

Article

L2 Chinese Learners' Lexical and Grammatical Development of Result-state Resultative Verb Compounds: A Usage-based Corpus Study

Jie Zhang

University of Oklahoma, USA

Abstract

Taking usage-based approaches to second language (L2) development and drawing on a written learner corpus, this study examined L2 Chinese learners' lexical and grammatical development of result-state resultative verb compounds (RVCSs). The lexical development was examined in their frequency of usage, compositionality, and accuracy. Grammatical development was analyzed by the interaction of RVCSs and the perfective aspect marker *le*. Findings showed that while the frequency of use, compositionality, and lexical range of RVCSs grew with learners' overall language proficiency, accuracy exhibited a certain degree of regression among high intermediate and advanced learners. L2 Chinese learners' lexical and grammatical development of RVCSs is a systematic yet complex and variable process. Factors affecting the acquisition of RVCSs are the nature of input, L1 blocking and learned attention, and the unique properties of RVCSs.

Keywords

Result-state resultative verb compounds, lexical and grammatical development, learner corpus, usage-based approaches to L2 development

1 Introduction

Resultative verb compounds (RVCs) are an important class of compound words that serve important grammatical functions in Chinese. As a morphological construction, RVCs are a condensed form of verb compounding, with the first component (conventionally referred to as V1) indicating an action and the second component (conventionally referred to as V2) signaling the result of the action (Chao, 1968; Li & Thompson, 1981; Packard, 2000). As a grammatical construction, RVCs are a primary resultative construction in Chinese that represent argument structure and change of state. Depending on the kind of result they denote, RVCs can be further divided into result-state resultative verb compounds, directional resultative verb compounds, and completive resultative verb compounds (Chao, 1968; Li & Thompson, 1981; Packard, 2000; Smith, 1990, 1997; Xiao & McEnery, 2004).

This study focuses on result-state resultative verb compounds (RVCSs). RVCSs are the most intensively researched resultatives by Chinese linguists due to its morphological and semantic complexities. Still, few studies have examined second language (L2) Chinese learners' development of this challenging type of verb compounds. While L2 research on RVCSs has investigated its syntactic and thematic configurations and interlanguage errors (Xing, 2003; Qiao, 2008; B. Zhang, 2008; Feng, 2017), little understanding has been gained about how L2 Chinese learners use RVCSs, the developmental paths across proficiency levels, and the factors affecting the acquisition of RVCSs. Adopting usage-based approaches to L2 development, this study examined the lexical and grammatical development of RVCSs in a written learner corpus while comparing learners' usage patterns against those by native Chinese speakers.

2 Review of the Literature

2.1 Result-state RVCs

A result-state RVC, also termed a stative resultative by Packard (2000) and Gu (2003), has a V2 indicating the resultant state of the action symbolized by the V1. The RVCS is compositional because its meaning is generally derivable from its components.

- (1) a. 看懂
kan-dong
'to understand through reading'
- b. 打扫干净
dasao-ganjing
'to dust and sweep until clean'

The V1s, such as *kan* 'look' and *dasao* 'dust and sweep' in (1), represent the open set of Chinese transitive verbs and adjectives. The V2s, such as *dong* 'understand' and *ganjing* 'dry and clean', are either members of an open class of adjectives or non-transitive stative verbs. The V1 and V2 can be a morpheme, a derived word, or a compound (Chao, 1968, p. 442). In (1b), the V1, *dasao*, is a coordinate verb compound, and the complement V2, *ganjing*, is a subordinate adjective compound. Surveying the *Dictionary of Verb Usage*, Xu (2000) reported that 488 (92%) of monosyllabic verbs can be used in the V1 position. Ma and Lu (1997, p. 156) identified 168 monosyllabic adjectives in the *Dictionary of Adjective Usage*. Apart from 15 adjectives, all other adjectives can take the V2 position of the RVCS. In contrast to adjectives, only a limited number of verbs can take the V2 position (Tang, 1989, p. 50).

RVCS formation is governed by lexical derivational rules and subject to conventionality and idiosyncrasy. According to Thompson (1973), RVCs consist of an open subset whose members are derived by lexical rules and a closed subset whose members are simply listed in the lexicon. She proposed the most general rule for creating RVCs as in (2).

$$(2) \begin{array}{ccccc} V & + & V & \rightarrow & [V - V]RV \\ \text{action} & & \text{intransitive} & & \text{action} \end{array}$$

Note: RV = resultative verb compound

(Thompson, 1973)

Meanwhile, many RVCSs cannot be accounted for by any general rules because their semantic properties cannot be predicated from those of the components. As such, they must be listed as lexical items. Examples of RVCSs listed as lexical entries include *zuo man* 'seat-be full', *zou guang* 'walk-be empty', and *shi tou* 'be wet-penetrate'. Deng (2010) discussed the role of conventionality and idiosyncrasy in

forming resultative compounds. She noted that many possible RVCSs derived from the lexical rules are not acceptable lexical items in Chinese. Examples include pairs such as *chang ku* ‘sing-cry’ and **chang xiao* ‘sing-laugh’, *ban huai* ‘move-broken’ and **ban po* ‘move-broken’, *xi ganjing* ‘wash-clean’ and **xi qi* ‘wash-angry’. Although the compounds marked by the asterisk share the same V1 and have V2s similar to the acceptable construction in the pair, they are not acceptable in Chinese; such exceptions to the lexical rules can only be explained by convention. The co-existence of lexical derivational rules and conventionality in forming RVCSs may make language acquisition difficult for learners of Chinese whose first language (L1) conveys results differently.

