Notes on Chinese grammar and ontology: the endurant/perdurant dichotomy and Mandarin D-M compounds

Chu-Ren Huang

Correspondence:
churen.huang@polyu.edu.hk
The Hong Kong Polytechnic
University, Hung Hom, Hong Kong

Abstract

Y. R. Chao's (1955) 'Notes on Chinese Grammar and Logic' illustrated how logical relations are encoded in Chinese Grammar and his Chinese grammar (Chao 1968) introduced the grammatical category of Measure (M) in Determiner-Measure (D-M) Compounds. Subsequent studies of Chinese typically adopt the general linguistic term of classifier (Aikenvald 2003) and either refer to Chao's M as a classifier (e.g. Li and Thompson 1981) or assume that it can be further subdivided into two categories: classifiers and measure words (Tai 1994). Many later studies tried to account for the classifiers/measure words contrast via semantic or syntactic tests without reaching a definite conclusion. This paper adopts and merges two lines of Chao's research to show that the ontological concept of endurant vs. perdurant is elegantly instantiated in Chinese grammar, and by the category of M in particular. By doing so I hope to follow Y. R. Chao's (1955) giant leap in studying logical relations in Chinese and to take the further step of exploring the significance of the Chinese language for ontological studies, including issues such as whether Quality should be ontologically dependent on entities or instead subsumed by them.

This paper is not concerned with Chinese logic as a part of technical Chinese philosophy, but rather, with the ways in which some elementary logical notions find expression in the Chinese language.

-Y.R. Chao 1955, First sentence of 'Notes on Chinese Grammar and Logic'
http://www.jstor.org/stable/1397106

1 Introduction

In the way of Chao's (1955) seminal paper on Chinese logical relations, this paper focuses on how two foundational ontological notions find expression in the Chinese language. Ontology in its modern form is the study of how knowledge is organized and represented in knowledge systems (Prévot et al. 2010). As such, recent studies on ontology have focused mostly on digital knowledge representation systems, especially web-based systems. Such studies, however, also involves the knowledge systems of human language and hence lead to crucial research issues in the interface between ontology and natural language lexicon and in how languages conventionalize knowledge representation systems (OntoLex, Huang et al. 2010a).

One important issue in ontology and OntoLex in particular is whether the ontological conceptual primes are also linguistically expressed. The focus of this study will be on one of the most fundamental concepts for the knowledge classification: the endurant/perdurant dichotomy for classification of entities. This concept dichotomizes entities according to whether they are dependent on time or not. To paraphrase the position taken in DOLCE ontology (Descriptive Ontology for Linguistic and Cognitive Engineering, Gangemi et al. 2010), an endurant, is (the concept of) an entity that has spatial components but does not depend on a specific time of occurrence. In other words, it can exist at any point in time and perceived to be identical at different temporal locations. A perdurant is (the concept of) an entity which has a time element crucially associated with its meaning. In other words, to define (the concept of) a perdurant, we need to take into consideration the variations of its instantiation at different time points. Rigid designators such as people and objects are the most typical endurants. For instance, Y. R. Chao in 1955 and in 1968 is the same entity in spite of physical changes. Processes and activities are the most typical perdurants. A perdurant, such as the process of writing, exists as the sum of different stages at different times. At any snapshot of time, it is possible to find instantiations of different aspects of the same process of writing.

As Chinese is a language that has been shown to explicitly encode ontology with its radical-based writing system (Chou and Huang 2010, Huang et al. 2013b), it is natural for us to ask whether the endurant/perdurant dichotomy is also represented in Chinese. To answer this question, the classifier system, which marks linguistic classifications of objects, should be the first system to be examined. In other words, we will be concerned with the issue of whether the linguistic system of classifiers have ontological basis. Classifiers are given the grammatical category of Measure (M) in Determiner-Measure Compound (D-M Compound), a grammatical category specific to Chinese introduced in Y. R. Chao's (1968) Chinese grammar. Although we adopt Chao's term of D-M, we follow subsequent studies (e.g. T'sou 1976, Mo et al. 1996, among others) in treating D-M as a classifier phrase. It is also important to note that Chao (1968) listed 9 different M's, including those measuring activities in a verbal phrase. The current study focuses on noun phrase M's that have been typically treated in Chinese linguistics as part of the linguistic system of classifiers (Aikhenvald 2003). The literature, however, does vary in how Chao's M should be further analyzed and whether all subclasses of M are in fact classifiers. Li and Thompson (1981) uses classifier as a covering term to include measure words; while Tai (1994) stipulate that M contains two distinct categories: classifiers and measure words, and in *A Reference Grammar of Chinese* (Huang and Shi 2016), the classifier category name is retained but differentiated into two distinct categories: sortal classifiers and measure words (Ahrens and Huang 2016). Many studies (e.g. Huang et al. 黃居仁等 1997, Her and Hsieh 2010) have tried to account for the classifiers/measure words contrast via semantic or syntactic tests without reaching a definite conclusion. Wiebusch (1995), in fact, studied the classification of Chinese classifiers in relation to the radical systems, underlining the conceptual basis of the linguistic representation of classification in Chinese.

The linguistic expression of the classifier system of Mandarin Chinese has two characteristics that make it a valued primary source for ontological studies. First, it is unique among classifier languages in the world to have classifiers for events and kinds

in addition to individual objects (e.g. Huang and Ahrens 2003, Huang et al. 黃居仁等 1997). This broad conceptual coverage provides a comprehensive coverage for onto-logical studies. Second, it has been shown in cognitive studies that the use of classi-fiers is semantically motivated (e.g. Ahrens 1994) and that there is neurological evidence for speakers to use classifiers to predict the semantic classes of nouns (e.g. Chou et al. 2014, Wang and Zhang 2014). Lastly, Huang et al. (1998) demonstrated that a Chinese noun class system could be automatically extracted based on the collocation of noun and classifiers. In sum, Chinese classifier system has both the conceptual robustness and the corresponding linguistic expressions needed to provide direct evidence of study of a shared knowledge representation. This paper adopts and merges two lines of Chao's research to show that the ontological concept of endurant vs. perdurant is elegantly instantiated in Chinese grammar, and by the category of M in particular.

In what follows, I will first introduce ontology as an emergent discipline studying how human knowledge system is represented, as well as illustrate the fundamental dichotomy of endurant/perdurant. This is followed by a brief introduction of recent studies in ontology with Chinese as a target language. I will then recapture the linguistic generalizations of Mandarin Chinese D-M compounds. This is followed by evidence and argumentation show-ing that D-M compounds is a linguistic system which expresses the endurant/perdurant dichotomy. The paper concludes with a summary of the results as well as their implications for the ontological studies of linguistic systems.

2 Ontology as knowledge system and the endurant/perdurant dichotomy

2.1 Ontology and knowledge system

Ontology studies the system for knowledge representation in terms of basic concepts and how these concepts are organized in terms of relations, especially in the context of computational representation (Gruber 1995). With the web becoming the primary source for information, which causes both the supply of information and desire for that information to increase exponentially, the need to directly process the semantics of web-based content has become urgent (i.e. the semantic turn of the world wide web). Ontology is the proposed solution to allow computers to process the semantic content of a web page by explicitly stipulating the knowledge representation system of that web site (Berners-Lee et al. 2001). Given that each web-site may present different knowledge systems (hence different ontologies), the construction of a common upper ontology for all ontological systems then become a foundational task in the study of ontology (e.g. SUMO, Niles and Pease 2001, DOLCE, Gangemi et al. 2003, and BFO, Smith and Grenon 2004). And since human beings access information and represent knowledge with different languages, the interface between lexica as knowledge representation systems for languages and ontology (Huang et al. 2010a), as well as among web content, is represented in different languages (Builtelaar and Cimiano 2014). The interface between different domains and among different languages is among the most challenging issues linking studies on language and ontology (Bond et al. 2014).

