



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Exemplar-based compounds: The case of Chinese

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Arcodia G.F., Mauri C. (2020). Exemplar-based compounds: The case of Chinese. *LANGUAGE SCIENCES*, 81, 1-21 [10.1016/j.langsci.2019.06.002].

Availability:

This version is available at: <https://hdl.handle.net/11585/772601> since: 2020-09-24

Published:

DOI: <http://doi.org/10.1016/j.langsci.2019.06.002>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Arcodia, Giorgio Francesco, and Caterina Mauri. "Exemplar-based compounds: The case of Chinese." *Language Sciences* 81 (2020): 101232.

The final published version is available online at:

<https://doi.org/10.1016/j.langsci.2019.06.002>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

Exemplar-based compounds: the case of Chinese*

Giorgio Francesco Arcodia (*University of Milan – Bicocca*), Caterina Mauri (*University of Bologna*)

ABSTRACT

The aim of this paper is to investigate a specific naming strategy, which is based on compounding and exemplification, examining data from Chinese. We will focus on what we will label ‘exemplar-based compounds’, i.e. compounds consisting of at least one lexeme denoting an exemplar of the category referred to by the whole compound. We propose that ‘exemplar-based’ compounds in Chinese be divided into two macro-types: (1) [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY}, in which the exemplars may or may not exhaustively list the members of the category denoted by the compound (e.g. *dāoqiāng* ‘sword-spear, sword and spear > ‘swords, spears and similar things = weapons’); (2) [EXEMPLAR-CLASS]_{CATEGORY}, in which the first constituent exemplifies the class denoted by the second one; this type includes compounds in which the second constituent is a classifier (e.g. *niǎozhī* ‘bird’, *chuánzhī* ‘ship’, with *zhī* ‘CL’). After a detailed discussion of exemplar-driven category naming and of compounding and classifiers in Chinese, we will present the results of a corpus-based study, based on data of Premodern and Modern Chinese. We will show how the exemplar-driven abstraction characterising these constructions evolved into systematic reference to a category and to its individual items, revealing a change from a procedural category construction to a naming concept label.

Keywords: co-compounding; classifiers; Chinese; ad hoc categories; naming; exemplification

1. Introduction

The aim of this paper is to investigate the development of a specific naming function of NN compounds in Chinese,¹ consisting of at least one lexeme denoting an exemplar of the category referred to by the whole compound. As a cover term for these compounds, we propose the label ‘exemplar-based compounds’, to convey the fact that they work as procedural concept labels, that is, as labels denoting a given category by evoking the process of category construction itself, which in this case is a bottom-up process of exemplar-driven abstraction.

The use of exemplification to denote categories is typical of what has been called indexical categorization (Mauri, 2017; Mauri and Sansò, 2018), which is characterized by a bottom-up, context-dependent abstraction process that, starting from the mention of one or more exemplars, leads to the identification of a context-relevant category (see also Mauri and Sansò, this issue). This way of communicating category construction in discourse is frequently, though not necessarily, employed to indicate what Barsalou (1983) terms “*ad hoc* categories”, namely goal-driven categories that are not stored in long-term memory and depend on context both for their construction and for their interpretation.

* This article is the result of a continuous collaboration between the two authors. For the purposes of Italian academia, the two authors are responsible for writing Section 1, Caterina Mauri is responsible for Sections 1, 2.1, 4, and Giorgio F. Arcodia for Sections 2.2 and 3. This research was developed within the SIR project “LEADhoC: Linguistic expression of ad hoc categories”, coordinated by Caterina Mauri (University of Bologna; prot. RBSI14IIG0).

¹ The *Pinyin* romanisation system and traditional Chinese characters have been used as a default throughout the article. The glosses follow the general guidelines of the Leipzig Glossing Rules.

Exemplar-based compounds can be ascribed to the set of linguistic strategies that convey indexical categorization, both because they may refer to context-dependent sets, and because they may involve classifiers/class nouns denoting heterogeneous, non-predictable, ad hoc categories (Huang and Chen, 2011).

Based on corpus data from Premodern and Modern Chinese, we will analyse how the naming function of exemplar-based compounds emerged through the conventionalization of listing patterns, originally enumerating relevant exemplars of the denoted category, and of classifying strategies, involving the use of class nouns and classifiers. We will especially focus on the peculiarities of a strategy which employs a procedural mechanism such as exemplification to convey a non-procedural, labelling function such as naming context-dependent categories.

In Section 2 we will provide the theoretical background of this study, starting with a discussion on exemplification and on its role in the communication and naming of (ad hoc) categories (Section 2.1). Section 2.2 will sketch the general picture of Chinese compounds and classifiers, identifying and defining the two types of compounds that constitute the object of this analysis. Adopting a cognitive perspective, we will subsume under the label ‘exemplar-based compounds’ words that have been classified in the literature as belonging to different types: additive and collective co-compounds, synonymic co-compounds, and attributive compounds (Wälchli, 2005). We propose that exemplar-based compounds be divided into two macro-types, as represented in (1) and (2) below.

(1) [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY}

In type (1), the exemplars may or may not exhaustively list the members of the category denoted by the compound. The compound itself may therefore acquire the meaning of ‘EXEMPLAR₁-EXEMPLAR₂, and similar things’ (as e.g. 刀槍 *dāoqiāng* ‘sword-spear, sword and spear > ‘swords, spears and similar things = weapons’). This includes compounds usually classified as ‘synonymic’ (as e.g. 朋友 *péngyou* ‘friend(pupil of the same master)-friend(person who shares one’s aspirations), friends’).

(2) [EXEMPLAR-CLASS]_{CATEGORY}

In type (2), the first constituent exemplifies the class denoted by the second one. This type includes compounds where the second constituent is a classifier (Loke, 1997; Zhang, 2013). Let us consider the case of compounds including the classifier 隻 *zhī*, originally meaning ‘a bird’, now a classifier for birds, small animals, boats (and others): 鳥隻 *niǎozhī* ‘bird’, 船隻 *chuánzhī* ‘ship’. In these cases, the function of the second constituent is in-between a class noun and a classifier: it indicates a broad category of entities, of which the first constituent is a relevant exemplar, so that the meaning of the compound identifies a subclass of the second constituent. Our data show that, especially in their early stages, type (1) and, to a lesser extent, type (2) compounds have a strong preference for generic reference to categories, rather than for countable entities (i.e. typically, they are not quantified).

After a synchronic description of these compounds, we will turn to the diachronic perspective (Section 3), in order to answer the question of how the exemplar-driven abstraction characterising these constructions evolved into systematic (more or less lexicalised/conventionalised) reference to a category (i.e. the one denoted by the compound). To this aim, we will consider data collected from corpora of Premodern Chinese (ctext.org, *Academia Sinica Tagged Corpus of Early Mandarin Chinese*) and Modern Chinese (*Academia Sinica Balanced Corpus of Modern Chinese*); we will also rely on Google searches to test the conclusions drawn from the (Modern) corpus data.

In Section 4 we will conclude by providing a unitary account for the observed patterns, which will lead us to consider exemplar-based compounds as procedural labels, at least until they retain some semantic transparency. We will show how the examination of the referential properties of these compounds may reveal their degree of conventionalization, with less conventionalized compounds being limited to generic reference to categories and more conventionalized compounds being employed also for specific reference to (possibly quantified) category items.

2. Theoretical foundations: category naming and Chinese word formation

Word formation, mainly compounding and derivation, is known to be a device for the creation of names for entities, while syntax can be analysed as a device to describe those entities and say something about them (Bauer, 2003: 135; Schlücker and Hüning, 2009). In the following sections we will first address the naming function, especially with respect to the creation of category labels, highlighting the role of compounding in the light of the recent literature on naming, and addressing the role of context (Section 2.1.1). We will then turn to exemplification, the main discourse strategy for inductive category construction, examining the mechanisms underlying the bottom-up abstraction that it triggers. We will be able to identify and define exemplar-driven naming strategies, that is, strategies employing exemplification to provide a category label (Section 2.1.2).

In section 2.2 we will focus on exemplar-driven naming strategies in Chinese. After a brief overview on the rise of coordinating compounds (Section 2.2.1), we will describe compounds involving an exemplar followed by a classifier or a class noun, providing a structural and semantic account in the light of the general system of classifiers in Chinese (Section 2.2.2). The picture sketched will highlight differences and similarities between the two types of exemplar-based compounds that constitute the core of this study (Section 2.2.3).

2.1 Exemplar-driven naming strategies

2.1.1 Naming concepts and category labels

As clearly stated by Booij (2009: 219), the linguistic expressions *par excellence* for the naming function are words, or better lexical units, whose main function is to provide names for concepts. Booij recalls the definition of naming provided by Koefoed (1993: 3), according to which ‘naming is creating a link between an expression and a concept’ and the expression is frequently a single word, but may also consist of a wider lexical unit or a multi-word expression.

The function of naming is opposed to that of description (Bauer, 2003: 135), although there is no neat division between the linguistic expressions performing the two, and both are associated in discourse to the general function of introducing or identifying a given object or discourse referent. According to Bhat (1994: 114), conventionalization is the key feature distinguishing between naming and describing strategies, whereby naming is characterized by a conventional relation between the expression and the object referred to, while in description such relation is not conventional, but is creatively built through similarity – i.e. the expression is internally composed in such a way to denote properties that can be observed in the referred object.

As extensively shown by Schlücker and Hüning (2009) and by the papers collected in the special issue of *Word Structure* they edited, naming strategies include a great variety of constructions, ranging from word formation to phrasal expressions, including root and derivative lexemes, compounds, and phrases, all showing some degree of conventionalization (cf. Wray, 2002). According to De Caluwe (1990), naming may also be performed through

semantic extensions, acronyms, borrowing phenomena, and clipping from phrases. The relation between morphological and syntactic naming constructions is accounted for by Booij (2009) and Masini (2009) within the framework of Construction Grammar, which allows to identify the commonalities underlying the various types of naming strategies and explain the non-compositionality of fixed naming expressions (cf. Wray, 2002). All these theories share the idea that naming expressions are stored in the Mental Lexicon of native speakers and each name is employed as a whole, referring to a specific concept.

Compounds are one of the most widespread word formation strategies that provide names for entities, properties and actions, together with derived words. The classificatory function of compounding has been widely acknowledged in the literature, having been noted already by Jespersen (1942), and has been the object of a considerable amount of research covering both compounds stored in the lexicon and nonce compounds, i.e. so-called deictic compounds created on the fly for specific purposes (cf. Downing's well known example *apple-juice seat*, Downing, 1977). According to Schlücker and Hüning (2009: 149), nonce compounds do not always have a naming function, but can be used in discourse as descriptions. This can be clearly observed in newspaper headlines (e.g. Dutch *het Cruyff-interview* 'the Cruyff-interview', Booij 2009: 220), which constitute one of the stylistic domains where we can observe a high degree of naming innovation and creativity (see Dardano, 2009: 228 on the diffusion of Italian NN compounds).

The literature on naming has mainly focused on the identification of the semantic and structural features characterizing fixed naming expressions, i.e. expressions that are stored in the Mental Lexicon in association to specific concepts (Sprenger, 2003: 4). Yet, within more cognitive approaches, the presence of a naming strategy does not imply a truly stable connection with a specific concept.

Wilson and Carston (2007; see also Carston, 2010) developed a theory of lexical pragmatics within Relevance Theory, according to which the meaning of words is always adjusted to context, so that their contribution to the proposition is different from their lexically encoded sense. According to them, semantic processes of narrowing and broadening lead speakers to interpret words as referring to 'ad hoc concepts', namely context-dependent concepts that are narrower or broader than the encoded word meaning. The typical example they provide is the interpretation of *drink* in the sentence *he used to drink too much*: the concept associated to *drink* is narrower than lexically encoded one and refers to 'drink alcohol', rather than the generic act of drinking. The sentence *I need a Kleenex* provides an instance of broadening, where the concept associated to *Kleenex* in this context is broader than the one stored in the lexicon, and refers to a handkerchief, independently of the brand. In Wilson and Carston's account, words (and in general lexical units) always convey ad hoc concepts, which depend on the specific speech situation. As a consequence, every abstract category conveyed by a lexical item is translated into a concrete, ad hoc category that is anchored in the situational context.

A similar perspective can be found in the work by Lakoff and Sweetser (1994), and Croft and Cruse (2004). Croft and Cruse (2004: 92) argue in favor of a dynamic construal of lexical meaning, because "neither meanings nor structural relations are specified in the lexicon, but are construed 'on-line,' in actual situations of use" (Croft and Cruse, 2004: 97-98). According to them, the linguistic properties of words and phrases do not act as naming labels, but rather as clues guiding speakers towards the intended concepts (Mauri and Sansò, 2018: 4). In this view, lexical semantics can be seen as having the same status as non-linguistic knowledge, contextual information, and any relevant information that may be retrieved in memory.

The notion of ad hoc concepts introduced within Relevance Theory and the notion of construal elaborated by Croft and Cruse (2004) are crucially connected to the naming function of lexical units, because they address the stability of the relation between an expression and the object it refers to. Both approaches underline the role of context in the interpretation of naming,

calling into question the process of category construction and the stability vs. context-dependency of categories. It is not a coincidence that both Wilson and Carston and Croft and Cruse recall the results of the psychologist Barsalou (1983), who conducted several experiments showing the existence of context-dependent, ad hoc categories.