The RVCS encodes the state-change event with the V1 representing the cause and the V2 representing the change of state. According to Vendler’s four-way categorization of verbal event types (Vendler, 1967), RVCSs fall into the categories of accomplishment and achievement (Tai, 1984, 2003). Accomplishments characterize situations with duration and natural endpoints (e.g., *chi bao* ‘eat-full’, *xie cuo* ‘write-wrong’). Achievements encode situations with natural endpoints but no duration (e.g., *sha si* ‘kill-dead’, *da kai* ‘hit-open’). Due to the inherent event structure of RVCSs, they frequently co-occur with the perfective aspect marker *le* to denote a completed event, as shown in (3a). However, the RVCS is not bound to the perfective aspect. The perfective aspect marker *le* is not used with the RVCS in negative sentences, sentences in the progressive form, sentences expressing indefinite past with the suffix *guo*, and sentences with modal verbs (Chao, 1968). Example (3b), for instance, is a negative sentence in which *le* is not permitted. In example (3c), the modal verb *yinggai* ‘should’ is used, suggesting obligation rather than completion, and for that reason, *le* cannot co-occur with the RVCS *wen qingchu* ‘ask-clear’.

(3) a. 我写完作业了。

Wo xie-wan zuoye le

I write-finish homework PFV

‘I finished my homework.’

b. 门没关紧。

Men mei guan-jin

door not close-tight

‘The door was not shut tightly.’

c. 你应该问清楚这件事儿。

Ni yinggai wen-qingchu zhe jian shi’er

you should ask-clear this-CL-matter

‘You should clarify this matter.’

2.2 L2 acquisition studies on RVCSs

Although RVCs are highly frequent in Chinese, L2 Chinese learners use them scarcely (Wang et al., 1987). A challenging aspect of RVC acquisition is its lexical nature. Using a written learner corpus, J. Zhang (2014) investigated the development of L2 learners’ lexical knowledge of three types of RVCs: directional RVCs, completive RVCs, and RVCSs. The learners’ proficiency levels ranged from lower-intermediate, higher-intermediate to advanced. Through analyzing the frequency, component versatility, and accuracy of RVCs produced by these learners, J. Zhang proposed three developmental stages of RVCs: the whole-word formula stage, the emergence of compound awareness stage, and the solidified compound awareness and lexical development stage. While J. Zhang (2014) pointed out the different patterns of development for the three types of RVCs, she did not expound on the unique features of RVCSs that made them challenging to acquire.

Research on interlanguage errors has shed light on lexical errors related to RVCSs. Using the Chinese Interlanguage Corpus developed by Beijing Language and Culture University, Xing (2003) examined lexical errors of Chinese compound words. Five categories of compound errors were identified: (1) substituting a morpheme with a semantically similar one; (2) under-using, over-using, or misusing a morpheme; (3) constructing a compound that does not exist in Chinese; (4) misplacing a morpheme; and (5) others. Xing found that learners who studied Chinese in an immersion environment had strong morphological awareness of compounds in terms of compositionality, morphological structure, and semantic relationships. Learners used two different methods in acquiring compounds: a morpheme-based approach and whole-word approach, with the former being the dominant method. Also drawing on a large-scale interlanguage corpus, B. Zhang (2008) analyzed English-speaking learners' use of verbs, and identified several categories of verb-related errors, including collocational errors, confusion between synonymous words, rhythm-induced errors, stylistically induced errors, and so on. She found that many easily confused words for L2 Chinese learners do not necessarily have semantic relevance, which means the easily confused words for L2 Chinese learners are not necessarily synonyms to L1 Chinese speakers. The sources of confusion may have come from orthographical similarities of those characters or L1 transfer. While such research may explain certain lexical errors of RVCSs, they fail to also examine the correct usages of RVCSs, and more importantly, by only looking at learners at a specific proficiency level, in both cases the advanced level, these studies have little to offer regarding the developmental patterns of RVCSs.

Another strand of research has focused on L2 learners' mastery of the event structure of RVCSs. In the generative paradigm, Qiao (2008) investigated L2 learners' identification and acquisition of RVCS telicity marking. Using a story comprehension task, participants were asked to choose an appropriate description of the story that contained either an RVCS denoting a change-of-state event or a single action verb describing a no-change-of-state event. The results revealed that the L1 parameter of telicity affects learners in the initial stage of L2 acquisition, but as overall language proficiency improves, the parameters can be reset, and learners can approximate native speakers in their judgments regarding the appropriateness of telicity marking. Feng (2017) investigated L2 learners' processing of the event structures of *ku-shi* and *tui-dao*. The E-Prime experiment showed that learners had the ability to process the event structure but failed to achieve automatic processing and needed more time to process.

These studies contribute to our understanding about how the argument and event structures of RVCSs are comprehended by L2 Chinese learners. However, researchers only examined the internal structure of RVCSs. Successful acquisition of RVCSs also lies in the interaction between event type and aspect. Research on L1 Chinese acquisition has found that at age three, children used RVCSs independent of aspect markers, indicating they were able to treat the event type and aspect as two different categories at a very early stage of language development. However, children still had problems with the state-change event structure encoded by RVCSs. They overused the perfective marker *le* when RVCSs co-occur with the progressive aspect marker, the negative form, and modal verbs (Deng, 2010, 2019). As few studies have looked into how L2 Chinese learners develop their grammatical awareness of RVCS, this paper fills the gap by investigating the co-occurrence of RVCS and the perfective marker *le* in learners' language.

2.3 Usage-based approaches to L2 development

This study subscribes to usage-based approaches to L2 development in explaining the learners' usage patterns of RVCSs. Usage-based approaches, as an overarching term for several theories, are built upon two fundamental assumptions about language learning. First, linguistic input is the primary source of language learning. Second, learners employ the same general cognitive mechanisms in acquiring a language as they do with other learning (Bybee & Hopper, 2001; Tomasello, 2003). Usage-based approaches regard construction as the basic unit of acquisition. Construction is the pairing of form and

meaning or function. It can range from simple words, combination of words, grammatical particles, to complex syntactic structures (Goldberg, 2006; Ellis & Wulff, 2015; Wulff & Ellis, 2018). Based on input, learners build mental representations of constructions of different abstraction and complexity through associative learning and “statistical tallying of its frequency of usage and probabilities of form-function mapping” (Wulff & Ellis, 2018, p. 51). The acquisition of constructions is susceptible to frequency and salience of the input, prototypicality of forms, contingency of form and function, and learner factors such as attention and memory (Ellis, 2002, 2012).

