2016

A Study On the Mutual Replacements of Three des in Chinese Blogs

Hui Sha
sandywisdom0909@gmail.com

Follow this and additional works at: http://scholarworks.umass.edu/masters_theses_2

Part of the Chinese Studies Commons

Recommended Citation
http://scholarworks.umass.edu/masters_theses_2/445

This Open Access Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses May 2014 - current by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
A Study On the Mutual Replacements of Three des in Chinese Blogs

A Thesis Presented

by

HUI SHA

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

September 2016

Chinese
A Study On the Mutual Replacements of Three des in Chinese Blogs

A Thesis Presented
by
HUI SHA

Approved as to style and content by:

_______________________________________________________________
Zhongwei Shen, Chair

_______________________________________________________________
Zhijun Wang, Member

_______________________________________________________________
David K Schneider, Member

_______________________________________________________________
Stephen Miller, Program Head
Asian Languages and Literatures, LLC

_______________________________________________________________
William Moebius, Chair
Department of Languages, Literatures and Cultures
DEDICATION

I would like to dedicate my thesis and give thanks to my family especially my mother, who has always supported and encouraged me whenever it is needed. I also want to give thanks to my friends, who have helped me with the data collection in parts of this work.

Without the support of my family and friends, this work would not have made possible. Many thanks, to my family and friends.
ACKNOWLEDGMENTS

First, I would like to give sincere gratitude and many thanks to my advisor, Professor Zhongwei Shen. His great linguistic knowledge and guidance have kept me improving throughout my pursuit of graduate studies. Professor Zhongwei Shen has given me a lot of creative ideas, insightful advices and support during the period of writing my thesis.

Second, I express my sincere thanks to my thesis committee: Professor David Schneider and Professor Zhijun Wang, for their encouragement, support, and comments.
ABSTRACT

A STUDY ON THE MUTUAL REPLACEMENTS OF THREE DES IN CHINESE BLOGS

SEPTEMBER 2016

HUI SHA, B.A., MINZU UNIVERSITY OF CHINA
M.A., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor ZHONGWEI SHEN

Three des, as structural particles in Chinese, are phonologically the same, but written differently. Through analyzing the written forms of these three homophonous particles, the research has come to some valuable conclusions that cannot be obtained by only observing the speaking language of Chinese.

This paper studies the relationships among three “des” (de1 as “的”; de2 as “地”; de3 as “得”), which function as structural particles in the written language of Chinese, by examining their mutual replacements in blogs. The research regards every living language as a Complex Adaptive System (CAS) with continuing changes. So this study’s perspective not only helps us understand more deeply the structures with three des, but also opens a new window to explore the variation of Chinese on the cognitive linguistic layer, including syntactic and semantic aspects.

Through analyzing the authentic data, which is obtained from a corpus built of articles in personal blogs including 400,000 Chinese characters, there are several worthy findings. First, the mutual replacements are asymmetric along with the generalization of de1. Secondly, there is a positive correlation between the frequency of replacements and the linguistic positive relevance among the three des, especially the syntactic and semantic aspects. Finally, the replacements among the
three des present a diverse and complicated situation when investigating the written forms of idiolects. The syntactic factor plays the main role in the replacements among the three des. The related degree of de1 & de2 is significantly higher than the one of de1 & de3, which is especially obvious on the writers with relatively frequent replacements.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
</tbody>
</table>

## CHAPTER

### 1. INTRODUCTION

1.1 Introduction To Three *des* in the Written Language of Chinese         1
1.2 Previous Studies on Three *des*                                      2
1.3 New Perspective of Studying Three *des*                                5

### 2. RESEARCH PROBLEMS AND HYPOTHESES

2.1 Issue Concerning Three *des*                                          8
2.2 Theoretical Analysis on Relationships Among Three *des*               9

#### 2.2.1 relationship between *de1* and *de2*

#### 2.2.2 relationship between *de1* and *de3*

#### 2.2.3 relationship between *de2* and *de3*

#### 2.2.4 relationships among *de1*, *de2* and *de3*

2.3 Research Problems and Hypotheses                                     15

#### 2.3.1 what—the first problem and hypothesis                           15
#### 2.3.2 why—the second problem and hypothesis                           16
#### 2.3.3 how—the third problem and hypothesis                            16

### 3. METHODOLOGY AND RESULTS

3.1 Corpus Building                                                      18

#### 3.1.1 target twenty writers                                          19
#### 3.1.2 gather 160 samples                                             21

3.2 Data Processing                                                     21

3.3 Data Analysis                                                       23
3.3.1 results ..................................................................................................................23
3.3.2 asymmetrical replacements and generalization of de1 .......................26

4. DISCUSSIONS .............................................................................................................29
   4.1 de1 Replaces de2 ....................................................................................................29
   4.2 de1 Replaces de3 ...................................................................................................35
   4.3 Replacements Among de1, de2 and de3 In Idiolects .......................................42
      4.3.1 zoom in on groups .........................................................................................43
      4.3.2 Complex Adaptive System applied in linguistics .......................................44
      4.3.3 zoom in on individuals ..................................................................................45

5. SUMMARY ..................................................................................................................50

BIBLIOGRAPHY ............................................................................................................52
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: First letters of the twenty bloggers' nick names</td>
<td>20</td>
</tr>
<tr>
<td>2: List of fixed words with $de3$ appearing in the corpus</td>
<td>22</td>
</tr>
<tr>
<td>3: Individual-total ratio of six replacements</td>
<td>23</td>
</tr>
<tr>
<td>4: Replace-replaced ratio of six replacements</td>
<td>24</td>
</tr>
<tr>
<td>5: Kinds of complement after $de3$</td>
<td>36</td>
</tr>
<tr>
<td>6: First letters of the twenty bloggers’ nick names</td>
<td>43</td>
</tr>
<tr>
<td>7: The percentage of replacements of twenty writers</td>
<td>45</td>
</tr>
<tr>
<td>8: Grouping situation when taking 10% as boundary</td>
<td>46</td>
</tr>
<tr>
<td>9: Grouping situation when taking 30% as boundary</td>
<td>47</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Relevance of three <em>des</em></td>
<td>15</td>
</tr>
<tr>
<td>2: Hypothetical replacements</td>
<td>15</td>
</tr>
<tr>
<td>3: Percentage of individual-total ratio of six replacements</td>
<td>24</td>
</tr>
<tr>
<td>4: Comparisons of replace-replaced ratio of six replacements</td>
<td>25</td>
</tr>
<tr>
<td>5: Actual replacements among three <em>des</em></td>
<td>26</td>
</tr>
<tr>
<td>6: Generalization of <em>de1</em></td>
<td>27</td>
</tr>
<tr>
<td>7: Kinds of complement</td>
<td>36</td>
</tr>
<tr>
<td>8: Distribution of replacements among three <em>des</em> in twenty writers</td>
<td>49</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Introduction To Three *des* in the Written Language of Chinese

Generally speaking, written language is the reflection of oral language. In Mandarin Chinese, what is particularly worth mentioning, three *des* are used as structural particles in the written language. However, actually they are phonologically the same in oral language. The written forms of these three homophonous particles were clearly not processed naturally, but were instead kind of man-made contrivances. As times goes, the usages of three *des* have infused into people’s natural written language system. The actual difference between oral language and written language is well worth studying. First of all, through investigating these three structural particles in written materials, we can access people’s cognitions on Chinese syntactic structures in oral language. Secondly, by studying how people apply the rules to practice, we can get several common characteristics and discover linguistic reasons for it. As well, we can discuss the individual differences while using three *des*. Thirdly, the relationships among three *des* plus some syntactic and semantic features of Chinese also can be dug from studying this special difference between oral language and written language of Chinese. Therefore, the paper will focus on studying the pivot—three *des*—between oral language and written language of Chinese and try to obtain some valuable and peculiar linguistic conclusions.
Since Chinese is not an inflectional language, as a supplement, applying to the forms of three homophonous particles is to clearly mark syntactic structures. Three des appear in the middle of two phrases (which can be words or minor sentences) to indicate relationships between the two parts and deliver syntactic and semantic information of the whole sentence. Three des (Although three des can be used as other parts of speech in Chinese, “three des” as a phrase in this paper refers solely to their usage as structural particles.) are the common structural particles in Chinese—“的” (referred as de1 in this paper), “地” (referred as de2 in this paper) and “得” (referred as de3 in this paper). They share the same pronunciation [tə] with neutral tone, thus they sound exactly same in colloquial speech. However, in written Chinese, de1, de2 and de3 are in charge of different divisions. Taking the pattern “A + de1/de2/de3 + B” as the general one, structures with de1, de2 and de3 can be depicted briefly and roughly as below:

A1 (attributive) + de1 + B1 (nominal head)
A2 (adverbial) + de2 + B2 (verbal/adjective head)
A3 (verbal/adjective) + de3 + B3 (complement)

1.2 Previous Studies on Three des

As mentioned above, studying the three des has unique linguistic value, especially in discussing the relationships between oral language and written language of Chinese. Therefore, there are many published articles discussing the three des.
Referring to *de1*, Xing Fuyi (Xing 1996) says *de1* is used between an attributive and a nominal head. It is the mark of attributive. Zhang Guoxian (Zhang 1994) classifies *de1* into three categories. From syntactic perspective, *de1* assists to form modification-center structure. For example, *xiǎozhào shīfu* (Master Xiaozhao) originally is a double-reference phrase Master Xiaozhao. But it becomes a modification-center structure after adding *de1*, that is *xiǎozhào de1 shīfu* (Xiaozhao's master). From semantic perspective, *de1* makes difference between with it and without it. Like, *huáqiáo yīyuàn* (Overseas Chinese Hospital) and *huáqiáo de1 yīyuàn* (the hospital of overseas Chinese) have different semantic references. From the pragmatics angle, it works as modification, as *hóng de1 shùyè* (red leaf). In the cognitive point of view, Shen Jiaxuan (Shen 2000) holds the opinion that *de1* changes unbounded into bounded. He explains why *xuěbái yīfu* (snow-white clothes) cannot establish but *xuěbái de1 yīfu* (snow-white clothes) can establish. That is because *xuěbái* (snow-white) is unbounded but *yīfu* (clothes) is bounded, they cannot match with each other on the semantic level. Lu Bingfu (Lu 2008) points out that the basic function of *de1* is the description on the semantic level.

As for *de2*, Zhang Bin and Zhang Yisheng (Zhang and Zhang 2000) conclude two roles that *de2* plays in the adverbial-center phrase. One is *de2* can make a non-adverbial phrase to function as adverbial. For example, *nánguò de2 kūle qǐlái* (begin to cry sadly) and *hàixiū de2 dīxiàle tóu* (bent the head shyly). The other function is to enhance description of phrases, such as *báibái de2 làngfèi* (waste idly) and *gèngjiā de2 nánguó* (more sad). Zhu Jingsong (Zhu 2007) classifies *de2* into two kinds: one
is used after adverbial to enhance its descriptive ability; the other is to endow some nouns, verbs and adjectives with descriptive ability to act as adverbial.

Likewise, many scholars have studied *de*3. Ding Shengshu (Ding 1997) states that “*de*3” structure belongs to complimentary structure. He thinks the part following *de*3 indicates the result or tendency for the part before *de*3. So the structure with *de*3 is complement. Fan Xiao (Fan 1992) clarifies *de*3 is structural auxiliary on the syntactic level. On the semantic level, it can show the completion of an action or behavior; it also can indicate affirmation.

The use of three *des* in written language of Chinese was stipulated by the authority and has been accepted by the public for more than fifty years (Tian 2004). However, there is controversy among Chinese linguists over whether using three *des* all the time. The representative, Lv Shuxiang (Lv 1981) insists that *de*2 should be entirely replaced by *de*1. At first, he argues *de*1 and *de*2 do not refer to different meanings while functioning as structural auxiliary. That is to say, *de*1 and *de*2 is one particle in essence. Secondly, if we use *de*1 to replace *de*2, it will bring many conveniences for Chinese writers, especially for Chinese students. Regarding to *de*3, Lv Shuxiang (Lv 1981) supports reserving it. But the Contemporary Chinese Dictionary (Chinese Academy of Social Science 2007) still retains the usages of three *des*. It states three *des* have different syntactic meanings. When they combine with other phrases, three *des* indicate different structural relationships—*de*1 is the mark of attributive, *de*2 is the sign of adverbial and *de*3 is the symbol of complement. Therefore, the authority keeps three different written forms—*de*1, *de*2 and *de*3—to make distinction of the homophonous form in written language.
1.3 New Perspective of Studying Three *des*

Based on the syntactic function summarized above, choosing to use one *de* from the three in a specific structure seems like a quite easy job for Chinese native speakers since the rule of how to use three *des* looks so clear. But it is not the case at all in the practical application of written Chinese. That’s because there exist sizable polysemous words, which can function as different parts of speech in sentences. What’s more, these words keep the same morphological appearance all the time.

In fact, even for native speakers, it is sometimes a tough task to decide which *de* should be used when facing a specific structure. Numerous usages, those do not obey the syntactic rules presented above, have been noticed in materials written by native speakers nowadays. The phenomenon used to be called “misuse of three *des*,” or even viewed as arbitrary errors. However, in fact, language is a living system with continuing changes everyday. Confronting such an actual living system at present, observing it objectively and then describing it in a faithful way is more valuable for linguistic research than subjectively making judgment of right or wrong.

Therefore, the uses of writing forms in the real situation happening in a living language are worth being paid attention to by linguistic researchers instead of simply being seen as misuse. One word, if the amount of the so-called “misuse” emerging among native speakers is large enough, deserves to be treated as an ongoing variation of the language. Hence it is more scientific and objective to name the phenomenon as “replacement” rather than “misuse”.
However studying the actual situation of using $de_1$, $de_2$ and $de_3$ by native speakers is a quite challenging project. The most time-consuming part is to collect enough first-hand reliable linguistic materials. Before the popularity of computers, writing materials accessible to linguistic researchers mostly have already been processed by professional editors. That is to say, most acquirable materials are not directly produced from the writers. According to the standard requirements of gaining actual data, the distance between those edited materials and the raw linguistic materials cannot be ignored. In other words, it becomes an obstacle to study the authentic usage of three $des$.

Fortunately, however, due to the widespread use of the Internet, every individual is endowed with the right to make his or her own voice. Thus people can bypass editors and publish their original essays to the world. Blog, an Internet product, was created for individuals to share their personal diaries online. Therefore, the linguistic materials collected from personal blogs are supposed to be viewed as first-hand reliable materials.

Apart from improving the reliability of data, more importantly the studying the linguistic materials from blogs lets us know more about the actual natural language. Several respectable linguistic scholars have already recognized the unique value of researching Internet languages. David Crystal holds the opinion that what makes Internet language so interesting, as a form of communication, is the way it relies on characteristics belonging to both sides of the speech/writing divide. (Crystal 2001, 28) Jill Walker Rettberg points out that blogs are conversational and social, they are constantly changing and their tone tends to be less formal and closer to everyday
speech than is the general tone of print writing. (Rettberg 2008, 33) Therefore, the linguistic materials collected from blogs offer a new perspective to study the mutual replacements of three des and the findings would lead us to comprehend more deeply our natural language.
2.1 Issue Concerning Three *des*

As the structural particles mentioned above, marking different relationships between the parts placing before and after them is the main function of three *des*. But their syntactic functions are not quite equally distinct from each other. So the actual usage situation among native speakers becomes an interesting issue because it will show some linguistic features of written Chinese. Much more interesting, the creative research perspective of this study—taking three *des*’ mutual replacements as—will at least guide us to explore the issue on two levels.

The first level is getting to know more particularly about the relationships of different parts in one sentence. The second one is to understand more profoundly the subtle relationships among three *des*. In reality, the first level is the foundation of the second level. Accordingly, characteristics of the first level would present on the second level. Consequently, the first level will be comprehended fully when investigating the second level. In view of that, this paper is organized based on the second level’s structure embedded the first level’s content.

In conclusion, through the perspective we can uncover complicated relationships within one sentence, which contains one or more than one of three *des*, from the syntactic and semantic aspects. Additionally, it is able to dig out some people’s cognitive features during the process of using three *des* as particles in written Chinese.
2.2 Theoretical Analysis on Relationships Among Three des

2.2.1 relationship between de1 and de2

Generally speaking as mentioned above, both de1 and de2 indicate A modifies B. Comparing the importance of information conveyed by A and B, B plays a more central role than A. There are diverse subordinate relationships belonging to the “modifier (A)–center (B)” relationship. Chinese linguists adopt many standards to do the classifying job, but more or less they have already reached the consensus on the general “modifier (A)–center (B)” relationship.