According to Barsalou (2010: 86), ad hoc categories are novel categories constructed on the fly to achieve specific discourse goals. For instance, the category [THINGS TO BUY ON A WINTER SATURDAY MORNING] may be relevant in a specific situation, such as planning a free morning in January, so as to make a number of purchases left pending for some time. Ad hoc categories are dependent on context, are not stored in long-term memory, and once the goal is achieved, they are dismissed. Contrary to common categories, such as [DOG] or [BOOK], which are stored in long-term memory, ad hoc categories are inherently volatile. Recent linguistic research on this topic shows that they are very frequently communicated in discourse, and languages display a broad range of linguistic constructions for their expression (cf. Mauri and Sansò, 2018, Barotto and Mauri, 2018).

Interestingly, Barsalou himself addresses the issue of how ad hoc categories are referred to and named. He notes that, while stable categories are typically expressed by fairly short conventional linguistic means (typically words or lexical units), ad hoc categories do not come with ready-made linguistic labels, but rather tend to be described by means of complex expressions, involving relative clauses or even lists (e.g., *clothing to wear while house painting and cleaning*, etc.). Basically, although he does not use the distinction between naming and description, he suggests that while stable categories can be named through lexical labels, ad hoc categories are more likely to be described by syntactic means.

To sum up, we observe, on the one hand, a literature on naming which is mainly concerned with the distinction between naming and description, as a heuristic tool to distinguish between lexical and syntactic strategies. On the other hand, there are more cognitive approaches which question the idea that naming is a function performed by accessing the Mental Lexicon and raise doubts on the very idea that some stability between a given expression and a specific concept may exist at all. These approaches assign great importance to context and rely on a context-dependent theory of categorization, where word meaning is analyzed not as a concept name, but as a clue towards a context-dependent construal of category. Finally, there is psychological evidence for different types of categories, with more stable categories being expressed by fairly typical naming strategies (e.g. lexical units, words, fixed phrases) and more volatile, ad hoc categories being referred to through syntactic, more descriptive-like strategies.

Should we conclude that naming is only possible for stable categories, and even for those categories it does not act as labeling, but rather as a clue? According to Barotto's (2017) corpus-based study of Japanese, ad hoc categories are very frequently referred to through some category label, which is typically preceded or followed by exemplification, in a pattern that alternates naming and description. Barotto and Mauri (2018) look at discourse data from Italian and show that speakers tend to be somehow redundant, when they refer to categories in discourse, typically recurring to lexical labels (compounds or lexical units), reformulations and exemplification, in a 'camel hump' pattern of top-down and bottom-up referential strategies. In other words, for ad hoc categories we do observe the use of labels, but these are typically enriched by descriptions, such as relative clauses or lists of examples, aimed at refining the category borders and anchoring the process of category construction to the context.

Barotto (2017: 49 ff.) provides a detailed discussion and typology of the semantic and syntactic types of labels employed for ad hoc categories. She discusses examples such as (3):

(3) *Relaxing drinks such as water, herbal teas, smoothies and such* (Barotto, 2017: 51)

In her analysis, *relaxing drinks* is used to name the category through a label, which is then further clarified by a list of exemplars, so that the hearer is guided to abstract a set of relaxing drinks around the one explicitly mentioned, ending up with [HEALTHY RELAXING DRINKS]: although alcoholic drinks may have relax as one of their possible effects, the ad hoc category conveyed in (3) is likely to exclude them. The concept denoted by the category label thus cannot be argued to coincide with the ad hoc category communicated by the speaker, which can only be inferred through the integration of a naming strategy followed by a descriptive one.

Cognitive approaches and the analysis of ad hoc categorization aim the lights on the fact that naming is a top-down process through which an abstract label is used to identify some given item(s), but this top-down process is successful to the extent to which it is interpreted in the light of the specific context of utterance. This anchoring process is frequently reinforced in discourse by complementing the top-down activity of naming with a bottom-up process of exemplification, and this is especially true for ad hoc categories. Specific items are indeed used as arrows towards the abstraction of some higher-level category, for which a conventionalized naming construction may not be easily available.

While the use of a label clearly fulfils a naming function, exemplification is to be ascribed to description. However, like just about any clear-cut border, there are cracks in this distinction: what about naming strategies based on exemplification, such as the compounding phenomena that are the object of this research? In the next section, we will first describe the mechanisms underlying an exemplar-driven process of abstraction, and will then argue that, when this process is triggered by a word formation strategy like derivation or compounding, we observe a descriptive process at the service of a naming function.

2.1.2 Bottom-up, exemplar-driven naming

The connection between naming and exemplification is more direct than it may seem. We defined naming as a stable relation between an expression and a concept, but, as Birk (2006: 5) points out, ‘the domestic life of a concept is a series of examples’.

Exemplification has been an object of research in many fields, ranging from rhetorical studies to cognitive experiments (see Barotto, 2017: 7-21 for a comprehensive overview of different approaches to exemplification). Exemplification as a linguistic phenomenon and its role in the construction of categories has been mainly analysed within research on vagueness (see Channel, 1994, among others) and ad hoc categorization (see Barotto, 2018 on Japanese and Lo Baido, 2018 on Italian). Based on a wide cross-linguistic analysis, Mauri and Sansò (2018) identify exemplification as the key mechanism underlying the linguistic communication of indexical categorization.

Mauri and Sansò defined indexical categorization as a bottom-up process through which, starting from one or more exemplars, speakers rely on frames and shared knowledge to abstract a context-relevant category. The term ‘indexical’ refers to the high-context dependency of this process and to the use of linguistic expressions having an indexical semantics, which implies reference to ‘further items sharing a context-relevant property P’ (see Barotto and Mauri, 2018). The identification of a value for P allows to abstract the relevant category. Let us consider example (4) from the *EnTenTen15* Corpus of English:

(4) *Jon is there any way these can be passed along using ['drop box' or email or something similar]?*

The utterer of (4) uses of a list of examples to refer to the higher-level category [WAY OF TRANSFERRING FILES BETWEEN COMPUTERS].² The label for such a concept is not conventionalized, but necessarily relies on description, leading to an abstract formulation that, despite being precise, does not prevent from possible ambiguities. The speaker thus chooses to construe the category on-line, rather than naming it, and follows a bottom-up process of indexical categorization: through a list of the most accessible and salient exemplars of the intended concept, the hearer is led to abstract the relevant property ('tools for transferring files') and build a larger set, including the explicit exemplars, together with possible additional members (e.g. 'WeTransfer').

Indexical categorization can be described as bottom-up process of construing and communicating a category in a context-dependent way. As shown by Mauri and Sansò (2018), the set of constructions triggering an indexical categorization process includes associative and simulative plurals, reduplication, non-exhaustive listing (by means of non-exhaustive connectives and general extenders), and also word formation strategies, such as collective derivatives and compounds.

Just as the distinction between naming and description cannot be drawn uniquely on the basis of structural parameters (Booij, 2009) and we find both compounds and phrases for both naming and description, a complex border can be described also between category labelling and category construction. In particular, we argue that there are strategies in which a process of indexical categorization through exemplification becomes a label for the category itself, i.e. an originally descriptive, procedural, exemplar-based linguistic strategy is employed with a naming function. Two types of constructions, among those described by Mauri and Sansò, can be analysed as exemplar-based names: certain kinds of derivate collectives and compounding. In these strategies, exemplification is not only the route towards the identification of the indexical category, but it also becomes an end point, i.e. the label.

Let us consider derivatives first. Italian shows two collective suffixes (-*ame* and -*aglia*) that can be attached to common nouns to derive collectives. As argued by Mauri (2017) and Magni (2018), these collectives can be suffixed to proper nouns of celebrities or politicians, to denote the set of persons, ways of acting, and situations whose pivot and prototypical exemplar is denoted by the proper noun. Example (5) shows the use of the suffix -*aglia* attached to the noun *destra* 'right-wing', to derive the noun *destraglia* 'set of right-wingish persons', and to the proper name Salvini, the right-wing Italian Deputy Prime Minister. The collective noun *salvinaglia* is coined to refer to a set of politicians characterized by a context-relevant way of acting, of which Salvini is a prototypical exemplar (something like 'Salvinish politicians').

(5) *A voi le parole di padre Alberto Maggi, invisibile naturalmente dalla vasta e rabbiosa destraglia e salvinaglia varia diffusa in ogni dove, pronta a "bastonare" in qualsiasi modo*

'To you the words of Father Alberto Maggi, naturally unpopular for the vast and angry *destraglia* and *salvinaglia* diffused everywhere, ready to "belabour" in any possible way'

We can call this type of derivatives 'exemplar-based', to the extent that they involve an exemplar-driven, context-dependent process of abstraction towards the identification of the relevant property defining the named category. *Salvinaglia* is an exemplar-based derivative noun used to name a specific category of politicians, namely the category that includes Salvini as a prototypical and highly accessible exemplar. The great visibility of Salvini in Italian politics makes him an efficient and accessible exemplar, to be used as a starting point to abstract

² See Ariel and Mauri (2018: 19-26) for a discussion on the various readings of English *or*, based on a corpus-based study of spoken American English. Ariel and Mauri show that, of the six main readings identified for *or*, the most frequent one is referring to a higher-level category.

the context-relevant properties defining a specific class of politicians. This class is characterized by conservative right-wing views, xenophobic ideas and an aggressive style of public communication, properties that are exemplified in a prototypical way by Salvini. Access to the specific context, namely recent Italian politics, is necessary to correctly use and interpret the derivative in *-aglia*, and thus to identify the concept it names. In other words, an exemplar-driven, context-dependent abstraction process constitutes the semantic core of a naming strategy. Exemplification becomes labeling, and the bottom-up process of category construction is exploited to fulfill a naming function.

A similar mechanism can be observed in compounding phenomena, such as the ones that will be analyzed in this paper, where at least one of the lexemes forming the compound can be analyzed as an exemplar of the category denoted by the compound as a whole. Typical instances of exemplar-based compounds are what Arcodia *et al.* (2010) call hypernymic coordinating compounds ('co-compounds' in Wälchli's 2005 terminology), where the referent of the compound is in a superordinate relationship to the meaning of the parts (as Mandarin *dao-qiang* lit. 'sword-spear', meaning 'weapons', or Arapesh *buwul nubat*, lit. 'pig-dog', meaning 'domestic animals'). They oppose these compounds to hyponymic compounds, such as Spanish *lanza-espada* 'spear - sword, a spear with a blade, a spear which is a sword at the same time' (Arcodia *et al.*, 2010: 178), where the compound refers to an item which is a hyponym of at least one of the lexemes forming the compound, i.e. a specific type of spear.

According to Arcodia *et al.*'s analysis (see also Arcodia 2018), hypernymic compounds are especially attested in East Asia, South East Asia and New Guinea, while hyponymic ones are frequent in languages belonging to the Standard Average European area. As said earlier, in this paper we will focus on compounding phenomena attested in Mandarin Chinese, a language showing a high number of hypernymic compounds, where compounding is the most frequent and productive word formation strategy. Moreover, since Mandarin Chinese has a very long documented written tradition, we will undertake a diachronic study aimed at understanding the mechanisms underlying the development of exemplar-driven compounds.

As said in the introduction (Section 1), we will consider two types of exemplar-based compounds: [EXEMPLAR1-EXEMPLAR2]_{CATEGORY} compounds and [EXEMPLAR-CLASS]_{CATEGORY} compounds. The former can be ascribed to the hypernymic class as defined by Arcodia *et al.* (2010): the category is abstracted by comparing the two exemplars and identifying the property they have in common, along a bottom-up process of indexical categorization (e.g. Vietnamese *cha mẹ* lit. 'father mother', meaning 'parents'). The latter involve a class noun or classifier as the second element, and look like hyponymic compounds (but see the discussion below, Section 2.2): the category referred to by the compound is a subset of the class denoted by the classifier/class noun, and the identification of the relevant subset is made possible by anchoring it to the first element, which acts as a prototypical exemplar of the category. Let us consider 紙張 *zhǐ-zhāng* lit. 'paper-CL' meaning 'paper': the category denoted by the compound, 'paper', constitutes a subset of the entities which fall under the scope of the classifier 張 *zhāng* 'classifier for flat objects'; in, particular a subset exemplified by 紙 *zhǐ* 'paper'. We may represent the exemplification process internal to [EXEMPLAR-CLASS]_{CATEGORY} compounds as a top-down and bottom-up process of indexical categorization, because the meaning of the compound is 'a subset of CLASS which is exemplified by EXEMPLAR', so that the relevant subset of CLASS denoted by the compound can be abstracted from EXEMPLAR.

Exemplar-based compounds have a clear naming function and are in many cases non-compositional, i.e. they are conventionalized, and stored in the lexicon. Yet, they involve an exemplification process and imply a bottom-up category abstraction. Once again, we observe an exemplification process which serves a naming function, with an indexical category construction that becomes the category label itself. To what extent is the exemplification process

still accessible? How did the naming function develop? Are there traces of the exemplification process in the semantic properties of exemplar-based compounds? In the next section, the two types of compounds will be described in detail against the general mechanisms of word formation in Chinese, leaving the diachronic analysis to Section 3.

