華語文教學研究 14.1:1-20, 2017

Semantic Variation on Chinese Classifiers for Vehicles: A Case Study on *Liang*（輛）, *Bu*（部）, and *Tai*（台）*

Chenhsuan HUANG
National Chung Cheng University

Abstract

The present study explored the Chinese classifiers *liang*（輛）, *bu*（部）, and *tai*（台）for vehicles by using the Qualia structure (Pustejovsky 1995) as the major criterion. Four independent variables, including Constitutive, Formal, Telic, and Agentivity, are tested in the *Rbrul* program to account for the multiple cross-cutting and interesting factors that influence language usage. The major findings of the study are as follows. First, *liang* serves as the standard usage for vehicles, which favors vehicles with at least four wheels (Constitutive) and non-personal use (Telic). Second, *bu* is salient for its personal use for ownership (Telic) and tendency toward an external appearance description (Formal). Third, *tai* tends to emphasize a vehicle with its appearance (Formal), origin (Agentivity), and wheels that are less than four (Constitutive). The study demonstrated that the use of vehicle classifiers *bu*, *liang*, and *tai* is not arbitrary, but instead follows cognitive-based categories.

Keywords: classifier, Qualia structure, *Rbrul*, Chinese pedagogy

* The author would like to thank the two anonymous reviewers for their helpful and constructive comments that greatly contributed to improving the final version of the paper. This article is part of my doctoral dissertation under the guidance of D. Victoria Rau to the Institute of Linguistics, National Chung Cheng University. The research was supported by Ministry of Science and Technology (R.O.C.) grants: ST-105-2420-H-194-008-DR. All remaining errors are my own.

1. Introduction

Classifiers are one of the most striking features of Mandarin Chinese. They are also morphemes that characterize nouns. Generally speaking, classifier like *tiao*（條）is for long, strip objects like ropes or roads, and *zhang*（張）is for flat objects such as sheets of paper or blankets. For learners of Chinese, knowing the semantic property of classifiers can make their learning more effective. Unfortunately, not every classifier shows salient properties as shape classifiers do. In addition, the association between classifier and noun is not fixed at all, as shown in the examples below from United Daily News（聯合報）[1].

(1) shiqi liang che (2016/06/12/a08)
seventeen CL car
'seventeen cars'

(2) zhe bu che (2016/03/09/d04)
this CL car
'this car'

(3) na tai che (2016/04/14/c07)
which CL car
'which car ?'

In (1) - (3), the noun *che* ('car') can be denoted by three different classifiers *liang*, *bu*, and *tai*. However, in current TCFL textbooks,[2] it is taught that *liang* is the only classifier for vehicles, without mention of the other two. Apparently, there is a usage gap between the textbook and actual use in society. Except for *liang* serving as a vehicle classifier, both *bu* and *tai* can be used for electronic equipment or machines. To some extent, a car in itself can be seen as a kind of machine. Thus, we wonder under what situation would a vehicle be viewed as a machine? Furthermore, what machine or transportation properties would these three classifiers tend to hold for? In order to learn more about this classifier variation, the present study therefore uses a

[1] http://udn.com/news/index

[2] TCFL refers to Teaching Chinese as a Foreign Language. The textbooks we consulted include *Xinban Shiyong Shiting Huayu*（新版實用視聽華語）'Practical Audio-Visual Chinese' and *Yuandong Shenghuo Huayu*（遠東生活華語）'Far East Everyday Chinese'.

logistic regression analysis to investigate the variation of Chinese classifiers *liang*（輛）, *bu*（部）, and *tai*（台）in relation to the issue of human categorization.

The organization of the paper is as follows. Section 2 provides a brief introduction of previous research on Chinese classifiers. Section 3 discusses a preliminary observation on the semantic items of the three classifiers. Section 4 states the methods of the study. Section 5 displays the statistical results and further discussions. Finally, section 6 concludes with some theoretical and methodological implications and suggestions for future studies of Chinese classifier variation.

