. *Ji yu zhong-jie-yu yuliaoku de lianhe guanxi lianci pianwu fenxi* (Error analysis of coordinative conjunctions based on the Interlanguage corpus). Master thesis. Shanghai: National Shanghai Normal University.

Zhang, Li-ping. 2012. The construction of a learner corpus based on the Top-Writing Test. National Science Council midterm project report (NSC100-2631-S-003-004).

Yi-Chun Tseng
Mandarin Training Center
National Taiwan Normal University
Taipei, Taiwan 106, ROC
derekya@gmail.com

Miao-Ling Hsieh
Mandarin Training Center
Department of English
National Taiwan Normal University
Taipei, Taiwan 106, ROC
mlhsieh@ntnu.edu.tw

APPENDIX1: EXAMPLE OF CONTRASTIVE ER[D]

er[D]: contrastive

細究民進黨變天成功的因素，除了超級助選員的明星效應，強棒候選人形象不錯，由強調台獨意識到民生議題為主的訴求，是十分成功的策略運用。而另一方面，近年黑金、治安、災難等問題影響執政黨的整體形象，對選情自然有重大影響。(擷取自光華雜誌: 跨世紀百里侯之爭——民進黨大捷 (陳妙鈴，1997))

Xijiu　　　　　Minjingdang　　　　　　　biantian
deep:probe　　Democratic:Progressive:Party　overthrow
chenggong　　de　　yinsu,　　chule　　chaoji　　zhuxuanyuan de
successful　　DE　　factor　　beside　　super　　advance:man DE
mingxing xiaoying, qiangbang　　houxuanren　　xingxiang bucuo,
superstar effect　strong　　　　candidate　　image　　not:bad
you qiangdiao　　Taidu-yishi　　　　　　　dao minsheng
from emphasize　Taiwan:independency:awareness to livelihood
yiti wei zhu　　de　　suqiu,　　shi shifen　　chenggong de
issue as principle DE　appeal　BE very　　successful DE
celüe　　yunyong. Er　ling-yi-fangmian,　jinnian　　heijin,
strategy exertion ER　on:the:other:side　recent:year corruption
zhi'an,　　　zainan　　deng　　wenti　　yingxiang
social:security disaster　so:on　　problem　affect
zhizhengdang　　de　　zhengti　xingxiang, dui
government:party　DE　overall　image　　toward
xuanqing　　　ziran　　you　　zhongda　yingxiang.
election:situation　naturally have　severe　impact

'Apart from the superstar effect and candidates' own image, the DPP made good strategic use of such themes as Taiwanese independence and living standards. Meanwhile the KMT suffered heavily from an image tarnished by its handling of such problems as corruption, law and order, accidents and natural disasters.'

(An excerpt from 'DPP Wins Big in Local Election' by Miaoling Chen, translated by Robert Taylor from Guanghua Magazine, Dec. 1997)

Yi-Chun Tseng and Miao-Ling Hsieh

APPENDIX2: **CORRELATORS COMMONLY PAIRED WITH ER^S**

The list, including 106 instances of *er*^{S} and ordered by descending frequency, is generated from a randomized extract of 500 instances from the *Academia Sinica Balanced Corpus of Modern Chinese* (*Corpus-ASB*).

Type of Correlators		Pinyin	English	Instances	Percentages
Cause		*yinwei/ yin youyu*	because (of)	27	25.5 %
Reason		*weile/ wei*	In order to, for	22	20.8 %
Peripheral	Contingency	*sui suizhe*	with, contingent	14	13.2 %
	Method/ Manner	*jing jingyou tongguo kao*	via, by way of	14	13.2 %
		you zi	from	12	11.3 %
		yi (third tone)	by	8	7.5 %
		yi (first tone) yizhao	according to	5	4.7 %
		yanzhe/yan	along	2	1.9 %
		zhendui	targeting on	2	1.9 %
Sum				106	100%

漢語「而」的二語習得：一個基於語料庫的研究

曾逸群 [1]、謝妙玲 [1, 2]
國語教學中心 [1]
英語系 [2]
國立臺灣師範大學

本文研究漢語「而」的習得，將「而」劃分為從屬標記類 (subordinator *er*)、補語標記類(complementizer *er*)、對等連接類(coordinator *er*)，以及篇章銜接類(discourse marker *er*)等四大類十一小類。研究結果發現，相較於漢語母語者，中級語言水平的學習者傾向使用大量從屬連接功能的「而」來連接具主從屬關係的子句，而其以「而」來銜接不同句子或段落的比例則偏低；另一個重要的結果是我們的研究顯示了一個中介語的特色，也就是隨著漢語水平的提升，學習者連接的成分也由句內成分的連接，發展到句子與句子或更大成分的連接。至於偏誤產生主要是來自於學習者未能顧及不同連接成分所應受到的語意和語法的限制。

關鍵字： 連接詞、而、第二語言習得、語料庫研究