2.2 The endurant/perdurant dichotomy as the primary bifurcation of entities

One of the most fundamental issues in knowledge representation and in providing the-oretical foundation for the construction of an ontology is the first binary bifurcation for

entities (i.e. beings that exist, and not limited to referential objects). This is a seemingly simple decision that will dictate the fundamental design of the knowledge system, that is, the underlying conceptual or informational criteria for creating different branches in the knowledge system. Hence, before committing to any structure, builders of upper ontology (the shared upper parts of ontologies systems) often engage in extensive discussion in philosophical, logical, linguistic, and cognitive theories before making commitment to this bifurcation (e.g. Guarino 1998, and Guarino and Welty 2002 for DOLCE, Niles and Pease 2001 for SUMO, and Grenon and Smith 2004 for Basis Formal Ontology (BFO)). Interestingly, many upper ontologies adopt the endurant/perdurant dichotomy for this primary classification, although in somewhat different ways. This fundamental classification of entities roughly corresponds to what is called continuant and occurrent in philosophy (Gangemi et al. 2003, Grenon and Smith 2004). To put it somewhat simplistically, an endurant is an entity which is fully present at any time; while a perdurant is an entity which may have only parts of it present at any specific time, i.e. its presence as captured by 'snapshots' at different time may vary, and its existence is defined by sum of these 'snapshots'. Hence the implication is that it is NOT the shape or other perceivable physical properties, but rather the entity's continuity of existence in time that plays a central role in the classification of entities in our knowledge systems. Different upper ontologies, however, do implement this bifurcation differently. BFO, for instance, has a straightforward bifurcation of continuant vs. occurrent, and allows quality and other properties to be subsumed under either type of entities (Grenon and Smith 2004, Smith and Grenon 2004)[a]. DOLCE, on the other hand, apply the endurant/perdurant dichotomy to entities only (Gangemi et al. 2003), and treats Quality as a separate ontological category[b]. The expression of endurant/perdurant dichotomy can be illustrated by the DOLCE upper ontology (adapted from Gangemi et al. 2003, Gangemi et al. 2010) and given Figure 1 below. A different representation is BFO's basic bifurcation of continuant/occurrent, as illustrated Figure 2 (adapted from Smith 2012).

Figures 1 and 2 present two alternatives to incorporate the endurant/perdurant dichotomy in ontology. BFO's view is that these are simply two views to represent our knowledge. If we take a three-dimensional view focused on the continuant, we could describe the independent (i.e. referential) part of the continuant as well as the dependent part of the continuant (i.e. the disposition and quality of the continuant). DOLCE, on the other hand, restrict the endurant/perdurant classification for entities only, and identifies quality as a separate unique beginning in ontology. Anticipating that the classifier system will involve quality of the entity, we can also compare these two views to see which is better suited to describe this linguistic system.

Given the prominent role of the time and variation driven endurant vs. perdurant dichotomy in ontology, it will be interesting to find out if it is expressed in linguistic systems and how. Intuitively, by the definition of endurant/perdurant and the DOLCE ontology example, we can see that noun is a part of speech (PoS) which is typically adopted for endurants; while verbs are typical PoS's adopted for perdurants. However, the similarity stops at broad conceptual motivation as most linguistic systems are far more complex. The link is fairly straightforward for proper nouns as rigid designators, as their references do not change over time. Similarly, the meaning of common nouns, such as 'book' or 'soldier', cannot be fully interpreted unless we assume the presence of the whole entity at any time where the existence of that entity is confirmed. 'A book with
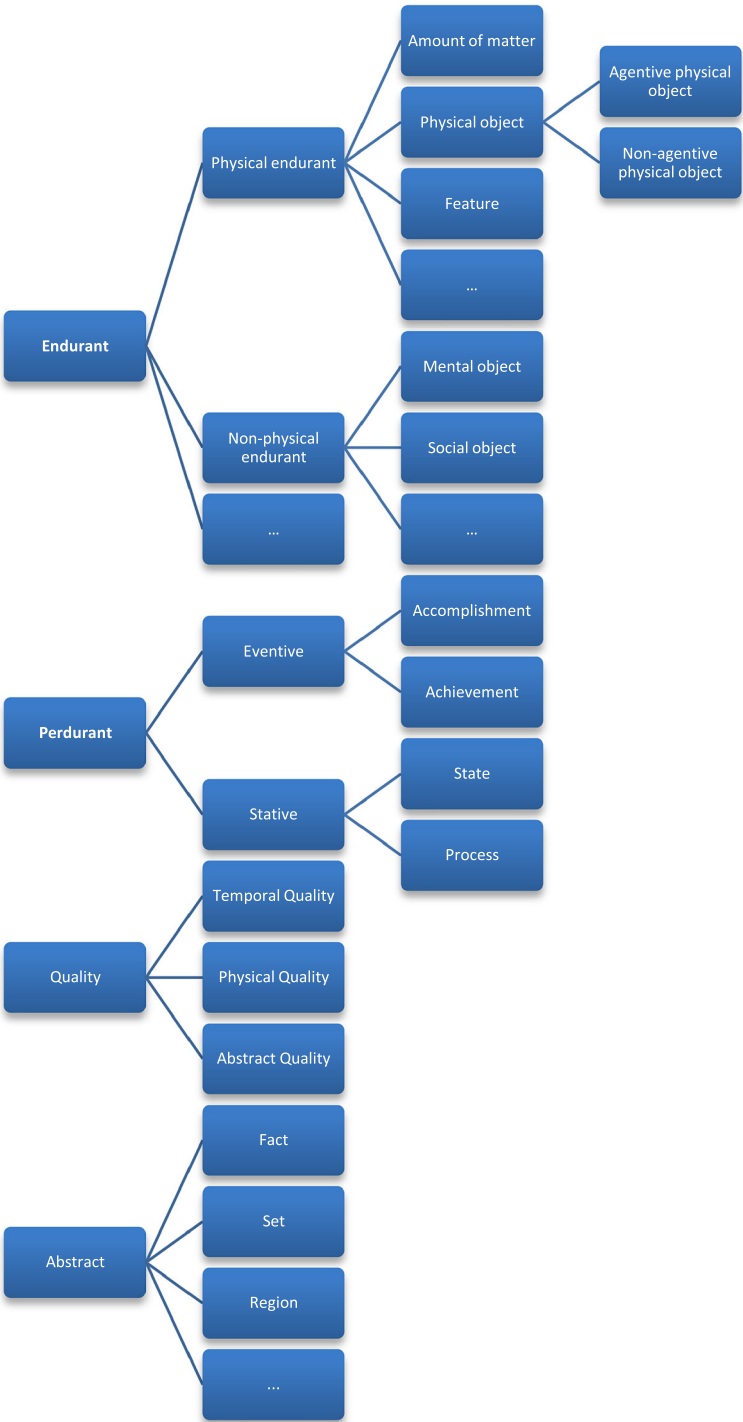
Figure 1 DOLCE upper ontology: entities.

it cover missing' or 'a soldier who lost an arm during World War II' can be understood and the entities can be recognized as 'the same' as before the loss of their parts because the conceptual whole may be invoked at any time. Verbs, on the other hand, refer to a process that is carried out in a dynamic way over time. With enough temporal granularity, one can see that the presence of an event entity must vary from one time point to the

Figure 2 Basic Formal Ontology (BFO) upper ontology.
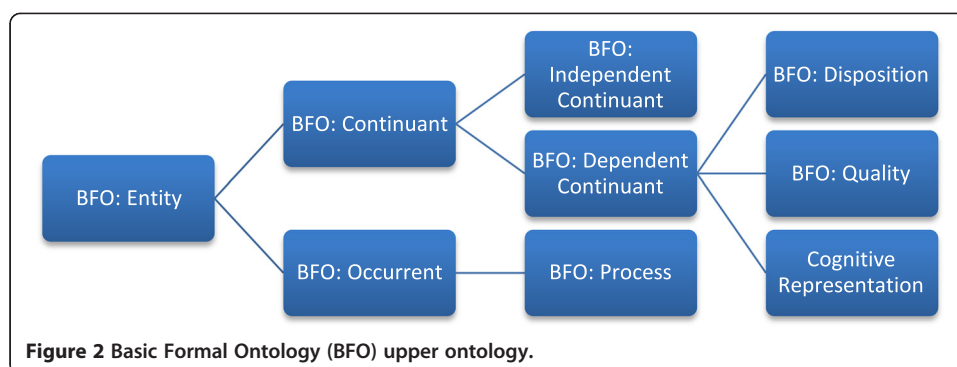
other. 'To run, running,' for example, can be envisioned as a series of snapshots where a foot is on or off the ground, or on an upward or downward trajectory. It is even more obvious for complex events such as accomplishment and achievement that at any given time, only part of the full event as entity is present. In other words, the endurant/ perdurant dichotomy seems to provide conceptual motivation behind the nominal/verbal dichotomy adopted in linguistic systems.