2. Literature Review

Categorization is one of the basic cognitive activities conducted by human beings in everyday life, and the use of classifiers is one of the devices that show the categorization. Tai and Wang (1990) adopted the prototype theory for Chinese classifiers from their development history and claimed that the use of *tiao* is based on an imputed salient perceptual property of 'extension in length'. Tai and Chao (1994) looked at the classifier *zhang*（張）and further found that classifiers *zhang*（張）, *fu*（幅）, *mian*（面）, and *pian*（片）are interchangeable in certain cases. This usage variation is owing to the different classifier systems across Chinese dialects and even register (spoken and written Chinese) and social variations. Tai (1994) then proposed the first cognition-based classifier system across Chinese dialects, displaying four relevant cognitive categories that underline Chinese classifier systems: animacy (human, animal, or inanimate), shape (long, flat, or round), consistency (hard, flexible, or non-discrete), and size (big or small). For instance, the classifier *tiao*（條）is often used for nouns with one-dimensional extension in length, while *gen*（根）is for three-dimensional long, rigid objects. He indicated that the classifier systems across Chinese dialects contain great differences, and these differences can be understood by the four cognitive-based categories. However, the proposed four categories are limited to shaped-based classifiers. For vehicle classifiers like *liang*, *bu*, and *tai*, the object appearance may not be the prior choice, and thus some other semantic facets would be neglected under the system.

Ji, Zhang, and Nisbett (2004) argued that speakers of different dialects/languages and divergent non-linguistic experiences may influence both language and concepts.

Kuo, Lee, and Tai (2008) studied the categorization patterns of classifiers in Taiwan Southern Min via an experiment. They asked the respondents to describe photos, including object naming, counting items, and describing the situation or event in the photos. This design was meant to find the central member of each classifier as *bue*（尾）'tail', *nia*（領）'collar', *tiao*（條）, *ki*（枝）, *tiu*（張）, *liap*（粒）, and so on. The results showed that meronomy (part-whole relation), shape, arrangement, and composition are all important cognition bases in Taiwan Southern Min. Kuo *et al.*'s study pointed out the variation among the participants and that the variation reflects that human beings are endowed with the ability to conceptualize the same physical world differently. Similarly, Zhang (2007) agreed that shape categories constitute a major cognitive base for classification, lexical taxonomy, and linguistic convention; however, stylistic creativities, like formal vs. informal, written vs. colloquial, educated vs. uneducated, positive vs. negative, and common use vs. local dialects, also play a role in determining the use of Chinese classifiers. Kuo and Sera (2009) further found that Mandarin speakers rely more heavily on shape than English speakers do when classifying objects. Moreover, reliance on shape by speakers of Chinese was correlated with the amount of exposure to Chinese.

Tai (1992; 1994) mentioned that different classifiers may take the same object. In terms of this issue, Rovira-Esteva (2007) conducted a cognitive-based study of Chinese classifiers *shuang*（双）, *dui*（對）, *fu*（副）, and *tao*（套）, using the prototype theory to show the distinction with respect to the noun classes they select and to identify the central, natural extension and metaphorical extension members. Huang and Ahren (2003), on the other hand, claimed that Chinese classifiers do not simply agree with the noun, but instead coerce a particular meaning from it, as seen in the examples below.

(4) bangongshi li you san ju dianhua
office in has three CL telephone
'There are three telephone sets in the office'

(5) bangongshi li you san xian dianhua
office in has three CL telephone
'There are three telephone lines in the office'

In (4) and (5), both *ju*（具）and *xian*（線）are individual classifiers, but they emphasize different meaning facets as *ju* stresses its machinery function, while *xian*

tends to highlight the more abstract meaning of lines of communication. To untie this knot, Huang and Ahren adopted the concept of the Qualia structure from Pustejovsky (1995) to interpret these semantic constraints or facets. According to Pustejovsky (1995:85-86), these constraints include Constitutive, Formal, Telic, and Agentive. The definitions are as follows. Constitutive constraints involve the relationship between an object and its parts. Under this category, the subgroups include material, weight, and components. Formal constraints distinguish the object within a large domain involving orientation, magnitude, shape, dimensionality, color, and position. Telic constraints refer to the purpose and function of the object. Lastly, Agentive constraints describe the origin of an object. Based on the Qualia structure, *dianhua* ('telephone') with the classifier *ju* in (4) has the Formal role of telephone as an object, while *xian* in (5) represents the Telic role of telephone as a tool to connect to the telephone line. Hence, the different meaning facets of this noun can be derived.

Variation in language use is seen to be structured, not random, and this structure can be discovered by studying the correlation of linguistic variables and social factors. The aforementioned studies demonstrated the association between human cognition and the use of classifiers, but little attention has been paid to the overlapping classifiers and quantitative linguistics. *Rbrul* (Johnson 2009) and its earlier version called VARBRUL (Sankoff, Tagliamonte, and Smith 2005) are the most used statistical tools for analyzing sociolinguistic variation. Sociolinguists like to use them to do the variable analysis. As Robinson (2010) pointed out that there is a need to merge cognitive and sociolinguistic approaches in the future language research. Therefore, in order to bridge this research gap, the current study serves as a pilot research for the vehicle classifiers *liang* (輛), *bu* (部), and *tai* (台) in order to see how these three classifiers interact with vehicle types in Taiwan Mandarin Chinese. The research questions of the study are addressed as below.