2.2 A background for exemplar-based compounds: word formation and classifiers in Chinese

Modern Mandarin Chinese fits very well in the ‘traditional’ definition of the isolating language type, as it has virtually no inflection (no obligatory expression of grammatical categories), few (uncontroversial; see Arcodia, 2012) derivational affixes, no cumulative exponence, no or little blurring of morpheme boundaries, and no allomorphy or suppletion (see Bisang, 2004; Packard, 2006). Most morphemes are realised by a syllable, which in turn corresponds to a character in writing, as e.g. 火 *huǒ* ‘fire’, 天 *tiān* ‘sky’, 水 *shuǐ* ‘water’, although a minority of plurisyllabic morphemes are also attested (mostly loanwords), as e.g. 玻璃 *bōli* ‘glass’, 阿司匹林 *āsīpīlín* ‘aspirin’. While morphemes as those seen just above also correspond to (syntactic) words, i.e. they may occupy a syntactic slot, more than half of the morphemes in the Modern Chinese lexicon (70% according to Packard, 2000) are bound, i.e. they have to be combined with some other morpheme in order to be used in a sentence, as e.g. 國 *guó* ‘country’ (as in 愛國 *àiguó* ‘love-country, patriotic’).³

Given the prevalence of bound morphemes, it is perhaps unsurprising that the modern Chinese lexicon is dominated by complex words, i.e. words made of more than one morpheme. The most common device for word formation is compounding, broadly understood as the juxtaposition of (free or bound) lexical morphemes:⁴ compounds make up the vast majority of the Modern Chinese lexicon (see Xing, 2006), and up to 95% of newly-coined words are compounds (according to a corpus study in Ceccagno and Basciano, 2007). In Table 1, we summarize the different subtypes of compounds in Modern Chinese, based on Bisetto and Scalise’s (2005) and Ceccagno and Basciano’s (2007) classification.

³ However, note that the distinction between ‘bound’ and ‘free’ morphemes may be blurry in some cases, as there are bound morphemes which actually do appear to be able to occupy a syntactic slot in specific styles/registers, or in specific constructions; for further details, see Yang, 2003; Packard, 2015.

⁴ There is at present no universally accepted definition of ‘compound’ for Chinese; the treatment of words made of bound roots as compounds is particularly controversial (for an overview, see Arcodia and Basciano, 2017). Since both bound roots and free roots are equally active in word formation, we chose to adopt this broad definition (*contra* e.g. Packard, 2000).

Table 1. *The classification of Chinese compounds according to Bisetto and Scalise's (2005) and Ceccagno and Basciano's (2007) taxonomy*

Subordinate		Attributive		Coordinate	
Endocentric	Exocentric	Endocentric	Exocentric	Endocentric	Exocentric
雞毛	鎮紙	斑马鱼	花心	酸辣	刀槍
<i>jī-máo</i>	<i>zhèn-zhǐ</i>	<i>bānmǎ-yú</i>	<i>huā-xīn</i>	<i>suān-là</i>	<i>dāo-qiāng</i>
chicken-feather	press-paper	zebra-fish	false-heart	hot-sour	sword-spear
'chicken feather'	'paperweight'	'zebrafish'	'unfaithful'	'hot and sour'	'weapons'

For the purposes of the present study, we will be concerned firstly with exocentric coordinate compounds, i.e. compounds in which the constituents are in a relation of coordination, but none of them may be regarded as the head of the compound. This is because these compounds are often used to express categories by combining two exemplars, as in the above-mentioned example 刀槍 *dāo-qiāng* 'sword-spear, weapons': they belong to the [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} class. The second class of compounds at issue here, namely the [EXEMPLAR-CLASS]_{CATEGORY} class, appears to be intermediate between the attributive endocentric and the coordinating exocentric types, as we shall see below.

The preference for compounding goes hand in hand with another trend in the Modern Chinese lexicon, namely the dominance of disyllabic (and, mostly, bimorphemic) words, which could represent up to 80% of the lexicon (according to figures in Shi, 2002). This, however, is a relatively recent development: the Old Chinese lexicon, especially before the Han Dynasty (206 BCE – 220 CE), was prevalently monosyllabic, and only about 20% of the words in the (written) language before 200 BCE were made of two syllables (Shi, 2002: 72). In point of fact, most scholars today agree on the fact that Old Chinese did have subsyllabic affixes, as the often-quoted 'causative' prefix, reconstructed as *s- (e.g. *mələk 'eat' vs. *s-mələk 'feed'; Baxter and Sagart, 2014: 230). Together with the loss of subsyllabic affixation, the single most important development of the lexicon in the evolution towards Modern Chinese is the creation of a large number of complex words, mostly disyllabic, leading to the situation sketched above for the modern language.

2.2.1 *The genesis of coordinating compounds and [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds*

How and why where new compounds created? Wang (1989) provides abundant examples of complex disyllabic words from the Archaic period (X-II cent. BCE), both coordinate and attributive, as e.g. 賓客 *bīn-kè* 'visitor-guest, visitors' and 天子 *tiān-zǐ* 'heaven-son, emperor' (Wang, 1989, pp. 233, 236). While an attributive compound as 天子 *tiān-zǐ* 'heaven-son, emperor' provides a label for a referent, compounds as 賓客 *bīn-kè* 'visitor-guest, visitors', made of (near-)synonymous constituents, seem to be redundant. In fact, coordinate compounds, often made of (near-)synonymous constituents, were very common in the early stages of development of the Chinese disyllabic lexicon (Feng, 1998), and later even grew in relevance; actually, they remained the dominant pattern for compounding until much later, perhaps even up to the XIX century (see Masini, 1993). Feng (1998) shows how, in Archaic texts from the VI century BCE, non-coordinative compounds had a slightly higher token frequency than

coordinative compounds; the balance, however, tipped in favour of coordinating compounds, which already have a higher number of tokens in a later text as the *Lunheng* (I cent. CE).

One explanation for this (perhaps) unusual phenomenon has been proposed by Feng (1998), who argues that the simplification of syllable structure in the history of Chinese has led to excessively ‘light’ and simple syllable types, which were prosodically ‘insufficient’; in order to achieve prosodic autonomy (i.e. to form a ‘Minimal Prosodic Word’), then, syllables had to be joined in pairs. To this end, arguably the easiest thing to do was to juxtapose morphemes in a relation of coordination, especially synonymous ones, since you can combine (quasi-)synonymous morphemes without any relevant changes in meaning, i.e. without necessarily creating a new label (in the absence of a real pragmatic need to have one). For instance, a coordinate compound as 殺戮 *shā-lù* ‘kill-slaughter, kill’ is not that different from 殺 *shā* ‘kill’ alone (Feng, 1998: 243). The fact that these coordinating compounds originate from mere syntactic juxtapositions is supported also by the observation that, in the earliest attestations, we may see both orders of constituents, as e.g. 圖書 *túshū* ‘picture-book, books, publications’ and 書圖 *shūtú* (Feng, 1998: 223); as pointed out by Feng (1998: 224), “(...) the function of two-syllable units was a fundamental need of the language, regardless of whether the outcome was a word or a phrase”.

However, many of these two-word combinations eventually did become compounds, with a fixed order and, sometimes, a specialised meaning. In some cases, the resulting compound had the meaning of only one of the constituents, as in 動靜 *dòng-jìng* ‘active-quiet, activity’ (Feng, 1998: 215); these are termed ‘ornamental compounds’ in Wälchli (2005). However, in the overwhelming majority of cases the meaning of the compound was obtained by the semantic integration of the constituents, either in a compositional or in a non-(strictly-)compositional construction; this holds even for many compounds which are made of (near-)synonymous constituents. See the following examples (from Wang, 1998: 126):

(6)朋友

péng-you
friend-friend
‘friend’

(7)疾病

jí-bìng
disease-disease
‘disease’

Based on the current meaning of their constituents, the compounds in (6) and (7) are easily classified as synonymic. However, if one looks at their original meaning, 朋 *péng* and 友 *yǒu* are not, strictly speaking, synonymous: 朋 *péng* refers to ‘disciples of the same master/school’ and 友 *yǒu* to ‘people with the same ideals, aspirations’; that is, they are more specific notions than ‘friend’. By combining them, one obtains a reading which is intermediate between an additive (compositional) and a collective (non-compositional) reading: ‘friends’ as the sum of 朋 *péng* and 友 *yǒu*, or ‘friends’ as the sum of 朋 *péng*, 友 *yǒu*, and other types of people which share a property *P* with them (see above, Section 2.1.2). As to 疾病 *jí-bìng* (7), nowadays both 疾 *jí* and 病 *bìng* mean just ‘disease’; however, originally, 疾 *jí* indicated ‘light and normal diseases’, while 病 *bìng* indicated ‘serious diseases’. In this case, the construction is more clearly additive (‘light, normal and serious diseases’ = ‘any disease, diseases’).

In fact, the overlap between synonymic and additive/collective compounds has already been pointed out by Wälchli (2005), who shows how synonymic compounds are closely related to, and often the evolution of, additive or collective compounds. He provides examples of synonymic compounds which look very similar to the Chinese examples seen above, as Uzbek *qadr-qimmat* ‘value-dignity, dignity’ (Wälchli, 2005: 144). According to Wälchli, the main characteristic which differentiates synonymic compounds from additive and collective ones is that “[s]ynonymic co-compounds [...] express homogeneous collection complexes in which (ideally) every element contained in them can be referred to by both parts of the co-compound” (Wälchli, 2005: 143); this is perhaps not true, strictly speaking, for constructions as 朋友 *péngyǒu*, before they became fully lexicalised and the semantic difference between the constituents was neutralised. Wälchli also stresses the fact that synonymic compounds have a strong affinity with plurality: he points out that Vietnamese synonymic compounds, as *bạn hữu* ‘friend-friend, friends’, are mostly used for plural reference. As hinted at in the Introduction, [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds in Chinese seem to have a strong preference for generic reference to categories, rather than to individual entities, in their early stages (see below, Section 3); also, note that when [EXEMPLAR-CLASS]_{CATEGORY} compounds indicate a collection, as is typically the case for non-compositional constructions, they are often translated as plurals in Chinese-English dictionaries or in English-language linguistic works (e.g. 書刊 *shū-kān* ‘book-periodical, reading materials, books and periodicals’; Packard, 2000: 89; see the next section).

2.2.2 Classifiers, class nouns and compound words: [EXEMPLAR-CLASS]_{CATEGORY} compounds

Moving now to the [EXEMPLAR-CLASS]_{CATEGORY} of compounds, a few remarks on classifiers and classification in Chinese are in order first. Modern Chinese is a language with a fully grammaticalised category of classifiers. Classifiers are required whenever a noun is associated with a numeral or a demonstrative, and also with some quantifiers:

- (8) 三*(本)書
*sān *(běn) shū*
 three CL book
 ‘three books’

As shown in (8), the omission of the classifier 本 *běn* results in the ungrammaticality of the phrase.

Chinese classifiers are associated with shape, dimension, function and other characteristics of the noun:⁵ compare 張 *zhāng* ‘classifier for flat objects’, 條 *tiáo* ‘classifier for long and thin objects’, 架 *jià* ‘classifiers for machinery’. Some classifiers can be highly specific, as e.g. 艘 *sōu* ‘classifier for boats/ships’; some nouns may be associated with more than one classifier, as e.g. 汽車 *qìchē* ‘car’, which allows both 部 *bù* and 輛 *liàng*; also, the categories identified by the classifier may be blurry and idiosyncratic. For instance, the classifier 部 *bù* is used with novels, movies, and vehicles; 件 *jiàn* is used for matters, affairs, presents, weapons and items of clothing; 匹 *pǐ* is used for horses, donkeys, but also for bolts of cloth. While a diachronic investigation

⁵ Actually, the category commonly used in the Chinese linguistic tradition of 量詞 *liàngcí*, lit. ‘measure word’, includes both classifiers *stricto sensu*, which categorise nouns by selecting some (physical or functional) property, and measure words, as e.g. 杯 *bēi* ‘cup’, which denote a quantity of the entity denoted by the noun (Peyraube, 1991; Chappel and Peyraube, 2011). The former have also been termed ‘sortal numeral classifiers’, whereas the latter have been termed ‘mensural numeral classifiers’ (Gil, 2013); it is only sortal classifiers we are concerned with here, since they have a categorising function. Hence, here by ‘classifier’ we mean ‘sortal numeral classifier’.

could tell us more about the motivations behind these unexpected associations of referents, from the synchronic perspective, these appear as non-conventionalised, ad hoc categories.

The ad hoc, non-conventionalised nature of the categories identified by (at least) some classifiers has been already suggested by Huang and Chen (2011), who performed a psycholinguistic experiment to find out whether speakers of Chinese actually categorised objects (also) by means of the classifiers they are associated with (see the source for the details). Their results show that speakers apparently rely on established, conventional categories, as ‘city’, ‘musical instrument’, or ‘vegetable’, rather than on the sets identified by the use of the same classifier, which had a limited positive effect on their ability to recall lists of nouns; this is interpreted as a feature of ad hoc categories, which are generally less effective than established categories for tasks as clustering and recalling. However, Huang and Chen also point out that this may depend on the individual classifier, with some obviously being closer to established categories (as e.g. 輛 *liàng* ‘classifier for road vehicles’; see below). On the other hand, the categories identified by the [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} class compounds seen above appear to be more stable, conventional categories. We will get back to this below, when discussing the differences between the two classes of exemplar-based compounds at issue here.