It is well known, however, that the intuitive nominal/verbal distinction can be easily blurred with many categorical change devices in language as well as with atypical members of each PoS: such as event nouns, deverbal nominal, denominal verbs etc. Hence for the verby/nouny bifurcation, the endurant/perdurant dichotomy seems to be a default motivation rather than a conceptual must and is not systematically expressed. Hence, we need to look further for clear evidence of if and how a linguistic system, such as Chinese, expresses the endurant/perdurant dichotomy.

2.3 Chinese as a knowledge system: recent studies on ontology and Chinese

Ontological studies of language and lexicon (e.g. Gangemi et al. 2003, 2010, Pease and Fellbaum 2010) have focused on the mapping of the linguistic system built on lexical semantic relations (e.g. hypernym) to the ontological system built on logical relations (i.e. ISA relation). Although similar studies have been carried out on Chinese (e.g. Huang et al. 2010b), the fact that Chinese orthography explicitly designate radicals as semantic primitives provides a new perspective on the possible relation between language and ontology. It has also previously been demonstrated that Chinese orthography is a conventionalized knowledge representation system (Hantology, Chou and Huang 2010). In addition, Huang et al. 黃居仁等 (2013a) and Huang et al. (2013b) showed that the ontological system with radicals representing semantic primitives is driven by Aristotle's four causes and have the generative power similar to Pustejovsky's (1995) qualia structure. This is illustrated by Figure 3, adapted from Chou and Huang (2010).

What Figure 3 illustrates is that all the Chinese characters sharing the same radical 艸 cao 'grass,' instantiated as the double cross components on top of each character, incorporates the conceptual primitive of 'plant'. How this differs from a typical taxonomy has to do with the fact that the relation between the semantic primitive and derived concepts is far richer than what is usually found in a typical IS-A relation. For characters with the radical 艸 cao 'grass,' the conceptual relations include IS-A, IS_Part_Of, Telic, and Event_descriptive. This

Figure 3 Ontology of the radical 艸 for plants.
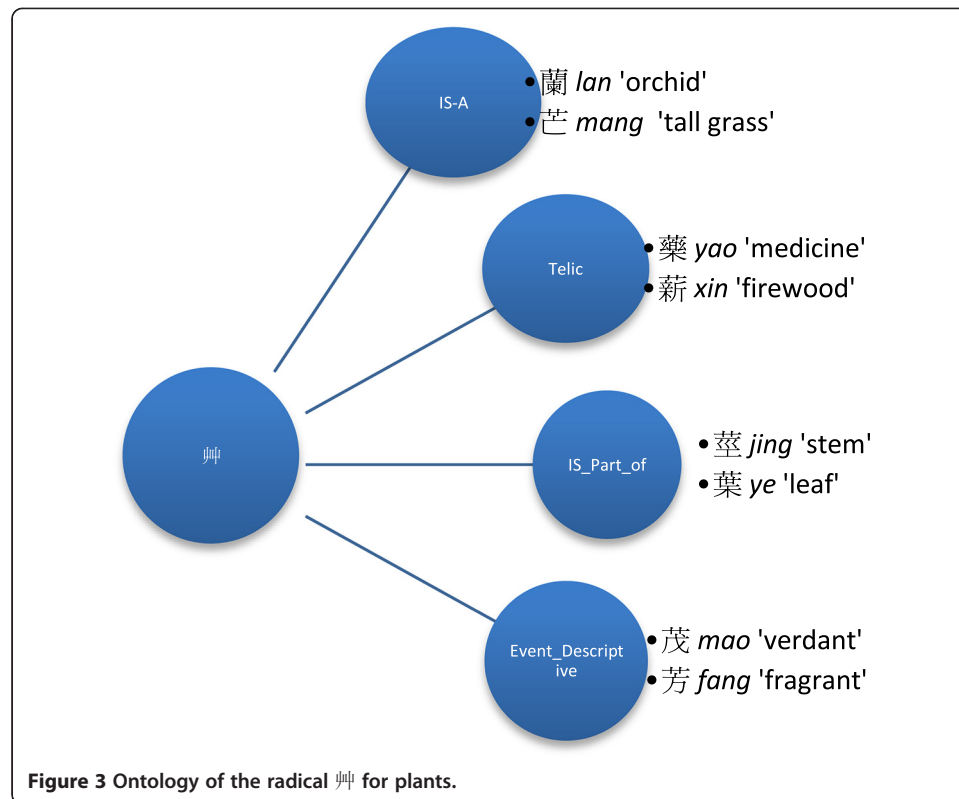
maps well to Aristotle's four causes (material, physical, agentive, and telic) as well as Puste-jovsky's (1995) qualia structure. Huang et al. (2013) takes this argument one step further when they point out that the Chinese orthography is indeed a knowledge system organized by radicals which each represent a conceptual primitive but are organized according to eventive relations similar to the Four Causes or the four qualia. Huang et al. (2013b) showed that in fact this analysis can be extended to all radicals in Chinese and that Chinese orthography is indeed a conventionalized knowledge representation system. This ontological interpretation of the Chinese orthography laid a foundation for accounts of its conceptual robustness and representational versatility as the shared writing system through historical changes (Chou and Huang 周亞民, 黃居仁 2006) and for typologically divergent languages (Huang and Chou 2015).

3 Classifiers as an ontological system

3.1 The Chinese classifier system

M in a D-M compound (including sortal classifiers and measure words) individuates the entity represented by NP to allow it to be quantified. It does so by selecting some properties of that entity as the basis for units of individuation and enumeration Aikhenvald (2003).

In what follows, I investigate the conceptual basis of the properties selected by M in order to attempt to differentiate the type of properties selected by different sub-types of M's in light of the endurant/perdurant dichotomy. In particular, I look at whether M selects a time-invariant (endurant) or time-variant (perdurant) property of that entity.

The current study adopts the sub-categories of classifiers (M) given in Ahrens and Huang (2016), which largely follow Chao's (1968) grammar with updates to reflect later studies on Chinese grammar and classifiers. As seen in Figure 4, classifiers (M's) are divided into two major categories based on their linguistic behaviors: sortal classifiers and measure words. Intuitively, sortal classifiers select the semantic class of their head nouns, while measure words do not have strict selectional preferences. Sortal classifiers consist of three sub-categories: individual classifiers, event classifiers and kind classifiers. Measure words also consist of three categories: container measure words, standard measure words, and temporary measure words.

Note that even though I use constructed examples for clear explication, they are constructed to be representational of generalizations attested and extracted from the 5 million word version of Sinica Corpus (Chen et al. 1996) and accounted for in Huang et al. 黃居仁等 (1997). It also important to note individual variation is a hallmark of human languages (Fillmore et al. 1979). Hence it is expected that some speakers may have differences in interpreting or usage of some of the examples presented. It is important to ensure that such variations are not in conflict with the basic expression of the ontological bifurcation. In addition, the aim of this paper is not to describe all linguistic variations, but to capture the systemicity of the expression of the ontological notions, as well as the robustness of the conceptual motivation of the linguistic system.