(1) How many semantic categories do the vehicle classifiers *liang*, *bu*, and *tai* contain?
(2) What semantic facet(s) can account for the usage variation among *liang*, *bu*, and *tai*?

3. Preliminary Observation

Before we set up the research, let us look at the semantic items of *liang*（輛）,

bu（部）, and *tai*（台）from historical and modern uses to see if we can find some clues to differentiate them. Chen (2014) compiled a historical dictionary of Chinese classifiers. According to Chen (2014:197), *liang*（輛）is another form of *liang*（兩）'two', which semantically meant two wheels in ancient Chinese. It is often used for vehicles; however, it can be referred to spinning wheels since it possesses two wheels.[3] In modern use, *liang*（輛）still serves as a classifier for vehicles, however, the number of wheels has been expanded into four or more. As a result, the number of wheels can be one of the characters for the choice of *liang*（輛）.

Tai（台/臺）is meant as the base of an object. A rectangle base is called *tai*,[4] whereas other shapes are called *guan*（觀）(Gao 1989:288). Moreover, it also refers to high and flat land, like a stage in *yitaixi*（一台戲）'a drama' (Chen 2014:35). As we can see, the usage of *tai* in ancient Chinese mostly depended on the shape of the object. In modern uses, this rectangle property becomes salient in the usages for electronics devices (e.g. TVs), instruments (e.g. pianos), and so on. Although we cannot find any direct evidence that *tai* is used for vehicles in both historical and modern dictionaries, the use of *tai* is inclined to be shape-oriented based on the above references. Thus, we can infer that its usage on vehicles likely come from the appearance.

When *bu*（部）is used as a classifier, it can be used for objects such as: a) books; b) big and strong man; c) parts or units; and d) vehicles. For vehicle uses, *bu* represents the protection cover（蓋斗）, and the protection cover is actually a part of the carriage; thus it can denote vehicles in expression as *yi bu mache*（一部馬車）'one carriage' (Chen 2014:34-36). In contemporary uses, *bu* serves as a classifier for: a) movies and books; b) electronics devices (e.g. cameras, TVs); c) instruments (e.g. bulldozers); d) parts or units; and e) vehicles based on the Ministry of Education (hereafter, MOE) online dictionary (2015). It is worth noting that the use for vehicles is not included in

[3] 里婦不能紡者，授紡車八百餘輛。<明史・陳幼學>
li fu bu neng fangzhe, shou fangche ba bai yu liang.
'For village women who cannot weave, they would be granted over eight hundred spin wheels'.
<*History of the Ming Dynasty*. Chen, Youxue>

[4] 三台印信都掌權，誰敢居功吾上。<紫釵記：延媒勸贅>
san tai yin xin dou zhang quan,shui gan ju gong wu shang.
'I am in charge of the three official seals, who dare to take credit from me'.
<*Zichaiji: Yanmeiquanzhui*>

previous editions of MOE dictionaries. However, comparing the historical and modern uses, we can see that part of the items have been changed and even disappeared. In addition, the use of *bu* in ancient history emphasizes its function or purpose for preventing wind and rain. Later, this property for function and purpose has seemed to transfer over to electronic devices and machines. Thus, we can infer that the use of *bu* is to stress the function or purpose of the vehicle.

After the preliminary observation on historical and modern uses, we learned that the number of wheels, appearance, and function/purpose could be the possible categorization to differentiate *liang*（輛）, *bu*（部）, and *tai*（台）. However, the next section shall display the detailed categorization and method.

4. Method

This study is based on a usage-based perspective, which views that the usage patterns, frequency of occurrence, variation, and change are all taken to provide direct evidence about cognitive representation (Bybee and Beckner 2010). Thus, in order to find out the cognition model of classifiers *liang*, *bu*, and *tai*, the data first need to be classified by their semantic meanings, and a quantitative analysis can then be applied to validate the data.

4.1 Data Collection

The data used herein are based on Sinica corpus (Academia Sinica Balanced Corpus of Mandarin Chinese) by the Chinese Word Sketch Engine.[5] The greatest advantage in using this corpus is that the data are already tagged by their linguistic features. The notable feature of the data is that they mainly come from news and magazines, which comprise about 68% of all the data; however, there is a little percentage from spoken sources. As for the styles and genre of the data, 90% are regarded as formal in the written form (Chinese Knowledge and Information Processing 1998).