As for the differences between de1 and de2, the parts of speech of A1 and A2, B1 and B2 are all different. In short, they can be demonstrated as below:

\[
\begin{align*}
A1 \text{ (attributive)} & \quad + \quad de1 & \quad + \quad B1 \text{ (nominal head)} \\
\quad \text{(modifies)}& \\
A2 \text{ (adverbial)} & \quad + \quad de2 & \quad + \quad B2 \text{ (verbal/adjective head)} \\
\quad \text{(modifies)}& 
\end{align*}
\]

The main reason of utilizing de1 and de2 instead of only one structural auxiliary is the existence of polysemous words. They have different syntactic functions in sentences without any morphological change. A case in point is some words can be used as both attributive and adverbial—the parts of speech respectively accord with A1 and A2 in order—without any morphological change. For example,

(1) ānquán de1 dìfāng
    
safe de1 place
    ‘a safe place’

(2) ānquán de2 dìdá
safely  de2 arrive
‘arrive safely’

“ānquán” (safe) is an attributive in Phrase (1), meanwhile acts as an adverbial in Phrase (2). In this case, the structural particles de1 and de2 point out the different parts of speech of “ānquán”—used as “safe” or “safely” in respective structures but with the same form. Yet at least the parts of speech of B1 and B2 are different and can be obviously distinguished. To some degree, the part of speech of B serves as hint while determining to use de1 or de2.

What’s more, the polysemous words appear in front of de1 and de2, as well as after them. Especially for the words that can both function as nouns and verbs. Such as,

(3) Tā zuòle yígè míngzhì de1 xuǎnzé.
he made a wise de1 choice
‘He made a wise choice.’

(4) Tā míngzhì de2 xuánzéle chūguó liúxué.
he wisely de2 chose abroad study
‘He wisely chose to study abroad.’

By using de1 and de2, it becomes much clearer for audiences to recognize the lexical category of the words standing before and behind de1 and de2. On this occasion, “xuǎnzé” acts as nominal—choice—in Sentence (3), while as verbal—choose—in Sentence (4). Correspondingly, “míngzhì” serves as attributive—wise—
in (3) and adverbial—wisely—in (4). Moreover, \textit{de}1 and \textit{de}2 indicate the roles of the chunks with them are within the sentence. Take “\textit{míngzhi de xuǎnzé}” as example, in Sentence (3) it functions as NP (noun phrase) and in Sentence (4) it is a VP (verb phrase). In this way, thus \textit{de}1 and \textit{de}2 mark not only the parts of speech of specific words before and after them, but also the syntactic function of the whole structure with them. In fact, the syntactic functions of the structures decide the parts of speech of central words in this case, not the other way around. The central words stand in the place of B in the three \textit{des’} pattern above.

According to the theoretical analysis, it looks like \textit{de}1 and \textit{de}2 help deliver quite different but important syntactic information for people’s communication. But it is noticeable that actually it sets a relatively high requirement for expressers to apply the rule. They are supposed to keep conscious all the time and be sure of all the functions of every word in the sentence. For native speakers, this cognitive process can be done on a conscious or subconscious level. This study will observe the actual situation of native speakers via researching the mutual replacements between \textit{de}1 and \textit{de}2.

2.2.2 relationship between \textit{de}1 and \textit{de}3

As the structural auxiliary, \textit{de}3 introduces complement following verbal or adjective. Referring to the complement standing after \textit{de}3, it can be adjective, verbal, verb phrase, minor sentence or particle, or sometimes it even can be nothing—leaving it as a vacancy. From the point of semantic perspective, the complement can be divided into describing situation, manifesting degree and pointing out result.
Just like its syntactic name—complement, apparently it serves for the center, as A3 in the pattern below, instead of being the center itself. In this sense, compared with the pattern of de1, the one of de3 is quite different. De3 marks an opposite syntactic direction between A and B, shown as below. Therefore, the replacements between de1 and de3 should be less on the basis of the theoretical analysis. This study will investigate it through real data and then verify whether the deduction conforms to the actual situation or not.

\[ A1 \text{ (attributive)} + de1 + B1 \text{ (nominal head)} \]

\[ A3 \text{ (verbal/adjective)} + de3 + B3 \text{ (complement)} \]

2.2.3 Relationship between de2 and de3

\[ A2 \text{ (adverbial)} + de2 + B2 \text{ (verbal/adjective head)} \]

\[ A3 \text{ (verbal/adjective)} + de3 + B3 \text{ (complement)} \]

Above are the general structures while using de2 and de3. When adverbial acts as the complement, the pattern of de3 just seems like an inversion of the one with de2, for instance:

Group I

(5) fēikuài de2 pāo

pretty fast de2 run
‘run pretty fast’

(6) pǎo de3 fēikuài  
run de3 pretty fast  
‘run pretty fast’

Group II

(7) xiàng zhū yīyàng de2 huó  
like pig de2 live  
‘live like a pig’

(8) huó de3 xiàng zhū yīyàng  
live de3 like pig  
‘live like a pig’

Group III

(9) liǎnsè rútóng zhǐ yīyàng de2 bái  
face as paper de2 pale  
‘face is as pale as paper’

(10) liǎnsè bái de3 rútóng zhǐ yīyàng  
face pales de3 as paper  
‘face is as pale as paper’

In Example (6), (8) and (10), adverb and prepositional phrases respectively function as complements. In the case of Group I, II and III, the patterns of de2 and de3 can be showed as below:
Therefore, the patterns above are acting as mirror images of each other. They can be presented like that:

\[
\begin{align*}
A2 &= B3 + de2 + B2 = A3 \\
A3 &= B2 + de3 + B3 = A2
\end{align*}
\]

Because the syntactic structures are symmetrically opposite, the structural difference between \(de2\) and \(de3\) is obvious. That is to say \(de2\) doesn’t relate to \(de3\) very much syntactically. Distinguishing \(de2\) and \(de3\) in written language should not be a very confusing job.

### 2.2.4 relationships among \(de1\), \(de2\) and \(de3\)

\[
\begin{align*}
A1 \text{ (attributive)} &= + de1 + B1 \text{ (nominal head)} \\
A2 \text{ (adverbial)} &= + de2 + B2 \text{ (verbal/adjec}tive \text{ head)} \\
A3 \text{ (verbal/adjec}tive) &= + de3 + B3 \text{ (complement)}
\end{align*}
\]

As shown above, compared with the patterns of \(de1\) and \(de2\), A is the syntactic center of the whole structure marked by \(de3\) instead of B. In this sense,
considering the degree of similarity of three *des*, the syntactic relationship between *de1* and *de2* is closer than the relationship between *de3* and *de1 / de2*. In regard to their structural difference, it can be showed roughly in image as figure 1. Since there are a lot of mutual replacements among three *des*, presenting the relationships like figure 2 would be more appropriate.

![Figure 1: Relevance of three *des*](image1.png)

![Figure 2: Hypothetical replacements](image2.png)

### 2.3 Research Problems and Hypotheses

This study focuses on the replacements among three *des* (the overlapped parts in the sketch), and discusses what linguistic factors are affecting the replacements. In sum, here are the research problems and their respective hypotheses, which are made according to the known linguistic theories.

#### 2.3.1 what—the first problem and hypothesis

Problem 1: Let the real data tell us what the actual situation is of mutual replacements among three *des*. What is the frequency specifically of the six replacements listed below? In the sketch, which overlapped part is the largest one
and which is the smallest? Are the mutual replacements between two *des* equal in quantity or not?

1. *de*1 replaces *de*2;
2. *de*2 replaces *de*1;
3. *de*1 replaces *de*3;
4. *de*3 replaces *de*1;
5. *de*2 replaces *de*3;
6. *de*3 replaces *de*2.

Hypothesis 1: Based on the syntactic functions and the usage history of three *des*, it is assumed that the replacement of ① appears most frequently and ⑤, ⑥ appear least frequently. The mutual replacements between two *des* are not equal in quantity, presenting a tendency instead.

### 2.3.2 why—the second problem and hypothesis

Problem 2: What are the linguistic factors influencing these mutual replacements?

Hypothesis 2: It is presumed that the relevance in syntactic aspects among three *des* is the main reason. Further, the higher degree of relevance, the more replacements happened between two *des* and vice versa.

### 2.3.3 how—the third problem and hypothesis

Problem 3: How do the factors affect mutual replacements? Is there a factor playing the key role? Do the frequency of replacements keep the same on different
groups, divided by economic status and literacy level, of writers of this study? What will we get when zooming in on every individual? How does individual language perform on the point of the issue—mutual replacements among three *des*?