3.2 Sortal classifiers denote endurant properties
3.2.1 Individual classifiers
First, as the most typical sortal classifier, individual classifiers typically select common nouns, which of course are endurant entities. However, the property the classifier selects depends on the meaning referred to. For instance, 張 *zhang* as an individual classifier has multiple senses, and two of the most frequent senses are exemplified by 一張紙 *yi zhang zhi* 'a piece of paper', and 一張椅子 *yi zhang yizi* 'a chair'. 張

Figure 4 Chinese classifier system (Ahrens and Huang 2016).
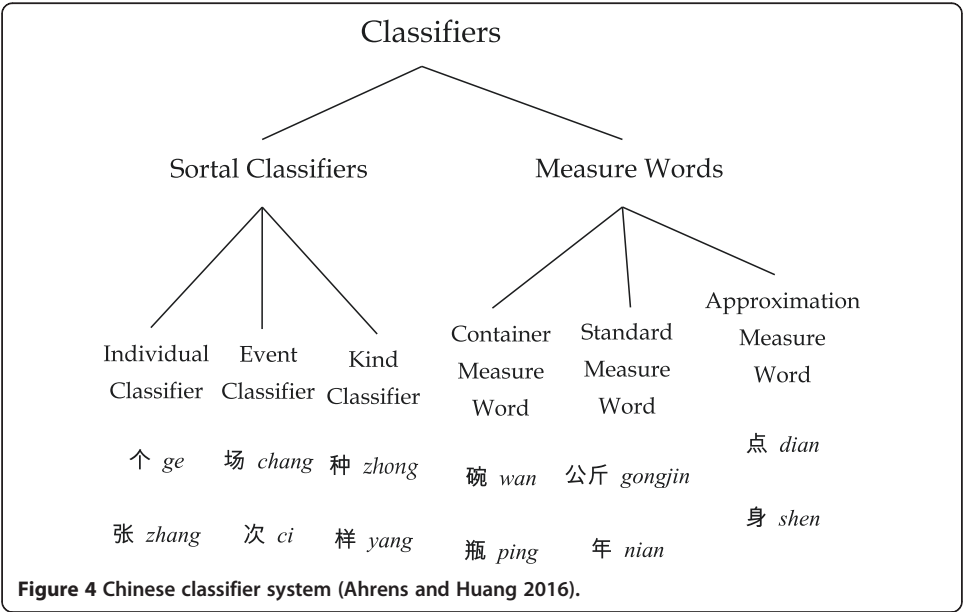
zhang in 一張紙 *yi zhang zhi* 'a piece of paper' represent the sheet-like property of paper, which is an endurant property of paper (or any other common noun selected by 張 *zhang*). If this type of common noun exists in this form, it cannot lose this property. Similarly for 一張椅子 *yi zhang yizi* 'a chair', the classifier 張 *zhang* selects the endurant property of a flat surface where people can put their body on (e.g. beds and chairs). Note that this is a telic property defined by the designed purpose, instead of direct description of physical properties of an object. The fact that the properties denoted by these individual classifiers are time-invariant can be shown in real world situations where that property seems to be challenged, as in 1a-b below.

(1) a. 一張破破爛爛的紙
yi__zhang__popolanlan__de__zhi
one__CL__tattered__DE__paper
one piece of tattered paper

  b. 那張缺腿的椅子
na__zhang__quetui__de__yizi
that__CL__leg-missing__DE__chair
that chair with a missing leg

1a-b show that the property denoted by individual classifiers endures at all time as long as that entity exists, regardless of the actual physical state of the entity. In 1a, as long as an entity's existence as paper is confirmed, its linguistic expression with the 張 *zhang* classifier is not affected by how tattered and un-sheet-like it is at a certain specific time. Similar for the furniture with flat surface 張 *zhang* in 1b, as long as the existence of the entity is confirmed, the classifier can be used to express that enduring property regardless of whether the object is capable of serving its furniture function at the specific time.

The individual classifier that is most difficult to analyze is perhaps the generic classifiers 個 *ge*, as the property it selects is famously difficult to capture precisely. We could in general describe the property as 'individualizable'. I.e. 個 *ge* typically selects common nouns that can be selected by one of the individual classifiers. In this sense, the classifier denotes a generic endurant property that is the common property shared by the set of all endurant properties denoted by each individual classifier.

3.2.2 Kind classifiers

The second type of sortal classifier is what is called the kind classifier (Huang et al. 黃居仁 等 1997, Ahrens and Huang 2016). However, the term 'type classifier' is probably an even more appropriate name as they differ from individual classifiers by selecting types instead of tokens (i.e. individuals). They coerce type reading from the common noun. That is, a kind/ type classifier denotes a property to select a sub-set of the referents of the common noun instead of individuals that share that property. For instance, in 2, the kind/type classifier 款 *kuan* selects among all sweaters three particular styles (e.g. by design or by brand).

(2) 這三款毛衣, 這個冬天很流行。
zhe__san__kuan__maoyi, zhege__dongtian__hen__liuxing
this__three__style(CL)__sweater, this__winter__very__fashionable
These three styles of sweater are very fashionable this winter.

Similarly 樣 *yang* 'type (CL)' in 3 states that the subject bought three types of objects, without committing to the number tokens of each type[c].

(3) 他買了三樣東西。
    ta__mai__le__san__yang__dongxi
    s/he__buy__PERF__three__type(CL)__things
    S/he bought three different kinds of stuff.

As mentioned, 款 *kuan* refers to properties of members of a type sharing the same style, such as referring to iPhone 6.1. as 這款手機 *zhe kuan shouji* 'this model cell-phone'. Similar to individual classifiers, the type selected share properties that are invariant through time. That is, the existence of the type denoted is continuant over different time. Last, but not the least, similar to the generic individual classifier 個 *ge*, the generic kind classifier 種 *zhong* selects a under-specified type that can be identified in context. In this usage, 種 *zhong* is the most generic of all classifiers as it select virtually all common nouns. This is because there are fewer semantic constraints on which entities can be referred to as types (that which can be referred to as individuals).

It is important to note that the use of kind/type classifier must denote time-invariant enduring properties. For instance, it would be appropriate to use 這一款手機 *zhe yi kuan shouji* 'this model cell-phone' to refer to iPhone 6, Samsung, android cell phones, etc. However, it would not be appropriate to use it to denote the sub-set of cell-phones that are bundled with a service contract. Being bundled with a service contract will change over time and is not an enduring property that is independent of time.

For example, even though any product can be sold with a discount (or mark-up) at a price different from its published price, it is simply infelicitous to use that situation dependent price to refer to that particular type of product. Hence, given attested 'snapshot' at a specific time of 4a, it is still appropriate to refer to that kind of cell phone as 4b, as this linguistic expression refers to a distinct model of cellphone with a published price of 2,500 dollars. It can only be referred to as 這一款三千元手機 *zhe yi kuan sanqian yuan shouji* 'the 3,000 dollar model cellphone'. In other words, the kind classifier clearly selects only time-invariant properties and is incompatible with properties whose instantiation is only true for some specific time and situation[d].

(4) a. 這一款三千元手機只賣兩千五
      zhe__yi__kuan__sanqian__yuan__shouji__zhi__mai__liangqianwu
      this__one__style(CL)__three-thousand__dollar__cellphone__only__sell__two-thousand-five
      This 3,000 dollar model cellphone was sold for only 2,500 dollars.
   b. 這一款兩千五百元手機
      zhe__yi__kuan__liangqianwubai__yuan__shouji
      this__one__style(CL)__two-thousand-five-hundred__dollar__cellphone
      this 2,500 dollar model cellphone

Last, but not the least, please also note that even though D-M compounds with kind classifiers can receive kind readings, (e.g. such as in 'Dogs are bigger than cats.'), they

should be treated simply as a semantic alternation of the construction, rather than as the meaning of the classifier.

3.2.3 Event classifiers

Lastly, the event classifier selects event nouns, which are names used to refer to events as entities (similar to Chierchia's account of English gerunds). In gerund usage of 'Swimming is fun,' Chierchia (1984) argued that 'swimming' refers to the event in whole as an entity that can be assigned some constant property, instead of the typical process with time-variant aspects of instantiations. As such they select entities that are continuant through time. Huang and Ahrens (2003) argued that in Chinese, such usage is further supported by the existence of event classifiers. An event classifier can either select event nouns (such as 會議 *huiyi* 'meeting' or 比賽 *bisai* 'race, competition') which directly encode events as entities or coerces referential eventive meanings from common nouns. For instance 三場電影 *san chang dianying* 'three (scheduled) showings of the movie' and 兩班公車 *liang ban gongche* 'two scheduled running of public bus'. These two classifiers, 場 *chang*, and 班 *ban*, select event nouns which are names of events that occur according to a schedule. 班 *ban* selects public transportation while 場 *chang* selects drama-like events (i.e. events which progress according to a script). These classifiers select event individuals, and allow them to be enumerated. In other words, once an event noun co-occurs with an event classifier, it loses its meaning of referring to the events themselves as occurrants. It now refers to the more abstract concept of the existence of each event as an individual. These individual entities, unlike the actual (non-) occurrence of the event, are not bound to any temporal point, such as in 5.