Usage-based researchers view L2 development as a complex, dynamic, and adaptive system that demonstrates both systematicity and variability (Ellis & Larsen-Freeman, 2009). Learners’ output often follows predictable paths with more or less fixed stages in the acquisition of a given structure, which can be explained by input regularities. On the other hand, variability permeates L2 development. Learners may master certain structures or forms before they acquire others within any area of the language. In restructuring the L2 representation of the construction, learners may develop better mastery of certain sub-dimensions than others (Ellis & Wulff, 2015).

In the acquisition of RVCSs, we predict that learners would first acquire the most frequent RVCSs they encounter in the instructional input as non-analytical compound words. The meaning of these words is paired with its forms as a whole unit, rather than two individual morphemes. Because learners are mainly producing RVCSs straightly based on input in the initial stage, the accuracy of their production will be high without many errors of creative use. When learners are exposed to sufficient amount of RVCSs, they would gradually develop an awareness of RVCSs as compounds and start coining new compounds. This stage will see an increase in the variety of RVCSs that learners can produce. But due to the immaturity of learners’ grasp of the underlying lexical formation rules, learners at this stage are prone to lexical errors, thus demonstrating a lower accuracy rate. Also due to the highly lexical nature of RVCSs, we predict that learners will continue to encounter problems in producing accurate RVCSs because it will take much longer for them to acquire enough input-based frequency information to form the right mental representation of synonymous RVCS pairs.

Usage-based approaches explain L1 interference as learned attention and L1 blocking. In acquiring the L1, learners have been attuned to noticing certain linguistic cues than others, which is termed learned attention. Although such selected attention is efficient in L1 learning, it may become a barrier in L2 acquisition. Learners’ L1 experience has trained them to block the less frequent and less salient cues in the L2, making L2 acquisition a less successful endeavor than their L1 counterpart (Ellis, 2006). In the case of RVCSs, the linguistic means of conveying resultative meaning in Chinese require the morphological form of a verb-complement compound. Because English takes linguistically different means in denoting the same change-of-state event, English L1 learners of Chinese have been accustomed to the forms in their L1, i.e. their attention is directed that way, making it difficult for them to notice the L2 morphological construction. Due to the L1 experience, learners may block the new compounding formation of V1-V2 in expressing resultative meaning. We predict that learners in the initial stage may omit one of the components of a RVCS. This tendency should be overcome when they move to the stage of creatively combining RVCS components.

3 Research Questions

Drawing data from a written corpus, this study looks at the lexical and grammatical development of RVCSs among lower-intermediate, higher-intermediate, and advanced learners of L2 Chinese. It addresses the following research questions:

1. How do L2 Chinese learners develop the lexical aspect of RVCSs, as manifested in their frequency of usage, compositionality, and accuracy?
2. What is the co-occurrence of RVCSs and the perfective aspect marker *le* used by L2 Chinese learners?

4 Method

4.1 Corpus data

The learner corpus consisted of 784 essays written by L2 Chinese learners at three levels: lower-intermediate level (LIL), higher-intermediate level (HIL), and advanced level (AL). 100 essays written by native Chinese speakers were used as a baseline for gauging the overall language performance of the learners. Table 1 summarizes the corpus composition.

Table 1

Composition of the Corpus

Proficiency Level	Number of Essays	Average Length (Chinese Characters)	Total Number of Word Tokens	Total Number of Word Types	Total Size (Chinese Characters)
LIL	409	231	45,615	3,232	94,461
HIL	204	326	32,137	2,486	66,408
AL	171	390	41,603	4,354	66,659
NS	100	873	62,552	9,014	87,320
Total	884	455	181,907	19,086	314,848

Note. LIL = low intermediate level; HIL = high intermediate level; AL = advanced level; NS = native speakers

The LIL essays were collected from students enrolled in the third course of a six-course Chinese language sequence in the fall semester of 2009 at a comprehensive North American public university. The HIL data were collected from students enrolled in the fifth Chinese language course at the same university. Most students in both courses were second- or third-year undergraduates aged 18-22. As required by the course curricula, students wrote a short essay at the end of each unit. Most writing assignments were based on the textbooks, with a few supplemental topics provided by the instructors. All but one essay were completed as homework assignments, so students had no time constraints and could use reference materials. Over the 15-week data collection period, 57 LIL students who gave their consent produced 409 essays with an average of 231 Chinese characters per essay, and 30 HIL students who gave their consent produced 204 essays with an average of 326 Chinese characters per essay. Appropriate to their proficiency levels, the genres of their writing were mostly narrative or descriptive.

The AL data were retrieved from the *Hanyu Shuiping Kaoshi (HSK) Dongtai Zuowen Yuliaoku* (Chinese Proficiency Test Dynamic Composition Corpus) Version 1.1. The HSK Dynamic Composition Corpus has a collection of 11,569 essays (totaling approximately 4.24 million Chinese characters) written by L2 learners of Chinese who took the HSK Advanced test between 1992 and 2005. To ensure that the backgrounds of the test takers were comparable to the LIL and HIL groups, only essays written by test-takers who registered their nation of origin as either the United States or Canada were retrieved. Altogether 171 essays were retrieved. The essays had an average length of 390 Chinese characters, the longest of the three learner groups, and were narrative, descriptive, or argumentative by genre.

The baseline data, the native speaker (NS) subset of the corpus, were a collection of essays written by Chinese high school students taking or preparing for China's National Matriculation Test. The essays were downloaded from the official educational websites *Zhongguo Jiaoyu Zaixian* (China Education Online) and *Renmin Wang* (People's Daily Online). Several genres were represented: narrative, argumentative, expository, and prose. This collection of NS essays provided a reasonable benchmark against which to compare learners' language use, as they were written by students of a similar age for institutional purposes covering similar genres.