Classifiers as a category grammaticalised relatively late in Chinese: while some elements with a ‘proto-classifier’ function are already attested in the Early Archaic period (i.e. XI – VI cent. BCE; see below, ex. 10), at this stage numerals normally appeared before a noun without any intervening element; true classifiers did not emerge until the Pre-Medieval period (II cent. BCE – III cent. CE), and became prominent only much later, in Late Medieval Chinese (VII – XIII cent. CE). Nearly all classifiers derive from nouns (rarely from verbs), as e.g. 頭 *tóu* ‘head’ > ‘classifier for animals’ (Peyraube, 1991; Chappell and Peyraube, 2011); as Bisang (1999: 166) put it, “[t]he diachronic development of classifiers started from a specific lexical use to the formation of prototypes which formed the basis for further generalizations”. For instance, Bisang (1999) illustrates how a noun denoting a tree, i.e. 箇 *gè* ‘bamboo tree’, later became a classifier for bamboo trees and arrows, then for other objects, and even for humans, turning later into the generic classifier it has become today (currently associated with the character 個). The ‘borrowing’ of nouns (or verbs) as classifiers was generally based on one of the following semantic relationships (Liu, 1965, qtd. in Loke, 1997: 5):

- a. shape-based similarity, as e.g. 片 *piàn* ‘thin chip of wood’ > ‘classifier for thin and flat objects’
- b. meronymy, as e.g. 根 *gēn* ‘root’ > ‘classifier for plants, etc.’ (> ‘classifier for long, slender objects’)
- c. synonymy, as e.g. 艘 *sōu* ‘big ship’ > ‘classifier for boats/ships’
- d. related action, as e.g. 乘 *chéng* ‘mount’ > ‘classifier for carriages, chariots, etc.’
- e. metaphor, as e.g. 張 *zhāng* ‘stretch open’ > ‘classifier for bows’ > ‘classifier for bows and stringed instruments’ > ‘classifier for objects made of a flat flexible sheet stretched over a frame, etc.’ (> ‘classifiers for flat objects’)

Furthermore, one important change in the evolution of classifier constructions in Chinese is the shift in word order, from ‘Noun – Numeral – Classifier’ to the current ‘Numeral – Classifier – Order’. Compare (exx. from Peyraube, 1991: 109, 110, 116, 119):

(9) 子七人

zǐ qī rén
son seven person
'seven sons' (Early Archaic Chinese)

(10) 車三百兩⁶
chē sān bǎi liàng
chariot three hundred CL?
'three hundred chariots' (Early Archaic Chinese)

(11) 因下玉鏡台一枚
yīn xià yù-jìng tái yī méi
then send jade-mirror frame one CL
'(he) then sent a jade mirror with frame' (Early Medieval Chinese, III - VI cent. CE)

(12) 乘一朵黑雲
chéng yī duǒ hēi yún
ride one CL black cloud
'(they) rode on a black cloud' (Late Medieval Chinese)

In example (9), 人 *rén* 'person' is but a noun retaining its full lexical meaning, rather than a classifier for 子 *zǐ* 'son'; 'person' indicates the broad category in which 'son' may be inserted. In our view, nouns as 人 *rén* in this type of construction are close to the notion of 'class noun', i.e. nouns representing "a rather high level of abstraction from which more concrete subcategories can be derived by further determination" (as *tree* → *apple tree*; Bisang, 1999: 167). In point of fact, Bisang (1993; 1999) suggests that class nouns may evolve into classifiers. Note also that there may be identity between the two nouns in this construction, as in (Peyraube, 1991: 108):

(13) (...) 羌一百羌
Qiāng yī bǎi Qiāng
Qiang one hundred Qiang
'one hundred Qiang (an ethnic group)'

This repetition of the counted noun in the position of the classifier is attested in several languages of East and Southeast Asia (see Peyraube, 1991; Bisang, 1999), and has been analysed as an indication of the strong association between class nouns and classifiers. However, Bisang (1999) points out that most class nouns in Modern Chinese are not also used as classifiers, and the development of classifiers in Chinese seems to have followed an item-oriented, rather than category-oriented pathway. The item-oriented development of classifiers "starts from the context of counting individual items which are of particular cultural importance", and then the classifier construction "spreads over to a wider range of nouns"; the category-oriented development, on the other hand, "starts from a categorial system already existing in the language" (Bisang, 1999: 158-159), as the above mentioned class nouns. We will get back to this below.

According to Peyraube's (1991) analysis, in constructions as (9) the numeral modifies the second noun (here, 人 *rén*), and nothing can intervene between them; following Ōta (1958), Peyraube argues that 七人 *qī rén* 'seven people' is a predicate, rather than a modifier, of 子 *zǐ* 'son'; that is, the sentence is to be understood as something like 'as to sons, (there are) seven of

⁶ This morpheme later became associated with the character 輛, which is used elsewhere in this article.

them'.⁷ The following example (10) is similar; however, 輛 *liàng*, whose lexical meaning is 'two-wheeled chariot', will later become a classifier for road vehicles (as said above). It is thus unclear whether in (10) 輛 *liàng* is still to be regarded as a lexical noun, or as a (proto-?)classifier.

On the other hand, in (11), an example from an Early Medieval Chinese text, 枚 *méi* has undoubtedly developed into a (generic) classifier, from its original lexical meaning 'tree trunk' (Peyraube, 1991). Here, however, the numeral-classifier complex still follows the quantified noun, while in the following example, which reflects modern usage (see ex. 12 above), it precedes the noun. Peyraube (1991) suggests that the change in position of the numeral-classifier complex from postnominal to prenominal has to do with its reanalysis as a modifier of the noun; in Chinese, modifiers of the noun almost invariably occur in prenominal position (for an additional explanation in terms of information structure, see Bisang, 1999). Once the numeral-classifier complex appears before the noun, it forms a 'classifier phrase' directly associated with the noun it modifies; it is at this stage, according to Peyraube's analysis, that they shed all their lexical properties and become 'true', fully grammaticalised classifiers.

How is the [EXEMPLAR-CLASS]_{CATEGORY} class created, then? As mentioned in the introduction, in this type of compounds the first constituent exemplifies the class denoted by the second one; the second constituent may be a class noun (14), or a classifier (15; exx. from Bisang, 1999: 175):

(14) 楊樹
yáng-shù
poplar-tree
'poplar tree'

(15) 房間
fángjiān
room-CL
'room'

In (14), the righthand constituent 樹 *shù* 'tree' acts as a class noun, indicating the class to which 楊 *yáng* 'poplar' belongs. However, 樹 *shù* 'tree' in itself has not developed into a classifier, and this type of nouns appear to have no special features, when it comes to countability. They are not unlike other right-headed attributive compounds (an extremely common construction in Modern Chinese).

The subtype of [EXEMPLAR-CLASS]_{CATEGORY} compounds on which we want to focus here is the one represented by (15), in which the second constituent is a classifier, as 間 *jiān* in 房間 *fángjiān*. Whereas noun-class noun compounds as (14) may only designate an entity within the set identified by the class noun, in noun-classifier compounds, depending from the origin of the classifier (and from its diachronic evolution), the second constituent may be on the same taxonomic level as the exemplar (as is the case when the classifier is a near-synonym), or may actually indicate a broader category, i.e. that identified by the set of referents which use the classifier (often, a seemingly inconsistent category; see the examples above). Thus, as said above, they appear to be intermediate between the attributive endocentric and the coordinating

⁷ An anonymous reviewer pointed out that a numeral-classifier complex might be as well be analysed as a (selecting) head, rather than as a modifier of the noun. This is the position defended e.g. by Cheng & Sybesma (2005), in which classifiers are analysed as the head of the Classifier Phrase (CIP), with an NP as its complement; the CIP may in turn be embedded into a Numeral Phrase (the difference between the two lying in their definiteness status; see Cheng & Sybesma 2005 for the details). However, we tend to agree with Peyraube and analyse the numeral-classifier complex as a modifier of the NP, mainly because it is the noun which selects the classifier, and not *vice versa*.

exocentric types, whereas noun-class noun compounds are invariably right-headed. Noun-classifier compounds may be said to be right-headed when the classifier constituent is not on the same taxonomic level of the lefthand nominal constituent, and when there is a difference in semantics or distribution between the noun by itself and the noun-classifier compound.

However, the underlying attributive (modifier-modified) structure of noun-classifier compounds may be hidden, as in the many cases in which the compound seemingly has the same meaning of the lefthand nominal constituent, as the above mentioned 紙張 *zhǐ-zhāng* ‘paper-CL, paper’. Compare:

(16) 三張紙
sān zhāng zhǐ
three CL paper
‘three sheets of paper’

(17) 三張紙張
sān zhāng zhǐ-zhāng
three CL paper-CL
‘three sheets of paper’ (Google search)⁸

In this case, the behaviour and semantics of the simple noun and of the corresponding noun-classifier compound appear to be virtually the same; a noun-classifier compound like 紙張 *zhǐ-zhāng* looks like an ‘ornamental’ compound, as the above mentioned 動靜 *dòng-jìng* ‘active-quiet, activity’. This may be interpreted, in our opinion, as the endpoint of a process of lexicalisation and opacisation of the compound, in which the second constituent gives almost no contribution to the meaning, features and distribution of the compound.

In this connection, note that Loke (1997: 10-11) indicates as having only “generic or indefinite reference”, and hence rejecting a numeral-classifier phrase, some of these ‘ornamental-like’ noun-classifier compounds, as the above mentioned 紙張 *zhǐ-zhāng* ‘paper-CL, paper’ or 書本 *shū-běn* ‘book-CL, books’; however, these are actually attested also with numerals and classifiers, i.e. they may be individuated and counted. Once again, this points towards a relatively recent development; on the other hand, those noun-classifier compounds which resist individualisation and counting are more clearly right-headed, as these semantic and distributional features obviously derive from the righthand classifier morpheme.

According to Loke (1997), noun-classifier compounds began to flourish towards the end of the Late Medieval period. Earlier studies (e.g. Liu, 1965, qtd. in Loke, 1997) suggest that noun-classifier compounds derive from the lexicalisation of the ‘noun-numeral-classifier’ pattern (see ex. 11 above) when the numeral was ‘one’, in which case it could be omitted. Compare:

(18) a. 馬一匹
mǎ yī pǐ
horse one CL
‘one horse’

b. 馬匹
mǎ-pǐ
horse-CL
‘horse’

⁸ <http://support-tw.canon-asia.com/contents/TW/TC/8000135902.html> (last access: 17.8.2017).

However, Loke adds that other factors favouring the creation of this type of compounds include the relatively transparent meaning of classifiers, which were still identical to the nouns they derive from, and whose meaning “enabled native speakers to use them as bound morphemes to form new nouns and verbs or compounds for generic reference” (1997: 8); also, he suggests that the preference for disyllabic words may have encouraged the formation of this type of compounds (as said above for coordinating compounds). In point of fact, noun-classifier compounds are always disyllabic (to the best of our knowledge), despite the prevalence of disyllabic words in the Modern Chinese lexicon (and, hence, one would expect to find also disyllabic word – classifier compounds).

In Loke’s (1997) treatment, the diffusion of noun-classifier compounds occurred only when classifiers had become a fully established category of Chinese grammar, and is interpreted by him as a case of ‘regrammaticalisation’, from classifier to bound morpheme in word formation. Bisang (1999) concedes that noun-classifier compounds do resemble a stage in the category-oriented development of classifiers, but he also stresses the late appearance of this type of compounds, when classifiers were already established, as pointed out above. Bisang (1999) also points out that classifiers which are also class nouns, as 所 *suǒ* ‘place; classifier for buildings’, emerged rather late. These two facts point towards a later category-oriented pathway of evolution for some classifiers, after the item-oriented development of the basic system of classifiers had already occurred (possibly also due to contact with languages to the South of the Chinese-speaking world; Bisang, 1999: 175).

Zhang (2013) analyses classifiers in noun-classifier compounds as singulatives, and she suggests that the construction is productive, i.e. classifiers are “systematically able to occur in an N-CL compound” (2013: 258). She also points out that noun-classifier compounds, despite claims to the contrary (Li and Thompson, 1981; Loke, 1997; see also above) are mostly countable, and may refer to singular entities, not necessarily collective/plural entities, even when used without a quantifier (Zhang, 2013: 261):

- (19) 他想要花朵
tā xiǎngyào huā-duǒ
 he want flower-CL
 ‘He wants a flower / flowers / the flowers’

Note, however, that her interpretation of noun-classifier compounds as singulatives seems to be based mostly on *stricto sensu* measure words, i.e. *mensural* numeral classifiers (see fn. 4), as e.g. 滴 *dī* in 水滴 *shuǐ-dī* ‘water-drop, water drop’; given the frequent use of noun-classifier with plural/collective reference, we doubt that her characterization of the (sub-)class as singulatives holds. In point of fact, Chinese-English dictionaries, as well as English glosses in linguistic works, often translate many noun-classifier compounds as plurals (e.g. 船隻 *chuán-zhī* ‘boat-CL, boats, vessels’; see Li and Thompson, 1981: 82), as already said for [EXEMPLAR1-EXEMPLAR2]_{CATEGORY} compounds (Section 2.2.2). Moreover, as said above, the quantification of noun-classifier compounds might be a recent development, made possible by the opacisation of the classifier constituent. We will get back to this below (Section 3.2 and 4).