(5) a. 10:49 那班火車, 11:23 才到。
     10:49__na__ban__huoche, 11:23 cai__dao
     10:49__that__CL__train, 11:23 just__arrive
     The 10:49 train has just arrived at 11:23.
    b. 三場電影,兩場滿座,一場取消。
     san__chang__dianying, liang__chang__manzuo, yi__chang__quxiao
     three__CL__movie, two__CL__full, one__CL__cancel
     Of the three showings of this film, two were full and one got cancelled.

In the two example sentences in 5, I show that event classifier coerce the reading of event as individuals and hence free from specific temporal reference and eventive instantiation. Hence a 10:49 train for a particular day/morning is always a 10:49 train, regardless of when it arrives, even if it gets cancelled. Similar, each scheduled showing of a film is treated the same, regardless of how many people see the film or even whether it was actually shown or not. Since event classifiers refer to the time-invariant aspect of an event, the endurant/perdurant contrast can be teased apart, as illustrated by 6.

(6) 10:49 那班飛機, 11:23 才起飛。
     10:49__na__ban__feiji, 11:23 cai__qifei
     10:49__that__CL__airplane, 11:23 just__take-off
     The 10:49 flight did not take off until 11:23
    a. 請問10:49 那班飛機, 什麼時候抵達?
     qingwen__10:49__na__ban__feiji, shenme__shihou__dida

please-ask__10:49__that__CL__airplane, which__when__arrive
Can you tell me when will the 10:49 flight arrive?

b. 請問11:23 那班飛機, 什麼時候抵達。

qingwen__11:23__na__ban__feiji, shenme__shihou__dida

please-ask__11:23__that__CL__airplane, which__when__arrive

Can you tell me when will the 11:23 flight arrive?

Given an attested flight delay in 6, even with the knowledge of the actual time of taking off, 6b will be an inappropriate query for the arrival time of the flight. 6a instead is the appropriate query sentence. This is because the event classifier 班 *ban*, similar to individual classifiers, selects a time-invariant property shared by this type of events. The property 班 *ban* selects is 'having the same scheduled time'. For any scheduled event, the scheduled time is an enduring property that will not be affected by the actual event time. In this particular example, the property of having a specific scheduled departure time of a flight will not change regardless of whether the flight is on time, delayed, early, or cancelled on a specific date. The actual departure time of a flight, however, is associated with a specific event instantiation and is not an enduring property of that flight, and cannot be used to identify that particular type of event. 6b can only be an appropriate query if there is a flight scheduled for 11:23.

The generalization that event classifiers select endurant eventive properties in fact can supplement the coercion account of event classifiers turning concrete object entities to event entities given in Huang and Ahrens (2003). For example, in 7a-b, as well as 6, the original meaning of the head nouns refers to the entities of the telephone set and the rain water respectively.

(7) a. 兩通 (未接) 電話

liang__tong__(weijie)__dianhua

two__CL__(unanswered)__telephone

two (unanswered) calls

b. 好幾陣雨

hao__ji__zhen__yu

good__several__CL__rain

quite a few rain showers

Huang and Ahrens (2003) argued that the event readings are coerced by the classifier but did not explicate how the coercion happened. Based on the generalization observed so far, a sortal classifier serves as a linguistic device to express a defining property of a type of time-invariant entities. To serve this function to conceptualize events as endurant entities, the most likely properties that an event classifier can pick up are properties of event structures. For instance, the classifier 通 *tong* has the original verbal meaning of 'connecting, going through'. As an event classifier, it picks up the property of individuating a single successful connection as the starting point to define a calling event as an endurant. This can be shared by all telecommunication events and indeed 通 *tong* is an event classifier for other telecommunication events including 電報 *dianbao* 'telegraph' or even the newly introduced 短訊 *duanxun* 'short message, SMS'. Similarly 陣 *zhen*'s original meaning refers to an episode of a meteorological events. As an event classifier, it

picks up the holistic feature of that episode from onset to ending as well as the shared feature of a non-volitional 'happening'. Intuitively, we could view the function of event classifiers as expressing the 'shapes of event structure', as described by Huang et al. (2000).

To sum up, our discussion showed that event classifier selects a time-invariant property. I also showed that by assuming a sortal classifier must express an 'enduring' property shared by the entity type, I can predict that event classifiers must refer to 'shapes of event structures' and furthermore, it is this expression of shapes of event structures that allows event classifiers to coerce event entity reading from nouns denoting concrete entities.

3.3 Measure words denote perdurant properties
3.3.1 Standard measure words
Measure words also have three sub-classes. The proto-typical measure word is a standard measure word, such as 公斤 *gongjin* 'kilogram' in 三公斤肉 *san gongjin rou* 'three kilograms of meat', which are part of the standard measurement system. The standard measurement system is a property introduced independent of the entity. It is not a property that is inherent to the entity and independent of time. Instead, it refers to perdurant information, as the measurement is only true at a particular point in time of that particular instantiation of the entity and may vary in time.

(8)　一公斤肉, 煮熟後只剩不到600公克
　　　yi__gongjin__rou, zhushou__hou__zhi__sheng__budao__600__gongke
　　　one__kilogram__meat, cooked__after__only__left__less__600__gram
　　　One kilogram of meat only weighs less than 600 grams after being cooked.

Example 8 shows that the same entity can take different measurements or measure words at different times. The fact that standard measure words stand for time-variant properties can also be illustrated by the fact that an entity can take as many standard measure words as long as the situation context allows it to be measured by the standard.

It is important to note, however, although the property denoted by a standard measure is inherently perdurant (i.e. time-dependent and not linked to any specific independent entity), other linguistic devices can be used to coerce an endurant-continuant reading of a noun modified by a standard measure.

(9)　這一公斤肉, 煮成了三道菜
　　　zheyi__gongjin__rou, zhucheng__le__san__dao__cai
　　　this__one__kilogram__meat, cook-make__LE__three__CL__dish
　　　This (piece) of one kilogram meat was made into three dishes.

In 9, the weight of the meat is used to establish the identity of the entity (rather than providing measurement). Hence it is considered to be an enduring property and used to refer to the same entity even though, as we know, the weight of the meat after being cooked has already changed. This interpretation is consistent with the BFO view that the same entity can be described either in terms of SNAP or SPAN ontology to focus

on different properties. It is also important to note that the perdurant reading in 8 allows DE-insertion, while the endurant/continuant reading in 9 in is resistant to DE-insertion. This issue will be explicated in section 3.4.

3.3.2 Container measure words

The same can be said of the container measure words as the second type of measure words. Container measure words, such as 包 *bao* 'package', 箱 *xiang* 'case', etc. can in principle measure any entity as long as the real world context allows that entity to be put inside that particular container. In other words, a container measure word denotes a time-variant state where the entity is (envisioned to be) contained inside the type of container specified. The interpretation of the following D-M compounds are situation and context dependent: 三包糖 *san bao tang* 'three packs of sugar', 三箱糖 *san xiang tang* 'three cartons of sugar', 三包筆 *san bao bi* 'three packs of pens', 三箱筆 *san xiang bi* 'three cartons of pens'. There is no way to ascertain the actual quantity of objects in each container without explicit knowledge of the particular situations. Like standard measure words, container measure words' perdurant property is shown by its high versatility in measuring and denoting properties of all types of entities. In addition, the interpretation of the property (both of volume and ways contained) of each container is also dependent on the container or the (partially conventionalized) way of packaging involved in defining the container. Again, the note on the possibility of borrowing SPAN ontology concept for description a SNAP ontology, discussed in the last section on measure words, also applies here. In other words, when required by real world context, the language does allow a speaker to select a perdurant property described by a container classifier to treat it as an endurant.

