

*Taiwan Journal of Linguistics*  
Book Series in Chinese Linguistics  
doi:10.6533/NCCULING-9789861475233

# Lexical Derivation in Mandarin Chinese

漢語的派生構詞法

馬振國

*Giorgio Francesco Arcodia*

# LEXICAL DERIVATION IN MANDARIN CHINESE

In the Chinese language, morphologically complex words have been attested since the remote past of the language, including both stem-modifying processes and agglutination of morphemes, mostly lexical and free in the classical language (see Baxter & Sagart 1998). Chinese word-formation has received much attention in the literature in recent times, but most descriptions and theoretical work on the topic are focussed on compounding (see e.g. Packard 1998, 2000, Lin 2001, Ceccagno & Basciano 2009a-b), and it is still a matter of debate whether compounding and derivation are two distinct phenomena in Modern Mandarin Chinese (see, among others, Pan, Ye & Han 2004).

In this monograph we intend to analyse Chinese word formation patterns which may be candidate to derivational status, according to the definition of such process of word formation which we find in the morphological literature (as e.g. Beard 1998, Naumann & Vogel 2000, Olsen 2000): they are patterns such as X—學 ‘the study of X’ (心理學 *xīnlǐxué* ‘psychology’) or X—性 ‘the property of (being) X’ (重要性 *zhòngyàoxìng* ‘importance’). The characteristics of the morphemes around which those patterns are built which sets them close to derivational affixes are that they appear in a fixed position, seem to form new words productively and convey a different, “emptier” meaning than that of the corresponding lexical morph (see Ma 1995). The apparent phonological (and, needless to say, orthographical) identity between a “would-be affix” and its lexical counterpart (as, say, 學 used as a verb, ‘to study’) is not surprising, since grammaticalization without alteration in the sound shape of a morph is a characteristic feature of languages belonging to the East and South-East Asian area (Bisang 1996, 2004). Therefore, the notion of “affixoid”, coined to describe word formation elements in European languages which are bound but phonologically identical to a free form in the language (such as Dutch *boer*, meaning ‘farmer’ as a word and ‘dealer’ when used as a bound form), proves to be unnecessary for Chinese.

---

*Giorgio Francesco Arcodia has a Ph.D. in linguistics from Pavia University in Italy. During his PhD studies, he spent one year at the People's University of China, Beijing (中國人民大學) and later conducted research at the National Taiwan University, Taipei (國立台灣大學). He currently teaches General Linguistics and Language Typology at the University of Milan-Bicocca, Italy. His main research interests are the morphology of Mandarin Chinese and Chinese dialects, grammaticalization in East and South-East Asian languages, Japanese linguistics, language typology and the acquisition of Italian as a foreign language.*



文鶴出版有限公司

CRANE PUBLISHING CO., LTD.

書號：002-0910

ISBN 978-986-147-523-3



9 789861 475233

Taiwan Journal of Linguistics  
Book Series in Chinese Linguistics  
Editors: One-Soon Her & Kawai Chui

LEXICAL DERIVATION IN MANDARIN  
CHINESE  
漢語的派生構詞法

Giorgio Francesco Arcodia  
馬振國



文鶴出版有限公司  
CRANE PUBLISHING CO., LTD.



獻給雷昱智先生。

：一日為師，終身為父；



## ACKNOWLEDGEMENTS

This book is a revised edition of my 2008 Italian-language monograph *La Derivazione Lessicale in Cinese Mandarino* ('Lexical Derivation in Mandarin Chinese'). Firstly, I would like to thank Prof. One-Soon Her (何萬順) of National Chengchi University (國立政治大學) for encouraging me to work on a new edition of my first monograph, to make it accessible to a broader readership. Since the original monograph was based on my (homonymous) doctoral dissertation, defended in 2008 at the University of Pavia (Italy), I would like to thank Prof. Anna Giacalone Ramat, Elisa Roma, Emanuele Banfi (雷昱智), Giuliano Bernini, Pierluigi Cuzzolin, Nicola Grandi, Gianguido Manzelli and Sergio Scalise for reading the original manuscript and providing numerous valuable suggestions, criticisms, and for the many enlightening discussions. Also, I would like to thank Prof. Federico Masini (馬西尼) of the University of Rome "La Sapienza" for the outstanding contribution to the discussion of the thesis. A special thank goes to Tommaso Gorni of Franco Angeli Publishers, for his kindness and patience.

I had the privilege of sharing my time as a doctoral candidate with many extraordinarily bright fellow students as Sergio Catanea (雷雷), Filippo Alessandro Torti (大力), Caterina Mauri, Federica Da Milano, Emanuele Raini (馬諾) and many others, but I am afraid it will not be possible to mention them all here. I want to express my heartfelt thanks for their presence and friendship throughout these years.

Doing extensive research work on Chinese would not have been possible without the help of the teachers I met during my period as a student in the P.R.C. and, later, as a researcher in Taipei; my special thanks go to Prof. Yang Cuntian (楊存田), Xu Yumin (徐玉民) and Dong Zhengcun (董正存) of Renmin University of China (中國人民大學), Prof. Hong Bo (洪波) of Nankai University (南開大學), Prof. Karen Steffen Chung (史嘉琳) of National Taiwan University (國立台灣大學), Prof. Claire Chang (張郇慧) of National Chengchi University (國立政治大學), Prof. Wu Taoyuan (吳桃源) of the Mandarin Training Centre (國語教學中心) at the National

Taiwan Normal University (國立台灣師範大學) and, again, Prof. One-Soon Her. Moreover, I would like to thank the staff of the Centre for Chinese Studies (漢學研究中心) at the National Central Library of Taiwan for their hospitality and for the many cheerful ‘tea breaks’ during my stay in Summer 2010: Keng Li-chun (耿立群), Jane Liao (廖箴), Tsai Ch’ing-lang (蔡慶郎), Wang Yu-ching (王玉琴), Ma Yu-chen (馬佑真), Chen Hsu-mei (陳栩梅), and the kindest librarian in the world, Wu Lily (吳莉莉).

Last, but not least, I want to thank my wife Bianca (白夏儂) for being with me all the time, even when a whole continent is between us.



## ABSTRACT

In the Chinese language, morphologically complex words have been attested since the remote past of the language, including both stem-modifying processes and agglutination of morphemes, mostly lexical and free in the classical language (see Baxter & Sagart 1998). Chinese word-formation has received much attention in the literature in recent times, but most descriptions and theoretical work on the topic are focussed on compounding (see e.g. Packard 1998, 2000, Lin 2001, Ceccagno & Basciano 2009a-b), and it is still a matter of debate whether compounding and derivation are two distinct phenomena in Modern Mandarin Chinese (see, among others, Pan, Ye & Han 2004).

In this monograph we intend to analyse Chinese word formation patterns which may be candidate to derivational status, according to the definition of such process of word formation which we find in the morphological literature (as e.g. Beard 1998, Naumann & Vogel 2000, Olsen 2000): they are patterns such as X-學 ‘the study of X’ (心理學 *xīnlǐxué* ‘psychology’) or X-性 ‘the property of (being) X’ (重要性 *zhòngyàoxìng* ‘importance’). The characteristics of the morphemes around which those patterns are built which sets them close to derivational affixes is that they appear in a fixed position, seem to form new words productively and convey a different, “emptier” meaning than that of the corresponding lexical morph (see Ma 1995). The apparent phonological (and, needless to say, ortographical) identity between a “would-be affix” and its lexical counterpart (as, say, 學 used as a verb, ‘to study’) is not surprising, since grammaticalization without alteration in the sound shape of a morph is a characteristic feature of languages belonging to the East and South-East Asian area (Bisang 1996, 2004). Therefore, the notion of “affixoid”, coined to describe word formation elements in European languages which are bound but phonologically identical to a free form in the language (such as Dutch *boer*, meaning ‘farmer’ as a word and ‘dealer’ when used as a bound form), proves to be unnecessary for Chinese.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS.....</b>	<b>V</b>
<b>ABSTRACT.....</b>	<b>VII</b>
<b>LIST OF TABLES.....</b>	<b>X</b>
<b>LIST OF FIGURES.....</b>	<b>XI</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>XII</b>
<b>PREFACE.....</b>	<b>XIII</b>
<b>CHAPTER 1 INTRODUCTION: THE CHINESE LANGUAGE, DERIVATION AND GRAMMATICALIZATION THEORY.....</b>	<b>1</b>
1.1 ON THE SUBJECT OF OUR STUDY: MANDARIN CHINESE.....	1
1.1.1 What is ‘Mandarin Chinese’? .....	1
1.1.2 The Phonology of Standard Modern Mandarin.....	4
1.1.3 Aspects of Mandarin Morphology and Syntax.....	5
1.1.4 The Chinese Script .....	9
1.2 DERIVATION IN WORD FORMATION .....	10
1.2.1 Lexical Derivation .....	14
1.2.2 Derivation and Compounding: Delimiting the Borders .....	17
1.3 DERIVATION AND GRAMMATICALIZATION.....	32
1.3.1 Lexical Derivation and Grammaticalization.....	34
1.3.1.1 The ‘Parameters of Grammaticalization’ .....	39
1.3.1.2 Grammaticalization and ‘bleaching’ .....	46
1.3.1.3 Grammaticalization and lexicalization.....	59
1.3.2 Grammaticalization in East and South-East Asian Languages ..	62
1.3.2.1 Comparing ‘Western’ and ‘Eastern’ grammaticalization phenomena.....	69
1.3.2.1 Affixoids reconsidered.....	77