2.2.3 Interim summary: exemplar-based compounds in Chinese

To sum up, let us elaborate on the similarities and differences between the two classes of compounds introduced in this section. Both [EXEMPLAR1-EXEMPLAR2]_{CATEGORY} and [EXEMPLAR-CLASS]_{CATEGORY} compounds involve an exemplar-based process of word formation which joins two constituents, also (or mostly?) for prosodic reasons, creating a form which often appears to

be semantically redundant: this is the case particularly for synonymic coordinating compounds (see above, exx. 6-7) and compounds which contain a classifier with a meaning very close to that of the lefthand nominal constituent, as e.g. 書本 *shūběn* ‘book-CL’, in which the classifier 本 *běn* also means ‘book’.

As to the differences between the two classes of exemplar-based compounds at issue here, the most obvious structural one is the fact that classifier [EXEMPLAR-CLASS]_{CATEGORY} compounds may actually design a set which is, in a sense, more specific than that designed by the class, as already noted in Section 2.1.2: in 紙張 *zhǐ-zhāng* ‘paper-CL, paper’, the *designatum* ‘paper’ is but a subset of the entities which fall under the scope of the classifier 張 *zhāng* (see above). Also, whereas [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds always have a coordinate structure, and both compositional and non-compositional instances are widely attested, noun-classifier compounds may be right-headed or (pseudo-)coordinate, and seem to be evolving towards opacisation of their structure (i.e. towards ‘ornamental’ compounds).

As to semantics, [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds seem to be mostly limited to established categories, whereas [EXEMPLAR-CLASS]_{CATEGORY} compounds are open to non-conventionalised categorisation, as we hinted at above. This is easily explained by the observation that established categories are obviously more likely to find lexical expression in languages (see above, Section 2.1); on the other hand, classifiers are sometimes based on ad hoc, heterogeneous categories, and this may be reflected in noun-classifier compounds (compare e.g. 鳥隻 *niǎo-zhī* ‘bird-CL, birds’ and 船隻 *chuán-zhī* ‘ship-CL, vessels’). Noun-classifier compounds made of (near-)synonymous constituents normally designate established categories, pointing again towards their close connection with [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds; on the other hand, coordinate constructions (seemingly phrasal, mostly) designing ad hoc categories are attested in the history of Chinese, but they were mostly not lexicalised and did not survive in the modern lexicon (as expected). Compare (exx. from the *Zuozhuan*, Late Archaic Chinese, IV cent. BCE):

(20) 羽毛齒革

yǔ-máo-chǐ-gé

feather-fur-ivory-leather

‘Raw materials needed in everyday life and for building weapons’

(21) 玉帛

yù-bó

jade-silk

‘jade and silk goods presented as state gifts’ > ‘friendship’ (also, ‘property’)

Both examples (20) and (21) designate ad hoc categories; the former did not survive (perhaps also because of prosodic reasons?), whereas the latter did, perhaps also because of its newly acquired metaphorical meaning (‘friendship’).

In the next section, we will present our diachronic data, which provide support for our view of an evolution for the constructions at issue from exemplar-driven abstraction to reference to a category, and then to individual instantiations of the category, in many cases.

3. Diachronic analysis: exemplar-based compounds in the history of Chinese

In order to bring to light the pathways of evolution for exemplar-based compounds, we chose to look into the behaviour of eight representative items for each of the two types of constructions (i.e. coordinating and noun-classifier); for the first group (exemplar-exemplar),

we chose four synonymic and four non-synonymic compounds. The words were chosen among those which appeared relatively early and survived into the modern language, in order to have a complete picture of their history. The items are listed in Table 2.

Table 2. *Exemplar-based compounds used in the present study*

[EXEMPLAR1-EXEMPLAR2] _{CATEGORY}		[EXEMPLAR-CLASS] _{CATEGORY}	
WORD	GLOSS AND TRANSLATION	WORD	GLOSS AND TRANSLATION
朋友 <i>péng-you</i>	friend-friend 'friend'	馬匹 <i>mǎ-pǐ</i>	horse-CL 'horses'
疾病 <i>jí-bìng</i>	disease-disease 'disease'	花朵 <i>huā-duǒ</i>	flower-CL 'flower'
牙齒 <i>yá-chǐ</i>	tooth-tooth 'tooth'	紙張 <i>zhǐ-zhāng</i>	paper-CL 'paper'
皮革 <i>pí-gé</i>	skin-skin 'leather, hide'	車輛 <i>chē-liàng</i>	vehicle-CL 'vehicles'
玉帛 <i>yù-bó</i>	jade-silk 'gifts; friendship; property'	船隻 <i>chuán-zhī</i>	boat-CL 'vessels'
干戈 <i>gān-gē</i>	shield-dagger/axe 'weapons; war'	鳥隻 <i>niǎo-zhī</i>	bird-CL 'bird'
禽獸 <i>qín-shòu</i>	bird-quadruped 'birds and beasts'	燈盞 <i>dēng-zhǎn</i>	lamp-CL 'lamp'
刀槍 <i>dāo-qiāng</i>	sword-spear 'weapons'	房間 <i>fāng-jiān</i>	room-CL 'room'

The diachronic language data used for the present research was collected from the *Chinese Text Project* digital library, an open-access searchable database of Chinese texts ranging from the V cent. BCE to the early XX century; the corpus is split into two parts for the purposes of text search, namely up to the Han dynasty (i.e. to the III cent. CE), and post-Han texts.⁹ The post-Han section of the corpus was expanded with seven texts from the *Academia Sinica Tagged Corpus of Early Mandarin Chinese*; the additional texts were chosen in order to have a more representative sample of the colloquial-oriented vernacular literature (as e.g. the Dunhuang 'transformation texts').¹⁰ Data for contemporary usage was extracted from the *Academia*

⁹ URL: ctext.org (last access: 17.8.2017). See the website for the full list of texts included in the collection.

¹⁰ URL: <http://lingcorpus.iis.sinica.edu.tw/early/> (last access: 21.8.2017). The list of texts from the *Academia Sinica Tagged Corpus of Early Mandarin Chinese* we made use of includes: a collection of Dunhuang 'transformation texts' (敦煌變文集新書 *Dūnhuáng Biànwéngjí Xīnshū*), the *Zutangji* (祖堂集 *Zǔ Táng Jí*), the *Collection of Thirty Yuan Dramas* (元刊雜劇三十種 *Yuánkān Zájù Sānshízhǒng*), the *Laoqida Yanjie* (老乞大諺解 *Lǎoqìdà Yànjiě*), the *Piaotongshi Yanjie* (朴通事諺解 *Piáo Tōngshì Yànjiě*), the *Water Margin* (水滸傳 *Shuǐhù Zhuàn*) and the *Story of a Marital Fate to Awaken the World* (醒世姻緣傳 *Xǐngshì Yīnyuán Zhuàn*).

Sinica Balanced Corpus of Modern Chinese (Version 4.0), a balanced collection of written and oral texts in Modern Chinese consisting of more than 11 million words;¹¹ however, we also used Google searches to complement the data, especially to check for occurrences of quantified compounds. Also, note that the temporal depth for the two types of compounds is inevitably different: as said above, many [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds are attested much earlier than [EXEMPLAR-CLASS]_{CATEGORY} compounds.

In the next section, we will present our data and analysis. Our discussion of the words belonging to the two classes will focus on different aspects: for [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds, we will take into consideration the context of occurrence (i.e. independent or as part of a broader construction), cohesiveness (i.e. reversibility of the constituents), and quantification (quantifiers, numeral-classifier constructions, demonstratives; see Cheung, Li and Barner, 2010). We will focus on quantifiers that select only countable items, as 無數 *wúshù* ‘countless, innumerable’, since some quantifiers may apply also to mass nouns, as e.g. 多 *duō* ‘much, many’ (see Ex. 26); the former, but not necessarily the latter, are indicative of the fact that the construction may be used to refer to individual instances of the category. For [EXEMPLAR-CLASS]_{CATEGORY} compounds, we will focus on reference (generic vs. specific; see Krifka *et al.*, 1995) and countability, since reversibility does not apply here (i.e. in the reverse order, they become just ordinary words preceded by a classifier).

3.1 [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds

Let us start our presentation from [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} class compounds. The first item in our list, 朋友 *péng-yǒu* ‘friend’, has been already discussed above (2.2.1) as an example of a synonymic compound with an additive/collective origin (‘disciples of the same master/school’ and ‘people with the same ideals, aspirations’). In our corpus, 朋友 *péng-yǒu* appears to be an independent construction (namely, a word) starting from the earliest attestations; it is most often found as it is, rather than as part of a list of coordinated items (compare ex. 20 above). If we limit our search to the Han dynasty, we find 170 occurrences of 朋友 *péng-yǒu*, but only 7 instances of the reverse order (i.e. 友朋 *yǒu-péng*), pointing towards early lexicalisation of this item (although sporadic attestations may be found until the XVIII century). As to quantification, the first clear instances of 朋友 *péng-yǒu* modified by a numeral we could find date to the late XIII century (in the 朱子語類 *Zhūzi Yǔlèi*, a Pre-Modern Chinese text); in earlier examples, 朋友 *péng-yǒu* is never quantified, and is often used for generic reference. See the following example from the *Book of Rites* (禮記 *Lǐjì*, Archaic Chinese):

- (22) 曾子曰：朋友之墓，有宿草而不哭焉
Zēngzǐ yuē péngyou zhī mù yǒu sù-cǎo bù kù yān
 Zengzi say friend of grave there.be old-grass not wail for.him
 ‘Zengzi said: when there is old grass on a friend’s grave, we do not wail for him anymore’

Here, 朋友 *péng-yǒu* refers to a generic notion of ‘friend’, not to a specific person. Such examples are very common in early texts.

The second item, 疾病 *jí-bìng* ‘disease’, has been also discussed before as an instance of an additive compound turned synonymic; its characteristics and distribution are very similar to 朋友 *péng-yǒu* ‘friend’. Firstly, it is overwhelmingly more common in the current order, although some occurrences of the reverse order (病疾 *bìng-jí*) are attested (in the pre-III century corpus,

¹¹ URL: <http://asbc.iis.sinica.edu.tw/> (last access: 17.8.2017).

210 vs. 9),¹² again pointing towards early lexicalisation. Also, it is normally found as an independent item, and it is seemingly never quantified in all of our Premodern corpus; nowadays, it is normally counted and associated with the classifiers 個 *gè* and 種 *zhǒng*. Note, however, that it is sometimes used with specific reference even in earlier texts.

The third item, 牙齒 *yá-chǐ* ‘tooth’, has the same origin as the two other synonymic compounds: originally, 牙 *yá* referred to the back teeth, and 齒 *chǐ* to the front teeth (incisors), but they then came to be both used for just any tooth. 牙齒 *yá-chǐ* has just two occurrences before the III cent. CE; curiously, the version with the opposite order of constituents has 16 occurrences, and the two versions have exactly the same number of occurrences (51) in the post-Han section of the corpus, but 齒牙 *chǐ-yá* is never attested in the Modern corpus. As to countability, examples of 牙齒 *yá-chǐ* with a numeral and with 皆 *jiē* ‘all’ may be found in Late Medieval texts (e.g. in the *Dunhuang Bianwen*). Also, 牙齒 *yá-chǐ* is often used with specific reference.

The last synonymic compound in our list, 皮革 *pí-gé* ‘leather, hide’, is again made of formerly non-synonymous constituents (‘skin with fur’ and ‘skin without fur’), nowadays both used for ‘leather’. 皮革 *pí-gé* is always attested in this order in texts up to the Han Dynasty, but (oddly) the opposite order is marginally attested (three occurrences) in later texts. In the *Book of Rites* (see above), one finds 皮革 *pí-gé* as part of a list of coordinated items, in which the two constituents have separate reference:

- (23) [...] 審五庫之量：金鐵，皮革筋，角齒，羽箭干，[...]
shěn wǔ kù zhī liàng jīn tiě pí gé jīn jiǎo chǐ yǔ jiàn
 inspect five storehouse of quantity metal iron skin hide sinew horn ivory feather arrow
gān
 wood
 ‘[...] to inspect the materials in the five storehouses: those of iron and other metals; of skins and hides and sinews; of horn and ivory; of feathers, arrows and wood, [...]’¹³

In this excerpt, 皮 *pí* and 革 *gé* are associated with 筋 *jīn* ‘sinew’ as exemplars in a list of related items; in another passage, they are associated with 筋 *jīn* ‘sinew’ and 角 *jiǎo*, again in the context of a list of items. There are several examples of this usage in our corpus, always in the same type of list constructions (see Masini, Mauri and Pietrandrea 2018), as e.g. lists of presents. At the same time, 皮革 *pí-gé* is frequently used as an independent word(-like) construction; incidentally, the almost absolute fixedness of the order points towards early conventionalisation/lexicalisation. One apparent difference between the other compounds seen above and 皮革 *pí-gé* is that the latter could be associated with a numeral and a measure word (石 *shí*), in post-nominal position, and hence could be quantified, at least since the I century CE. In Modern Chinese, it may be associated with the classifier 張 *zhāng* (used for flat objects) to indicate one or more pieces of leather.