3.3.3 Temporary measure words

Lastly, temporary measure words are often called 'pseudo-classifiers' they typically only take the numeral 一 *yi* 'one' (Chao 1968). There are two sub-types of temporary measure words. The first is derived from a familiar cognate concrete common noun (with body parts being some the more prominent ones), as in (10).

(10) a. 一/滿身（的）灰
        yi/man__shen__(de)__hui
        one/full__body__(DE)__dust
        a body-ful of dust
     b. 一/滿屋子（的）灰
        yi/man__wuzi__(de)__hui
        one/full__room__(DE)__dust
        a roomful of dust

10 describes a result state where the existence of certain entity is prevalent at a certain location. Here a container metaphor is invoked to describe the result state, with the filled-in location being used as the temporary measure word. The situation describe occurs at a specific time-location and is not enduring for either the entity or the measure word. When used as an object, it refers to the extent of a result state, as in 11, again underlining the perdurant property.

(11) 汽車濺了他一身的水

      qiche__jian__le__ta__yi__shen__de__shui

      car__splatter__PERF__s/he__onel__body__DE__water

      The car splattered water all over him/her.

Another sub-type of temporary measure words involve denoting the prevalence of certain mental activities, as in 12:

(12) a. 這是我的一點心意

        zhe__shi__wode__yi__dian__xinyi

        this__be__mine__one__point__heart

        This represents my gratitude/heart-felt appreciation.

      b. 他心中有一堆問題

        ta__xin__zhong__you__yi__dui__wenti

        s/he__heart__middle__have__one__pile__question

        S/He has so many unanswered question on his/her mind.

In 12, the temporary measure word denotes the extent of mental state. This is again a time-specific occurrent and thus, a perdurant property.

3.4 Linguistic expression of ontological notions

3.4.1 The correlation between DE-insertion and perdurant properties

Interestingly, the endurant/perdurant contrast observed above corresponds to the morphosyntactic behavior of DE-insertion between the D-M compound and the NP it modifies. This, in fact, leads to the common practice in the field of Chinese linguistics to use the lack of potential for DE-insertion as a litmus test for classifiers (Huang et al. 黃居仁等 1997, Her and Hsieh 2010, Ahrens and Huang 2016).

(13) a. 三本書

        san__ben__shu

        three__CL__book

        three books

      b. 三款手機

        san__kuan__shouji

        three__type__cell phone

        three types of cell phones

      c. 三場電影

        san__chang__dianying

        three__CL__movie

        three showings of movie

(14) a. 一/滿身（的）灰

        yi/man__shen__(de)__hui

        one/full__body__(DE)__dust

        a body-ful of dust

　　b. 三公斤（的）書
　　　san__gongjin__(de)__shu
　　　three__kilogram__(DE)__book
　　　three kilograms of books
　　c. 三包（的）書
　　　san__bao__(de)__shu
　　　three__package__(DE) book
　　　three packages of books

There is a clear contrast between endurant M, i.e. sortal classifiers in 13, and perdurant M. i.e. measure words in 14, which demonstrate that DE-insertion is allowed only when the M selects perdurant properties and that in general, DE-insertion does not change the meaning of perdurant D-M compounds.

3.4.2 When DE-insertion applies to sortal classisifers

This generalization in fact can be extended to even seeming exceptions with when DE-insertion occurs with sortal classifiers, as the examples in 15 allow perdurant readings only.

(15) a. 一百回的水滸傳
　　　　yibai__hui__de__shuihuzhuan
　　　　one-hundred__chapter__DE__Water-Margin
　　　　the one hundred chapter edition of Water Margin
　　b. 一百二十回的水滸傳
　　　　yibaiershi__hui__de__shuihuzhuan
　　　　one-hundred-and-twenty__chapter__DE__Water-Margin
　　　　the one hundred and twenty chapter edition of Water Margin

回 *hui* in 15 is in fact an event classifier for literary works, referring to both scenes in play and chapters in classical vernacular novels (which typically originated from 評書 *pingshu* 'oral storytelling'). As a sortal classifier, it should not allow DE-insertion. In 15, with DE-insertion, the interpretation is, in fact, perdurant. That is, instead of the enumerating function of a typical D-M compound, 15a and 15b are used to differentiate distinct editions of *Water Margin,* which is known to have multiple editions containing different numbers of chapters. In other words, 水滸傳 *shuihuzhuan* 'Water Margin' here is not longer a single time-invariant entity, it is now viewed as a collection of endurant entities (i.e. each different edition of *Water Margin* is considered a separate entity). These entities are, however, differentiated by the situation specific property of the number of chapters they contain.

Similarly, in 16a and 16b, the examples involving internal modification of individual classifiers unexpectedly allow DE-insertion. A small set of adjectives is allowed to occur before sortal classifiers to describe the situation specific properties of the entity (paper in our examples). In such usages, the property the classifiers denote becomes time-specific and perdurant. And linguistically this is clearly marked by the allowance of DE-insertion. In addition, comparing 16c-d with 16a-b underlines the assignment of perdurant properties.

(16) a. （一）大張（的）紙

   (yi)__da__zhang__(de)__zhi

   (one)__big__sheet__(DE)__paper

   a sheet of big paper

b. （一）小張（的）紙

   (yi)__xiao__zhang__(de)__zhi

   (one)__small__sheet__(DE)__paper

   a sheet of small paper

c. 一張大紙

   yi__zhang__da__zhi

   one__ sheet__big__paper

   a sheet of big paper

d. 一張小紙

   yi__zhang__xiao__zhi

   one__sheet__small__paper

   a sheet of small paper

16c, d shows that it is possible to have these adjectives directly modify the common noun. However, in such usages, it follows the typical individual classifiers for not allowing DE-insertion. In other words, the direct modification of the entity does not change the endurant property of the entity. This contrast nicely illustrate that the linguistic expression of by DE-insertion in fact marks the perdurant property of the D-M compounds. In sum, seeming exceptions of sortal classifiers allowing DE-insertion are instances where the DE-insertion coerces a perdurant reading. Hence, I have showed that *de*–insertion is allowed in D-M construction when M selects perdurant properties of the entity it modifies. This generalization about the semantic function of   DE-insertion is consistent with another observation made by Chao (1968) between 白花油 *baihuayou* 'Pak Fah Yeow' 白花的油 *baihua de you* 'a(n) (essential) oil made from a white flower', as illustrated in 17.

(17) a. 白花油

   baihuayou

   white__flower__oil

   Pak Fah Yeow

b. 白花的油

   baihua__de__you

   white__flower__DE__oil

   a(n) (essential) oil made from a white flower

Example 17 shows that 白花油 *baihuayou* 'Pak Fah Yeow', a proper name for a product with time-invariant referent, does not allow DE-insertion, while 白花的油 *baihua de you* 'a(n) (essential) oil made from a white flower', which refers to time-variant referent depending on which kind of flower is used to produce the (essential) oil on each occasion, must be used with 的 *de* 'DE' inserted. Following the generalization obtained so far, we can account for this contrast observed in Chao (1968) by hypothesizing that the insertion of 的 *de* 'DE' in a compound or noun phrase requires a time-variant/ perdurant interpretation of the pre-head element.

3.4.3 Does DE-insertion mark time-variant property?

的 *de* 'DE', as the most frequent word and character in Chinese, accounts for up to 5% of word frequency in a corpus (e.g. Chen et al. 1996), and remains one of the most challenging function words to be accounted for in Chinese. Contrary to pervasive literature in Chinese linguistics, following Zhu 朱德熙 (1961), trying to differentiate a range of different functions and meaning of 的 *de* 'DE', Huang (1987) argued that all 的 *de* 'DE' in Chinese has one single syntactic function: to mark the unit following it as syntactic head. In addition, Huang 黃居仁 (2013) suggested that such head marking functions could be treated as a construction. Based on the occurrence of DE-insertion with D-M phrases, and supported by examples involving other compounds, it seem that perhaps 的 *de* 'DE' may have a single uniform function of marking the property denoted by preceding element as time-variant and perdurant. This seems to be a plausible account given the emergent account that all relative clauses, marked by 的 *de* 'DE' before its head noun, are all restrictive (Shi 2016). It seems that the different accounts attempting to give a uniform linguistic function to 的 *de* 'DE' can in fact be unified by the conceptual motivation that 的 *de* 'DE' is a linguistic expression of the ontological notion of perdurant in Chinese. That is, the phrase before *de* introduces time or situation dependent property, which intersects with the endurant and/or perdurant entity represented by the head noun to establish a more restrictive meaning.