<b>CHAPTER 2 PREVIOUS TREATMENTS OF LEXICAL DERIVATION IN CHINESE LINGUISTICS .....</b>	<b>81</b>
2.1 SOME KEY NOTIONS OF MORPHOLOGY IN CHINESE LINGUISTICS ....	81
2.1.1 The notion of ‘morpheme’ and related issues.....	81
2.1.2 ‘Root’ vs. ‘Affix’, ‘Derivation’ vs. ‘Compounding’ .....	89
2.2 LEXICAL DERIVATION IN CHINESE LINGUISTICS.....	96
2.2.1 Overview.....	96
2.2.2 Recent works.....	101
2.2.3 Summary .....	112
<b>CHAPTER 3 DERIVATION OR COMPOUNDING? THE MANDARIN CASE .....</b>	<b>113</b>
3.1 THE BOUNDARY BETWEEN DERIVATION AND COMPOUNDING IN MODERN MANDARIN CHINESE .....	113
3.1.1 Methodological Issues .....	113
3.1.2 Our Sample .....	114
3.2 ANALYSIS OF THE SAMPLE .....	117
3.2.1 Class Nouns in East and South-East Asian Languages .....	117
3.2.1.1 Some Remarks on the Role of Japanese on the Development of Chinese Word Formation.....	124
3.2.1.2 Concluding Remarks on Class Nouns .....	136
3.2.2 Some ‘Newly-coined Affixes’: on Analogy and Productivity .	142
3.2.3 ‘Chinese’ vs. ‘European’ affixes? .....	158
3.2.4 Between Morphology and Syntax: –者 – <i>zhě</i> and –式 – <i>shì</i> ....	167
3.2.5 Two models for prefixation.....	184
3.3 SUMMARY AND CONCLUDING REMARKS .....	200
<b>CHAPTER 4 CONCLUSION .....</b>	<b>203</b>
4.1 LEXICAL DERIVATION IN GRAMMATICALIZATION THEORY .....	203
4.2 ‘INNER’ AND ‘OUTER’ FORCES DRIVING THE EVOLUTION OF CHINESE WORD FORMATION.....	204
4.3 LEXICAL DERIVATION AS A CROSS-LINGUISTICALLY VALID CATEGORY .....	206
4.4 SUGGESTED AREAS FOR FURTHER RESEARCH .....	207
<b>BIBLIOGRAPHY .....</b>	<b>210</b>

## LIST OF TABLES

Table 1.1 The parameters of grammaticalization (Lehmann 1995) .....	39
Table 2.1 A classification of Chinese morphemes (Dong X. 2004).....	86
Table 2.2 A classification of Chinese morphemes (Packard 2000).....	88
Table 2.3 Affixes and affixoids in Pan, Ye & Han (2004) .....	97
Table 3.1 Our sample.....	116
Table 3.2 The evolution of syllable structure in the history of Chinese....	125
Table 3.3 Complex nouns with 學 <i>xué</i> as the head in Masini (1993) and in NXCD (2005) .....	134
Table 3.4 Sample of words containing 吧 <i>bā</i> as the right-hand constituent (from Arcodia 2011) .....	146
Table 3.5 Prefixes, prefixoids and related categories in nine sources (Xu & Cai 2007, Yang Y. 2007) .....	186
Table 4.1 Inflection, derivation and compounding: salient formal and semantic features.....	206

**LIST OF FIGURES**

Figure 3.1 The minimal syllable in Old Chinese (ca. 1000 BCE) and Middle Chinese (ca. 800 CE).....	127
Figure 3.2 Radial network for 吧 <i>bā</i> .....	151

**LIST OF ABBREVIATIONS**

ADJ	Adjective
Chin.	Mandarin Chinese
CLF	Classifier
COLL	Collective
CONJ	Conjunction
DAT	Dative case
DET	Determiner (的 <i>de</i> )
Du.	Dutch
EMPH	Emphatic particle
Eng.	English
Fr.	French
Ger.	German
It.	Italian
Jap.	Japanese
LOC	Localizer
M	Masculine
N	Noun
NMLZ	Nominalization
PART	Final particle
PFV	Perfective
PRON	Pronoun
PRS	Present
PTCP	Participle
Q	Question marker
Rus.	Russian
SG	Singular
V	Verb

## PREFACE

This book is a revised edition of my 2008 Italian-language monograph *La Derivazione Lessicale in Cinese Mandarino* ('Lexical Derivation in Mandarin Chinese'). The original monograph was based on my (homonymous) doctoral dissertation, defended in 2008 at the University of Pavia (Italy). In this revised edition, the core of the research, which is the analysis of language data, is very similar to the first edition, albeit it has been modified to include newer data and analyses which emerged from further research carried out after writing the Italian-language manuscript. Also, new references have been included, in order to provide an up-to-date coverage of the relevant literature and to enrich the theoretical background of the research (see CHAPTER 1). The first two chapters, providing background information on Chinese and on the existing research on the subject of our study, have been significantly shortened and amended, following the suggestions of the reviewers of the original manuscript. The numbering of sections has changed, and many of them have been almost entirely rewritten.

The aim of this study is to analyse, both in a synchronic and in a diachronic perspective, possible phenomena of derivation in Mandarin Chinese, in order to gain a better understanding of the processes of word formation in Chinese and to contribute to a cross-linguistically consistent characterization of derivation. Also, the issue of how grammaticalization works in different language types will be dealt with extensively.

It is important to stress the fact that, although general remarks on 'derivation' as a class of morphological phenomena will be made, our research will deal specifically only with 'lexical derivation' (cf. Kuryłowicz 1936), i.e. patterns of derivation which alter the lexical meaning of the word they are applied to, as It. *-eria* in e.g. *gelateria* 'ice cream parlour' or *pizzeria* 'pizza parlour'. Also, our 'lexical derivation' will include processes of derivation which have the sole function of assigning a word class to the base word, as the German suffix *-heit / -keit* in e.g. *Freundlichkeit* 'friendliness', turning the adjective *freundlich* 'friendly' into a noun (cf. Beard 1998).

The Mandarin data for our study came from a variety of sources, including dictionaries of Classical Chinese, pre-Modern and Modern

Mandarin (近代漢語 *Jīndài Hànyǔ* and 現代漢語 *Xiàndài Hànyǔ*), reverse lexica, corpora offering data from different historical stages of the language (as the *Academia Sinica* family of web corpora), raw web data (from Google searches) and the countless examples which may be found in the literature on Chinese word formation. Since our research is qualitative in nature, rather than quantitative, we believe that the choice of collecting data from various, non-homogeneous sources was appropriated.

In CHAPTER 1, we shall devote some space to the definition of the subject language of our study, i.e. Mandarin Chinese, and then we shall discuss extensively the issue of the definition of ‘derivation’ and the boundary of such class of word formation processes with compounding. Here reference will be made mostly to recent approaches to the problem (Amiot 2005, Bauer 2005 and 2006, Booij 2005, 2007, 2009 and 2010 among others). We shall introduce the theoretical framework which we shall adopt for our research, namely *Construction Morphology* (as in Booij 2005, 2007, 2009, 2010); in *Construction Morphology*, which is an offspring of *Construction Grammar* (Goldberg 1995, Michaelis & Lambrecht 1996, Goldberg 2006), both word formation patterns and syntactic patterns are treated as constructions (“form-meaning-function complexes”; Michaelis and Lambrecht 1996:216). The other main topic of the first chapter will be grammaticalization, and especially the status of derivation and, more specifically, of lexical derivation in grammaticalization research. Contrastive examples of grammaticalization phenomena in Indo-European languages and Chinese will be provided; we shall argue that the semantic processes commonly accepted as characteristic of grammaticalization, as metaphor, metonymy and abstraction, operate also in the evolution of lexical morphemes into derivational affixes.

The subject of CHAPTER 2 will be the treatment of some notions from Western Linguistics in Chinese linguistics, such as ‘morpheme’, ‘root’, ‘derivation’, ‘compounding’, etc. The most influential works on the issue of (lexical) derivation in Chinese linguistics will be discussed, focussing on some recent approaches to the question (as e.g. Ma Q. 1995, Sun Y. 2000, Dong X. 2004).

The core of this book is CHAPTER 3, where our treatment of lexical derivation in Mandarin Chinese will be illustrated, through the analysis of a sample set of Chinese morphemes which may be (and, often, have been)



regarded as instances of grammaticalized (or partially grammaticalized) derivational affixes. We shall first identify some (non-homogeneous!) subclasses of possible derivational affixes, for the sake of simplicity, and we shall then analyze them both in a diachronic and in a synchronic fashion.

In CHAPTER 4, we shall attempt at summarizing the main findings of our research. Apart from restating our reasons for advocating in favour of the cross-linguistic consistence of processes of grammaticalization of derivational affixes, we shall also point out areas for further research, in order to bring to light further evidence for the universality of the processes of morphological change illustrated in the present study.

In the present book, traditional Chinese characters have been chosen as a default. However, in order to be consistent with the sources quoted, we shall also be employing simplified characters in examples when they were found as such. As to the glosses of examples, we adopted the Leipzig set of abbreviations, when applicable (Url: [http://www.eva.mpg.de/lingua/pdf/LGR09\\_02\\_23.pdf](http://www.eva.mpg.de/lingua/pdf/LGR09_02_23.pdf)); however, we had to add a few more abbreviations for labels which may not be found in the Leipzig set. The romanization system used here for Mandarin Chinese is 漢語拼音 *Hànyǔ Pīnyīn*, which is the standard virtually in all the Chinese-speaking world. Also, we shall give the modern reading of characters even when writing about earlier stages of the language, as is common practice in sinological studies; reconstructed pronunciations will be provided only when necessary.



# CHAPTER 1

## THE CHINESE LANGUAGE, DERIVATION AND GRAMMATICALIZATION THEORY

In this chapter we shall first define the subject language of our study, i.e. Mandarin Chinese, and we shall provide a brief description of some of its salient features. We shall then deal with derivation, providing an overview of the literature on the topic, focussing on the issue of delimiting the phenomenon of ‘derivation’ and ‘compounding’. Lastly, we shall introduce grammaticalization theory or, rather, the aspects of it which are most relevant for the purpose of our research. Some space will be also devoted to contrastive analysis of phenomena of grammaticalization in ‘Western’ languages and in Chinese.

### 1.1 On the Subject of Our Study: Mandarin Chinese

#### 1.1.1 What is ‘Mandarin Chinese’?

In many languages of Europe, the adjective ‘Chinese’ is also used as to refer to the standard language of the People’s Republic of China and of the Republic of China (Taiwan), as e.g. French *chinois*, German *Chinesisch*, Italian *cinese*, etc. The standard language mentioned here is often called ‘Mandarin’ or ‘Mandarin Chinese’ in the English-speaking world; both ‘Chinese’ and ‘Mandarin (Chinese)’ are terms which deserve further clarification.

The term ‘Chinese’ is used even in English, especially by non-specialists, to refer to Modern Mandarin Chinese, i.e. the official language both of the P.R.C. (中華人民共和國 *Zhōnghuá Rénmín Gònghéguó*) and of the R.O.C. (中華民國 *Zhōnghuá Mínguó*)<sup>1</sup>. As a matter of fact, ‘Chinese’ could be used to refer to any language (or

---

<sup>1</sup> The political status of Taiwan is a very sensitive issue. Since the Mainland and Taiwan are *de facto* controlled by different governments, we believe it is appropriate to mention both areas as far as the question of the standard language is concerned. This is not a political endorsement of either the P.R.C. or the R.O.C.