Hypothesis 3: The syntactic characteristics would be the key factor. Different groups may perform differently on three *des’* mutual replacements. Individual languages might present quite differently on the issue since the variation of individual language is not synchronous.
CHAPTER 3

METHODOLOGY AND RESULTS

3.1 Corpus Building

This study builds 400,000 Chinese characters corpus, which is collected from twenty people’s personal blogs. The twenty bloggers come from two groups: Ten are scholars and the other ten are “grass-roots,” indicating that these people come from the bottom of society. So the standards of dividing writers in this study are economic status and literacy level.

Blogs were already known by Chinese people in 2005 and attract more and more users since 2006. In order to reflect the sustaining situation of the replacements, articles written by the same bloggers from 2006 to 2013 are gathered into the corpus chronologically.

With the purpose of managing the corpus systematically, each sample has been given a serial number. Four digits form a number. The first digit represents which group the sample comes from, “1” means scholars’ group and “2” indicates grass-roots group. The second digit is the sequence number within the group, starting from zero to nine. The latter two digits show the writing date, only specific to year. Take “2209” as an example, it tells that the sample is extracted from the blogs written by the third writer (the second “2” indicates) in the grass-roots group (the first “2” shows). In addition to, the sample with 2,500 characters is written in the year of 2009 (as “09” shows).
3.1.1 target twenty writers

SINA, a famous Chinese Internet corporation, has already divided bloggers into several groups—the scholars and the grass-roots are among them. The bloggers are listed alphabetically based on the first letter of their nick names. Under the principle of extracting random samples, to understand the real situation of Chinese blogs, this study obeys three regulations while seeking twenty qualified writers. The qualification is that bloggers must have published enough original articles (blogs) every year from 2006 to 2013.

Here are the three regulations:

1. Fixed point rule

Since it is necessary to find ten qualified bloggers in each group, this study regards “A, C, F, H, K, M, P, R, U, W” as the fixed points. The process is examining the bloggers listed under these fixed points from the first to the end until finding the qualified one. For example, in the scholars’ group, the first blogger under C list is “Cài Tiānxīn”, but he just opened his personal blog in 2007. So he is not qualified to be the study writer. The next step is continuing to look for the next blogger under C list.

2. Preceding rule

The Internet offers the freedom for everyone to express himself or herself. It increases opportunities for linguistic researchers to be more easily exposed to the authentic language materials, especially written language or Internet language. But the freedom for users also brings some troubles for researchers when doing serious
scientific study. The targeting process cannot be done if only adhering to the first rule. There is no qualified blogger under certain fixed points’ lists.

So the second rule comes into play—the preceding rule. For example, if there is no qualified blogger under K list, the next step is going back to J list. Then, look for the bloggers from the first one listed under J. If it fails, even go back to I list.

3. Adjacent rule

During the process of data collection, for the most cases, the two principles above are enough to find the qualified writers. But they cannot solve all the problems. Therefore, the last rule applied by this study is the adjacent rule. That is to say, after examining all the bloggers listed under and before the fixed points, but failing, widening the range after the fixed point is the next step. For instance, the first step is to look for the blogger under R list, but there is no qualified one. Then search the list under letter Q according to the second rule. Unfortunately, there is still no qualified writer. If continuing to search ahead, the one under the fixed point P would be collected repeatedly. So this method does not work. It is time to apply the third rule—seeking for the bloggers listed under S, and perhaps even T.

As a result, below is the table of first letters of the twenty bloggers’ nick names.

<table>
<thead>
<tr>
<th>Group</th>
<th>First Letters of the Twenty Bloggers’ Nick Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Points</td>
<td>A       C       F       H       K       M       P       R       U       W</td>
</tr>
<tr>
<td>Serial No.</td>
<td>0       1       2       3       4       5       6       7       8       9</td>
</tr>
<tr>
<td>1. Scholars</td>
<td>A       C       F       H       J       M       P       R       T       W</td>
</tr>
<tr>
<td>2. Grass Roots</td>
<td>A       C       F       K       L       Q       S       T       W       Y</td>
</tr>
</tbody>
</table>

Table 1: First letters of the twenty bloggers’ nick names
3.1.2 gather 160 samples

After determining the twenty writers of study, the next step is to extract individual samples from writers' blogs every year. Each sample includes 2,500 Chinese characters. It is eight years from 2006 through 2013, so every blogger ought to be sampled eight times chronologically. Normally the first 2,500 characters in the articles listed ahead of that year are extracted. When coming across the articles, forwarded from newspapers or magazines, they would be rejected for the corpus because these articles almost certainly have been edited by others instead of coming directly from the writers. This kind of situation happens more frequently in scholars’ group than the grass-roots group. Fortunately, it is pretty easy to identify. In consideration of copyright, bloggers usually declare it in the title or at the end of the article.

In a word, sampling randomly and scientifically under the feasible premise is the general guideline. Considering the complicated circumstances of blogs online, the whole standard procedure of data collection is a creative and effective trial in sampling Internet language. Eventually 160 samples with totally 400,000 Chinese characters form the corpus for this study.

3.2 Data Processing

There are three steps in processing every sample. The first step is locating sentences containing the three Chinese characters (的, 地, 得) with the help of the search function built into Microsoft Word. However, the three Chinese characters
not only are used as structural particles, but also have other usages in Modern Chinese. After picking out sentences with the three characters, it is necessary to get rid of the examples—in which the three characters are not used as structural particles, like “的确 díquè (indeed), 地点 dìdiǎn (location), 得到 dédào (obtain)” and so on.

When dealing with de3, there is a tricky question. In Modern Chinese, some combinations of a verb plus de3 have already evolved into fixed words. In this case, de3 should be seen as the morpheme rather than the structural auxiliary. It is an ongoing process to form the fixed words. So when coming across the structure with de3, the Contemporary Chinese Dictionary (Chinese Academy of Social Science 2007) is the standard to determine whether to count it into the sum of using de3 as structural auxiliary or not. That is to say, if the structure has already been recorded as an entry in the dictionary, it won’t be counted into the sum of using de3 as structural auxiliary.

Below is the table listing the fixed words, which appear in the corpus, already taken down into the dictionary as entries.

<table>
<thead>
<tr>
<th>Chinese</th>
<th>Meaning in English</th>
<th>Chinese</th>
<th>Meaning in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>shìde</td>
<td>make; cause; render</td>
<td>juéde</td>
<td>feel</td>
</tr>
<tr>
<td>shěngde</td>
<td>so as to avoid</td>
<td>dōngde</td>
<td>understand; know</td>
</tr>
<tr>
<td>xiānde</td>
<td>look; seem; appear</td>
<td>rènde</td>
<td>recognize; identify</td>
</tr>
<tr>
<td>luòde</td>
<td>end up</td>
<td>bùyóude</td>
<td>cannot help</td>
</tr>
</tbody>
</table>

Table 2: List of fixed words with de3 appearing in the corpus
The last step is examining the sentences with these three des one by one to observe whether there is a replacement happening or not. If there is, make a record of the sentence for studying in depth.

3.3 Data Analysis

3.3.1 results

There are totally 308 replacements happening among 400,000 Chinese characters. Also, de1 appears 14,398 times, de2 appears 460 times and de3 appears 390 times. Below are the tables and figures demonstrating the detailed data of the six replacements.