Moving to the second part of the list, 玉帛 *yù-bó* ‘jade and silk as state gifts; friendship; property’ was introduced above (ex. 21). Its attested meanings are the product of a metaphorical shift, following which ‘jade and silk’ are taken as typical exemplars of state gifts, and thus, by extension, ‘friendship’; ‘jade and silk’ have also been reinterpreted as ‘valuables, property’, again by taking them as exemplars of (valuable) goods. It is always attested only as 玉帛 *yù-bó* in the whole corpus, never with the opposite order. Just as seen above for 皮革 *pí-gé*, 玉帛 *yù-bó* is

¹² Note that 疾病 *jì-bìng* in Archaic Chinese is often used as a predicative element, having thus a status closer to a verb or a predicative adjective (‘be sick’).

¹³ Translation by James Legge (retrieved from ctext.org).

also found in the context of lists of items, as in the following examples (from the *Records of the Grand Historian*, I cent. BCE):

(24) 羽毛齒角玉帛，君王所餘 [...]

yǔ máo chǐ jiǎo yù bó jūnwáng suǒ yú
feather fur ivory horn jade silk king REL abundance

‘Feathers, fur, ivory, horns, jade and silk, the king has those in abundance [...]

This list has many items in common with those seen above in (20) and (21); in essence, they all contain valuable items which often cooccur in these contexts (e.g. gift listing). From early on, 玉 yù and 帛 bó are juxtaposed also, for instance, as items for sacrificial offerings. It appears that 玉 yù and 帛 bó became associated quite early, given the fixed order and their common occurrence outside list constructions; semantic shift strengthened the conventionalisation of this word form. In Modern Chinese, 玉帛 yù-bó belongs to the formal, classic-sounding stratum of the vocabulary, and is still almost never counted (no occurrence in the *Academia Sinica* corpus). With a cursory Google search, we did find some instances of 玉帛 yù-bó quantified by a numeral and the generic classifier 個 gè, but these appear to be found only in the context of video gaming, in which 玉帛 yù-bó is a quantifiable resource which may be used in a game (arguably best translated as ‘gems’).

The following item, 干戈 gān-gē ‘weapons; war; military affairs’, is virtually always attested as such in the whole corpus, except for an isolated attestation of 戈干 gē-gān in a X century text. Following the usual pathway of evolution, 干 gān ‘shield’ and 戈 gē ‘dagger/axe’ are chosen as representative exemplars to designate the whole category of ‘weapons (of war)’, and then are metaphorically extended to indicate ‘war, military affairs’; for instance, in the *Records of the Grand Historian* one finds 干戈之事 gān-gē zhī shì ‘the matters of shield and dagger/axe’ to indicate ‘war’, a possible step in the metonymical evolution of the construction. Differently from 玉帛 yù-bó, 干戈 gān-gē is rarely found as part of a broader list; the only example we could find comes from an ode in the *Classic of Poetry*, an Early Archaic Chinese text:

(25) 弓矢斯張、干戈戚揚 [...]

gōng shǐ sī zhāng gān gē qī yang
bow arrow all stretch shield dagger/axe axe raise

“With bows and arrows all ready, shields, daggers and axes raised [...]

This, and the nearly absolute fixedness of the order of constituents, are indicative of early conventionalisation of this construction. As to countability, it is generally not counted in the diachronic corpus (we found no clear instance), and neither in the *Academia Sinica* corpus. With a Google search, we did find a few scattered instantiations of 干戈 gān-gē quantified with a numeral and the classifiers 個 gè or 根 gēn; the pattern is clearly not very frequent, and its significance is arguably limited.

The next item, 禽獸 qín-shòu ‘birds and beasts’ is a fairly typical additive compound, albeit an ornamental reading is also attested (i.e. ‘beasts, walking animals’). The construction is almost always found in this order, apart from three exceptions, all in the post-Han portion of the corpus; however, in these examples 禽 qín and 獸 shòu retain their separate reference (i.e. they are used to refer separately to ‘birds’ and ‘beasts’). Examples of 禽 qín and 獸 shòu in list constructions

may be found, for instance, in the *Book of Rites* (禽獸魚鱉不中殺 *qín shòu yú biē bù zhōng shā* ‘birds, beasts, fishes and turtles not fit to be killed’).¹⁴

Quantification was possible from relatively early on, albeit with 多 *duō* ‘many’ and not with a numeral, as in the following example (from the *Zhuangzi*, III cent. BCE):

(26) 古者禽獸多而人少

gǔ zhě qín shòu duō ér rén shǎo

past NMLZ bird beast many and people few

‘In the past, birds and beasts were many, but men were few’

In later examples, 多 *duō* ‘many, much’ is found also as a prenominal modifier; note, however, that 多 *duō* may be used also with mass, non-count nouns (as e.g. 水 *shuǐ* ‘water’), as hinted at above, and hence the significance of (26) is limited. In the Modern language, 禽獸 *qín-shòu* may be quantified by a numeral with the classifier 個 *gè* or 頭 *tóu*, although it is more often used without quantification (again, no quantified occurrence in the *Academia Sinica* corpus).¹⁵

Differently from the items seen hitherto, the last construction in the [EXEMPLAR1-EXEMPLAR2]_{CATEGORY} class, 刀槍 *dāo-qiāng* ‘sword and spear; weapons’, appeared relatively late in Chinese. There is no occurrence of this construction in the earlier part of the corpus, and the occurrences in the post-Han corpus are relatively few; also, it is only slightly more frequent as 刀槍 *dāo-qiāng*, rather than as 槍刀 *qiāng-dāo* (37 vs. 34 occurrences). Interestingly, both versions are very often found in list constructions: 槍刀 *qiāng-dāo* is found five times in a coordinating construction with 劍戟 *jiàn-jǐ* ‘sword-halberd’, and 刀槍 *dāo-qiāng* two times, suggesting again some conventionalised association of items. 刀槍 *dāo-qiāng* is also found as part of longer lists of items, which may even include also non-weapon tools. There is but one instance of 槍刀 *qiāng-dāo* quantified by 許多 *xǔduō* ‘many’, but as part of a list in which reference is made to each item separately (from the *Journey to the West*, XVI cent.):

(27) 牆根頭插著許多槍刀叉棒

qiáng gēn-tou chā-zhe xǔduō qiāng dāo chā bàng

wall root-SUF stick-DUR many spear sword trident club

‘At the foot of the wall many spears, swords, tridents and clubs were put’

One may easily hypothesise that the collective meaning of 刀槍 *dāo-qiāng* emerged from the use of the two constituent lexemes as exemplars of the class, either alone or in recurring combinations with other items in the same category, as seen above. When 刀 *dāo* and 槍 *qiāng* are used to convey the meaning ‘weapons’, they are not quantified, at least in our diachronic corpus (although, in this specific case, this may be an artefact of the limited data). As to countability, even in the Modern language it seems that 刀槍 *dāo-qiāng* is seldom counted and, when it is, reference is generally made to the two items separately; when it is associated with the measure word 副 *fù*, reference is made to sword and spear as a pair of items, rather than as the class of weapons.

¹⁴ Incidentally, this particular list of coordinands (‘birds, beasts, fishes and turtles’) is found also elsewhere in our corpus, and hence was possibly conventionalised to some degree.

¹⁵ 禽獸 *qín-shòu* also has another meaning acquired by metaphorical extension, i.e. ‘bestial person’. In this sense, it is normally quantified in Modern Chinese, and even in earlier texts (e.g. in the *Story of a Marital Fate to Awaken the World*, XVII cent.). In point of fact, in most of the quantified occurrences of this word which we found with a Google search 禽獸 *qín-shòu* is used with the meaning ‘bestial person’.

In short, all of the items analysed here seem to follow a similar pathway of evolution in the earliest stages; namely, they are created as associations of related items, both as such and as part of broader lists, and they evolve into names for a category. As categorial names, they seem to resist quantification; however, (some) synonymic compounds, if compared to the non-synonymic constructions we considered, seem to allow quantification from an earlier stage. In point of fact, some non-synonymic compounds are not often quantified even in Modern Chinese. We may hypothesise that the more transparent category-building structure of non-synonymous compounds, in which the exemplars are still evident, may have contributed to retaining a tendency to be used as category labels, rather than as individual members of the category; this is not the case for synonymic compounds, in which the original semantic difference between the constituents is no longer evident, and they no longer look like exemplars, arguably.

3.2 [EXEMPLAR-CLASS]_{CATEGORY} compounds

As hinted at earlier (Section 2.2.2), noun-classifier compounds appeared relatively late in the Chinese lexicon, and flourished towards the end of the medieval period; in fact, genuine examples may be found only in the post-Han section of our corpus. The first item in our list, 馬匹 *mǎpǐ* ‘horses’, is attested three times in the earlier portion of our corpus; however, here 匹 *pǐ* is not yet a classifier (see Peyraube 1991), and hence these examples are not relevant for our discussion. In the post-Han texts, the overwhelming majority of the occurrences of 馬匹 *mǎpǐ* are found in two novels, the above-mentioned *Journey to the West* and the *Romance of the Three Kingdoms* (a XIV cent. text). According to Loke, in Modern Chinese 馬匹 *mǎpǐ* is one of those noun-classifier compounds which are not counted and “usually denote generic or indefinite reference” (1997: 11); in dictionaries, it is often translated as a plural (in English) or as a collective noun. However, in the *Journey to the West* 馬匹 *mǎpǐ* is frequently used with specific reference to a singular referent, i.e. a specific horse. In the slightly earlier *Romance of the Three Kingdoms*, 馬匹 *mǎpǐ* is also occasionally used for specific reference to individual instances, although it is more often used to refer to a plural collectivity; in the latter text, 馬匹 *mǎpǐ* may be quantified by 無數 *wúshù* ‘countless, innumerable’ and 許多 *xǔduō* ‘many’, albeit always in the context of a list of related items, as in the following example:

- (28) 搶奪馬匹衣甲無數
qiǎngduó mǎpǐ yī jiǎ wúshù
 pillage horse clothes armour countless
 ‘[they] snatched many horses, clothes and armour’

However, no occurrence with a numeral and/or a classifier is attested in the diachronic corpus. No quantified occurrence may be found in the *Academia Sinica* corpus either, but examples may be easily found with a Google search (*contra* Loke, 1997) as in the following excerpt from the novel *The Deer and the Cauldron* (1969-1972) by Jin Yong:

- (29) [...] 只見道旁倒斃了兩匹馬匹
zhī jiàn dào pang dǎo-bì-le liǎng pǐ mǎ-pǐ
 only see road side fall-dead-PFV two CL horse-CL
 ‘[one] could see only two dead horses at the side of the road’

The second item, 花朵 *huāduǒ* ‘flower’, is also found only in the post-Han corpus, never with a numeral and/or a classifier; however, it is once modified by 數 *shù* ‘several’ (in a poem from

the Late Medieval Period). It is seemingly used also with specific reference (i.e. to refer to specific occurrences of ‘flowers’), but it is unclear whether it is ever used to refer to a single individual flower; this is possible in Modern Chinese (see ex. 19), in which 花朵 *huāduǒ* is also counted with the generic classifier 個 *gè* or with 朵 *duǒ* itself.

The compound 紙張 *zhǐzhāng* ‘paper’ has already been discussed above (2.2.2), and we suggested that, in Modern Chinese, it should probably be analysed as (akin to) ornamental compounds. From the point of view of counting and reference, it seems to be virtually identical to 紙 *zhǐ* ‘paper’ alone. 紙張 *zhǐzhāng* occurs only three times in the diachronic corpus; in the XVIII cent. novel *The Scholars* (儒林外史 *Rúlín Wàishǐ*), it is used with specific reference:

- (30) 看見紙張白亮 [...]
Kàn-jiàn zhǐ-zhāng bái-liàng
 look-see paper-cl white-bright
 ‘seeing that the paper was bright white [...]

The following item, 車輛 *chē-liàng* ‘vehicles’, is made of the constituents 車 *chē* ‘car, wheeled vehicle’ and 輛 *liàng*, which, as said above (Section 2.2.2), originally meant ‘two-wheeled chariot’ and is now used as a classifier for road vehicles. We already mentioned that one finds examples of 車 *chē* counted with 輛 *liàng* already in Early Archaic Chinese, but here 輛 *liàng* most likely had not yet become a classifier. In the post-Han corpus, 車輛 *chē-liàng* is used both with generic reference and with specific reference to a singular referent; there is one instance of 車輛 *chē-liàng* quantified by 無數 *wúshù* ‘countless, innumerable’. However, differently from the other items seen above, there is also one instance of 車輛 *chē-liàng* followed by a numeral, (車輛四百 *chē-liàng sì-bǎi* ‘four hundred chariots’), in the XVI cent. novel *Investiture of the Gods* (封神演義 *Fēng Shén Yǎnyì*). In Modern Chinese, 車輛 *chē-liàng* is normally counted with the classifier 輛 *liàng*, just as 車 *chē*.