4.2 Corpus annotation and tagging

The essays were first word-segmented and parts-of-speech tagged using the Chinese Lexical Analysis System developed by the Institute of Computing Technology at the Chinese Academy of Sciences (Xiao, Rayson, & McEnery, 2009). Two coders then manually coded all instances of RVCs identified in the corpus. Using the coding scheme developed specifically for analyzing RVCs (Zhang, 2011), the two coders individually analyzed 40 essays (10 from each group), discussed and revised the coding. The inter-rater reliability for the raters was Kappa = 0.945 ($p < 0.001$). Then, they each coded half of the data independently. The coding consisted of two procedures. First, the coders applied tags that distinguished the different types of RVCs (RVCSs being a major type tagged as <RVCS>). Inappropriate uses of RVCSs and missing RVCSs in obligatory situations were then identified and given an error tag. The coders then double-checked all tagging, discussed different opinions, and resolved disagreements through discussions. After annotation, AntConc 3.2.2.1 (Anthony, 2011) was used to extract every instance of RVCSs and instances of missing RVCSs in obligatory contexts. The extracted concordance lines were exported to Excel spreadsheets for further analysis.

5 Results

5.1 Lexical development of RVCSs

5.1.1 Frequency of usage

The distribution of RVCSs was measured by raw frequencies and normalized frequencies per 1,000 Chinese characters. For each index, both tokens (instances of RVCSs) and types (unique RVCSs) were calculated as shown in Table 2. Both the token and type measures showed a clear, steady positive relationship between the frequency of RVCSs and their overall language proficiency, suggesting token and type frequencies are effective measures of productive knowledge of RVCSs. Chi-square tests were performed to see whether significant differences existed overall and between groups. For token frequencies, the overall chi-square is $\chi^2=43.265$, $df=3$, $p < 0.0001$, meaning that there were significant differences in the token frequencies of RVCSs. Between groups, significant differences were found between the LIL and all other groups ($\chi^2 = 23.626$, $df = 0.0017$, $p < 0.0001$; $\chi^2 = 36.870$, $df = 0.002$, $p < 0.0001$; $\chi^2 = 40.043$, $df = 0.0018$, $p < 0.0001$, respectively), but no significant difference was found between the HIL, AL, and NSs. Type frequencies also showed significant differences, the overall chi-square being $\chi^2 = 14.093$, $df = 3$, $p < 0.01$. Due to the relatively lower frequencies of RVCS types, no significant differences were found between groups. Compared with the NSs, the learners seemed to have a limited lexical repertoire of RVCSs, as even the AL group was not able to approximate the performance of NSs in terms of RVCS tokens and types, but the difference was not statistically significant.

Table 2

Frequencies of RVCSs

	LIL	HIL	AL	NS
# of RVCS tokens	63	100	142	199
# of RVCS tokens per 1,000 characters	0.67	1.51	2.13	2.28
# of RVCS types	27	49	65	120
# of RVCS types per 1,000 characters	0.29	0.74	0.98	1.37

5.1.2 Compositionality

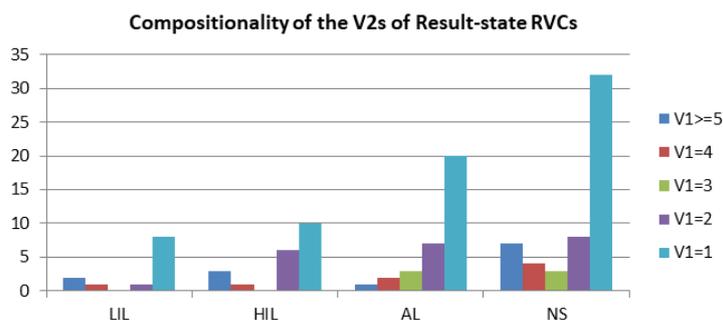
The numbers of V1s and V2s (by type) produced by the learners and NSs are presented in Table 3. Comparing the two components, V1 is lexically richer -- the number of V1s is about twice the amount of V2s for both learners and NSs. In comparison, V2 seems to be the more productive element because on average, a given V2 can be paired with two different V1s. Developmentally, both numbers increased with language proficiency, with the NSs producing the most V1s and V2s, and the LIL producing the least.

Table 3

Number of V1 and V2 Types Produced by Learners and Native Speakers

	LIL	HIL	AL	NS
Number of V1 Types	23	41	46	91
Number of V2 Types	12	20	33	54
V1/V2 Type Ratio	1.92	2.05	1.39	1.69

The compositionality of RVCSs was examined in the number of components a V1 or V2 can take. Figure 1 (a) displays the distribution of V2s that can take various numbers of V1s. The NSs' production of the V2 exhibited a roughly U-shaped distribution with a sizeable number of V2s taking four or more V1s on the left end, a small set of V2s taking two or three V1s in the middle, and a considerable number of V2s taking one V1 on the right end. The sizeable number of V2s taking four or more V1s in the NSs' data suggests the productivity of V2s in forming compounds with different V1s. For example, the NSs produced 8 different compounds sharing the same V2 adjective *man* 'full or filled', 7 different compounds taking the adjective *jin* 'exhausted or finished' as the V2, and 5 different compounds with the same adjective *qing* 'distinct or clarified' as the V2. The considerable number of V2s taking only one V1 suggests the NSs' extensive lexical range of the RVCS morphemes at their command. NSs used numerous V2s that the learners did not; some examples include the adjectives, *san* 'scattered', *xiang* 'loud', *shu* 'ripe', *shi* 'wet', and *tong* 'painful', and the verbs, *kua* 'collapse', *chuan* 'penetrate', and *li* 'leave'. As language proficiency grew, the frequency with which learners used V2s that take just one V1 increased, suggesting an expanding lexical repository of productive use. Unlike the NSs, learners' inventory of V2s taking four or more V1s was still low, but the number of V2s taking two or three V1s expanded steadily. For example, LILs used the V2 adjective *cuo* 'wrong' only with the V1 verb *nong* 'make', whereas ALs used *cuo* 'wrong' with the verbs *zou* 'walk' and *zuo* 'do'. In addition, LILs produced only one compound, *kai zou* 'drive-leave' using the V2 verb *zou* 'leave', while HILs produced nine compounds sharing the same V2 *zou* 'leave': *dai zou* 'bring-leave', *cuan zou* 'scurry-leave', *fei zou* 'fly-leave', *kai zou* 'drive-leave', *na zou* 'hold-leave', *qi zou* 'ride-leave', *chu zou* 'exit-leave', *tao zou* 'flee-leave', *tou zou* 'steal-leave', and *xia zou* 'frighten-leave'.