‘dialect’) belonging to the Sinitic branch of the Sino-Tibetan family of languages; this, however, is rarely done. Typically, Modern Mandarin is taken as the Chinese language *par excellence*, and other Sinitic varieties are just termed e.g. ‘Cantonese’, ‘Southern Min’, ‘Hakka’ and so on.

On the other hand, even the usage of ‘Mandarin’ could, in principle, be questioned. The term ‘Mandarin’ is the English rendering of 官話 *guānhuà*<sup>2</sup>, which is actually used both to refer to a group of Northern Chinese dialects (北方話 *Běifānghuà*, ‘northern speech’), and to the *koiné* language spoken by government officials and educated people (Chen P. 1995:205, endnote 4). In the latter sense, 官話 *guānhuà* is a dead language, having been replaced by Modern Standard Mandarin.

In what follows, we shall use the term ‘Mandarin (Chinese)’ to refer just to Modern Mandarin Chinese as a standard language; such system is usually referred to as 普通話 *Pǔtōnghuà* (‘common language’) on the Mainland and as 國語 *Guóyǔ* (‘national language’) or 華語 *Huáyǔ* (‘Chinese language’) in Taiwan. The ‘sociolinguistically neutral’ terms for such language are 中文 *Zhōngwén* (‘Chinese language’) and 漢語 *Hànyǔ* (‘language of the Han people’), the former seeming particularly appropriate to refer to the written form of Modern Mandarin. The term 漢語 *Hànyǔ* as ‘Standard Mandarin’ is also opposed to 漢語方言 *Hànyǔ Fāngyán*, ‘dialects of Chinese’, i.e. Sinitic languages.

Needless to say, the term ‘Chinese’ will also be employed throughout the present work, when making a statement which holds for Chinese languages as a whole, or which does not refer specifically to the Modern language only, but rather to previous historical stages of the language, or to just any of its historical stages. For instance, we shall be dealing with the ‘Chinese script’ (1.1.4) and not with ‘Mandarin script’; also, we might say that “in Chinese, modifiers have always been placed before the modified element”, since such a statement holds for the language regardless of the diachronic stage considered.

As far as the history of Chinese is concerned, there are several known partitions of the language into historical stages. As remarked by Norman

---

<sup>2</sup> Literally, ‘officials’ language’.

(1988:23), such issues are problematic especially since the remotest phases proposed are so long that an adequate description becomes rather difficult. In the present work, we opted for Wang Li's partition (1980:35), not among the most recent ones but, still, quite apt for our purposes:

- a. Old Chinese (上古漢語 *Shàngǔ Hànyǔ*), spanning from the oldest attestations of the language (ca. 1200 BCE) up to the end of the Han Dynasty (3rd cent. CE);
- b. Middle Chinese (中古漢語 *Zhōngǔ Hànyǔ*), from the 4th to the 12th centuries;
- c. Old and Middle Mandarin (近代漢語 *Jìndài Hànyǔ*, lit. 'Modern Chinese'), from the 13th century up to the First Opium Wars (19th cent.);
- d. a transition period, from 1840 to 1919;
- e. Modern Chinese (現代漢語 *Xiàndài Hànyǔ*, lit. 'Contemporary Chinese'), up to the present day.

This is one of the simplest possible partitions of Chinese into historical stages and, also, appears to be the most entrenched in the lexicographic tradition. Among other proposals, one may quote e.g. Sun C. (2006:17-18), which sets the end of Middle Chinese around 960 CE (the first year of the Song Dynasty) and terms the following stage 'Early modern Chinese'; see Shi Y. (2002:20-21) for further proposals.

We said above that Wang Li's partition is quite apt for our purposes; this is because we are not concerned with some particular change in the syntax of the language and, thus, a more detailed subdivision would prove to be superfluous. Also, we chose not to consider a Proto-Chinese language at all, as our research is concerned with developments in word formation which are strongly related to the written language.

Apart from the diachronic stages mentioned above, we shall also be using the term 'Classical Chinese', a familiar one for anyone involved in Sinological studies. 'Classical Chinese' (文言 *wényán* 'literary language' or

古文 *gǔwén* ‘classical Chinese prose’) is a conventional term used when referring to the written language since the 5th century BCE, which had become a model for writing also for the times to come (Norman 1988:83, Pulleyblank 1995:3-4). It is not to be regarded as a *stricto sensu* synonym for ‘Old Chinese’; the Classical language was modeled after the Confucian classical texts, and was used for the major literary genres in Confucian culture, i.e. poetry and essays, and it was practically never used for oral communication (see Pulleyblank 1995:3-4, Biasco, Wen & Banfi 2003:38). Thus, Classical Chinese is based upon writings in Old Chinese, but this obviously does not entail that the whole stage of Old Chinese can be represented by the classical language; indeed, Classical Chinese was used as an official language in China until the beginning of the XX Century, but it was by no means “completely static and uniform” (Pulleyblank 1995:4); rather, one finds differences between different historical periods and different authors, and also between different styles.

### 1.1.2 The Phonology of Standard Modern Mandarin

In Chinese, as it is known, virtually each graphematic unit, namely a Chinese character (漢字 *Hànzi*) corresponds to a syllable<sup>3</sup>. The syllable, in turn, tends to correspond to the morpheme, and thus the syllable represents the “foundation” of Chinese words, which are made up of one or more syllables: “[t]he foundation of a Chinese word is the set of monosyllables available to the language. All words in the vocabulary are built on these monosyllables” (Yip P. 2000:20).

Traditionally, Chinese syllables are divided into initial (聲母 *shēngmǔ*) and rhyme (or final; 韻母 *yùnmǔ*). Modern Mandarin Chinese has a set of 21 initials and 35 rhymes; these, however, cannot be combined freely (see Yip P. 2000:24-25), syllable structure is quite simple. The only possible combinations of sounds in a syllable are (Yip P. 2000:20; “V” stands for “vowel”, “C” for “consonant”):

---

<sup>3</sup> The only exception in this respect is represented by 兒 *ér*, a character which is also used to represent the subsyllabic *r* sound; in such usage, the character is sometimes written in a lower case, to avoid confusion with the ‘canonical’ reading *ér*.

- a. V
- b. CV
- c. VC
- d. CVC

Moreover, the only possible coda consonants are [n] and [ŋ]. Mandarin syllables are thus short and simple; the whole inventory of Mandarin syllables amounts to 405. The four tones of Standard Mandarin add some more distinctions; however, not all syllables are attested in all of the four tones (for instance, *kan* is apparently never uttered in the second tone). Even if all the 405 syllables were actually attested in four different tones, the total number of distinct syllables would be little more than 1,200; this is a relatively low number, if compared, for instance, to English, for which estimates are around 8,000 (Lin H. 2001:27-9; cf. DeFrancis 1984:15).

The (relatively) simple structure of the syllable in Modern Mandarin has been a relevant factor in the evolution of the domain of word formation, as we shall see in CHAPTER 3. As far as the modern language is concerned, an obvious consequence of the low number of distinct syllables, given also the abundance of morphemes, is the phenomenon of diffuse homophony. We shall get back to this point in 1.1.4.

### 1.1.3 Aspects of Mandarin Morphology and Syntax

Although Mandarin Chinese belongs to the isolating language type, this does not mean that the language is devoid of morphology<sup>4</sup>. Modern Mandarin lexicon is rich in multi-morphemic words, which amount to around 80% of the total, according to one estimate (Xing J. 2006); someone went so far as to define Mandarin as “a language of compounded word” (Lin

---

<sup>4</sup> Interestingly, Old (or Classical) Chinese is often regarded as the prototype of the isolating language. However, the progress made in the reconstruction of the Old language has made possible to demonstrate that Old Chinese *had* morphology, and the typology of morphological processes was indeed richer than in Modern Chinese (see, among others, Baxter & Sagart 1998, Sagart 1999, Pulleyblank 2000).

H. 2001:62; cf. Arcodia 2007). Thus, multi-morphemic words are often regarded as compounds in the literature (*contra* Packard 2000). ‘Compounds’ make up the great part of Mandarin lexicon, and compounding is apparently the most productive means of word formation; as suggested by Ceccagno and Basciano (2007:208), “[i]n Chinese compounding seems to be the rule in the formation of new words”.

Here, however, we shall take a neutral stance on the issue of the compound status of multi-morphemic words in Chinese, and we shall refer to any word which is made of more than one morpheme as a ‘complex word’. This is also because one of the main points in our research will be to set a distinction between compounding and (productive) derivation in Chinese word formation.

Needless to say, to give a complete description of Mandarin morphology is far beyond the scope of this introductory paragraph; here we shall just provide a few representative examples of the kind of complex words which are attested in the Modern Language, even though not all of them have been built with a productive pattern<sup>5</sup>.

(1) A word is made of morphemes in a coordinate relation, often (quasi-)synonymous:

長短	寒冷	能夠
<i>chángduǎn</i>	<i>hánlěng</i>	<i>nénggòu</i>
long-short	cold-cold	can-be.up.to
‘length’	‘cold’	‘can, to be able’

(2) A word is made of morphemes in a modifier-modified relation:

大人	淺藍	復習
<i>dàrén</i>	<i>qiǎnlán</i>	<i>fùxí</i>
big-man	light-blue	again-learn
‘adult’	‘light blue’	‘to revise’

---

<sup>5</sup> Here we shall adopt the classification of compounds proposed by Bisetto & Scalise (2005) which identify the three subclasses of coordinate, attributive and subordinate, according to the relation which holds among the constituent morphemes.



(3) A word is made of morphemes having a subordination (argument-head) relation:

毒販	桌腳	扶手
<i>dúfàn</i>	<i>zhuōjiǎo</i>	<i>fúshǒu</i>
drug-vendor	table-leg	support-hand
‘drug dealer’	‘table leg’	‘handrail’

According to the traditional classification of morphological types, an isolating language should have a very low index of synthesis, i.e. morpheme and word should tend to a 1:1 ratio. The data presented above clearly demonstrates that this is not always the case and that, on the contrary, it is by far more common for a word to be composed of more than one morpheme. Also, as already mentioned, in Modern Chinese the creation of multi-morphemic words to provide new ‘labels’ is the standard, as new morphemes (i.e. new characters) are never (or seldom) created. This is not to say that Mandarin is not an isolating language; as reminded by Goddard (2005:6),

[i]t’s important to point out that a language can be isolating and still have complex word forms. Being an isolating language is not a matter of the internal complexity of words, but rather of how words behave when they occur in different grammatical contexts. Some of the classic isolating languages of Asia, such as Mandarin Chinese and Vietnamese, have a high proportion of complex words formed by compounding or by reduplication.