<table>
<thead>
<tr>
<th>NO.</th>
<th>Replacements</th>
<th>Times</th>
<th>Individual-Total Ratio of Replacements</th>
</tr>
</thead>
<tbody>
<tr>
<td>①</td>
<td>de1 replaces de2</td>
<td>180</td>
<td>58.44%</td>
</tr>
<tr>
<td>②</td>
<td>de2 replaces de1</td>
<td>10</td>
<td>3.25%</td>
</tr>
<tr>
<td>③</td>
<td>de1 replaces de3</td>
<td>107</td>
<td>34.74%</td>
</tr>
<tr>
<td>④</td>
<td>de3 replaces de1</td>
<td>10</td>
<td>3.25%</td>
</tr>
<tr>
<td>⑤</td>
<td>de2 replaces de3</td>
<td>1</td>
<td>0.32%</td>
</tr>
<tr>
<td>⑥</td>
<td>de3 replaces de2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>308</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 3: Individual-total ratio of six replacements
<table>
<thead>
<tr>
<th>NO.</th>
<th>Replacements</th>
<th>Times</th>
<th>Times of de should be</th>
<th>Replace-Replaced Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>①</td>
<td>de1 replaces de2</td>
<td>180</td>
<td>(de2) 629</td>
<td>28.62%</td>
</tr>
<tr>
<td>②</td>
<td>de2 replaces de1</td>
<td>10</td>
<td>(de1) 14131</td>
<td>0.07%</td>
</tr>
<tr>
<td>③</td>
<td>de1 replaces de3</td>
<td>107</td>
<td>(de3) 488</td>
<td>21.93%</td>
</tr>
<tr>
<td>④</td>
<td>de3 replaces de1</td>
<td>10</td>
<td>(de1) 14131</td>
<td>0.07%</td>
</tr>
<tr>
<td>⑤</td>
<td>de2 replaces de3</td>
<td>0</td>
<td>(de3) 488</td>
<td>0.20%</td>
</tr>
<tr>
<td>⑥</td>
<td>de3 replaces de2</td>
<td>0</td>
<td>(de2) 629</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 4: Replace-replaced ratio of six replacements

Based on the data showing in Table 3 and 4, the relationships of six replacements in quantity can be observed visually in the figures below.

Figure 3: Percentage of individual-total ratio of six replacements
Figure 4: Comparisons of replace-replaced ratio of six replacements

The high percentage of replacements, especially replacements ① (28.62%) and ③ (21.93%), proves that the mutual replacements among three des cannot be simply regarded as word misuse. Moreover, it is worth studying carefully and deeply. In particular, there is a big disparity between replacements ① and ②, ③ and ④.

In contrast, the replacements between de2 and de3 (⑤ and ⑥) are so rare that they can almost be seen as non-existent. The phenomenon has value to be researched as well, particularly comparing with replacements ① and ③. The situation can be showed roughly with this image.
Besides, the reasons for the gap (the bracket points out in Figure 4) between ① and ③ are waiting for further investigation.

In sum, the first hypothesis is largely in line with the real data. In absolute figure terms, the replacement of ① appears most frequently and ⑤, ⑥ appear least frequently. In terms of percentage, the replacement of ① also appears most frequently and ⑥ appears least frequently. But actually there is only one replacement of ⑤, so to some degree it can be ignored.

### 3.3.2 asymmetrical replacements and generalization of \textit{de1}

Based on the data, the mutual replacements of three \textit{des} present an obvious asymmetric situation instead of an equal one. The frequencies of \textit{de1} replacing \textit{de2} (28.62\%) and \textit{de3} (21.93\%) are much higher than \textit{de1} being replaced by \textit{de2} (0.07\%) and \textit{de3} (0.07\%). It indicates there is a tendency of the mutual replacements between \textit{de1} and \textit{de2}, \textit{de1} and \textit{de3}. Most of these replacements direct to \textit{de1} instead of \textit{de2} or \textit{de3}. In other words, \textit{de1} is in the strong position in the word competition of three \textit{des}. Therefore, the generalization of \textit{de1} has formed. It is depicted as below:
The generalization of *de1* is due to three reasons. First at all, as the structural auxiliary, *de1* (appearing 14,398 times in the corpus) is used much more frequently than *de2* (appearing 460 times in the corpus) and *de3* (appearing 390 times in the corpus). As mentioned above, when there is no morphologic change, it is difficult even for native speakers to decide which *de* should be used in some situations. Thus, for improving the accuracy in using three *des*, native speakers are inclined to choose the most used one when they are not quite sure. Besides, from the word competition perspective, the word, used frequently would have advantage in the competition.

Secondly, the previous usage habit has some effects. Back to the history, the three *des* were not always used separately. To be strict, after the foundation of the People’s Republic of China, the explicit division of three *des* like today was just stipulated officially. Especially, before May Fourth, there is no *de2*. It is *de1* that took the duty of *de2*. For example, the outstanding novels *Dream of the Red Chamber* and *The Scholars* from the Qing Dynasty apply *de1* solely instead of distinguishing *de1* and *de2*. The previous usage habit sets an example for people and fuels the generalization of *de1* nowadays.
Thirdly, as a character, the main function of “的” is as the structural auxiliary but that is not the case for “地” and “得.” In written Chinese, one character usually possesses many meanings and assumes more than one function. It is kind of similar to one person undertaking multiple responsibilities in society. For “地,” the meaning “ground; field” is used more frequently than the structural auxiliary de2. Also for “得,” the usage of de3 is much less than using as verb “get; obtain.” That is to say, “的” always does strengthen the usage as structural auxiliary while “地” and “得” do not. From the cognitive perspective, the connection between “的” and structural auxiliary become tighter and tighter in people’s mind. Meanwhile, the connection between “地、得” and structural auxiliary is gradually weakening. De1 tends to take over the duties of de2 and de3 rather than the reversed direction. Hence, the asymmetrical replacements and generalization of de1 arise. In the following discussion, we will focus on de1 replacing de2/de3 instead of the reversal direction because of the asymmetry.

Additionally, the mutual replacements between de2 and de3 are quite less—both account for only 0.07%. It verifies that the linguistic distance between de2 and de3 is relatively far and the syntactic positive relevance of de2 and de3 is low. Therefore, the authentic linguistic materials agree with previous analysis concerning this case.
CHAPTER 4
DISCUSSIONS

4.1 \textit{de}1 Replaces \textit{de}2

There are 180 replacements of \textit{de}1 replacing \textit{de}2 in total 629 \textit{de}2. It accounts for nearly thirty percent of the whole usages of \textit{de}2. What’s more, it is the most frequent one among the six replacements. There are two main reasons for this. One is the ambiguity of the lexical category. A number of polysemous words increase the difficulty when distinguishing \textit{de}1 and \textit{de}2. That is to say an exact same word can act as verb, adjective or adverb without any hint of morphological changes. For example,

(11) 1607

\begin{verbatim}
Shēngmìng    \textit{de}1 yìyì    bú     zài yú yōngsú    \textit{de}1 (should be \textit{de}2) huózhe.
\end{verbatim}

\begin{verbatim}
life       \textit{de}1 meaning      not    lie in vulgarly    \textit{de}1 (should be \textit{de}2) living
\end{verbatim}

‘The meaning of life does not lie in living vulgarly.’

Here \textit{yōngsú} (vulgar) is originally an adjective but functions as adverb to modify the verb \textit{huózhe} (living). But the blogger still treats \textit{yōngsú} (vulgar) as an adjective, so the replacement happens.

(12) 2007
Tā rèqíng  
  de1 (should be de2) yāoqǐng  wǒ cānjiā

yígè wǎnyàn.

he passionately  
de1 (should be de2) invited  me to take part in

a banquet

‘He passionately invited me to take part in a banquet.’

Rèqíng can act as nominal, adjective and adverb as below:

Nominal: rèqíng sihuó (the enthusiasm is like a flame)

Adjective: rèqíng de1 yōngbào (a warm hug)

Adverbial: rèqíng de2 zhāodài (enthusiastically entertain)

Besides, yāoqǐng can also act as nominal (invitation) and verbal (invite). Thus, the appearance of de1 replacing de2 is reasonable.

(13) 2306

Wǒmen yīnggāi hěn yánsù hěn zhuāngzhòng de1 (should be de2) duidài zhèjiànshì.

we should quite seriously quite solemnly de1 (should be de2) treat the thing

‘We should treat the thing quite seriously and solemnly.’

Similarly, zhuāngzhòng (solemnly) acts as adjective, like yíwèi zhuāngzhòng de1 yīhūn fùnǚ (a dignified matron); also acts as adverb, such as zhuāngzhòng de2
qingzhù jiérì (celebrate the festival with dignity). Additionally, duidài can mean “treat” or “treatment” in different sentences.

From three sentences above, it is clear that the flexible parts of speech of the words before and behind de are the main reason for the high percentage of replacements between de1 and de2.

The other reason is on the syntactic layer, especially when de is used in a subordinate unit. As mentioned above, using de1 or de2 has a great relationship with parts of speech of the words before and after de. In a compound sentence, parts of speech of words are up to the syntactic functions they play. It requires the writers to make sure of the syntactic functions of each part in their mind.

Take an example from the research corpus,

(14) 1510
Nà wǒ jiù jiǎnmíng’èyào de1 (should be de2) zōngjié yíxià wǒ de1 èrlínglíngjiù ba.
then I briefly de1 (should be de2) sum up my
del 2009

‘Then I am going to sum up my year of 2009 briefly.’