This ontological meaning is also consistent with the type-shifting formal account proposed in Huang (1987) where *de* introduces type-shifting and will look for a situation specific (in the temporal-locational-event continuum) variable to bind to the preceding clause, as illustrated in 18 where the most typical 'light' head nouns to follow *de* are in fact time, situation and location (adopted from Huang 黃居仁 2013). In 18a-c, for example, the element preceding 的 *de* 'DE' each denotes a specific situation, which then intersects with the 'light' head noun indicating a situation type that stipulates a particular time-variant situation.

(18) a. 開會的時候
         kaihui__de__shihou
         convene-meeting__DE__time
         when meeting is held
     b. 去過的地方
         qu__guo__de__difang
         go__GUO__DE__place
         place(s) one has been to
     c. 沒有的事
         meiyou__de__shi
         NEG-exist__DE__matter
         Impossible!/Nonsense!

A full account of the all Chinese expressions involving 的 *de* 'DE' is clearly beyond the scope of the current paper. However, based on ontological interpretations discussed in this paper, there are two possible interpretations. The first, consistent with the upper ontology design of DOLCE, is that the insertion of 的 *de* 'DE' marks the preceding element as denoting perdurant properties. The second, consistent with the treatment of

continuant/occurrent contrast of BFO, is that the insertion of 的 *de* 'DE' marks the shift to a SPAN (i.e. four-dimensional spatiotemporal) ontological view, and hence underlines time-dependent properties. Either ontological account will have important implications for explanatory accounts of Chinese grammar.

4 Conclusion

I have shown in this paper that the Chinese classifier system offers robust linguistic expression of the ontological notions of endurant vs. perdurant. In particular, the dichotomy is encoded with the sortal classifier vs. measure words sub-systems of the Chinese classifiers. In addition, I have also shown that DE-insertion in D-M compounds is an explicit and reliable mark to underline time-variant properties, either marking the shift to a SPAN ontological view or to directly mark the preceding property as endurant. I have shown that DE-insertion not only applies to all D-M compounds involving measure words (which denotes perdurant properties), but also to specific sortal classifier constructions where a time-variant meaning is coerced. I have also given additional examples to show that the marking of perdurant properties/SPAN ontological view may be a semantic feature of many *de*-constructions in Chinese. Taking this into consideration in addition to the intuitive nouny/verby categorical dichotomy, I claim that ontological notions do find linguistic expression in Chinese, similar to what Chao (1955) found when looking for the linguistic expression of logical relations in Chinese.

In addition to potential extension of a unified conceptual account of 的 *de* 'DE' in Chinese, our study of the expression of endurant/perdurant ontological dichotomy in Chinese has implications for future studies on the relation between ontology and language as knowledge systems. For instance, event and kinds as endurant individuals are not specified in the current version of upper ontology of BFO, DOLCE (as well as many other competing ontologies), and it remains open for further research to determine if the evidence from Chinese classifiers requires the addition of such nodes. Moreover, as the classifier system involves both description and measurement of different qualities, a full explanatory account of the system must address the interaction between entity and quality. For example, further work needs to be done to determine if such qualities are better treated independently of entities (i.e. the DOCLE approach) or as dependent on entities in order to allow shift of ontological views (i.e. the BFO approach). This is a fundamental ontological decision and I hope that further exploration of the linguistic expressions of ontology in Chinese will shed light on this important issue.

Last but not the least, as mentioned earlier, the standard position of current studies on ontology is that the formal ontology is the rigorous and logically robust system which is the shared foundation of knowledge representation through either domain specific (and potentially inconsistent) local ontologies, as well as less rigorous and potentially conflicting language specific ontologies. However, as I have shown that a linguistic system such as Chinese can encode (and manipulate) basic ontological concepts, the notion of a formal ontological system existing *a priori* and independent of language usages needs to be challenged, as the results herein demonstrate that ontological notions can be verified by their expressions in linguistic systems. It also suggests that manipulations of linguistic expressions of ontological notions may reflect how ontological notions evolve.

5 Endnotes

[a]The actual design of BFO allows ontological dichotomy as well as reduction to either type of ontologies: three-dimensional SNAP ontologies without temporal dimension; versus four-dimensional SPAN ontologies incorporating spatiotemporal information Grenon and Smith (2004).

[b]Although Chinese classifier system does involve quality and our data poses interesting challenges to different ontological systems, it is beyond the scope of this paper to resolve this issue and we will simply note possible implications without attempting a full ontological account of Quality.

[c]In context, a reading of 'bought three things' referring to three separate objects is also possible, provided that these three objects belong to three separate types.

[d]This is an example where two alternative ontological views on how quality should be treated may lead to different accounts and predictions. If we take BFO's approach, which has Quality as part of a SNAP ontology, an intuitive account would be that the price specification is simply a Quality associated with a continuant/endurant. I.e. the kind classifier system allows additional quality description (such as the published price) of an enduring entity. The DOLCE view where Quality and Quantity are ontologically independent will require a more elaborate system to account for why one quality is considered endurant while the other is not.

References
Ahrens, Kathleen, and Chu-Ren Huang. 2016. Classifiers. In A reference grammar of Chinese, ed. Chu-Ren Huang and Dingxu Shi. Cambridge: Cambridge University Press.
Ahrens, Kathleen. 1994. Classifier production in normals and aphasics. Journal of Chinese Linguistics 22(2): 203–248.
Aikhenvald, Alexandra Y. 2003. Classifiers: A typology of noun categorization devices. Oxford: Oxford University Press.
Berners-Lee, Tim, James Hendler, and Ora Lassila. 2001. The semantic web. Scientific American 284(5): 28–37.
Bond, Francis, Christiane Fellbaum, Shu-Kai Hsieh, Chu-Ren Huang, Adam Pease, and Piek Vossen. 2014. A multilingual lexico-semantic database and ontology. In Towards the multilingual semantic web: Principles, methods, and application, ed. Paul Buitelaar and Philippe Cimiano, 243–258. Berlin Heidelberg: Springer-Verlag.
Chao, Yuen Ren. 1955. Notes on Chinese grammar and logic. Philosophy East and West 5(1): 31–41.
Chao, Yuen Ren. 1968. A grammar of spoken Chinese. Berkeley: University of California Press.
Chen, Keh-Jiann, Chu-Ren Huang, Li-Ping Chang, and Hui-Li Hsu. 1996. Sinica Corpus: Design methodology for balanced corpora. In Proceeding of the 11th Pacific Asia Conference on Language, Information and Computation, ed. Byung-Soo Park and Jong-Bok Kim, 167–176. Seoul: Kyung Hee University.
Chou, Chia-Ju, Hsu-Wen Huang, Chia-Lin Lee, and Chia-Ying Lee. 2014. Effects of semantic constraint and cloze probability on Chinese classifier-noun agreement. Journal of Neurolinguistics 31: 42–54.
Chou, Ya-Min, and Chu-Ren Huang. 2010. Hantology: Conceptual system discovery based on orthographic convention. In Ontology and the lexicon: A natural language processing perspective, ed. Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot, 122–143. Cambridge: Cambridge University Press.
Chou, Ya-Min, and Chu-Ren Huang 周亞民，黃居仁. 2006. Computational representation of character and lexical knowledge in Chinese–A historical change perspective 漢語文字和詞彙知識在計算機的表達一歷史變遷的觀點. In Mountain lofty, river long: Festschrift in honor of Professor Pang-hsin Ting on his seventieth birthday 山高水長：丁邦新先生七秩壽慶論文集, ed. Dah-an Ho, H. Samuel Cheung, Wuyun Pan, and Fuxiang Wu 何大安，張洪年，潘悟雲，吳福祥, 595–611, special issue, Language and Linguistics. Taipei: Academia Sinica.
Chierchia, Gennaro. 1984. Topics in the syntax and semantics of infinitives and gerunds. Ph.D. dissertation. University of Massachusetts.
Fillmore, Charles J, Daniel Kempler, and William S-Y Wang (eds). 1979. Individual differences in language ability and language behavior. Waltham, MA: Academic Press.