(a)

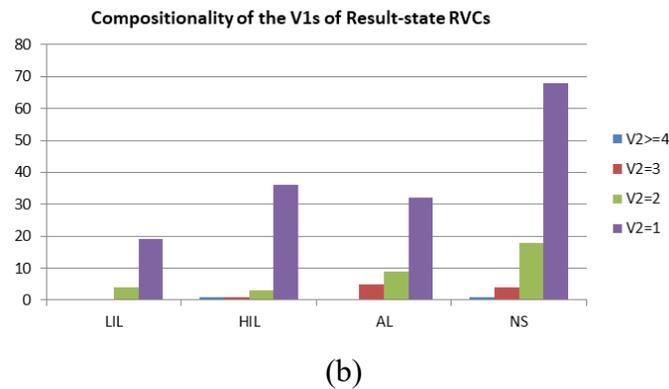


Figure 1. Compositionality of RVCs produced by learners and native speakers

The numbers of V2s taken by the V1 are presented in Figure 1 (b). The NSs' data is skewed towards the right end, suggesting that the majority of the V1s collocated with only one or two V2s. The learners show a similar tendency. Nevertheless, the learners' lexical range of the V1 seems to be much more limited than that of the NSs, with the HIL and AL producing about half the number of V1s that the NSs produced. Another disparity between the learners and NSs is that the learners had a much smaller set of V1s collocating with two or more V2s, suggesting a weaker capacity to form compounds using the V2s in their lexicon.

5.1.3 Accuracy

Lexical deviations were categorized into three types: V1-related deviations, V2-related deviations, and compound-related deviations. Table 4 shows the distribution of these lexical deviations, and several patterns can be observed. First, the total number of lexical deviations did not decrease notably with learners' overall language proficiency. The LIL had the lowest deviation rate of 14%, the HIL had the highest deviation rate of 26%, and the AL had a deviation rate of 23%, indicating that advanced learners still struggled with choosing the semantically and lexically appropriate components of RVCs. Second, all learner groups committed more deviations related to V2 than V1, implying that V2 may be the more challenging component for learners. Third, the different types of errors seemed to pose different acquisition difficulties for learners. Overuse, with only two instances, is a non-typical deviation type. In contrast, omission and especially misuse seem to be major error types that persist throughout the learning process.

Table 4

RVCs Deviations

	LIL	HIL	AL
V1-related Deviations			
Misuse	1	5	6
Omission	2	3	1
V2-related Deviations			
Misuse	4	7	6
Omission	0	3	9
Overuse	0	2	0
Compound-related Deviations			
Misuse	2	6	11
Total	9	26	33
Percentage	14%	26%	23%

Omitting a component of the RVCS, either the V1 or the V2, seems to be related to learners' partial understanding of the event structure of RVCSs. In (4a), the verb *chu* 'rid' was used to describe a tonsil being removed. *Chu* alone indicates the result of the action, but it is missing the action verb itself. The correct form should be *zhai chu* 'extract-rid', with the verbs *zhai* denoting the action and *chu* denoting the resulting-state that the tonsil was removed from the body. Interestingly, such errors of omitting the V1 or V2 component tended to occur with words that could find a seemingly direct translation in English. The English equivalents encapsulate the action and result, whereas Chinese requires the use of a RVCS for the same argument structure with the action denoted by a verb and the result denoted by a complement. In (4b), the direct translation of the English word 'learn' was used where a complement *hui* 'know' was required to denote the resulting state of the action *xue*. Notably, both intermediate and advanced learners committed component omission errors due to L1 interference. The analysis seems to indicate that at least for a set of RVCSs, the learners acquire them as non-compositional and non-analytical words.

(4) a. * 医生要除了我的扁桃腺。(HIL)

Yisheng yao chu le wode biantaoti

doctor will rid PFV my tonsil

'The doctor will remove my tonsil.'

b. * 孩子看着父母用嘴巴沟通就也学说话。(AL)

Haizi kan-zhe fumu yong zuiba goutong jiu ye xue shuohua

child look-DUR parent use mouth communicate then also learn speak

'By watching their parents communicate by speaking, children learn how to talk.'

Note. The asterisk * indicates an inappropriate use.

The misuse of either one component or the whole compound accentuates the lexical nature of RVCSs. A characteristic pattern of learners' language use is that they tend to use a general verb to describe a concrete action, suggesting a limited lexicon at command. In (5a), the word *nong* 'make' was used to denote the action of wiping with a towel in the context where a specific verb *ca* 'wipe' is more appropriate. In (5b) and (5c), the learners used RVCSs of concrete referents instead of those of abstract referents. What learners produced are not semantically wrong but sound strange to a native ear.

(5) Forms Produced by Learners

a.* 用毛巾弄干净梨子 (HIL)

yong maojin nong-ganjing lizi

use towel do-clean pear

'Use a towel to make the pear clean'

b.* 扔掉生命 (AL)

ren-diao shengming

throw-drop life

'Throw away life'

c.* 拿走自由 (AL)

na-zou ziyou

take-walk freedom

'Take away freedom.'

Appropriate Target Forms

用毛巾擦干净梨子

yong maojin ca-ganjing lizi

use towel wipe-clean pear

'Use a towel to wipe the pear clean'

放弃生命

fangqi shengming

give-up life

'Give up one's life'

剥夺自由

boduo ziyou

deprive freedom

'Deprive of freedom.'