Zōngjié can be a noun (summary) and verb (to sum up). Here zōngjié actually is used as a predicate for the whole sentence. After abbreviation of the sentence, actually the subject is wǒ (I), the predicate is zōngjié (sum up) and the object is
èrlínglíngjiǔ (2009). Therefore, zōngjié in the sentence is a verb, not a noun. It is supposed to use de2 instead of de1. But if the sentence is saying like below:

Nà wǒ jiù zuò yígè wǒ de1 èrlínglíngjiǔ de1 jiǎnmíng’èyào

de1 (should be de1) zōngjié ba.

Then I make a my de1 2009 de1 brief

de1 (should be de1) summary

‘Then I am going to make a brief summary of my 2009.’

In the edited sentence, the subordinate unit jiǎnmíng’èyào de zōngjié has not changed at all, but it functions differently. In the former sentence, which really appears in the corpus, the unit functions as predicate. In comparison, it functions as an object following the predicate zuò (make) in the rewritten sentence. In fact, to mark different syntactic functions is one of the original purposes to apply three des. However, the variety of syntactic functions, one same phrase can take, conversely increases the difficulty in choosing one of three des. Nearly thirty percent of the replacements demonstrate that when facing the multiple functional phrases, the syntactic role may sometimes present a lack of clarity in native speakers’ mind. Let’s investigate another example from the same blogger to analyze more concretely how linguistic factors affect the replacements between de1 and de2.

(15) 1507
Tāmen què hěn xiànsì de (should be de2) bāwò zhùle zìgěr

de1 mingyùn: niè bùyándùchuán de2 bā shì nònqdàle!
they but quite realistically de1 (should be de2) seized own
de1 destiny: without saying any word de2 obtaining the success

'They seized their own destiny quite realistically: obtaining the success without saying any word.'

In this sentence, de appears three times and two of them are used correctly. It means the author has already been aware of distinguishing de1 and de2 but cannot succeed in doing it all the time. The replacements of 1507 and 1510 are coherent, not just happening randomly. Both of them indicate that there is a correlation between the frequency of replacements and the linguistic relevance. Let us compare the three minor sentences with de in detail.

(15.1)

hěn xiànsì de (should be de2) bāwò zhùle
quite realistically de1 (should be de2) seized

'seized quite realistically'

(15.2)

zìgěr de1 mingyùn
their own de1 destiny

'their own destiny'

(15.3)
bùyánbúchuǎn  
without saying any word  

de2 bā shì nòngdàle  
de2 obtaining the success  

‘obtaining the success without saying any word’

In (15.2), the part of speech of mingyùn (destiny) is exclusively nominal. In (15.3), bā shì nòngdàle is a “ba” structure indicating to dispose the object, which is apparently verbal. But bǎwò (seize; certainty), showing in (15.1), can be verbal and nominal as well. Like, bǎwò jīhuì (seize the opportunity), here bǎwò acts as verbal; hǎowú bǎwò (have not the least certainty), here bǎwò acts as nominal. Specifically bǎwò in (15.1) acts as verbal because it functions as the predicate for the whole sentence and it follows a result complement zhùle as well. However the multiple parts of speech of bǎwò still confuse the writer.

In addition to, hěn xiànshí, the phrase standing before de, is another important factor resulting in the replacement. Actually, the syntactic structure of hěn xiànshí is adverb (hěn) modifying adjective (xiànshí). If only considering xiànshí, which can act as nominal and adjective, there is a tendency to use de1 instead of de2. Even regarding hěn xiànshí as a whole, it also can function as adjective, such as, hěn xiànshí de1 shèhuì (a very realistic society). But hěn xiànshí in the sentence above in example (15) is supposed to be viewed as an adverb, which modifies the action of seizing.

Combining the two factors above, it is quite understandable that the author replaces de1 by de2. Due to the high flexibility of syntactic function, actually the choice between de1 and de2 not only relates to the parts of speech but also is
connected with the syntactic structure of the whole sentence. It requires the native speakers to have a very clear vision of the whole structure of that sentence when using *de* in their mind. But the principle of proximity during the writing process sometimes plays a negative effect, just like the part of speech of *xiànsí*. From the cognitive perspective, that is to say the relationship between adjacent words may influence people’s grasp of the whole structure of the sentence.

In brief, the multiple parts of speech and the high flexibility of syntactic functions combined together lead to the high percentage of *de1* replacing *de2*.

### 4.2 *de1* Replaces *de3*

There are 107 replacements of *de1* replacing *de3* in total 488 *de3* in the corpus. It accounts for approximately twenty percent of the whole usages of *de3*. It is the second most-frequent replacement, less than the replacements of *de1* replacing *de2*. Two main reasons can explain the relatively high frequency. To some degree, it is consistent with the theoretical analysis in Chapter 2. The syntactic relationships between *de1* and *de2* are closer than *de1* and *de3*. The comparison of the related degrees of *de1 & de2* and *de1 & de3*, will be discussed specifically in Chapter 4.4.

According to the *Xiandai Hanyu Cidian* (Contemporary Chinese Dictionary), *de3*, as structural auxiliary, has three main functions: 1. used between a verb and its complement to indicate possibility, e.g. *ná de3 dòng* (be capable of carrying); 2. used after a verb or an adjective to introduce a complement of result, e.g. *tiào de3 gāo* (jump high); 3. used after a verb or an adjective to introduce a complement of degree, e.g. *kuài de3 hěn* (quite fast). Besides, as Chapter 3 mentioned, some
combinations of verb plus *de3* have already become fixed words, e.g. *xiăngde* (seem).

We call this kind of replacement as the fourth in the table. According to Lu Shuxiang’s statement, all replacements of *de1* replacing *de3* are classified into these four kinds. Below is the table and figure to specifically show the appearing times and percentages of these four kinds separately.

<table>
<thead>
<tr>
<th>Kinds</th>
<th>1. possibility</th>
<th>2. result</th>
<th>3. degree</th>
<th>4. fixed words</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Times</td>
<td>7</td>
<td>70</td>
<td>29</td>
<td>1</td>
<td>107</td>
</tr>
<tr>
<td>Percentage</td>
<td>6.54%</td>
<td>65.42%</td>
<td>27.11%</td>
<td>0.93%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 5: Kinds of complement after *de3*

Based on the data, the replacements of *de1* replacing *de3* happen most frequently in the situation that *de3* is used after a verb or an adjective to introduce a complement of result. The data is in accordance with the first reason for the high frequency of *de1* replacing *de3*. The reason is there exists one particular usage of
As we know, omitting unnecessary words is one of the significant characteristics of written Chinese. In the structure:

\[ \text{A1 (attributive)} + \text{de1} + \text{B1 (nominal)} \]

In some situations, B1 can be omitted if the context has already indicated what the nominal structure is. That is to say, the structure changes into:

\[ \text{A1 (attributive)} + \text{de1} \]

Mostly the comment will follow the structure above to form a “Topic-Comment” sentence. When the comment points out result, its format presents as exactly same as de3 used after a verb or an adjective to introduce a complement of result. But de1 and de3 indicate different syntactic structures. For instance,

(16) Tā chàng \text{de1 (“gē” omitted)} hěn hǎotīng.

she sings \text{de1 (“song” omitted)} very good

'The song singing by her sounds very good.'

(17) Tā chàng \text{de3 hěn} hǎotīng.

she sings \text{de3 very} well

'She sings very well.'

In example (16), actually gē (song) after de1 has been omitted but actually is the topic of the sentence. Hěn hǎotīng (very good) forms the comment to the topic. As for example (17), hěn hǎotīng (very well) acts as a complement for the sentence. Although (16) and (17) deliver the same semantic information, the syntactic structures are different respectively marked by de1 and de3. However, the existence
of this particular usage of *de*1 would leave an impression in native speakers’ minds that *de*1 and *de*3 are interchangeable under several situations. The wrong impression would cause confusion while determining using *de*1 or *de*3. It also accounts for the replacements happening most frequently when *de* introduces result complement. For example,

(18) 2206

\[ Fâte呼吸 \quad \text{bǎ} \quad \text{piànzi pāi} \quad \text{*de1 (should be de3) wéimēi.} \]

French made movie *de1 (should be de3) aesthetic

‘The movie made by French is aesthetic.’