There are 2995 tokens in total for *liang*, *bu*, and *tai* serving as classifiers (*liang*: 590; *bu*: 1897; *tai*: 508). However, we only code those nouns that contain a reference

[5] http://wordsketch.ling.sinica.edu.tw/

to motor vehicles. The excluded tokens include the usages as classifiers for movies, artwork, electronic devices, instruments, TV channels, organization units, and so on. Note that the coded tokens should go with either *liang* or *bu* or *tai*, so that we can see the variation among them. Token as *feiji*（飛機）'plane' is also a kind of vehicle, however, it is excluded since it possesses its own specific classifier as *jia*（架）, and thus it may skew our data distribution in certain classifier combinations.

4.2 Data Distribution

Table 1 shows the token numbers in detail. Out of 2995 tokens, about 28.54% (855 tokens) are coded as valid data. Of these, 100% of the instances of *liang* are coded since the use of *liang* is the most limited, as it can only be used for cars. Additionally, *bu* is 12.23% and *tai* is 6.49%. Thus, in the use of classifying vehicles, *liang* is the most typical one among the three. In the total token ratio, *bu* is 27.13%, *liang* is 69%, and *tai* is 3.86%. The ratio of *tai* takes only 3.86%, and such ratio would underestimate its influence in the analysis. The higher the number is, the greater the contribution of that factor to the probability of the form (Tagliamonte 2011). Thus, this unbalanced ratio may cause a low input probability[6] in the analysis since the "input" indicates the overall tendency of the dependent variable to surface in the data.

Table 1: Overview of Tokens

CLs	Valid tokens	Not coded	Total
bu	232	1665	1897
liang	590	0	590
tai	33	475	508
Total	855	2140	2995

4.3 Data Analysis

The use of classifiers may vary depending on the semantic facets. In section 3 we mentioned that the use of *liang*, *bu*, and *tai* could be categorized by the number of wheels, appearance, and function or purpose. To make our analysis more complete, we

[6] The input probability ranges from 0 to 1.

follow the constraints of the Qualia structure, and each constraint is further classified into several subgroups according to the observed data. The detailed semantic classifications are as follows.

(a) Constitutive: This constraint refers to the "internal" aspect of the object, which describes the relation between the object and its parts. In our case, we divide the vehicles based on the number of wheels, since a wheel is one of the main characteristics in old Chinese based on our preliminary observation earlier in section 3. In addition, the number of wheels is often used for differentiating vehicles in modern uses, such as *sanlun jiche*（三輪機車）'a three-wheeled motorcycle' and *sulun diandongche*（四輪電動車）'four-wheeled electric vehicle'. Noted that, for vehicles like *jiche*（機車）'scooter' and *qiche*（汽車）'car', though they do not contain "wheel" from their morphology, by default, their number of wheels refer to two and four respectively. The word *lun*（輪）'wheels' is only used for non-default cases like *sanlun jiche*（三輪機車）'a three-wheeled motorcycle', whose wheel is three-wheels rather than the default two-wheels. Thus, the data in this category are coded as either [+] for 4 wheels (or above) or [-] for less than 4 wheels by their default or non-default number of wheels. In addition, it is worth noting that botha power types (motor/manual) and steering wheel can also be the possible candidates; however, they did not show any statistical significance in our initial analysis.

(b) Formal: In contrast to constitutive, the constraint Formal encodes taxonomic information about the lexical item. Pustejovsky (1995) provided a list of Formal attributes for the class of nouns denoting concrete entities. These include physical characteristics such as the following: spatial characteristics, intrinsic orientation, size and dimensional properties, shape and form, color and position. Within this constraint, we group the data into two as [+] specific and [-] specific. For nouns with a specific description on the appearance of vehicles, examples include *xinche*（新車）"new car", *jiaoche*（轎車）'sedan', which literally means palanquin-like car, and *xiangxingche*（箱型車）'van', which literally means box-shaped car.

(c) Telic: The Telic constraint refers to the nouns that specify certain activities or functions. According to the corpus data, vehicle purposes can be grouped into four: individual/family uses, commercial uses, social uses, and non-specific.