More instances of misuse occurred due to confusion of synonymous RVCSs that are similar in form and meaning with differences in semantic reference and collocational convention. For instance, in (6a), the AL learner collocated *zeng jia* 'increase-add' with *jiaqian* 'price', which generally denotes an increase in quantity. The appropriate form for price inflation is *ti gao* 'lift-high'. Similarly, in (6b), the AL learner's

collocation of *zeng jia* ‘increase-add’ with *xingqu* ‘interest’ and *zhishi* ‘knowledge’ is inappropriate; the correct forms are *ti gao* ‘lift-high’ *xingqu* ‘interest’ and *jilei* ‘accumulate’ *zhishi* ‘knowledge’. Such misuses tend to occur with learners at the higher-intermediate to lower-advanced proficiency levels when they have accumulated a sizeable bank of lexical forms but have not yet developed full mastery of the form-meaning pairs to the extent that they can use them in a linguistically appropriate context.

(6) Forms Produced by Learners	Appropriate Target Forms
a. * 增加烟的价钱 (AL)	提高烟的价钱 (AL)
zeng-jia yan de jiage	ti-gao yan de jiage
add smoke GEN price	lift-high smoke GEN price
‘Add the price of cigarettes’	‘Raise the price of cigarettes’
b. * 增加兴趣和知识 (AL)	提高兴趣、积累知识
Zeng-jia xingqu he zhishi	Ti-gao xingqu, jilei zhishi
increase-add interest and knowledge	lift-high interest, accumulate knowledge
‘Increase interest and knowledge’	‘Improve interest and accumulate generate’

5.2 Co-occurrence of RVCs and *le*

This section examines the co-occurrence of RVCs and the perfective aspect marker *le* to see how RVCs independently or collaboratively contribute to aspect marking in Chinese. Instances of RVCs produced by the NSs and learners, occurring either independently or together with the perfective aspect marker *le*, were counted. Targetlike and non-targetlike usages of RVCs were distinguished. Table 5 shows the distribution of sentential RVCs and perfective aspect by token measures. While both usages are possible, the NSs used RVCs more often without the perfective aspect marker *le* (64%) than with *le* (36%). The ALs closely resembled the NSs in their use or non-use of *le*. The HIL learners, however, showed a completely opposite pattern, using RVCs without *le* 31% of the time, and with *le* 69% of the time. The LILs, unlike the HILs or ALs, had a more balanced distribution of both choices with 54% without *le* and 46% with *le*.

Table 5

RVCs with the Perfective Aspect Marker le (Token Frequencies)

Proficiency	Result-state RVC without <i>le</i>			Result-state RVC with <i>le</i>		
	Total	TL	Non-TL	Total	TL	Non-TL
LIL	34 (54%)	31 (49%)	3 (5%)	29 (46%)	27 (43%)	2 (3%)
HIL	31 (31%)	22 (22%)	9 (9%)	69 (69%)	67 (67%)	2 (2%)
AL	109 (77%)	106 (75%)	3 (2%)	33 (23%)	30 (21%)	3 (2%)
All Learners	174 (57%)	159 (91%)	15 (9%)	131 (43%)	124 (95%)	7 (5%)
NS		127 (64%)			72 (36%)	

Note. TL = targetlike; Non-TL = non-targetlike

The non-targetlike instances, both over-use and under-use of *le*, indicate learners’ limited understanding of the change-of-state event structure encoded by RVCs. In (7a), the sentence final particle *le* is missing, which truncates the sentence. A primary function of *le* is to indicate a currently relevant state (Li & Thompson, 1981). The sentence final particle *le* in (7a) is required to indicate that ‘I am now in a state of being busy’. Redundancy occurred when the learners failed to pick up the temporal references of

the sentences and broke a licensure by using RVCSs together with modal verbs, such as *yao* ‘want’ and *keneng* ‘possible’, durative adverbials, such as *yizhi* ‘always’, or negative words (Li & Thompson, 1981). Such instances of *le* redundancy suggest a limited understanding of the change-of-state event structure encoded by RVCSs and its incompatibility with on-going events or propositions that express modality or possibility. In (7b) *le* after *gua duan* ‘hang-broken’ is inappropriate because the compound is modified by the durative temporal adverbial *yizhi* ‘always’, which contradicts *le* in temporal reference.

(7) a. * 現在 我 在 寫 信 可 是 我 忙 死 ! (HIL)

Xianzai wo zai xiexin keshi wo mang-si

now I DUR write-letter but I busy-dead

‘Now I am writing a letter, but I am crazy busy.’

b. * 我 好 幾 天 要 給 小 美 打 電 話 , 可 是 她 一 直 挂 斷 了 , 讓 我 很 擔 心 。 (HIL)

Wo haojitian yao gei xiaomei da dianhua, keshi ta yizhi guaduan le, rang wo hen dan-xin.

I several-days will to Xiao-mei make phone-call, but she always hang-break PFV, let me very worried.

‘For several days, I tried to call Xiaomei, but she kept hanging up, making me very worried.’

6 Discussion

In this section, we first summarize the L2 developmental patterns of RVCSs as showed in our data. We then discuss three factors that may have contributed to such patterns and characteristics: the role of input, L1 influence, and unique properties of RVCSs as compared to the other types of resultative compounds.

6.1 L2 developmental patterns of RVCSs

Most our predictions have been met by the learner corpus data. A clear developmental pattern was observed. The learners’ data manifested a positive relationship between the frequencies of RVCSs and their overall language proficiency. Compositionally, an expanding lexical inventory of RVCSs was observed, with steady growth in the use of V2s taking one V1, and in the use of V2s taking two or three V1s. V2 was found to have stronger compositionality than V1 among learners at all levels and in the NSs. On the other hand, our data showed that the development of lexical complexity was in a weak competing relationship with that of accuracy. While the frequency of usage and compositionality of RVCSs grew with overall language proficiency, the accuracy of RVCSs did not improve noticeably among the higher-intermediate and advanced learners. As to the sources of learner errors, we found that different types of errors tend to occur at different developmental stages. Meanwhile, we confirm what were reported by Xing (2003) and B. Zhang (2008) as advanced learners still have difficulty in choosing appropriate RVCSs due to the highly lexical nature of RVCSs.