Banfi (2005) has gone so far as to suggest a typological drift for Chinese towards the agglutinating type, especially since some of those complex words of Mandarin Chinese may be regarded as derived words, containing grammaticalized (proto-)derivational affixes. One good instance of such phenomenon is the morpheme 性 *xìng* ‘nature, character, disposition’, which will be dealt with extensively in 1.3.1.2. Let us give but a few examples of complex words built having 性 *xìng* as a constituent:

- (4) a. 重要性                                    ‘importance’  
       *zhòngyàoxìng*  
       important-*xìng*

- b. 可能性                                   ‘possibility’  
     *kěnéngxìng*  
     possible-*xìng*
- c. 性格                                   ‘nature, disposition’  
     *xìnggé*

In exx. (4a-b), the morpheme 性 *xìng* ‘nature, character, disposition’ is the rightmost constituent of the word, and has the same function in both examples, namely, that of building a deadjectival noun. In (4c), the morphemes 性 *xìng* and 格 *gé* ‘pattern, style’ are conjoined to form a complex (compound?) word, with an attributive structure (cf. ex. 2); 性 *xìng* contributes with its proper, core meaning to the complex word. Hence, examples as (4a-b) are often treated as instances of derivation (cf. e.g. Chen R. 1986, Luo J. 2004), whereas words as (4c) are normally analysed as compounds; so, a morpheme like 性 *xìng* is seen as possessing more than one identity, being sometimes a lexical root and sometimes a derivational affix.

Since the distinction between compounds and derived words will be the core theme of the present book, we shall not deal with it here any further; let us now just give but a few remarks on the aspects of Mandarin syntax which are most relevant for the purposes of our argumentation.

Mandarin Chinese has as its basic order of constituents SVO, but it deviates in many respects from the ‘ideal’ Verb-Object type, having the AN and GN orders in the noun phrase and putting adverbs before the verb (Dryer 2003)<sup>6</sup>. Many linguists believe that the order of constituents in syntax and the kind of relations instantiated among them are reflected in compound formation (see e.g. Yip P. 2000:90 ff. and Beutel 2005); here we shall not discuss such issue, and we shall just remark that the order modifier-modified has apparently always been a stable feature of Chinese, both in syntax and in word formation (Bisang 2001).

---

<sup>6</sup> A general presentation of Mandarin syntax in a typological perspective may be found in Li & Thompson (1981).

The isolating character of Mandarin is clearly visible in its grammatical morphology. We have no obligatory marking of gender, number (except for personal pronouns) or case in nouns and adjectives; there is no verbal tense and only aspect is generally marked. The aspect markers of Mandarin may be regarded as clitic particles, but in Chinese linguistics they are often regarded as suffixes, as e.g. the perfective marker -了 *-le* and the progressive marker -著 *-zhe*. As we shall see in CHAPTER 2, it is this kind of markers, mostly, which have been the core issue in grammaticalization studies for Chinese, rather than ‘our’ lexical derivational morphemes (see e.g. Sun C. 1996, Shi & Li 2001). However, aspect markers and, generally speaking, all markers of inflectional categories fall beyond the scope of our study and, therefore, we shall not discuss them any further. Let us now introduce briefly the Chinese script and some issues related to the relationship between units of writing, sound and meaning.

### 1.1.4 The Chinese Script

The Chinese script is a system of logograms, usually referred to as ‘Chinese characters’ (漢字 *Hànzi*). A Chinese character, as mentioned above (1.1.2), corresponds (almost) always to a syllable; the character / syllable tends to correspond to a morpheme, as around 90% of Chinese characters represent a morpheme (Wang F. 1998:3). A word, as said before, may be made of one or more morphemes and, therefore, of one or more characters / syllables. Below are the possible relationship among morpheme, character and word (cf. Lin H. 2001):

(5) a character / syllable corresponds to a monomorphemic word

→ 書 *shū* ‘book’, 懂 *dǒng* ‘to understand’

(6) two or more characters / syllables correspond to a monomorphemic word

→ 葡萄 *pútáo* ‘grape’, 奧林匹克 *Àolínpǐkè* ‘Olympics’;

(7) two or more characters / syllables correspond to a multi-morphemic word

→ 手機 *shǒujī* ‘mobile phone’, 賽馬場 *sàimǎchǎng* ‘horse race ground’

The configuration in (6) is quite uncommon in Modern Mandarin, whereas that in (7) is the standard nowadays.

As briefly mentioned in 1.1.2, there is massive homophony among morphemes in the lexical inventory of Modern Mandarin, which is not surprising, given the relatively low number of distinct syllables in the language. Only 297 out of ca. 1200 (theoretical) syllables of Modern Mandarin correspond to only one morpheme, whereas well above 70% of the total number of distinct syllables have at least two meanings. Very often, different meanings correspond to different characters, eliminating the potential ambiguity in the written language (Lin 2001:9 and 85).

Let us take the syllable *yì* as an example:

(8)	億	易	譯	異	藝
	100,000,000	easy	translate	different	art

All of the five characters in (8) correspond to the same syllable, *yì*, but each of them has a different meaning, i.e. it represents a different morpheme. The same characters may also be used, sometimes, to write different morphemes: 易 *yì*, for instance, means ‘easy’ in 容易 *róngyì*, but conveys the meaning ‘change’ in 易經 *Yìjīng*, the original name of the ‘Book of Changes’, a Confucian classical text.

Having provided some basic information on the aspects of Mandarin Chinese which are relevant for the purposes of our research, in the next paragraph we shall introduce the phenomenon which we shall investigate, namely lexical derivation.

## 1.2 Derivation in Word Formation

As stated in the introduction, this book deals primarily with (proto-)derivation in Mandarin Chinese, both in a synchronic and in a diachronic perspective. In this section, we shall broadly define ‘derivation’ as a morphological process, before turning to the specific topic of lexical derivation.

We learned from the general linguistic literature that derivation is a morphological process which results in the creation of a new word from an

existing one (cf. e.g. Beard 1998:55). This is true also of compounding; the difference lies in the means, as compounding involves the combination of words or, rather, lexical morphemes, whereas compounding typically involves the adding of an affix to a lexical morpheme<sup>7</sup> (Naumann & Vogel 2000).

Derivation is a category with somewhat blurred boundaries, both with respect to inflection and to compounding. As Nauman & Vogel (2000:929) put it, “[i]nflection, derivation and the lexicon seem to merely represent central points on a more general underlying continuum, ranging from grammar to lexicon”; along such continuum, derivation is more closely related to the lexicon, whereas inflection is closer to grammar (cf. Bybee 1985:82).

The borderline between derivation and compounding will be the main topic of the present research; let us postpone the discussion on such issue to 1.2.2. The dividing line between derivation and inflection has been a major subject for research<sup>8</sup>; for instance, Beard (1998) begins his paper on derivation by discussing the issues related to the distinction between derivation and inflection. It is far beyond the aims of this book to contribute to the discussion on the borderline between derivation and inflection; here we shall just try to draw from the relevant literature insights on the nature of derivation, summarizing the main characteristics of such word formation process.

As mentioned above, any word-formation process which builds a new word by adding a non-lexical morpheme might be regarded as derivation (Beard 1998:55). Inflection, on the other hand, consists typically in the specification of grammatical information on a lexeme, as e.g. gender and number for nouns and adjectives, tense and mood for verbs, etc. Such definitions may be easily challenged; to give but one example, we have cases when an inflectional process alters the lexical category of the base

---

<sup>7</sup> This does not mean that affixation is the only attested formal device for derivation; other processes as *Ablaut*, as well as tone change and other suprasegmental alterations may be involved in derivation (for a list, see Naumann & Vogel 2000:934 ff.). However, as our research deals specifically with the topic of the grammaticalization of lexical morphemes into derivational affixes, only affixation will be considered.

<sup>8</sup> See e.g. Scalise (1988), Dressler (1989), Plank (1994), Beard (1998), Naumann & Vogel (2000), Haspelmath (2002).

word, thus building a new word (ex. adapted from Haspelmath 1996:44):

- (9) *Der im Wald laut singende Wanderer*  
 The in-DAT.SG.M forest loudly sing-PTCP.PRS wanderer  
 “The wanderer who sings loudly in the forest”

In (9), present participle inflection turns the verb *singen* ‘to sing’ into an adjective. Even though a ‘watertight’ separation of inflection and derivation appears to be a challenging issue, still much research has been oriented to provide criteria for that; this is what may be termed the “dichotomy approach”, as opposed to the “continuum approach”, whereby the prototypical cases of inflection and derivation are defined, with no clear boundary between those two phenomena (Haspelmath 2002:77-82).

In the present study, we are rather inclined towards a continuum approach; the phenomena which fall under the label of ‘lexical derivation’, anyway, are not among those borderline cases, especially since, as we shall see, mostly convey relatively ‘concrete’ meanings. Let us now provide a selection of the most relevant properties of (prototypical) derivation as opposed to (prototypical) inflection; the selection has been drawn from four relevant works on the topic (Scalise 1988, Dressler 1989, Plank 1994, Booij 2006)<sup>9</sup>:

- a. syntactic properties: derivational suffixes are heads, whereas inflectional suffixes are not (Scalise 1988:567-8);
- b. functional properties: derivational morphology has the function of enriching the lexicon, whereas inflectional morphology cannot do so (Dressler 1989:6);
- c. semantic properties: derivational morphology alters the conceptual meaning of the base word, whereas inflectional morphology adds grammatical information (Scalise 1988:563); the meaning conveyed by inflectional morphology is more abstract / relational than that conveyed by derivational morphology (Dressler 1989:7); the kind of meaning conveyed

---

<sup>9</sup> See, also, Haspelmath (2002:71ff.) for an overview on such issue.

by derivational categories<sup>10</sup> is relatively concrete and non-relational<sup>11</sup> (Plank 1994:1672-1673);

d. structural and formal properties: the competition among different rules is typical of derivation (e.g. Eng. *-ness* vs. *-ity*), but quite rare in inflection (Dressler 1989:6); cumulative exponence is rare for derivation (Plank 1994:1675); the internal structure of derivational marker is similar to that of free morphs in the language, whereas it is not so for inflectional markers (Plank 1994:1676);

e. openness vs. closeness of the class: inflectional categories constitute a relatively small, cross-linguistically quite common set, whereas the meanings conveyed in derivation are an open set and many among them are attested in one or few languages (Dressler 1989:6, Plank 1994:1676; Booij 2006; cf. Bauer 2002).