For individual/family uses, examples contain *jiatingche*（家庭車）‘family cars’, *ziyong xiaokeche*（自用小客車）‘private cars’, *wo de che*（我的車）‘my car’, etc. *jichengche*（計程車）‘taxi’, *huoche*（貨車）‘truck’, *kuaicanche*（快餐車）‘bustaurants’, and *youlanche*（遊覽車）‘tour bus’ are examples that stress their commercial uses. As to social purposes, examples include *jingche*（警車）‘police car’, *xunluoche*（巡邏車）‘patrol car’, *jiuhuche*（救護車）‘ambulance’, *zhechache*（偵查車）‘scout car’, and so on.

(d) Agentive. The constraint Agentive denotes the information about the origin the object, and it is a crucial role for differentiating the kinds of objects, properties and relations that exist in the world, since it provides a mechanism for distinguishing natural (default) and non-natural (non-default) entities (Pustejovsky and Jezek 2016). For this, in our cases for vehicles, the Agentive constraint can be referred to sources of the vehicle. Examples as *zangche*（贓車）‘stolen car’, *jinkouche*（進口車）‘imported car’, and *ribenche*（日本車）‘Japanese car’ would be labeled with Agentive value because their semantics represents the origin of the vehicle.

After coding the tokens, we adopt the *Rbrul* program for logistic regression analysis for the statistical testing of the data. The main reason that chooses *Rbrul* over VARBRUL is that *Rbrul* can handle more than two dependent variables at a time. To run the program, we need to divide the variables into two: independent variables and dependent variables. The independent variables are possible factors that determine the variation, including Constitutive, Formal, Telic, and Agentive. The depend variables are classifiers: *liang*, *bu*, and *tai*.

5. Results and Discussion

We first present the results of the distribution of three classifiers by cross comparison, based on the quantitative analysis of the corpus. Next, we offer the best combinations of factors that can account for the variability of the three classifiers.

Table 2 shows the token numbers in detail. Due to the unbalanced token number, the use of *liang* seems to have the most dominate tendency among all the constraints. Still, these are the results of the raw counts. A multivariate analysis allows us to take each factor into account in order to determine which factors can best explain the variation.

Table 2: Distribution of *Liang*, *Bu*, and *Tai* Across Factors in Overall Data (Raw Numbers)

Group Factors		*liang*	*bu*	*tai*	**Total**
Constitutive					
[+]4 wheels		534 (71.01%)	197 (26.20%)	21 (2.79%)	752 (100%)
[-]4 wheels		56 (54.37%)	35 (33.98%)	12 (11.65%)	103 (100%)
Total		590 (69.01 %)	232 (27.13%)	33 (3.86%)	855 (100%)
Formal					
[+] specific		135 (59.73%)	76 (33.63%)	15 (6.64%)	226 (100%)
[-] specific		455 (72.34%)	156 (24.80%)	18 (2.87%)	629 (100%)
Total		590	232	33	855
Telic					
[+]specific	Individual/ family uses	9 (32.15%)	17 (60.71%)	2 (7.14%)	28 (100%)
	Commercial uses	88 (69.84%)	33 (26.19%)	5 (3.97%)	126 (100%)
	Social uses	59 (68.60%)	25 (29.07%)	2 (2.33%)	86 (100%)
[-] specific		434 (70.57%)	157 (25.53%)	24 (3.90%)	615 (100%)
Total		590	232	33	855
Agentive					
[+]specific		77 (60.63%)	40 (31.50%)	10 (7.88%)	127 (100%)
[-]specific		513 (70.47%)	192 (26.37%)	23 (3.16%)	728 (100%)
Total		590	232	33	855

5.1 Interpretation of *Liang*

Table 3 shows the *Rbrul* results for *bu* and *tai*. Since *Rbrul* can only deal with binominal analysis, the results on *bu* and *tai* actually reveal the opposite result for the classifier *liang*. For instance, if the results of *bu* and *tai* have tendency in X, then it indicates that *liang* would dominate in the opposite factors of X. Conventionally, weight over 0.5 indicates the associated factor favors the use of certain patterns. Log odds are raw co-efficients for the regression model. The larger the number is, the bigger the effect size would be. Here, the factors that govern *bu* and *tai* from *liang* are constraints Constitutive, Formal, and Telic.

First, the combination of *bu* and *tai* favor vehicles with less than four wheels since the centered weight is over 0.5. Their common collocations include *jiche*（機車）‘scooters’, *jiaotache*（腳踏車）‘bike’, *sanlunche*（三輪車）‘tricycle’, and so on. For example: (extracted from Sinica corpus)

(6) Liang ming daitu gong cheng yi bu ji che
two CL criminals share ride one CL scooter
‘The two criminals rode a scooter.’

(7) Youde houxuanren, yi juan jiushi shi tai jiaotache
some candidates, once donate just ten CL bike
‘Some candidates donate ten bikes at once.’