In example (18), the author mistakes the sentence as being the same as

\[ fâte呼吸 \quad \text{pāi} \quad \text{de1 (could be de1 or de3) “piànzi” omitted wéimēi.} \]

(French filmed aesthetically.) However, in reality the object *piànzi* (movie) has been moved before the verb *pāi* (film) instead of being omitted.

(19) 2407

\[ 不过 他仅对这节目关心 de1 (should be de3) bijiào shāo. \]

\[ \text{but I about the program concerned about de1 (should be de3) less} \]

‘But I am less concerned about the program.’
Similarly, the object zhè jiémù (the program) is preceded by the preposition dui (about), not omitted. But the position change makes the language user confound with the sentence below:

Búguò běnrén guānzhù de1 (could be de1 or de3) (“zhè jiémù” omitted) bǐjiào shǎo.

but I concerned about de1 (could be de1 or de3) (“the program” omitted) less

'But I am less concerned.'

(20) 2508

Tā Shànghǎi huà shuō de1 (should be de3) gēn běndì rén yíyàng.

he Shanghai dialect speaks de1 (should be de3) like local people

'He speaks Shanghai dialect just like local people.'

In example (20), the object Shànghǎi huà (Shanghai dialect) appears before the verb shuō (speak). If the object is omitted, using de1 is correct. Tā shuō de1 (could be de1 or de3) (“Shànghǎi huà” omitted) gēn běndì rén yíyàng. (His Shanghai dialect is just like local people’s.)

Therefore, when the object appears before the verb instead of normally after the verb, it makes writers confused when determining whether to apply de1 or de3. This is the main reason for the high frequency of de1 replacing de3.
Another explanation is the similarity of $de1$ and $de3$ in the semantic perspective. Under several situations, the semantic centers both stand after $de1$ and $de3$. There is no doubt that the center is the phrase after $de1$ both from syntactic level and semantic level. Considering $de3$, as the paper analyzed in Chapter 2, the phrase standing after $de3$ functions as modifier rather than center in the view of syntactic perspective. But from the semantic facet, the phrase after $de3$ sometimes acts as semantic center. For example,

(21) 1507

Pingliang rén néng $de1$ (should be $de3$) hěn.

‘People in Pingliang are particularly competent.’

Compared with $Pingliáng rén hěn néng$ (People in Pingliang are very competent), the degrees of competent are different. Néng $de1$ (should be $de3$) hěn (particularly competent) emphasizes hěn (quite, very, particularly), meanwhile hěn néng (very competent) focuses on néng (competent). In this sense, the semantic orientations of $de1$ and $de3$ are in the same direction. It increases the linguistic positive relevance between $de1$ and $de3$. Let’s investigate other examples.

(22) 2612

Yǎnyuán jǐnzhāng $de1$ (should be $de3$) dōu páodiào le.

‘Performers nervous $de1$ (should be $de3$) all out of tune’
'Performers are so nervous that all of them are out of tune.'

In example (22), jǐnzhāng (be nervous) is the syntactic center without doubt. But on the semantic layer, since jǐnzhāng (be nervous) and pǎodiào (be out of tune) both can act as predicate, and pǎodiào (be out of tune) is the result of jǐnzhāng (be nervous), pǎodiào (be out of tune) can also be the semantic center of the sentence when the author intends to emphasize the result.

(23) 1808

Tā huó  de1 (should be de3) yào  bǐ  biérén hǎo.

he live  de1 (should be de3) would like to better than  others

‘He would like to live better than others.’

Here it is apparent that the semantic center is after de. The writer wants to express the meaning that bǐ biérén hǎo (better than others) rather than huó (live).

(24) 1810

Zhè  jiànshì xià  de1 (should be de3) wǒ  máogūsōngrán.

the  thing  scared  de1 (should be de3) me,  made my flesh creep

‘The thing scared me and made my flesh creep.’

(24.A)

zhè jiànshì xià wǒ

the thing  scared me
'the thing scared me'

(24.B)

zhè jiànshi rang wò máogǔsōngrán

the thing made my flesh creep

‘the thing made my flesh creep’

Comparing with (24.A), the meaning delivered by (24.B) is more important. So it plays a more central semantic role in the sentence than (24.A).

Therefore, the special usage of de1 is the main reason for the high frequency of de1 replacing de3. The similarity of de1 and de3 on the semantic orientation facet is the complimentary explanation.

4.3 Replacements Among de1, de2 and de3 In Idiolects

As data present, the generalization of de1 is quite obvious. The replacements of de1 replacing de2 and de1 replacing de3 are much more than the other four replacements. Regarding to the four replacements, they account for so few percent that they can be ignored while discussing the relationships among three des.

Therefore, comparing de1 replacing de2 and de1 replacing de3 becomes the point for investigating the linguistic relationships among three des.

For the degree of similarity on the syntactic layer, de1 ought to be closer to de2 rather than de3. The percent of these two replacements also support it, as the replacements of de1 replacing de2 (①) account for 28.62% and the ones of de1 replacing de3 (③) only account for 21.93%. ① is 6.69% more than ③. In observing
the relationships among three *des* in-depth, the focal point will be groups and individuals in the light of Complex Adaptive System (CAS) theory.

### 4.3.1 zoom in on groups

The proper use of three *des* is regulated by the authority in order to standardize the written language. In most cases, native speakers are taught how to distinguish three *des* at school or by parents. Therefore, the frequency of replacements may be related to the literacy level. The twenty research writers are equally divided into two groups—scholars and “grass-roots.” Relatively speaking, the scholars have higher literacy level than the “grass-roots.” Below is the table showing the replacements of these two groups in total.

<table>
<thead>
<tr>
<th></th>
<th>Ten Scholars</th>
<th>Ten Grass-Roots</th>
</tr>
</thead>
<tbody>
<tr>
<td>① de1 replaces de2</td>
<td>10.86%</td>
<td>41.71%</td>
</tr>
<tr>
<td>③ de1 replaces de3</td>
<td>8.05%</td>
<td>34.92%</td>
</tr>
</tbody>
</table>

Table 6: Replacements of two groups

If just taking each group as a whole, it could be concluded that the frequency of replacements among three *des* has a relationship with people’s literacy level for *p* = .016 < .05. That is to say, the higher the literacy level, the less frequency of replacements happens among three *des* and vice versa. Besides, the replacement ① is more than ③ both in scholars’ group (2.81%) and “grass-roots” (6.79%) group. It
seems that the syntactic structural similarity of de1 and de2 functions more on the writers with relatively higher frequency of replacements among three des. Here, the writers are the groups of “grass-roots” with more than 40% replacement ① and more than 30% replacement ③.

4.3.2 Complex Adaptive System applied in linguistics

Complex Adaptive System (CAS) was formally proposed by John Holland in 1999. Complex adaptive systems contain two perspectives—the macroscopic view and the microscopic view. Complex adaptive systems are a “complex macroscopic collection” of relatively “similar and partially connected micro-structures” – formed in order to adapt to the changing environment, and increase its survivability as a macro-structure. They are complex in that they are dynamic networks of interactions, and their relationships are not aggregations of the individual static entities. They are adaptive in that the individual and collective behavior mutate and self-organize corresponding to the change-initiating micro-event or collection of events.

A lot of pioneering scholars, including John Holland, believe that CAS can be used in the area of language research. Wang Shiyuan (Wang 2006) not only holds that language is a complex adaptive system, but also was the first to take Chinese materials as examples. Shen Zhongwei (Z. Shen 2014) states that language is a system composed of idiolects. The idiolect is equal to the microscopic “agent” in the CAS. Actually language is the synthesized summary of idiolects rather than an entity
existing independently. Language always lies in idiolects. Therefore, it is helpful to know language in-depth through studying idiolects carefully.

Inspired by the CAS theory, this study not only focuses on the macroscopic groups further but also zooms in on the microscopic individuals. In other words, to research the mutual replacements among three des more deeply and objectively, the next step is investigating these twenty writers one-by-one and analyzing the individual characteristics presented in this problem.

4.3.3 zoom in on individuals

Below is the table to show the performance of the twenty bloggers in the replacements among three des. It also includes the percent of replacement ① more than replacement ③.