The compound 船隻 *chuán-zhī* ‘vessels’ is much more common than the previous ones in the diachronic corpus (150 occurrences, all in the post-Han section). It is used with specific reference, and it occurs quantified by 皆 *jiē* ‘all’ and by a numeral, as in the following example from the *Journey to the West*:

- (31) [...] 望求一船隻渡河
wàng qiú yī chuán-zhī dù hé
 hope ask one boat-CL cross river
 ‘[I] was hoping to ask for a boat to cross the river’

In the *Romance of the Three Kingdoms*, 船隻 *chuán-zhī* is also modified by 多少 *duōshǎo* ‘how many’; this indicates that it was a countable noun at least since the XIV century (*contra* Loke 1997). Interestingly, however, when 船隻 *chuán-zhī* is found in a list of (asyndetically) related items, it normally refers to a plurality of items; the same is true for 車輛 *chē-liàng* ‘vehicles’.

The next compound, 布匹 *bù-pǐ* ‘cloth’, shares the classifier constituent 匹 *pǐ* with 馬匹 *mǎ-pǐ* ‘horse’, as mentioned above (Section 2.2.2).¹⁶ 布匹 *bù-pǐ* is attested only six times in our

¹⁶ This association is probably explained by the fact that among the attested lexical meanings of 匹 *pǐ*, one finds it both as a unit of measure for length (corresponding to four 丈 *zhàng*) and as a name for (ordinary) single horses or sets of horses.

(post-Han) corpus, and it is once modified by 多 *duō* ‘much; many’ (in the *Piaotongshi Yanjie*, written between the XIV and the XV centuries). It is nowadays counted with the classifier 匹 *pǐ*, but also with 張 *zhāng*, the classifier for flat objects, just as 布 *bù*.

The compound 燈盞 *dēng-zhǎn* ‘lamp’ contains the lexeme 燈 *dēng* ‘lamp’ and the classifier 盞 *zhǎn*, which as a lexeme meant ‘small cup’ or ‘container for lamp oil’. The compound has only seven occurrences in the (post-Han) diachronic corpus, but it is commonly found with numerals and classifiers, as in 一個燈盞 *yī gè dēng-zhǎn* ‘one lamp’ (from the novel *The Scholars*). It is clearly different from the others seen above, in that it seems to have been countable since the beginning.

The last item in our list, 房間 *fāng-jiān* ‘room’, contains the classifier 間 *jiān*, whose original meaning is ‘space; room’, and is now used as a classifier for (small) houses and rooms, including the word 房間 *fāng-jiān* ‘room’ itself. It has only five occurrences in our post-Han diachronic corpus; it is not quantified, but it can have specific reference.

To sum up, the claim that noun-classifier compounds cannot, and *could* not, be quantified does not seem to be supported by our data; however, for most of the compounds in our list, quantified examples are rare in the diachronic corpus, and only three out of eight are modified by numerals (and for one of them, 車輛 *chē-liàng*, we only have one example in which the numeral is not even a modifier). As to reference, all of the compounds considered seem to be able to convey non-generic (specific) reference. If we look at Modern usage, it is clear that all the noun-classifier compounds we considered can be quantified and can have specific reference. Arguably, some of them are most often used with generic reference, and are usually not quantified (as claimed by Loke, 1997, among others), but this is based mostly on impressionistic judgements; it is in fact not easy to provide a quantitative assessment of the specific vs. generic/unquantified usage for these items, since most of them occur too frequently in the Modern Chinese corpus to check each instance for the parameters at issue. However, even with a cursory look at our corpus data, the tendency to avoid quantification and/or express generic reference is not apparent at all at least for 船隻 *chuán-zhī* ‘vessel(s)’ and 房間 *fāng-jiān* ‘room’; interestingly, as said above, 船隻 *chuán-zhī* is one of the few constructions for which we found instances of quantification with a numeral in the diachronic corpus. We believe it is safe to suggest that compounds as 馬匹 *mǎ-pǐ* ‘horse’, 花朵 *huā-duǒ* ‘flower’ and 紙張 *zhǐ-zhāng* ‘paper’ developed countability fairly recently, as shown by our diachronic data and as confirmed by the relatively low frequency of quantified occurrences in Modern Chinese.

All in all, what emerges is that the behaviour with respect to countability and reference of noun-classifier compounds is much less consistent than what we saw in the preceding section for [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds. Just as for synonymic compounds, it seems that at least some (probably, most) noun-classifier compounds were reanalysed and their structure is no longer really transparent; they sometimes become almost like ornamental compounds, i.e. the classifier no longer plays a significant role in the compound. Interestingly, whereas the tendency for noun-classifier compounds to have generic reference and resist quantification had been pointed out previously in the literature, as said above (Section 2.2.2), it does not appear to be as strong as expected; it seems that non-synonymic [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds actually have a stronger tendency to express categorial meaning and resist quantification. If the hypothesis we sketched in the preceding section, namely that a more transparent category-building structure favours categorial use of an item, the blurry categorial significance of classifiers may have rendered their structure opaque, hence favouring their development as ‘ordinary’ referential nouns.

4. Concluding remarks: the conventionalization of procedural naming

4.1 The development of exemplar-based compounds and the conventionalization of procedural naming

The diachronic data discussed so far show that the naming function of exemplar-based compounds develops through a gradual conventionalization of an exemplar-driven abstraction process, which leads to the creation of a label for the category originally exemplified. Table 3 provides an overview of the analysis described in Section 3.1 for [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds, while Table 4 shows data on the [EXEMPLAR-CLASS]_{CATEGORY} compounds.

[EXEMPLAR ₁ -EXEMPLAR ₂] _{CATEGORY}									
SYNONYMIC 朋友 <i>péng-yǒu</i> ‘friend’, 疾病 <i>jí-bìng</i> ‘disease’, 牙齒 <i>yá-chǐ</i> ‘tooth’, 皮革 <i>pí-gé</i> ‘leather, hide’				NON-SYNONYMIC 玉帛 <i>yù-bó</i> ‘jade and silk as state gifts; friendship; property’, 干戈 <i>gān-gē</i> ‘weapons; war; military affairs’, 禽獸 <i>qín-shòu</i> ‘birds and beasts; animals’, 刀槍 <i>dāo-qiāng</i> ‘sword and spear; weapons’			TOTAL		
	PRE-HAN & HAN	POST-HAN	MODERN	PRE-HAN & HAN	POST-HAN	MODERN	PRE-HAN & HAN	POST-HAN	MODERN
Internal reversibility	3	2	---	---	2	---	3	4	---
Occurrence in a list	1	---	---	3	1	---	4	1	---
Occurrence as an independent word	4	4	4	3	4	4	6	8	8
Quantification	1	3	4	1	1	1	2	4	5
Specific reference	2	4	4	---	1	1	2	5	5

Table 3. Overview of the diachronic data on [EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds: for each period, we indicate the number of compounds attested with the relevant parameter.

Synonymic and non-synonymic compounds are kept separate in Table 3, to show their differences. As can be observed in the Table, synonymic compounds show a relatively early-stage conventionalization, as revealed by the fact that they are almost never attested as parts of lists, and show specific reference and even quantification in the post-Han period. This is likely to be due to the early loss of the original semantic differentiation between the two synonyms, leading to ornamental compounding. On the other hand, nearly all of the non-synonymic compounds we considered in our study are attested within lists in their early stages. They show conventionalization, as proved by the complete loss of internal reversibility, but they seem to have a lower degree of internal semantic bleaching, maintaining a more transparent exemplar-based generic reference to an abstract category, which makes quantification and reference to specific items infrequent even in Modern Chinese.

Synonymic compounds rarely occurred in lists and were highly favoured by the general shift of Chinese towards a disyllabic system (see Section 2.2.1), and this may motivate their ornamental, non-compositional reading from the very beginning. On the contrary, non-

synonymic compounds are the result of the conventionalization of frequent listing patterns, whereby frequently associated exemplars of a category come to be employed as a label to refer to the category itself. Given the semantic distinctness of the two lexemes involved, they retain an exemplification function even in the compound, providing clear instances of procedural, exemplar-based naming strategies.

Although they developed later, [EXEMPLAR-CLASS]_{CATEGORY} compounds show a number of similarities with synonymic compounds, as can be observed in Table 4: in Modern Chinese they are indeed all countable and quantified, a tendency that was already fairly apparent before the Modern Chinese period. As argued at the end of the preceding section, this may be due to the low internal transparency of these compounds, in which the blurry categorial significance of classifiers may have led to an internal semantic bleaching. Furthermore, as pointed out in Section 2.2.2, the rise of these compounds may have been favoured by the need for disyllabic words, just like synonymic compounds. As a consequence of their opacity, they quickly acquire the ability to name not only category and abstract concepts, but also specific items.

[EXEMPLAR-CLASS] _{CATEGORY}			
馬匹 <i>mǎpǐ</i> ‘horses’, 花朵 <i>huāduǒ</i> ‘flower’, 紙張 <i>zhǐzhāng</i> ‘paper’, 車輛 <i>chē-liàng</i> ‘vehicles’, 船隻 <i>chuán-zhī</i> ‘vessels’, 布匹 <i>bù-pǐ</i> ‘cloth’, 燈盞 <i>dēng-zhǎn</i> ‘lamp’, 房間 <i>fāng-jīān</i> ‘room’			
	PRE-HAN & HAN (no occurrences)	POST-HAN	MODERN
Generic quantification (e.g. with ‘many’, ‘several’, ‘all’)	---	5	8
Countability (with numerals and classifiers)	---	3	8
Specific reference	---	8	8

Table 4. Overview of the diachronic data on [EXEMPLAR-CLASS]_{CATEGORY} compounds: for each period, we indicate the number of compounds attested with the relevant parameter.

Despite all the differences that we just highlighted, we argue that in the development of both types of exemplar-based compounds in Chinese we observe, at least in the early stages, the persistence of a procedural, bottom-up category construction (see Section 2.1.2), which poses some limits to quantification and countability. In other words, at the beginning of their development, many of these compounds are associated to a collective, abstract meaning, which clashes with specific reference. It is only later that we see the conventionalization of a procedural naming function. In some cases, this conventionalization process led to a complete bleaching of the internal constituency of the compound, within which it becomes hard to identify two exemplars, or an exemplar and its class.

The successive stages in the development of exemplar-based compounds are summarized in Fig. 1:

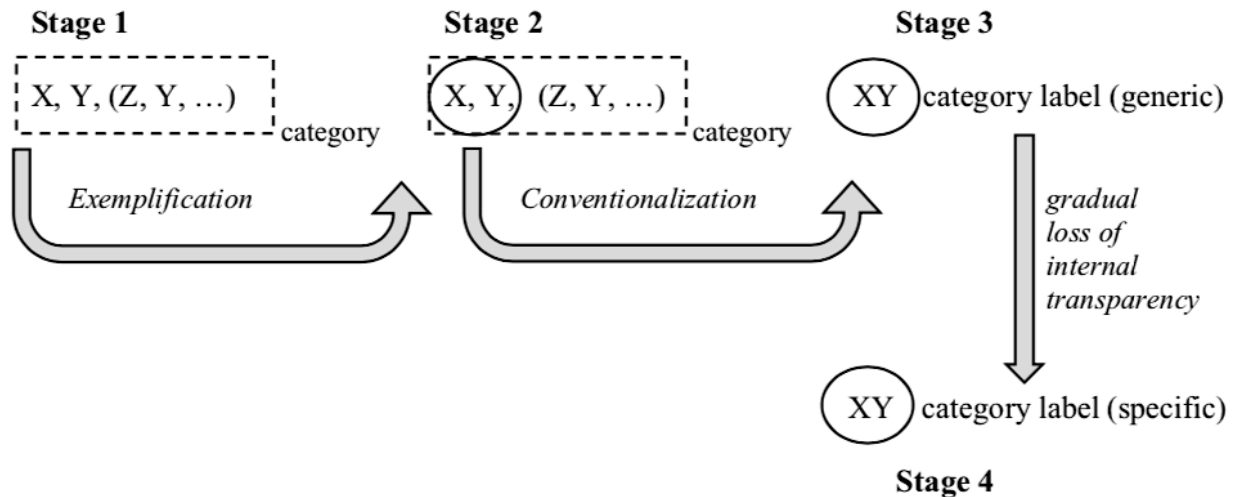


Figure 1. Four stages in the development of exemplar-based compounds in Chinese

After Stages 1 and 2, during which the sequence of two lexemes is conventionalized and the exemplification process acquires a labeling function, we observe two possible outcomes. When internal transparency is completely lost, the compound behaves as an ‘ordinary’ noun and acquires the ability to refer to specific referents, which can be counted and quantified (Stage 4). When conventionalization does not lead to the complete loss of internal transparency, the compound retains its categorial semantics and cannot be quantified (Stage 3). Stage 3 is where we find non-synonymic compounds, while synonymic compounds and classifier compounds reach Stage 4.

The role of conventionalization is crucial in the development of a naming function, because it creates a fixed expression (cf. Wray, 2002), i.e. a fixed relation between a form and a concept. However, a fixed expression is not necessarily opaque, it may fulfil a naming function while triggering a process of exemplar-based abstraction. Actually, the procedural component is crucial for the labeling function itself, at least in its emergence. The more the exemplification process becomes blurred, as we observed for synonymic and classifier compounds, the more the compound is interpreted as a unitary, non-procedural whole.

4.2 Conclusions

The aim of this paper was to examine the development of a particular type of compounds, characterized by an internal process of indexical categorization, i.e. a bottom-up, exemplar-driven abstraction. Being compounds, these constructions are mainly used to label concepts and categories. Yet, their ability to work as names depends on an initial descriptive, syntactic, procedural function of exemplification, which makes them procedural labels. After addressing the issue of naming, discussing what defines a naming function and what are the most frequent strategies for naming, we focused on the naming of categories, including ad hoc categories, and analyzed in detail how naming may emerge in an inductive way from exemplification.