On the contrary, the result also indicates that *liang* prefers vehicles with at least four wheels, such as *gongche*（公車）‘bus’, *qiche*（汽車） ‘car’, or so. Examples are as followed:

(8) San qian qi bai liang gongche neng anzhuang guangbo xitong
three thousand seven hundred CL bus can install broadcast system
‘The broadcasting system would be installed for the three thousand and seven hundred buses. ’

(9) Zheli shi keyi tingfang er bai liang qiche de tingchechang.
here is can park two hundred CL car NOM parking lot
‘This parking lot can park two hundred cars.’

As mentioned earlier in section 3, the meaning of *liang* is meant to be two wheels in ancient uses. *Liang* nowadays refers to four wheels or above as in (8) and (9). This reflects the changes of cognition. Second, descriptions on machines or equipments like to stress the appearances or functions. The results on the Formal constraint indicate that the combination of *bu* and *tai* tends to describe the external appearance of the vehicles as in (10) which stressed the type (2000 series sedan), and color (black) of the car and (11) which pictured the color (red) and size (50 c.c.) of the scooter while *liang* does not have this preference, just as (8) and (9) in which readers cannot receive further information about the vehicle. In other words, the use of *bu* and *tai* tends to make the vehicle in focus, and thus more descriptions on its appearance would be given.

(10) Yi bu yisi Yulong er qian xing heisei jiao che
one CL suspected Yulong two thousand type black sedan
'One suspected Yulong 2000 series black Sedan'

(11) Yi tai hongse Sanyang DIO wushi
one CL red Sangyang DIO fifty
'one red Sangyang DIO 50 c.c. scooter'

Third, in the Telic constraint, the results show that the combination of *bu* and *tai* favor vehicles for individual or family uses as in (12) and (13). On the other hand, *liang* prefers to be used for other uses, such as commercial or other social purposes as (14).

(12) Zhuang shang zhensuo qian de yi bu zixiaokeche
Bump on clinic front NOM one CL private car.
'Bumped a private car in front of the clinic'

(13) Quan jia geng keyi zu tai luyingche
family more can rent CL camping car.
'The family can even rent a recreational vehicle.'

(14) Na ming jingyuan da le yi liang xiangxing zhenfang che
that CL police ride ASP one CL van type investigation car
'That police officer rode in a police van.'

Although the uses of the three classifiers are categorized into different purposes, it seems that the use of *bu* and *tai* is regarded as being closer to the speaker since their

usages are for individual or family uses, while the use of *liang* is thus viewed as more subjective with less personal emotions. In context, *liang* is often used for reports, such as government vehicle reports or car sales reports while *bu* and *tai* are used more in descriptive articles. Therefore, we can conclude that the use of *liang* tends to be formal or with less personal opinions.

Table 3: *Rbrul* Report for *Bu* and *Tai*

Factor groups	Factor	Log odds	Tokens	Centered weight
Constitutive *p*=0.000403	[+]4 wheels	-0.398	752	0.402
	[-]4 wheels	0.398	103	**0.598**
Formal *p*= 0.000842	**[+] specific**	0.282	228	**0.57**
	[-] specific	-0.282	627	0.43
Telic *p*= 0.0161	**Individual/ family uses**	0.730	32	**0.675**
	Commercial uses	-0.229	126	0.433
	Social uses	-0.096	86	0.476
	Nonspecific	-0.405	615	0.4
Deviance=1024.882; df=6; Grant mean=0.31; Centered input probability =0.478. All factors are significant by Bonferroni correction.				

5.2 Interpretation of *Tai*

Table 4 displays the *Rbrul* results for the combination of *bu* and *liang*. The factors that distinguish *bu* and *liang* from *tai* are constraints Constitutive, Formal, and Agentive. First, in terms of the constraint Constitutive, *bu*, and *liang* favor vehicles with more than four wheels (centered weight = 0.6); on the contrary, it indicates that *tai* tends to co-occur with vehicles with fewer wheels, such as *jiche*（機車）'scooter'. It is interesting to note that the number of wheels affects the use of the verb for vehicles. For example, for wheels less than four, the verb *qi*（騎）would be used while *kai*（開）is for four wheels or more. Thus, what *tai* emphasizes are those light vehicles. Second,

tai prefers to collocate vehicles with appearance description, and this tendency echoes our observation from the historical use in modifying an object shape close to a rectangle. When *tai* is used as a vehicle classifier, it not only reserves the semantic facet in shaping objects, but its proto-shape has also been shifted from rectangle to other properties, such as color and size. Finally, regarding the constraint Agentive, *tai* has the tendency for showing the origin of the noun referents. When mentioning about the origin of the vehicle, the related concepts then emerge, such as price, country, brand, and so on. Example in terms of *tai* in Agentive constraint can be seen from (11) in which the brand name *Sanyang*（三陽）is a famous scooter company in Taiwan, and its name has become synonymous with scooters.