<table>
<thead>
<tr>
<th>Serial Number</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>de1 replaces de2</td>
<td>3%</td>
<td>6%</td>
<td>0</td>
<td>8%</td>
<td>0</td>
<td>71%</td>
<td>7%</td>
<td>3%</td>
<td>6%</td>
<td>4%</td>
</tr>
<tr>
<td>de1 replaces de3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25%</td>
<td>6%</td>
<td>20%</td>
<td>21%</td>
<td>0</td>
</tr>
<tr>
<td>(①) &gt; (③)</td>
<td>3%</td>
<td>6%</td>
<td>0</td>
<td>8%</td>
<td>0</td>
<td>46%</td>
<td>1%</td>
<td>-17%</td>
<td>-15%</td>
<td>4%</td>
</tr>
<tr>
<td>Serial Number</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>de1 replaces de2</td>
<td>75%</td>
<td>7%</td>
<td>6%</td>
<td>27%</td>
<td>80%</td>
<td>33%</td>
<td>100%</td>
<td>5%</td>
<td>12%</td>
<td>0</td>
</tr>
<tr>
<td>de1 replaces de3</td>
<td>50%</td>
<td>0</td>
<td>39%</td>
<td>0</td>
<td>85%</td>
<td>34%</td>
<td>97%</td>
<td>0</td>
<td>13%</td>
<td>0</td>
</tr>
<tr>
<td>(①) &gt; (③)</td>
<td>25%</td>
<td>7%</td>
<td>-33%</td>
<td>27%</td>
<td>-5%</td>
<td>-1%</td>
<td>3%</td>
<td>5%</td>
<td>-1%</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7: The percentage of replacements of twenty writers
In reality, the situations of replacements among three *des* are performed quite differently in the twenty writers. They can be roughly classified into four groups based on the two frequencies of the replacement $\text{	extcircled{1}}$ and $\text{	extcircled{3}}$. After investigating the percentages in Table 7, 10% can be regarded as the relatively high frequency, the grouping situation presents as below.

<table>
<thead>
<tr>
<th>Group No. &amp; Description</th>
<th>Serial Number of The Research Writers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\text{	extcircled{1}}$&amp;$\text{	extcircled{3}}$ &lt; 10%</td>
<td>10; 11; 12; 13; 14; 16; 19; 21; 27; 29</td>
</tr>
<tr>
<td>2. $\text{	extcircled{1}}$ &gt; 10% but $\text{	extcircled{3}}$ &lt; 10%</td>
<td>23</td>
</tr>
<tr>
<td>3. $\text{	extcircled{1}}$&amp;$\text{	extcircled{3}}$ &gt; 10%</td>
<td>15; 20; 24; 25; 26; 28</td>
</tr>
<tr>
<td>4. $\text{	extcircled{1}}$ &lt; 10% but $\text{	extcircled{3}}$ &gt; 10%</td>
<td>17; 18; 22</td>
</tr>
</tbody>
</table>

*Table 8: Grouping situation when taking 10% as boundary*

It can be concluded that the relatively high frequency replacements, where 10% is the boundary, happen in about half of the population’s writing in the Chinese language. Among them, one writer, whose serial number is 26, is worth being paid more attention. The writer never uses *de2* and rarely uses *de3*. This writer in most situations just uses *de1* to replace *de2* and *de3*. The generalization of *de1* takes on an extreme tendency in the writer’s blogs. But there is only one writer out of twenty having this situation, accounting for only 5%.

On the other hand, there are three writers (12, 14, 29), who never misuse three *des* in their articles. This indicates that approximately 15% of the population is able to clearly distinguish three *des* in their mind and correctly use three *des* in their
writing language. The two opposite sub-groups above precisely verify the diversity and complexity of replacements among three *des* happening in individuals. What’s more, the phenomena prove that language is a complex adaptive system as well.

To be honest, 10% replacements cannot be seen as a significant percentage, but 30% replacements definitely can represent a high frequency of replacements among three *des*. If taking 30% as the boundary, the table below shows the detailed groups.

<table>
<thead>
<tr>
<th>Group No. &amp; Description</th>
<th>Serial Number of The Research Writers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ① &amp; ③ &lt; 30%</td>
<td>10; 11; 12; 13; 14; 16; 17; 18; 19; 21; 23; 27; 28; 29</td>
</tr>
<tr>
<td>2. ① &gt; 30% but ③ &lt; 30%</td>
<td>15</td>
</tr>
<tr>
<td>3. ① &amp; ③ &gt; 30%</td>
<td>20; 24; 25; 26</td>
</tr>
<tr>
<td>4. ① &lt; 30% but ③ &gt; 30%</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 9: Grouping situation when taking 30% as boundary

Since 30% is a relatively high frequency of replacements, most of the writers belong to Group 1. This is especially true of the writers from the scholars’ group, in which nine out of ten belong to it. As for the whole writers, there are fourteen people out of twenty belonging to Group 1, which accounts for 70%. In other words, the high frequency of replacements, 30% is the boundary here, among three *des* happen in approximately 30% of the population. Replacement ① and ③ happen frequently at the same time to most of them, about 67% (four writers out of six). The high frequency of replacement ① can happen solely to bloggers, like No. 15. It is
the case with replacement ③, such as No. 22. Again, the replacements among three des present a diverse situation in individual writers.

As for the relationships among de1, de2 and de3, the Figure 8 displays the distribution of replacements among three des in the twenty writers of this study. The fine line in the figure means the frequencies of replacement ① and ③ are equal. That is to say, the linguistic distance between de1 and de2 is equal to the one between de1 and de3. The thick line is the actual trend line drawn on the basis of the authentic data. Compared with the fine line, the trend line reveals that to the same writer replacement ① happens more frequently than replacement ③. More specifically, de1 prefers to replace de2 rather than to replace de3. The tendency happens on the writers with relatively frequent replacements, more than 10% here. The results are precisely consistent with the syntactic characters of structures containing three des. Therefore, it can be concluded that the syntactic similarity is the main reason for the replacements occurring among three des. Moreover, the related degree of de1 and de2 is significantly higher than the one of de1 and de3. It is especially obvious on the writers with relatively frequent replacements.

The conclusion proves that as structural particles, when using three des, their syntactic functions play the key and main role in writers’ minds. So the similarity in the syntactic function, like de1 and de2 both acting as modification mark, facilitates the dominant part to replace the weak part during the language variation. Specially speaking, the replacement of de1 replacing de2 plays the leading role rather than the other way. Additionally, it is the syntax that acts as the main factor. Hence, the replacement of de1 replacing de2 happens more frequently than the replacement of
$de_1$ replacing $de_3$. That is to say, the linguistic relevance between $de_1$ and $de_2$ is closer than the one between $de_1$ and $de_3$.

Figure 8: Distribution of replacements among three $des$ in twenty writers
CHAPTER 5

SUMMARY

Language is a Complex Adaptive System and exists in idiolects in reality. It is impossible to collect all the idiolects as research samples, but distinguishing the concepts of language and idiolect are crucial for linguistic studies. What’s more, emphasizing the importance of studying idiolects during the process of investigating concrete linguistic problem brings incredible values. Applying the theory of Complex Adaptive System to linguistic studies inspires us to notice the diversity and complexity of language. Especially, it offers us a way to understand our language more scientifically and objectively.

This study focuses on personal blogs is because they supply an excellent opportunity to directly and reliably observe the written forms of idiolects. Like Rettberg (Rettberg 2008, 31) clarifies that we have moved from a culture dominated by mass media, using one-to-many communication, to one where participatory media, using many-to-many communication, is becoming the norm. That is to say by means of blogs, everyone has the right to make a voice and can publish their essays without editors. So the linguistic phenomena presented in the content of personal blogs are much closer to idiolects.

More interesting, the existence of three homophonous des in written Chinese provides the superexcellent chance to study the relationships among three des. The problem has its unique value because it cannot be observed by investigating the oral language. Therefore, the corpus collected from personal blogs is exactly suitable and convincing in studying this particular linguistic problem.
Through analyzing the authentic data, which is obtained from a corpus built of articles in personal blogs including 400,000 Chinese characters, there are several worthy findings. First, the mutual replacements are asymmetric along with the generalization of $de_1$. Secondly, there is a positive correlation between the frequency of replacements and the linguistic positive relevance among three $des$, especially the syntactic and semantic aspects. Finally, the replacements among three $des$ present a diverse and complicated situation when investigating the written forms of idiolects. The syntactic factor plays the main role in the replacements among three $des$. The related degree of $de_1 \& de_2$ is significantly higher than the one of $de_1 \& de_3$, especially obvious on the objects with relatively frequent replacements.
BIBLIOGRAPHY


Shen, Zhongwei. "Complex adaptive system and dynamic studies of Chinese."