As we shall see in the next section, most instances of lexical derivation apparently conform to the prototype of derivation sketched above.

As far as the distinction between inflectional and derivational morphology is concerned, we want to remark one last point, namely that inflectional morphology is organized in paradigms, whereas derivation is not. That is to say, grammatical information conveyed by inflectional morphology is organized in categories, as case, gender, tense, aspect (see footnote 9); one value of each category as, say, genitive case or past tense must be chosen, when required by the context. Inflection is, therefore, obligatory (cf. above, b.), whereas derivation is not; derivation is employed to build a new word (see Haspelmath 2002).

Although not all the attested inflectional categories are present in every language, needless to say, there is a set of categories which are cross-linguistically frequent and consistent, as number for nouns and adjectives and aspect for verbs (cf. above, point e.); inflectional

---

<sup>10</sup> Here Plank uses the term category in a rather different sense from that of grammatical / inflectional category; in the latter case, a category has its values; for instance, the category 'gender' has the values 'masculine', 'feminine' and 'neuter'.

<sup>11</sup> Here 'concrete' vs. 'abstract' is to be understood as a purely semantic notion, whereas 'non relational' vs. 'relational' are concepts related to syntagmatic relations.

morphology may therefore be compared across languages. The different meanings expressed by derivation, on the other hand, are virtually unlimited, i.e. one may imagine just about any meaning (albeit general enough<sup>12</sup>) to be expressed by an affix, or by other morphological means: one often-quoted example of this is the Polish affix *-ówka*, meaning ‘type of vodka made from NOUN’ (Carstairs-McCarthy 1992:187; cf. Bauer 2001b:208 and 2002:27).

Incidentally, these are the main reasons for the ‘preference’ of typological research for inflection, rather than derivation (cf. Ricca 2005:32). In other words, we have universals like the well-known “no language has a trial number unless it has a dual. No language has a dual unless it has a plural” (Greenberg’s universal n. 34; see Ricca 2005:34, Gaeta 2005:12), but it is very hard to formulate such implicatures for derivational morphology, as there are no paradigms, i.e. no categories and values. Only a few among derivational ‘categories’, i.e. meaning labels as ‘AGENT’, ‘FEMALE’ and the like, are suitable for cross-linguistic comparison; these are usually non-prototypical instances of derivation, as e.g. deverbal and expressive morphology, located at the borderline with inflection (Ricca 2005:32; see also Bauer 2002, Heine & Kuteva 2002).

Having given a broad definition of derivation, let us now turn to the delimitation of the subject of our research, namely lexical derivation.

### 1.2.1 Lexical Derivation

The term ‘lexical derivaton’ (*dérivation lexicale*) was introduced by Kuryłowicz (1936, quoted in Beard 1998:58) to label those word formation rules which add “features” to the base they are added to, as *-ery* in *bakery*; what is meant here by ‘features’ is actually ‘lexical meaning’. Our usage of the term ‘lexical derivation’ is broader, and includes also

---

<sup>12</sup> As remarked by Bauer (2002:37), “(...) nobody has found a language in which a derivational affix means ‘grasp NOUN in the left hand and shake vigorously while standing on the right foot in a 2.5 gallon galvanized pail of corn-meal-mush’ (as predicted by Rose 1973: 516)”.



other derivational processes, as we shall show<sup>13</sup>.

Basing on a review of the relevant literature, Beard (1998:57 ff.) proposes a classification of derivational phenomena into four types:

- a. “Featural derivation”, i.e. the processes which do not change the category of the base, but rather alter its “inherent features”, as e.g. gender in Rus. *student* → *studentka* ‘male student – female student’;
- b. “Functional derivation”, i.e. Kuryłowicz’s *dérivation lexicale*, namely those processes that alter the lexical semantics of the base, as It. *pizza* → *pizzeria* ‘pizza parlour’, or Eng. *employ* → *employer* (→ *employee*). As mentioned in the preceding section, this type of derivation may convey a virtually endless number of meanings, but it appears that a relevant part of such processes is based on grammatical case relations, as ‘locative’ (*pizzeria*), ‘nominative’ (*employer*), ‘accusative’ (*employee*), and so on;
- c. “Transposition”, namely “a simple change of category without any functional change”, as e.g. Ger. *freundlich*<sub>A</sub> → *Freundlichkeit*<sub>N</sub> ‘friendly – friendliness’;
- d. “Expressive derivation”, also known as evaluative morphology, i.e. those derivational processes which add meanings as GOOD, SMALL, BAD, etc., without assigning a part of speech to the base and without shifting its reference, as Rus. *dožd* → *doždik* ‘rain – light rain’; this is a rather peculiar type of derivation, since the same evaluative process may often be found in words belonging to different word classes, as e.g. the Italian diminutive suffix *-ino* in *tavolino* ‘small table’ and *giallino* ‘light yellow’ (see Scalise 1994, Grandi 2001).

In the present book, we shall regard ‘lexical derivation’, basically, as functional derivation, similarly to Kuryłowicz (1936); however, we shall also take into consideration transpositional processes, i.e. class-changing

---

<sup>13</sup> Beard’s own usage of the term ‘lexical derivation’ is even broader than ours, as it appears to include all major types of “regular grammatical derivation”; see Beard (1998:57 ff.) for further details.

morphology with no specific lexical content<sup>14</sup>. This is because we observed that many processes of transposition originate from functional derivation, i.e. there is sometimes a diachronic link between them; this leads to situations in which it is difficult to set a clear boundary between functional derivation and transposition. For instance, the German transpositional suffix quoted above, the nominalizer *-heit* (*/-keit*), is the product of the grammaticalization of an Old High German lexeme meaning ‘person, rank, manner, gender’ and other related meanings, as we shall see in further detail below (1.3.2.1). This is to say that at different point in time, or even in the present, a process of transposition may be close to functional derivation; albeit a distinction is normally possible, they are ‘neighbouring’ phenomena and it is worthwhile analysing both of them, in our perspective.

The two other types of derivation (according to Beard), namely featural and evaluative derivation, will not be considered in our study. They are both non-prototypical instances of derivation; expressive derivation does not assign a word class to the base and may be applied recursively to the same word, as in It. *tavol-in-ett-o* ‘table-SMALL-SMALL-MS.SG’. Also, evaluative morphology is apparently marginal in Mandarin<sup>15</sup>. Featural derivation involves meanings which are close to inflectional categories as e.g. gender, and we believe that it is not uncontroversial to say that such processes build a new lexeme, i.e. we are not sure that one may say that *studentka* ‘female student’ is a separate lexeme from *student* ‘male student’ (see the Russian example above, a.).

Having defined the object of our research, albeit in a sketchy fashion, let us turn to an overview of the literature on one of the core issues of our research, namely the borderline between derivation and compounding.

---

<sup>14</sup> In Kuriłowicz’s terms, transposition is ‘syntactic derivation’ (*dérivation syntaxique*).

<sup>15</sup> The well-known word forming suffixes *-子* *zi* and *-兒* *-er* originate from diminutives; this function, however, is almost completely lost in Modern Mandarin (see, among others, Wang L. 1989 and 1980). The only ‘true’ expressive affixes in Modern Chinese are *老*-*lǎo*- ‘old’ and *小*-*xiǎo*- ‘small, young’, used with surnames (老李 *lǎo Lǐ* ‘old Li’, 小王 *xiǎo Wáng* ‘young Wang’; see Dong X. 2004).

### 1.2.2 Derivation and Compounding: Delimiting the Borders

In the literature on word formation, a compound is often defined as a word made up of two or more words. This simple definition meets with two huge problems: on one hand, the problem of the input of compounding and, on the other hand, the definition of compounds as opposed to other “multi-word expressions” (Bauer 2001a:704). A further general problem is the definition of the ‘word’ itself, which is still a question open for debate; see e.g. Ramat (1990 and 2005) and Dixon (2002).

As far as the input of compounding processes is concerned, many authors do not mention the ‘word’ as the basic unit. Bauer (1998:404) defines compounds as “words (...) made up of two or more stems”, whereas according to Haspelmath (2002:85) “base lexemes” are the input of compounding. Haspelmath (2002:86), however, remarks that it is stems which combine in compounding: “Thus, we get English compounds such as *lipstick* (not \**lipsstick*), although it is used for both lips, and *child support* (not \**children support*), even if several children are supported (...)”. Lieber & Štekauer (2009:5) as well regard lexemes as the base units of compounding; the term ‘lexeme’ includes words, roots and stems, “uninflected parts of independent words that do not themselves constitute independent words”. The authors provide such contrastive examples, from Slovak:

(10) *rýchlovlak*  
‘express train’

(11) *rýchly vlak*  
‘fast train’

In the word in (10), the stem of the Slovak adjective *rýchly* ‘fast’ is used, with no inflectional morpheme and a linking element *-o* (cf. Eng. *Anglo-Russian*); also, the word as a whole has a specific meaning, ‘express train’, which has been lexicalised. In (11) we have a phrase, rather than a compound; this is because the adjective *rýchly* here is inflected for agreement and *rýchly vlak* refers to any train which goes fast,

i.e. it has no lexicalised meaning.

The opinion that uninflected bases, rather than fully-fledged words, are involved in compounding is therefore quite diffused. However, examples like *suggestions box* or *weapons inspector* (Bauer 2006:720), or It. *ufficio informazioni* ‘information office’ and *centro trapianti* ‘transplant centre’ (Terreni 2005), in which the non-head constituents are marked for plural, seem to contradict Haspelmath’s stance. In a recent paper, Bauer (2006:719) opts for the term “subword” as an all-encompassing term for the possible basic units of compounding: “(...) the forms in which the individual subwords appear may be differently defined in different languages; a citation form in one, a stem in another, a specific compounding form in yet a third, a word form in a fourth”. In other words, Bauer advocates for an idiolinguistic solution to the problem of the input of compounding; if this position has the advantage of putting no ‘Indo-European’ bias in the analysis of word formation in non-inflectional languages, it can also result in a definition way too large of a specific phenomenon, i.e. compounding.