6.2 The role of input

Learners’ usage patterns, especially in the early stage of acquisition, testify the strong influence of input they received in classroom instruction. We surveyed how RVCSs are presented as grammatical constructions and lexical items in two mainstream Chinese textbooks for basic Chinese language courses in American universities, *Chinese Link* by Wu, Yu, and Zhang (2008a, 2008b) and *Integrated Chinese* by Liu et al. (2010). Both books introduce the grammatical function of RVCSs followed by illustrations

of how individual complements, such as *hao* ‘good’ and *dong* ‘understand’, are used. Although their grammatical function is presented, RVCSs that appear in the book are listed as lexical items and treated as non-compositional whole words. For example, the RVCS *jian qing* ‘reduce-light’ is glossed as a single English word ‘lessen’ without information about its compounding nature. Moreover, it is not uncommon to see synonymous words translated into the same English word. The RVCS *xue hui* ‘study-know’, for instance, is glossed as ‘learn, master’, and the single verb *xue* ‘study’ as ‘study, learn’. The glosses do not take into account how the addition of the morpheme *hui* ‘know’ contributes to the meaning of the compound *xue hui* ‘study-know’, offering little in helping students notice the semantic composition of the forms. Students may have to learn *xue hui* as a non-compositional and un-analytical lexical chunk, which partly explains their lack of awareness in verb compounds and their inability to form new RVCSs.

6.3 L1 influence

As predicted, our findings showed that learners’ L1 plays an important role in acquiring RVCSs. It is manifested in terms of learned attention and L1 blocking, which can account for why omissions of a RVCS component, sometimes the action verb and more frequently the resultative complement, were common errors. We found that such omissions tend to occur with words with seemingly direct translations in L1 English, in which the English equivalents encapsulate the action and result into one word. Their learned attention resulted in overlook of the L2 construction. This influence was observed in both intermediate and advanced learners.

The instances of over- or under-use of the perfective aspect marker *le* in the presence of an RVCS also suggested learners’ partial understanding of the event structure in Chinese. Our findings corroborate Qiao (2008) and Feng (2017) in that learners are able to develop correct event structure representation of RVCSs, but developmentally it takes a long time to achieve the goal.

6.4 Properties of RVCS construction

The RVCS construction has its unique properties, making its acquisition different from acquiring other types of RVCs. According to J. Zhang (2014), compared with directional and completive RVCs, RVCS morphemes are less productive and more lexical. Moreover, RVCSs have the lowest frequencies of usage among the three. A necessary condition for acquisition is abundant frequent input. Due to the lexical nature and lower frequencies of RVCSs, accumulating sufficient input of individual RVCS forms seems more difficult. This may explain why accuracy rates remained low even for advanced learners. In addition to the sources of errors reported by Xing (2003) and B. Zhang (2008), our findings revealed that learners tended to use a common verb instead of a specific verb as the V1, and they frequently confused synonymous RVCSs that share some similarities in general meaning but have different semantic references or collocational preferences. This may translate into the fact that the likelihood for learners to have encountered individual RVCS forms is lower, making it harder to generalize its form-meaning mapping accurately.

7 Conclusion and Pedagogical Implications

Based on a written learner corpus, this study revealed that L2 Chinese learners’ lexical and grammatical development of RVCSs is a systematic yet complex and variable process. It is under the influence of multiple factors, mainly the nature of input, L1 blocking and learned attention, and the unique properties RVCS constructions.

In order to maximize the efficiency of explicit learning, we propose the following pedagogical suggestions. When RVCSs are first introduced, instruction should clearly explain its lexical composition,

the components' grammatical functions, and how each morpheme contributes to the compound structure. A contrastive approach that compares synonymous English and Chinese expressions is recommended to draw learners' attention to the different form-meaning pairings in Chinese. At the advanced level when learners have accumulated a good number of RVCSs, instruction should focus on deepening, widening, and refining their lexical knowledge of the forms. A corpus-informed pedagogy can be helpful as the concordance lines make it easier for learners to notice the language use patterns, the immediate linguistic environment, and idiosyncrasies in the use of RVCSs. A collocational approach is recommended to show learners preferable collocations of synonymous RVCSs, and how similar RVCSs can be distinguished by their semantic and collocational preferences.

Abbreviations

PFV = perfective aspect (-le)

CL = classifier

DUR = durative aspect (-zhe)

GEN = genitive (-de)

Acknowledgements

I would like to express my gratitude to the anonymous reviewers for their constructive comments on the early versions of the paper.

References

- Anthony, L. (2011). AntConc (Version 3.2.2.1) [Computer Software]. Tokyo: Waseda University.
- Bybee, J. L. & Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Deng, X. (2010). *The acquisition of the resultative verb compound in Mandarin Chinese*. Hong Kong: The Chinese University of Hong Kong MA thesis.
- Deng, X. (2019). The acquisition of resultative verb compounds in Mandarin Chinese. *The Journal of Chinese Linguistics*, 47(1), 42-81.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories in implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.
- Ellis, N. C. (2006). Selective attention and transfer phenomenon in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27, 1-31.
- Ellis, N. C. (2012). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.
- Ellis, N. C., & Larsen-Freeman, D. (2009). Language as a complex adaptive system (Special Issue). *Language Learning*, 59: Supplement 1.
- Ellis, N. C., & Wulff, S. (2015). Usage-based approaches in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 75-