Exemplar-based naming is well represented by exemplar-based compounds, which are quite common in Modern Chinese, and offer an interesting testbed for the relation between category labeling and category exemplification. After identifying the two macro-types of compounds which constitute the object of this research, we provided an historical background, suggesting possible external reasons underlying the rise of compounds in Chinese (including the development of a dominantly disyllabic lexicon), and describing the systems of Chinese classifiers.

[EXEMPLAR₁-EXEMPLAR₂]_{CATEGORY} compounds and [EXEMPLAR-CLASS]_{CATEGORY} compounds show many differences, but this did not prevent us from providing a unitary account, based on a qualitative diachronic study of their emergence in texts from the pre-Han and post-Han periods, up to Modern Chinese. We selected eight compounds for each type and examined their occurrences in the corpus, focusing on their reversibility, the surrounding context, their countability and quantification, and their referential properties.

The analysis of these parameters allowed us to observe the gradual conventionalization of the compounds, which started to occur as independent words, with a fixed order. Initially, they were characterized by a higher tendency to denote abstract concepts, as a likely consequence of the internal exemplar-based abstraction, and were not frequently quantified, nor counted. Gradually, the increase in conventionalization determined a bleaching of the exemplification process and of the categorial semantics of the compound. From generic reference to sets and categories, highly conventionalized exemplar-based compounds started to acquire the ability to identify specific objects, allowing for countability and quantification.

Quantification is nowadays still problematic for a subtype of exemplar-based compounds, namely non-synonymic ones, because their internal semantic differentiation somehow prevents the complete loss of transparency. Therefore, though conventional, these compounds still retain a generic, abstract semantics, and are mainly used to name categories, rather than items.

Through this study, we have shown that the process of category construction by exemplification may become a label for the category itself, thus acquiring a naming function. The existence of exemplar-based names, such as compounds, and the factors at play in their emergence, confirm the great permeability between functions and forms in language, which can hardly be described by means of neat divisions, and rather display dynamic patterns of usage and change even when it comes to ‘giving names to things’.

References

- Arcodia, G.F., 2012. *Lexical Derivation in Mandarin Chinese*. Crane, Taipei.
- Arcodia, G.F., 2018. Coordinating nominal compounds: universal vs. areal tendencies. *Linguistics* 56(6), 1197-1243.
- Arcodia, G.F., Basciano, B., 2017. Morphology, Modern Chinese. In: Sybesma, R., Behr, W., Gu, Y., Handel, Z., Huang, J.C.-T., Myers, J. (Eds.), *Encyclopedia of Chinese Language and Linguistics*, Vol. 3. Brill, Leiden, pp. 104-118.
- Arcodia, G.F., Grandi, N., Wälchli, B., 2010. Coordination in compounding. In: Scalise, Sergio & Vogel, Irene (Eds.), *Cross-disciplinary issues in compounding*, John Benjamins, Amsterdam / Philadelphia, pp. 177-197.
- Ariel, M., Mauri, C., 2018. Why use *or*? *Linguistics* 56(5), 939-993.
- Barotto, A., 2017. *Exemplification and categorization: the case of Japanese*. University of Bergamo, PhD dissertation.
- Barotto, A., 2018. The hedging function of exemplification: the case of Japanese. *Journal of Pragmatics* 123, 24-37.
- Barotto, A., Mauri, C., 2018. Constructing lists to construct categories. *Italian Journal of Linguistics* 30(1), 95-134.
- Barsalou, L.W., 1983. Ad hoc categories. *Memory & Cognition* 11 (3), 211-227.
- Baxter, W., Sagart, L., 2014. *Old Chinese: A New Reconstruction*. Oxford University Press, Oxford.
- Bauer, L., 2003. *Introducing linguistic morphology*, 2nd Edition. Edinburgh University Press, Edinburgh.

- Bhat, D.N.S., 1994. The adjectival category. Criteria for differentiation and identification. Benjamins, Amsterdam and Philadelphia.
- Birk, E., 2006. The Use of Examples. Paper presented at the Conference *Beyond Mimesis and Nominalism: Representation in Art and Science* (London School of Economics, 22-23. June 2006). <http://philsci-archive.pitt.edu/view/confandvol/bmanriaas.html> (last access on 23/08/2018)
- Bisang, W., 1993. Classifiers, quantifiers and class nouns in Hmong. *Studies in Language* 17 (1), 1-51.
- Bisang, W., 1999. Classifiers in East and Southeast Asian languages: Counting and beyond. In: Gvozdanovic, J. (Ed.), *Numeral types and changes worldwide*. Mouton de Gruyter, Berlin-New York, pp. 113-185.
- Bisang, W., 2004. Grammaticalization without coevolution of form and meaning: the case of tense-aspect-modality in East and mainland Southeast Asia. In: Bisang, W., Himmelmann, N.P., Wiemer, B. (Eds.), *What Makes Grammaticalization? A Look from its Fringes and its Components*. Mouton de Gruyter, Berlin-New York, pp. 109-138.
- Bisetto, A., Scalise, S., 2005. Classification of compounds. *Lingue e Linguaggio* 2, 319-332.
- Bonvino, E., Masini, F., Pietrandrea, P., 2009. List constructions: A semantic network. In: *Proceedings of Grammars in Construction(s) – 3rd International AFLiCo Conference* University, Paris Ouest – Nanterre – La Défense, May 27-29, 2009
- Booij, G., 2009. Phrasal names: A constructionist analysis. *Word Structure* 2(2), 219-240.
- Carston, R., 2010. Lexical pragmatics, ad hoc concepts and metaphor: from a relevance theory perspective, *Italian Journal of Linguistics* 22, 153-180.
- Ceccagno, A., Basciano, B., 2007. Compound headedness in Chinese: an analysis of neologisms. *Morphology* 17, 207-231.
- Channell, J., 1994. *Vague Language*. Oxford University Press, Oxford.
- Chappell, H., Peyraube, A., 2011. Grammaticalization in Sinitic languages. In: Bernd, H., Narrog, H. (Eds.), *The Oxford Handbook of Grammaticalization*. Oxford University Press, Oxford, pp. 784-794.
- Cheng, L.L.-S., Sybesma, R., 2005. Classifiers in four varieties of Chinese. In: Cinque, G., Kayne, R.S. (Eds.), *The Oxford handbook of comparative syntax*. Oxford University Press, Oxford, pp. 259-292.
- Cheung, L., Barner, D., 2010. Individuation and quantification: do bare nouns in Mandarin Chinese individuate? In: Clemens, L.E., Liu, L.C.-M (Eds.), *Proceedings of the 22nd North American Conference on Chinese Linguistics and the 18th Annual Meeting of the International Association of Chinese Linguistics*, Vol. 1. Harvard University, Cambridge (MA), pp. 395-412.
- Croft, W., Cruse, A.D., 2004. *Cognitive Linguistics*. Cambridge University Press, Cambridge.
- Dardano, M., 2009. *Costruire parole*. Il Mulino, Bologna.
- De Caluwe, J., 1990. Complementariteit tussen morfologische en in oorsprong syntactische benoemingsprocédé's. In: Johan De Caluwe (ed.), *Betekenis en produktiviteit*. Gentse bijdragen tot de studie van de Nederlandse woordvorming, Seminarie voor Duitse Taalkunde, Gent, pp. 9-24.
- Downing, P., 1977. On the creation and use of English compound nouns. *Language* 53, 810-842.
- Feng, S., 1998. Prosodic structure and compound words in Classical Chinese. In: Packard, J. (Ed.), *New Approaches to Chinese Word Formation*. Mouton de Gruyter, Berlin-New York, pp. 196-259.
- Gil, D., 2013. Numeral classifiers. In: Dryer, M.S., Haspelmath, M. (Eds.) *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

- Huang, S., Chen, J.-Y., 2011. The effects of numeral classifiers and taxonomic categories in Chinese speakers' recall of nouns. In: Carlson, L., Hoelscher, C., Shipley, T.F. (Eds.), Proceedings of the 33rd Annual Conference of the Cognitive Science Society. Cognitive Science Society, Austin (TX), pp. 3199-3204.
- Jespersen, O., 1942. A modern English grammar on historical principles. Vol. VI, Morphology, Munksgaard, Copenhagen.
- Koefoed, G., 1993. Benoemen. Een beschouwing over de faculté du langage, Meertens-Instituut, Amsterdam. (Publikaties van het P.J. Meertens-Instituut, 20).
- Krifka, M., Pelletier, F.J., Carlson, G.N., ter Meulen, A., Chierchia, G., Link, G., 1995. Genericity: an introduction. In: Carlson, G.N., Pelletier, F.J. (Eds.), *The Generic Book*. The University of Chicago Press, Chicago, pp. 1-124.
- Lakoff, G., Sweetser, E. 1994. What is a Conceptual System? In: Overton, W. (Ed.), Proceedings of the 1991 Meeting of the Piaget Society, Erlbaum, Norwood, N.J..
- Li, C.N., Thompson, S., 1981. *Mandarin Chinese: A Functional Reference Grammar*. University of California Press, Berkeley.
- Liu, S., 1965. 魏晋南北朝量词研究 Wèi-Jìn Nán-Běi Cháo liàngcí yánjiū (in Chinese). Zhonghua Shuju, Beijing.
- Lo Baido, C., 2018. Categorization via exemplification: evidence from Italian. *Folia Linguistica Historica* 39(1), 69-95.
- Loke, K.-K., 1997. The grammaticalisation and regrammaticalisation of Chinese numeral classifier morphemes. *Journal of Chinese Linguistics* 25 (1), 1-20.
- Magni, E., 2018. Collective suffixes and ad hoc categories: from Latin -ālia to Italian -aglia. *Folia Linguistica Historica* 39(1), 201-224.
- Masini, F., 1993. The Formation of Modern Chinese Lexicon and its Evolution towards a National Language: The Period from 1840 to 1898. Project on Linguistic Analysis, University of California, Berkeley.
- Masini, F., 2009. Phrasal lexemes, compounds and phrases: A constructionist perspective. *Word Structure* 2(2), 254-271.
- Masini, F., Mauri, C., & Pietrandrea, P., 2018. List constructions: towards a unified account. *Italian Journal of Linguistics* 30(1), 49-94.
- Mauri, C., 2017. Building and interpreting ad hoc categories: a linguistic analysis. In: Joanna Blochowiak, Cristina Grisot, Stephanie Durrleman & Christopher Laenzlinger (eds.), *Formal models in the study of language*, Springer, Berlin, pp. 297-326.
- Mauri, C., Sansò, A., 2018. Linguistic strategies for ad hoc categorization: theoretical assessment and cross-linguistic variation. *Folia Linguistica Historica* 39(1), 1-35.
- Ōta, T., 1958. 中国語歴史文法 Chūgokugo rekishi bunnō (in Japanese). Konan Shoten, Tokyo.
- Packard, J.L., 2000. *The Morphology of Chinese*. Cambridge University Press, Cambridge.
- Packard, J.L., 2006. Chinese as an isolating language. In: Brown, K. (Ed.), *Encyclopedia of Language and Linguistics (Second Edition)*. Elsevier, Amsterdam, pp. 355-359.
- Packard, J.L., 2015. Morphology: morphemes in Chinese. In: Wang, W., Sun, C. (Eds.), *The Oxford Handbook of Chinese Linguistics*. Oxford University Press, Oxford, pp. 263-273.
- Peyraube, A., 1991. Some remarks on the history of Chinese classifiers. *Santa Barbara Papers in Linguistics Volume 3: Asian Discourse and Grammar*, 106-126, 1991.
- Schlücker, B., Hüning, M., 2009. Introduction. *Word Structure* 2(2), 149-154.
- Shi, Y., 2002. *The Establishment of Modern Chinese Grammar. The Formation of the Resultative Construction and Its Effects*. John Benjamins, Amsterdam and Philadelphia.
- Sprenger, S., 2003. Fixed expressions and the production of idioms. *Max-Planck-Institut für Psycholinguistik*, Nijmegen.
- Wälchli, B., 2005. *Co-compounds and Natural Coordination*. Oxford University Press, Oxford.

- Wang, F., 1998. *La Formazione delle Parole nella Lingua Cinese Contemporanea*. Università degli Studi di Trieste, Trieste.
- Wang, L., 1989. *汉语语法史 Hànyǔ Yǔfǎshǐ* (in Chinese). Shangwu Yinshuguan, Beijing.
- Wilson, D., Carston, R. 2007. A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In: Noel Burton-Roberts (ed.), *Pragmatics*. London, Palgrave, pp. 230-259.
- Wray, A., 2002. *Formulaic language and the lexicon*. Cambridge University Press, Cambridge.
- Xing, J.Z., 2006. *Teaching and Learning Chinese As a Foreign Language: A Pedagogical Grammar*. Hong Kong University Press, Hong Kong.
- Yang, X., 2003. *汉语语素论 Hànyǔ Yǔsùlùn* (in Chinese). Nanjing, Nanjing Daxue Chubanshe.
- Zhang, N., 2013. *Classifier Structures in Mandarin Chinese*. Mouton de Gruyter, Berlin-New York.