The second issue mentioned above, namely the definition of compounds as distinct from other multi-word expression, is also an open question (see Lieber & Štekauer 2009 for an overview). To give but an example, the Italian examples quoted above, *ufficio informazioni* ‘information office’ and *centro trapianti* ‘transplant centre’, have been termed also ‘broad compounds’ (“composti larghi”; see Terreni 2005) as they are quite far from the prototypical word. Also, the contrastive examples from Lieber & Štekauer (2009) discussed above, exx. (10) and (11), may be best understood as defining a prototype; it is not clear, for instance, whether having a lexicalised, non-fully predictable meaning is a valid criterion for distinguishing compounds and phrases.

As remarked by Grandi (2006:32), the category of compound runs the risk of being turned into a ‘utility room’ (“categoria ripostiglio”), where one may store just any entity in the lexicon which resembles a syntactic object. This, however, is a secondary problem, in our perspective, since we are mostly concerned with the ‘other side’ of the border, i.e. the dividing line between derivation and compounding, rather than that between compounds and phrases. Therefore, in what follows we shall focus on the input of compounding and derivation; this will be a crucial

point in the discussion of the Chinese data.

We mentioned before (1.2; cf. footnote 7) that in our study we shall take into consideration only prefixation and suffixation as possible markers of derivation, since the phenomena of grammaticalization of Mandarin we intend to investigate here have prefixes and suffixes as their outcome. It appears that in the Chinese linguistic tradition the possibility of having derivation is strictly connected with the identification of affixes in the language, as we shall see in the next section.

In 1.2, we remarked that the distinction between processes of derivation and of inflection is a fundamental issue in the literature on morphology. In fact, in Indo-European languages we can often employ formal criteria to distinguish between affixes (bound forms) and words (free forms): above all, an affix cannot occupy a syntactic slot (i.e. it cannot act as a ‘word’); once we know what is an affix, we are basically left with the task of distinguishing inflectional affixes from derivational affixes.

In a language such as Mandarin, the greatest difficulty is met just when one wants to distinguish derivation from compounding (if at all); in Chinese, many lexical morphemes are bound (cf. Packard 2000) and both bound and free morphs have analogous formal and semantic features, generally speaking<sup>16</sup>. Inflection, as said above, is not an issue, since there are apparently no obligatory grammatical markers (with the possible exception of aspect markers). Given such a picture, it is clear that the core problem is the distinction between (possible) affixes and compound constituents.

We said above that the distinction between affixes and lexemes in Indo-European language is usually easy; however, it is not always so. As a matter of fact, even in Indo-European languages of Europe we have borderline issues as in languages like Mandarin, when it is difficult to

---

<sup>16</sup> This is not to say that in Mandarin there is never formal distinction between lexical and grammatical affixes. For instance, aspect markers lose their tone, as -过 *-guo* (experiential past) and -了 *-le* (perfective aspect). However, we have tone neutralization also in compounding: the phrase 打手 *dǎshǒu* ‘to hit the hand’ may be distinguished from the compound 打手 *dǎshou* ‘thug’ just because in the latter the second constituent is toneless (ex. from Anderson 1985:42-43). As the loss of tone is attested both in grammaticalization and in lexicalization, it is not a reliable test to determine whether a morpheme is grammatical or lexical in nature.

decide whether a morpheme is derivational or lexemic, i.e. whether it is an affix or a stem.

Such ambiguity may often be explained by looking at the origin of those bound morphemes: “since many derivational morphemes have developed from lexical morphemes, and since language is changing continuously, this clear distinction [the distinction between compounds and derivatives] is to some extent blurred” (Naumann & Vogel 2000:931). A free form like the German adjective *los* ‘free’ (as in *aller Verpflichtungen los* ‘free from all obligations’) may be a bound word-formation element in words as *hoffnungslos* ‘hopeless’. In Russian, the preposition *bez* ‘without’, as in *bez problemy* ‘without problem’, may act as a prefix in words like *bezopasnost* ‘security’, lit. ‘without danger’<sup>17</sup>.

Word-formation elements like *-los* have been defined by some authors ‘pseudoaffixes’ or ‘affixoids’ (see e.g. Naumann & Vogel 2000:931). The label ‘affixoid’ has been used also to include ‘neoclassical constituents’ (Bauer 1998); neo-classical compounding has been defined as: “a type of composition in which the units of the combination are not native stems but rather non-native roots (mostly from the classical languages Latin and Greek) such as *bio-*, *auto-*, *tele-*, *scope*, *-ology*, *phile* etc.” (Olsen 2000:901).

Booij’s definition of ‘affixoids’ (or ‘semi-affixes’) is “morphemes which look like parts of compounds, and do occur as lexemes, but have a specific and more restricted meaning when used as part of a compound” (2005:114); by such definition, therefore, neoclassical constituents are not affixoids, since they usually have no corresponding lexeme. Neoclassical constituents are bound roots, in English and in other Standard Average European languages as well, and they share the property of being bound with affixes.

However, Neoclassical compounds and ‘ordinary’ compounds have much in common: they are made of two constituents endowed with lexical meaning and despite the fact that neoclassical constituents are not free morphemes, they are easily recognised by speakers; an average speaker of English knows that *bio-* means ‘life’ in words like ‘biology’. Also,

---

<sup>17</sup> The difference between prepositional and prefixal usage is evident in governmental features: in the phrase *bez problemy*, *bez* triggers genitive marking in the noun *problema* ‘problem’, whereas, as expected, this does not happen in the word *bezopasnost*’.

neoclassical constituents, or at least some of them, may appear in different positions inside a word: compare ‘biography’ and ‘graphology’ (ten Hacken 2000:354); affixes, on the other hand, are characterised by their position in the word (prefixes, suffixes and infixes). One more characteristic which sets apart neoclassical constituents from affixes is that a neoclassical constituent may combine with a derivational affix to form a word, as *electric*, and it would be very challenging to posit the formation of a word through the combination of two derivational affixes and no lexical morpheme (ten Hacken 2000).

Another term which may be found in the literature to refer to neoclassical constituents is ‘semi-words’ (Scalise 1984). In fact, neoclassical constituents possess some word-like features, as they act as stems in word formation and they have a word class; on the other hand, they obviously cannot be classed as words, since they are not ‘syntactic words’, i.e. they cannot occupy a syntactic slot (they are not free). The notion of ‘semi-word’ has also been applied to Chinese (Ceccagno & Basciano 2009a and b), as we shall see in CHAPTER 3.

Although the separation between derivation and compounding may be a challenging issue even in the languages of Europe, it is much less of a problem than in Chinese. The above mentioned cases from SAE languages seem to be a minority, whereas in the majority of instances derivational affixes may be clearly identified; in Mandarin, borderline issues are the rule, rather than the exception. Compare the following complex words:

(12) 人當少年不努力

*rén dāng shàonián bù nǔlì*

person be youth not hard-working

‘When men are young, they are not hard-working’

(13) a. 老人

*lǎorén*

old-person

‘old man’

b. 工人  
*gōngrén*  
 work-person  
 ‘worker’

(14) a. 臺灣人  
*Táiwānrén*  
 Taiwan-person  
 ‘Taiwanese’

b. 北京人  
*Běijīngren*  
 Peking-person  
 ‘person from Peking’

The morpheme 人 *rén* ‘person’ is used as a word, a syntactically free from, in the sentence (12). In the two words in (13), 人 *rén* is the head of two compounds. The words in (14a-b) appear as formally identical to those in (13); however, complex words like 臺灣人 *Táiwānrén* and 北京人 *Běijīngren* have been regarded by some (e.g. Wang F. 1998, Yip P. 2000) as derived words (compare exx. 4a-c in 1.1.3). Typically, such treatment is motivated by the high productivity of a pattern, with a morpheme in a fixed position, contributing a consistent meaning: in the case of 人 *rén*, one could just build any noun with a place name.

If we were to look at this problem in an idiolingusitic and strictly synchronic perspective, the semantic and formal identity among the usages of 人 *rén* exemplified above would lead us to conclude that we are dealing with the same lexeme. However, analysing such question in a cross-linguistic perspective, making use of historical data, will enable us to gain a better understanding of the nature of derivation as a distinct phenomenon, and of the processes that lead to the genesis of derivational markers. We recognise the importance of compounding as the most relevant process of word formation in Modern Mandarin, both quantitatively and qualitatively (cf. 1.1.3); we also believe that it is worth



investigating phenomena which are ‘candidates’ for derivational status.

Such an endeavour may be fruitful only if we take carefully into consideration the peculiarities of Mandarin; above all, the fact that grammaticalized signs in languages of East and South-East Asia typically show no phonological alteration (Bisang 1996), as we shall see in detail in 1.3.2. In what follows, we shall review the recent morphological literature on the issue of the distinction of derivation and compounding in the World’s languages; in the following section, we shall briefly introduce the research done in Chinese linguistics on such topic.

At the beginning of this section, we pointed out that the definition of ‘word’ is an unresolved issue in general linguistics, and this has consequences on the question of the dividing line between derivation and compounding; in Bauer’s words (2005:106-7),

[g]iven the difficulty that there has been for many years in defining a word, it is not surprising that there should be difficulty with the borderline of compounding. Items which fit poorly into the category of word should also fit poorly in the category of possible compound element. (...) [I]t is items to which it is difficult or impossible to attribute a word-class which seem to cause the problems, and instances where items are in the process of gaining or losing the independence that goes with having a word-class.

Bauer discusses some items at the borderline between derivational affix and compounding constituent, Eng. ‘fishmonger’, ‘motorcade’, ‘seascape’, in which the right-hand constituent is not a lexeme of (Modern) English; the uncertainty in their status is a consequence of the lack of a precise definition of the ‘word’ (as opposed to the affix).

The problems in distinguishing between ‘words’ (or, better, stems) and affixes arise, according to Bauer, because word status and affix status are not eternal; we have many instances of free morphemes which become affixes and, less frequently, cases of affixes or ‘splinters’ of words which become fully-fledged words, as the often-quoted English examples ‘ism’ and ‘burger’. When an item is evolving towards (or away from) affixal status, it is endowed with ‘hybrid’ properties which make it hard to classify it as either a word (stem) or an affix.