Table 4: *Rbrul* Report for *Bu* and *Liang*

Factor groups	Factor	Log odds	Tokens	Centered weight
Constitutive *P* = 0.000109	[+]4 wheels	0.814	752	**0.693**
	[-]4 wheels	-0.814	103	0.307
Formal *P* = 0.000389	[+] specific	-0.398	228	0.402
	[-] specific	0.398	627	**0.598**
Agentive *P* = 0.0436	[+]specific	-0.424	142	0.396
	[-]specific	0.424	713	**0.604**
Deviance=255.156; df=4; Grant mean=0.961; Centered input probability =0.918. All factors are significant by Bonferroni correction.				

5.3 Interpretation of *Bu*

In Table 5, the main factors to differentiate *bu* from the other two are by the constraints Telic and Formal. Similar to the result in section 5.1 for the Telic constraint, the results below show that *bu* prefers individual or family uses. It is worth noting that for tokens coded as individual or family uses, their semantic meanings often emphasize the ownership of vehicles. The semantic of ownership is an extension of the concept of "part of the whole". In an earlier section, *bu* can refer to parts or units. For example, *bu*

'part' in the term *toubu*（頭部）'head' indicates that *tou* 'head' is a part of the body. When the concept is applied to vehicle use, it stresses on ownership or belonging to certain people or a unit. In addition, *bu* also prefers to co-occur with vehicles stressing their external appearances.

Table 5: *Rbrul* Report for *Liang* and *Tai*

Factor groups	Factor	Log odds	Tokens	Centered weight
Telic *p*= 0.0231	Individual/family uses	-0.760	32	0.319
	Commercial uses	0.310	126	**0.577**
	Social uses	0.098	86	**0.525**
	Nonspecific	0.352	611	**0.587**
Formal *p*= 0.0263	[+] specific	-0.193	228	0.452
	[-] specific	0.193	627	**0.548**
Deviance=984.246; df=5; Grant mean=0.729; Centered input probability =0.653. All factors are significant by Bonferroni correction.				

In summary, four factors are tested for their impacts on the use of *liang*, *bu*, and *tai*. As the results showed, the number of wheels (Constitutive), external appearance (Formal), vehicle purpose (Telic), and origin (Agentive) can account for the usage variation. For *liang*, it is the most typical vehicle classifier. The semantic property that *liang* possesses is a vehicle exhibiting at least four wheels and non-personal uses. In other words, the use of *liang* mainly portraits the dynamic facet of the vehicles. *Bu* and *tai*, on the other hand, view vehicles more as machines by describing their appearances, functions, and origins. To differentiate *tai* from *bu*, *tai* highlights those vehicles with less than four wheels and the sources, while *bu* tends to focus on the ownership of the vehicle.

6. Conclusion

This study has shown the usage variation of Chinese classifiers *bu*, *liang*, and *tai* for vehicles, in which the variation can be understood statistically by the constraints of

the Qualia structure. Theoretically, to the extent that this study has demonstrated that the use of vehicle classifiers *liang*, *bu*, and *tai* is not arbitrary, it does instead follow cognitive-based categories. Methodologically speaking, using logistic regression analysis enabled us to conduct the linguistic variation in a more precise way than previous studies in Chinese classifier analyses could have. The results of the study can thus serve as a reference to Chinese teaching. Since the study is a pilot research by corpus, the data are limited to style and source. In addition, the use of a classifier may be influenced by dialect mixtures, education, age, or so on. Therefore, more divergent and real-time data will be needed for a future study in Chinese classifier variation.

References

Bybee, Joan Lea, and Clay Beckner. 2010. Usage-based theory. *The Oxford Handbook of Linguistic Analysis.* Oxford : Oxford University Press.

Chen, Rong-an. 2014. *A Historical Dictionary of Chinese Classifiers.* Taipei: Crane Publishing Co. Ltd. 207-211.

Chinese Knowledge Information Processing Group. 1998. Content and Instruction of the Sinica Balanced Corpus. Technical report No. 95. 02, Taipei: Academia Sinica.