Gangemi, Aldo, Nicola Guarino, Claudio Masolo, and Alessandro Oltramari. 2003. Sweetening ontologies with DOLCE. AI Magazine 24(3): 13–24.

Gangemi, Aldo, Nicola Guarino, Claudio Masolo, and Alessandro Oltramari. 2010. Interfacing WordNet with DOLCE: Towards OntoWordNet. In Ontology and the lexicon: A natural language processing perspective, ed. Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot, 36–52. Cambridge: Cambridge University Press.

Grenon, Pierre, and Barry Smith. 2004. SNAP and SPAN: Towards dynamic spatial ontology. Spatial Cognition and Computation 4(1): 69–103.

Gruber, Tom. 1995. Toward principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies 43(5–6): 907–928.

Guarino, Nocola. 1998. Some ontological principles for designing upper-level lexical resources. In Proceedings of the First International Conference on Language Resources and Evaluation, Granada, 28–30 May 1998, ed. Antonio Rubio, Natividad Gallardo, Rosa Castro, and Antonio Tejada, 527–534. Paris: ELRA.

Guarino, Nocola, and Christopher Welty. 2002. Evaluating ontological decisions with ONTOCLEAN. Communications of the ACM 45(2): 61–65.

Her, One-Soon, and Chen-Tien Hsieh. 2010. On the semantic distinction between classifiers and measure words in Chinese. Language and Linguistics 11(3): 527–551.

Huang, Chu-Ren, and Dingxu Shi (eds.). 2016. A reference grammar of Chinese. Cambridge: Cambridge University Press.

Huang, Chu-Ren, and Ya-Min Chou. 2015. Multilingual conceptual access to lexicon based on shared orthography: An ontology-driven study of Chinese and Japanese. In Language production, cognition and the lexicon, ed. Núria Gala, Reinhard Rapp, and Gemma Bel-Enguix, 135–150. Berlin Heidelberg: Springer-Verlag.

Huang, Chu-Ren 黄居仁. 2013. On a functional uniformity of de 关于 "的" 的功能一致性. In Towards the contemporary cutting-edge science of modern Chinese grammar 走向当代前沿科学的现代汉语语法研究, ed. Yang Shen 沈阳, 129–135. Beijing: The Commercial Press.

Huang, Chu-Ren, Jia-fei Hong, Sheng-yi Chen, and Ya-Min Chou 黄居仁, 洪嘉馡, 陈圣怡, 周亚民. 2013a. Exploring event structures in Hanzi radicals: An ontology-based approach 汉字所表达的知识系统：意符为基本概念导向的事件结构. Contemporary Linguistics 当代语言学 2013(3): 294–311.

Huang, Chu-Ren, Ya-Jun Yang, and Sheng-Yi Chen. 2013b. Radicals as ontologies: Concept derivation and knowledge representation of four-hoofed mammals as semantic symbols. In Breaking down the barriers: interdisciplinary studies in Chinese linguistics and beyond. A Festschrift for Professor Alain Peyraube, ed. Hilary Chappell, Redouane Djamouri, and Thekla Wiebusch, 1117–1133. Taipei: Academia Sinica.

Huang, Chu-Ren, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari and Laurent Prévot (eds.). 2010a. Ontology and the lexicon: A natural language processing perspective. Cambridge studies in natural language processing. Cambridge: Cambridge University Press.

Huang, Chu-Ren, Ru-Yng Chang, and Shiang-bin Li. 2010b. Sinica BOW: Integration of bilingual WordNet and SUMO. In Ontology and the lexicon: A natural language processing perspective, ed. Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot, 201–211. Cambridge: Cambridge University Press.

Huang, Chu-Ren, and Kathleen Ahrens. 2003. Individuals, kinds and events: Classifier coercion of nouns. Language Sciences 25(4): 353–373.

Huang, Chu-Ren, Kathleen Ahrens, Li-Li Chang, Keh-Jiann Chen, Mei-Chun Liu, and Mei-Chih Tsai. 2000. The module-attribute representation of verbal semantics: From semantics to argument structure. In Chinese Verbal Semantics, ed. Yung O Biq, special issue, Computational Linguistics and Chinese Language Processing, 5(1): 19-46.

Huang, Chu-Ren, Keh-Jiann Chen, and Zhao-ming Gao. 1998. Noun class extraction from a corpus-based collocation dictionary: An integration of computational and qualitative approaches. In Quantitative and computational studies on the Chinese language, ed. Benjamin K T'sou, Tom BY Lai, Samuel WK Chan, and William S-Y Wang, 339–352. Hong Kong: City University of Hong Kong.

Huang, Chu-Ren, Keh-Jiann Chen and Ching-hsiung Lai 黄居仁，陳克健，賴慶雄 1997. Mandarin Daily dictionary of Chinese classifiers 國語日報量詞典. Taipei: Mandarin Daily Press.

Huang, Chu-Ren. 1987. Mandarin Chinese NP de. A comparative study of current grammatical theories. Ph.D. dissertation. Ithca, NY: Cornell University.

Li, Charles N, and Sandra A Thompson. 1981. Mandarin Chinese: A functional reference grammar. Berkeley: University of California Press.

Mo, Ruo-Ping Jean, Yao-Jung Yang, Keh-Jiann Chen, and Chu-Ren Huang. 1996. Determinative-measure compounds in Mandarin Chinese: Formation rules and parser implementation. In Readings in Chinese Natural Language Processing, Journal of Chinese Linguistics Monograph Series No. 9, ed. Chu-Ren Huang et al., 123–146. Berkeley: Journal of Chinese Linguistics.

Niles, Ian, and Adam Pease. 2001. Towards a standard upper ontology. In Proceedings of the International Conference on Formal Ontology in Information Systems, 2–9.

Pease, Adam, and Christiane Fellbaum. 2010. Formal ontology as interlingua: The SUMO and WordNet linking project and GlobalWordNet. In Ontology and the lexicon: A natural language processing perspective, ed. Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot, 25–35. Cambridge: Cambridge University Press.

Prévot, Laurent, Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, and Alessandro Oltramari. 2010. Ontology and the lexicon: A multi-disciplinary perspective. In Ontology and the lexicon: A natural language processing perspective, ed. Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot, 3–24. Cambridge: Cambridge University Press.

Pustejovsky, James. 1995. The generative lexicon. Cambridge: MIT Press.

Smith, Barry. 2012. On classifying material entities in basic formal ontology. In Proceedings of the Third Interdisciplinary Ontology Meeting, 1–13. Tokyo: Keio University Press.

Smith, Barry, and Pierre Grenon. 2004. The cornucopia of formal-ontological relations. Dialectica 58(3): 279–296.

Shi, Dingxu. 2016. Noun phrases. In A reference grammar of Chinese, ed. Chu-Ren Huang and Dingxu Shi. Cambridge: Cambridge University Press.

Tai, James H-Y. 1994. Chinese classifier systems and human categorization. In In honor of William S.-Y. Wang: Interdisciplinary studies on language and language change, ed. Matthew Y Chen and Ovid JL Tzeng, 479–494. Taipei: Pyramid Press.

T'sou, Benjamin K. 1976. The structure of nominal classifier systems. In Austoasiatic Studies, vol. 2, ed. Philip N Jenner, Stanley Starosta, and Laurence C Thompson, 1215–1247. Honolulu: University Press of Hawaii.

Wang, Ruijing, and Caicai Zhang. 2014. Effect of classifier system on object similarity judgment: A cross-linguistic study. Journal of Chinese Linguistics 42: 188–217.

Wiebusch, Thekla. 1995. Quantification and qualification: Two competing functions of numeral classifiers in the light of the radical system of the Chinese script. Journal of Chinese Linguistics 23: 1–41.

Zhu, Dexi 朱德熙. 1961. On de 说 "的". Studies of the Chinese Language 中国语文 12: 1–15.