Bauer (2005:106) suggests that semantic and distributional features may work as diagnostics for affixhood. He quotes Renouf & Baayen’s

treatment of Eng. ‘mock’ and ‘type’ as used in sentences like (15) and (16) (quoted from Renouf & Baayen 1998):

(15) flights of mock-literary dialogue

(16) a funky, regional blues-type version

In (15), ‘mock’ is used as the modifier of an adjective, adverbially, whereas it is normally an adjective in itself. In (16), ‘type’ is attached to a noun, rather than being used in structures as ‘a type of blues’. For both examples, there is at least distributional differentiation between their ‘standard’ use and their use as part of complex words. In other cases, as e.g. Eng. ‘-wise’ (‘resource-wise’), there is a semantic differentiation with the corresponding lexeme (Bauer 2005:100). This is not to say that we may certainly locate the above mentioned items in the domain of derivation; they “appear to be at different points along a potential diachronic development of the same kind [i.e. towards affixhood]. Although we cannot guarantee that the outcome in all of these cases will be an affix, we seem to have the relevant conditions for this to happen” (Bauer 2005:98).

In Amiot (2005), some French prepositions (*après, avant, contre, en, entre, sans, sous, sur*) for which a prefixal use is attested (*avant-guerre* ‘pre-war (years)’, *sur-exposition* ‘overexposure’) have been analysed. She holds that *contre, en, entre, sous e sur* are actual prefixes, since they never change the gender of the base, they can combine with words belonging to different classes, they form endocentric nouns and they are used to convey at least on meaning which is different from that (or those) of the corresponding preposition. For instance, she distinguishes the preposition *sur* ‘on, over’ from the prefix *sur-*, which adds a different meaning in complex words, ‘excessively, in excess’, as in *surcharge* ‘overload’ (Amiot 2005:186-187).

Booij (2005, 2007, 2009, 2010) highlights the analogies between derivation and compounding, in a synchronic perspective. He applies the basic principles of ‘Construction Grammar’ (Goldberg 1995, Michaelis and Lambrecht 1996) to word formation, treating both word formation patterns and syntactic patterns as constructions (“form-meaning-function

complexes”; Michaelis and Lambrecht 1996:216); this is the theoretical framework of ‘Construction Morphology’, to which we will subscribe in our research. In Construction Morphology (henceforth, CM), both ‘true’ derivational affixes and affixoids<sup>18</sup> are represented as ‘constructional idioms’, i.e. structures in which one slot is occupied by the affix(oid) and the other is a variable, containing semantic and categorial information, as we shall see below.

According to Booij (2007:34), “[w]ord formation patterns can be seen as abstract schemas that generalize over sets of existing complex words with a systematic correlation between form and meaning”. To give an example, the formal representation of the construction underlying all English and Dutch endocentric compounds is represented as follows (Booij 2009:201):

$$(17) [[a]_X [b]_{Y_i}]_Y \quad \text{‘}Y_i \text{ with relation } R \text{ to } X\text{’}$$

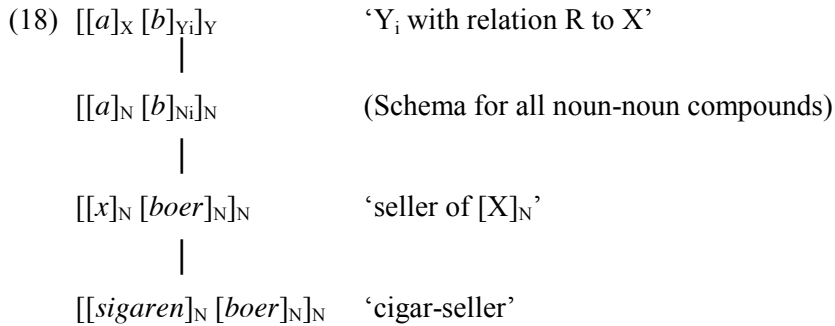
In (17), *a* and *b* stand for “arbitrary sound sequences” (Booij 2009:201), entailing that there are no phonological restrictions in the schema. *Y* is the word class variable which, of course, is the same for the head and for the whole compound, being the structure endocentric. The fact that the right-hand constituent is the head is also stated in the semantic specification of the schema: the compound IS A kind of *Y<sub>i</sub>* (and not a kind of *X*), and a relation *R* holds between the constituents; such relation cannot be predicted for all English and Dutch compounds. The identity of *Y* on the head and on the whole compound entails that the lexical category, but also other features (as e.g. gender for nouns or conjugation class for verbs) are shared.

In a CM approach, the lexicon is conceived as hierarchically ordered: “there are intermediate schemas in between the individual words and the most abstract word formation schemas, which express generalizations about subsets of complex words of a certain type” (Booij 2007:34). How would we represent the relationship between a “general” word formation

---

<sup>18</sup> We shall repeat here, for the reader’s convenience, Booij’s definition of affixoid quoted earlier: “morphemes which look like parts of compounds, and do occur as lexemes, but have a specific and more restricted meaning when used as part of a compound” (2005:114).

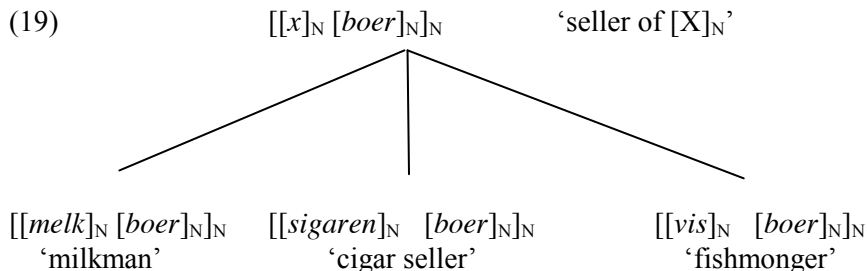
schema and a complex word as Du. *sigarenboer* ‘cigar-seller’? The first node could be that in (17), which is the one dominating all endocentric compounds; then, we would have more nodes, increasingly specific, the final node being the instantiation(s) of the actual complex word(s) (adapted from Booij 2005:125<sup>19</sup>):



The ordering of schemas reflects a hierarchy: “properties of higher nodes are percolated to lower nodes, unless the lower node bears a contradictory specification for the relevant property” (Booij 2009:206). A schema as  $[[x]_N [boer]_{N}]_N$  ‘seller of [X]<sub>N</sub>’, for instance, shares the properties of the higher nodes, as e.g. being right-headed, but has also additional specifications (the meaning ‘seller of [X]<sub>N</sub>’) which apply to what is ‘below’ it, i.e. words built according to the template, which represent the terminal nodes:

---

<sup>19</sup> Since Booij’s formalism has changed over the years, the representation in (18) has been modified to conform with his latest conventions (as in Booij 2009).



A template as that at the top of (19) is a “constructional idiom” (a notion first proposed in Jackendoff 2002); *melkboer*, *sigarenboer* and *visboer* share a common (head) constituent, *-boer*, and a specific meaning. The form *boer* is also a lexeme of the Modern Dutch lexicon, meaning ‘farmer’; only when used as the head constituent in complex words it conveys the meaning ‘seller of [X]<sub>N</sub>’, and thus is a good example of an affixoid in CM terms (cf. footnote 17)<sup>20</sup>.

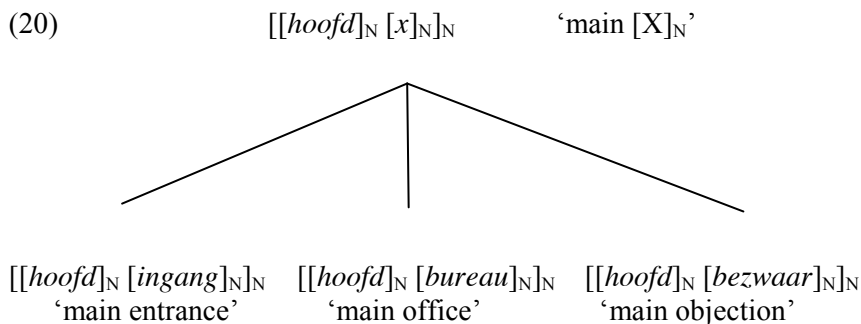
How are word-formation schemas and, thus, constructional idioms, created? In CM, both the abstract schemas and the individual (established) lexical items conforming to the schema are listed in the lexicon; the schemas actually arise from the words themselves (Booij 2009:207):

“(…) people acquire the morphological system of a language, that is, the abstract morphological schemas, on the basis of their knowledge of a set of words that instantiate these patterns. Once they have come across a sufficient number of words of a certain type, they can infer an abstract schema, and will be able to expand the relevant class of words. (...) the native speaker’s competence to create new compounds and derived words is based on abstractions over sets of existing complex words and the words that are paradigmatically related to them.”

In CM terms, the creation of a word formation schema on the basis of “paradigmatic relations” among words with a common constituent is “paradigmatic word formation” (Booij 2007:36). To give but one example,

<sup>20</sup> Note that the variable slot in [[x]<sub>N</sub> [boer]<sub>N</sub>]<sub>N</sub> ‘seller of [X]<sub>N</sub>’ can be occupied by a noun denoting a non-agricultural product (e.g. *kabelboer* ‘provider of broadband cable services’; see below, 1.3.1.2).

Du. *hoofd* ‘head’ is employed as a bound form, found as the left-hand constituent in a number of compounds, in which it bears the meaning ‘main’ (exx. adapted from Booij 2009:207):



Each of the complex words in (20) is an item of the Dutch lexicon, on the base of which the speaker ‘abstracts’ the constructional idiom  $[[hoofd]_N [x]_N]_N$  ‘main  $[X]_N$ ’. Such idiom represents a productive pattern, made of a fixed constituent (*hoofd*) and a variable slot, for which features may be specified, as e.g. noun class; the same holds for  $[[x]_N [boer]_N]_N$  ‘seller of  $[X]_N$ ’, illustrated above (19).

Two characteristic of those schemas are crucial: firstly, the fact that the schemas are productive, which is what sets a constructional idiom apart from occasional analogy; secondly, the fact that *hoofd* and *boer* are also used as words, as free forms, but with a different meaning (respectively, ‘head’ and ‘farmer’).

As for the distinction between a constructional idiom and occasional analogy, it is also worth remarking that, in the framework of CM, a particular model word is not even necessary for an idiom to be “created” (cf. Booij 2007:37, Booij 2010), whereas, as we know, a model is included in the very notion of analogical word formation.