Gao, Shufan (ed). 1989. *Zhengzhong Comprehensive Dictionary of the Form, Pronunciation, and Meaning [Zhengzhong Xingyinyi Zonghe Dacidian]*.Taipei: Zhengzhong shuju .

Huang, Chu-Ren. and Kathleen Ahrens. 2003. Individuals, kinds, and events: classifier coercion of nouns. *Language Sciences* 25.4: 353-373.

Ji, Li-Jun, Zhiyong Zhang, and Richard E. Nisbett 2004. Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology* 87.1: 57.

Johnson, Daniel Ezra. 2009. Getting off the Goldvarb standard: Introducing Rbrul for mixed effects variable rule analysis. *Language and Linguistics Compass* 3.1: 359-383.

Kuo, Jenny Y. C., May-ling Lee, and James H.-Y. Tai. 2008. Categorization patterns of classifiers in Taiwan Southern Min. *Chinese Linguistics in Leipzig*, ed. by Redouane Djamouri, Barbara Meisterernst, and Rint Sybesma, 207-221. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale.

Kuo, Jenny Y. C. and Maria D. Sera. 2009. Classifier effects on human categorization: The role of shape classifiers in Mandarin Chinese. *Journal of East Asian Linguistics* 18.1:1-19.

Pustejovsky, James. 1995. *The Generative Lexicon.* Cambridge: MIT Press.

Pustejovsky, James, and Elisabetta Jezek. 2016. Integrating Generative Lexicon and Lexical Semantic Resources. *A Guide to Generative Lexicon Theory.* Oxford : Oxford University Press.

Robinson, Justyna A. 2010. Awesome insights into semantic variation. *Advances in Cognitive Sociolinguistics.* Berlin: Mouton de Gruyter, 85-110.

Rovira-Esteva, Sara. 2007. A cognition-based study of 双 (shuang), 对 (dui), 副 (fu), and 套 (tao). *Journal of the Chinese Language Teachers Association* 42.2: 41.

Sankoff, David, Sali Tagliamonte, and Eric Smith. 2005. Goldvarb X: A variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto.

Tagliamonte, Sali. A. 2011. *Variationist sociolinguistics: Change, observation, interpretation* (Vol. 39). Oxford: Wiley-Blackwell.

Tai, James H-Y. and Lianqing Wang. 1990. A semantic study of the classifier *tiao*. *Journal of the Chinese Language Teachers Association* 25.1:35-56.

Tai, James H-Y. 1992. Variation in classifier systems across Chinese dialects: Towards a Cognition-based Semantic Approach. *Chinese Language and Linguistics* 1: 587-608.

Tai, James H-Y. 1994. Chinese classifier systems and human categorization. *In honor of Professor William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, Matthew Chen, and Ovid Tseng, ed. Pyramid Publishing Company, 479-494.

Tai, James H-Y., and Fangyi Chao. 1994. A Semantic Study of the Classifier *zhang*. *Journal of Chinese Language Teachers Association* 29.3:67-78.

Taiwan zhonghua shuju (ed.). 1990. *Cihai*（辭海）. Taipei: Zhonghua shuju.

Zhang, Hong. 2007. Numeral classifiers in Mandarin Chinese. *Journal of East Asian Linguistics* 16.1:43-59.

[審查：2016.10.27 修改：2016.11.2 接受：2016.11.26]

黃榛萱
Chenhsuan HUANG
62102 嘉義縣民雄鄉大學路 168 號　國立中正大學語言學研究所
Institute of Linguistics, National Chung Cheng University
No.168, University Rd., Min-Hsiung, Chiayi 62102, Taiwan
chenhsuanfun@gmail.com

華語車輛量詞之語意變異研究：以「輛、部、台」為例

黃榛萱
國立中正大學語言學研究所

摘要

本文旨在探討華語車輛量詞「輛、部、台」間的語意變異，依屬性結構（Pustejovsky, 1995）將詞意分為組成成份（Constitutive）、動作（Telic）、外在形式（Formal）以及生產方式（Agentivity）等四類，經由 Rbrul 羅吉斯迴歸(logistic regression analysis)找出最佳的影響因子。主要的研究發現如下：首先，「輛」較常使用於四輪以上及非個人使用目的之車輛類型。其次，「部」的語意偏好突顯車輛之所有權以及外觀。至於「台」則傾向強調車輛外觀和來源，同時常用來計算四輪以下之車輛。本研究說明了華語車輛量詞「輛、部、台」的使用非任意，而是依循認知模式的選擇。

關鍵詞：量詞 屬性結構 Rbrul 華語文教學