- 93). New York and London: Routledge.
- Feng, L. (2017). 汉语动结式事件结构加工的实验研究——以“哭湿”和“推倒”类为例 (An experimental study on the processing of Chinese VR constructions: A case study of Ku-shi and Tui-dao types). *语言教学与研究 (Language Teaching and Linguistic Studies)*, (1), 3-13.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Gu, Y. (2003). On the syntactic projection of causatives: Resultative compound predicates revisited. In J. Xu, D. Ji, & K. T. Lua (Eds.), *Chinese syntax and semantics* (pp. 27-62). Singapore: Pearson Prentice Hall.
- Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Liu, Y., Yao, T., Shi, Y., Ge, L., & Bi, N. (2010). *Integrated Chinese* (3rd ed.). Boston: Cheng & Tsui Company.
- Ma, Z., & Lu, J. (1997). A survey of adjectives as result complements. In A. O. Yue, & M. Endo (Eds.), *In memory of Mantaro J. Hashimoto* (pp. 155-172). Tokyo: Uchiyama Books.
- Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge, UK: Cambridge University Press.
- Qiao, Z. (2008). The acquisition of resultative verb compounds in Mandarin by English speakers. In R. Slabokova, J. Rothman, P. Kempchinsky, & E. Gavrusseva (Eds.), *Proceedings of the 9th generative approaches to second language acquisition conference (GASLA 2007)* (pp. 188-195). Somerville, MA: Cascadilla Proceedings Project.
- Smith, C. S. (1990). Event types in Mandarin. *Linguistics*, 28, 309-338.
- Smith, C. S. (1997). *The parameters of aspect* (2nd ed.). Dordrecht: Kluwer.
- Tai, J. H-Y. (1984). Verbs and times in Chinese: Vendler's four categories. In D. Testen, V. Mishar, & J. Drogo (Eds.), *Papers from the parasession on lexical semantics* (pp. 289-96). Chicago: Chicago Linguistic Society.
- Tai, J. H-Y. (2003). Cognitive relativism: Resultative construction in Chinese. *Language and Linguistics*, 4(2), 301-316.
- Tang, T. (1989). *Studies on Chinese morphology and syntax*. Taipei: Student Book.
- Thompson, S. A. (1973). Resultative verb compounds in Mandarin Chinese: A case for lexical rules. *Language*, 49, 361-79.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Boston: Harvard University Press.
- Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
- Wang, Y., Jiao, Q., & Pang, Y. (1987). *A dictionary of Chinese verbs and resultative complements collocations*. Beijing: The Beijing Language Institute Press.
- Wu, S., Yu, Y., & Zhang, Y. (2008a). *Chinese link: Intermediate Chinese, level 2 part 1*. Upper Saddle River, NJ: Prentice Hall.
- Wu, S., Yu, Y., & Zhang, Y. (2008b). *Chinese link: Intermediate Chinese, level 2 part 2*. Upper Saddle River, NJ: Prentice Hall.
- Wulff, S., & Ellis N. C. (2018). Usage-based approaches to second language acquisition. In D. T. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 37-56). Amsterdam/Philadelphia: John Benjamins.
- Xiao, R., & McEnery, T. (2004). *Aspect in Mandarin Chinese. A corpus-based study*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Xiao, R., Rayson, P., & McEnery, T. (2009). *A frequency dictionary of Mandarin Chinese: Core vocabulary for learners*. Routledge frequency dictionaries. London and New York: Taylor and Francis Group.
- Xing, H. (2003). 留学生偏误合成词的统计分析 (An analysis of the error compound words used by foreign learners of Chinese). *世界汉语教学 (Chinese Teaching in the World)*, (4), 67-78.
- Xu, D. (2000). 动补结构的上字和下字 (V-shang and V-xia in resultative construction). *语法研究和探索 (Grammar Research and Exploration)*, 112-120. 北京: 商务印书馆 (Beijing: The Commercial Press).
- Zhang, B. (2008). 基于中介语语料库的汉语词汇专题研究 (Interlanguage corpus based Chinese vocabulary research). 北京: 北京大学出版社 (Beijing: Peking University Press).
- Zhang, J. (2011). *Acquisition of the Chinese resultative verb complements by learners of Chinese as a foreign language: A learner corpus approach*. University Park: The Pennsylvania State University doctoral dissertation.
- Zhang, J. (2014). A learner corpus study of L2 lexical development of Chinese resultative verb compounds. *Chinese as a Second Language – The Journal of the Chinese Language Teachers Association, USA*, 49(3), 1-24.

Jie Zhang is Associate Professor of Chinese Pedagogy and Applied Linguistics in the Department of Modern Languages, Literatures, and Linguistics at the University of Oklahoma, USA. Her research interests include second language acquisition, foreign language pedagogy, corpus-based research, and sociocultural theory. She has published in the *Modern Language Journal*, *Language Testing*, *Chinese as a Second Language*, and *Language Teaching Research*.

汉语二语学习者结果状态式复合动词的词汇和语法发展： 一项基于用法的语料库研究

张洁

俄克拉荷马大学，美国

摘要

本文使用基于用法的二语习得理论，考察汉语学习者作文语料库，研究汉语二语学习者在习得结果状态式复合动词时的词汇和语法的发展过程。词汇发展从使用频率、成分多样性、准确度来考察；语法发展从结果状态式复合动词与“了”的共现来考察。研究发现结果状态式复合动词虽然在使用频率和成分多样性方面与学习者的语言水平成正比，但是准确度在中高级学习者中出现一定程度的退化。结果状态式复合动词的二语习得是一个既系统又复杂多样的过程，主要受输入、注意力阻滞、结果状态式复合动词自身特点的影响。针对结果状态式复合动词的具体习得特点，文章提出了相应的教学对策。

关键词

结果状态式复合动词，词汇发展，语法发展，学习者语料库，基于用法的二语习得理论

作者简介

张洁博士现任美国俄克拉荷马大学现代语言、文学和语言学系副教授。研究兴趣包括第二语言习得、外语教学法、学习者语料库研究、社会文化理论。在 *The Modern Language Journal*, *Language Testing*, *Chinese as a Second Language*, *Language Teaching Research*, 《世界汉语教学》等期刊上发表过文章。