Both *hoofd*- ‘main  $[X]_N$ ’ and *-boer* ‘seller of  $[X]_N$ ’ may be regarded as affixoids in CM, according to the definition quoted above (1.1.2), since, although they also occur as lexemes, they have “a specific and more

restricted meaning when used as part of a compound”. Affixoids resemble affixes in many respects. First, the particular meaning they convey depends on being part of a complex word, just like affixes, which cannot “exist” outside a complex word, except for a few cases of degrammaticalization/lexicalization (as Eng. ‘ism’; Booij 2009:208). Also, their behaviour is more “regular” than that of compound constituents (as for meaning interpretation and selectional properties), as we shall see below (Scalise, Bisetto & Guevara 2005). Why, thus, are they not termed just “affixes”? The difference between, say, *-boer* and a “proper” derivational affix is that for the former there is no formal difference between it and the “corresponding” lexeme in the language, i.e. *boer* ‘farmer’. Thus, affixoids in CM are conceived as bearing a word class; for true affixes, the category belongs to the schema itself rather than to a constituent, as in the one which underlies English and Dutch agentive deverbal nouns in *-er* (Booij 2007:34):

(21)  $[[x]_V \text{er}]_N$  ‘one who Vs’

Affixes, thus, are not lexical items; rather, “they only exist as parts of complex words, and as parts of abstract schemas for these complex words” (Booij 2007:34). Here we want to stress the fact that affixoids as well exist as part of schemas; they do have a lexemic counterpart, differently from affixes proper, but their occurrence with their specific meaning is limited to complex words. In our opinion, this is a relevant similarity between affixes and affixoids, as they are conceived in CM.

‘Affixoid’, however, is not to be taken as a new category between those of affixes and lexemes, but rather as a lexeme that occurs “in a subschema for compounds in which the other position is still a variable, that is, without a lexical specification.” (Booij 2005:130). Eventually, the connection between free usage and affix(oid)al usage of a form may be lost due to sound change, or because the lexeme falls out of usage: at this stage, we may just say that a new derivational affix is born. Several examples of such a development may be found in the history of individual languages: for instance, the English suffix *-dom* is generally no longer perceived by the speaker as related to the lexeme ‘doom’, although they both originate from Old Eng. *dom* (compare Booij 2010).

To sum up, according to Booij, an affixoid is a lexeme which is employed with a (consistent) different meaning in word formation, which is not available when used as a word (except for cases of degrammaticalization/lexicalization); the constructional idiom is the *locus* where the development into a derivational affix may occur. The label “affixoid” is thus descriptive in nature; in a hierarchical lexicon there are subschemas (constructional idioms) which generalise over a subset of complex words in which a particular meaning of a lexeme is used, consistently and productively (Booij 2010). We shall go back to the notion of affixoid in 1.3.2.2, discussing its relevance for the study of Chinese morphology.

What all of the approaches illustrated above on the issue of the separation between derivation and compounding have in common is their emphasis on meaning differentiation as a valid test for the grammaticalization of a free form into an affix; it is not to be given for granted, however, that such process will eventually end in producing a new affix, as pointed out by Bauer (2005) and Booij (2005).

Some more interesting remarks on distinctive features of derivational affixes and compound constituents in terms of their selectional properties may be found in Scalise, Bisetto & Guevara (2005). We shall not go into the details of their argumentation to save space; here is a summary of their proposal (2005:142-146):

- a. selection of the base by a derivational affix is fixed and constant, each affix requires the base to have certain features, whereas compound heads select the non-head in a more variable and flexible fashion;
- b. argument structure is not concerned in selection by a derivational suffix, whereas in compounding argument satisfaction is present; in subordinate compounds (as per Bisetto & Scalise 2005) as ‘taxi driver’, the non-head is the internal argument of the verb underlying the head constituent;
- c. while it is theoretically possible to calculate the number of possible (including unattested) derived words for a given process by applying the restriction on the range of bases, this is not feasible for a compounding process, as the kind of selection operated by the head is less stable;



d. a regular derived word always conveys the same meaning, while the interpretation of a compound is less predictable and can depend on the context; for instance, a compound as ‘dog bed’ would be probably interpreted by default as ‘bed for a dog’, but could also mean, in an appropriate context, ‘bed for human beings with a drawing of a dog’<sup>21</sup>;

e. typically, the base in a derived word does not receive a metaphorical reading, whereas this may happen in compounds; in a compound as ‘snail mail’, the non-head ‘snail’ stands for ‘slow’, and all other semantic features of the lexeme are irrelevant.

Now, we may reflect upon the status of affixoids in the light of the properties of derivation and in compounding outlined above. In fact, what emerges is that productive affixoids resemble more closely derivational affixes than compound constituents.

Let us take, once more, Du. *-boer* as an example. The kind of selection operated by *-boer* appears as constant, and the interpretation of the complex word follows the constructional idiom: a *kolenboer* is a seller of coal, a *tijdschriften-boer* is a seller of magazines (ex. from Booij 2007), etc. There is no metaphor involved as well: the full, literal meaning of the base (or non-head) is understood. Is it unclear to us whether it makes sense to calculate the number of possible complex words having *-boer* as the head; nevertheless, we believe that this is theoretically possible.

To sum up, the distinction between affixes and lexemes (as compound constituents) appears to be based mainly on formal criteria. An affix is a bound grammatical morpheme which may never act as a lexeme, i.e. it can never occupy a syntactic slot (even if marked with the required inflectional categories). Affixoids seem to have most of the core properties of affixes, in terms of selection, interpretation, etc.; the difference between them and ‘true’ affixes is just that the former still have a corresponding lexeme in the present stage of the language. The key assumption here is, again, that grammaticalization inevitably goes together with some alteration in the form of the sign (as per Bybee, Perkins & Pagliuca 1994);

---

<sup>21</sup> Note that, apparently, this is not a universal characteristic of compounding. In Italian, for instance, compounds typically have only one reading (Scalise, Bisetto & Guevara 2005:144).

since this is generally not true for Mandarin Chinese, as mentioned above, it is unclear whether it makes sense to distinguish affixes and affixoids. In **1.3.2.2**, we shall argue in favour of dispensing with the notion of affixoid altogether in Chinese.

In what follows, we shall deal with the diachronic aspect of the theme of our research, namely grammaticalization theory.

### **1.3 Derivation and Grammaticalization<sup>22</sup>**

As our research is concerned primarily with processes of evolution of lexemes into derivational affixes, we cannot avoid introducing the aspects of the theory of grammaticalization which are most relevant for our purposes. However, as we shall see, lexical derivation apparently has never been a central issue in grammaticalization studies; we believe that this is especially because the kind of meaning involved in lexical derivation is often too ‘concrete’, or ‘lexical’, to fit in the general picture of the genesis of grammatical morphs.

Also, we shall highlight that, as mentioned before, one of the best-known correlates of grammaticalization, i.e. the “dynamic coevolution of meaning and form” (Bybee, Perkins & Pagliuca 1994:20) does not seem to be a universal feature of such processes; in fact, it is not generally so in Mandarin Chinese, as well as in other languages of East and mainland South-East Asia (Bisang 1996, 1998 and 2004). We shall illustrate this point with a couple of contrastive examples of the grammaticalization of derivational affixes in English and Chinese. Lastly, we shall devote some space to the status of lexical derivation in historical Chinese linguistics.

What do we mean by ‘grammaticalization’, then? The term ‘grammaticalization’ is supposed to have been coined by Antoine Meillet, who defined it as “l’attribution du caractère grammatical à un mot jadis autonome” (Meillet 1958, qtd. in Hopper & Traugott 2003:19). However, speculations on the origin of grammatical categories are “almost as old as linguistics” (Heine, Claudi & Hünnemeyer 1991:5)<sup>23</sup>.

---

<sup>22</sup> Sections 1.3, 1.3.1, 1.3.1.2 and 1.3.1.3 are mainly based on Arcodia (2011).

<sup>23</sup> For an overview on grammaticalization research in modern times, see Lehmann (1995:1-8) and Hopper & Traugott (2003:19-38).

Grammaticalization as a domain of research is defined by Hopper & Traugott (2003:1-2) as:

that part of the study of language change that is concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions. (...) As a term referring to actual phenomena of language, “grammaticalization” refers most especially to the steps whereby particular items become more grammatical through time.

In grammaticalization studies, ‘grammatical’ means “that which belongs to, is part of, the grammar, as opposed to, e.g., what belongs to the lexicon, to stylistics or to discourse” (Lehmann 1995:9), rather than ‘well formed’; ‘grammaticality is understood as a gradual property, a sign may be ‘less grammatical’ or ‘more grammatical’.

For the purposes of our study, we shall not be dealing with grammaticalization in general, but rather specifically with ‘morphologization’, i.e. the expression of grammatical categories by morphological means, which is regarded as the last stage of grammaticalization, followed only by reduction to zero (Ricca 2005:29; cf. Bybee, Perkins & Pagliuca 1994). Not all grammatical categories must be morphologized in a given language, and they can be expressed, for instance, in syntactic constructions, as e.g. the English progressive construction.

Also, as we shall see in detail in the next section, a discussion on the genesis of (lexical) derivational affixes involves also the notion of ‘lexicalization’, which is sometimes seen, somehow, as the opposite of grammaticalization (Brinton & Traugott 2005:87)<sup>24</sup>.

Let us now turn to the most relevant issue, namely the status of lexical derivation in grammaticalization research.

---

<sup>24</sup> Cf. Aikhenvald (2007:58): “Grammaticalization focusses on how grammatical forms and constructions develop out of lexical items. Lexicalization involves the opposite phenomenon: the development of grammatical units into lexical items.”

### 1.3.1 Lexical Derivation and Grammaticalization

As Hopper & Traugott put it, “when long written histories are available, many bound morphemes can be shown to go back to independent words” (2003:141). This is normally the case for Modern Chinese, as we shall see throughout the work. However, not everyone agrees on the point that the development of a word into a bound derivational formant is to be regarded as grammaticalization, i.e. if such processes have much in common with ‘classic’ instances of grammaticalization, as e.g. the birth of the Romance inflectional future (Lat. *cantare habeo* ‘I have to sing’ > ‘I will sing’ > \**cantar’abeo* > It. *canterò*; Norde 2009:78).

To decide whether the evolution of a lexeme into a derivational affix is to be regarded as grammaticalization, lexicalization or even as some independent process requires, firstly, a better understanding of the very notions of grammaticalization and lexicalization, and of the relationship between them (see Himmelmann 2004, Brinton & Traugott 2005). Also, as mentioned before, the status of lexical derivation, i.e. derivational phenomena conveying lexical/content meaning, rather than grammatical/relational meaning, generates much controversy as to whether they are to be regarded as grammatical morphemes.

As said in the preceding section, grammaticalization as a phenomenon of language is defined by Hopper & Traugott (2003:2) as “the steps whereby particular items become more grammatical through time”; even though such a definition has raised objections in the literature, especially because of its focus on ‘items’, rather than on constructions (cf. Himmelmann 2004:31), it is commonly accepted, and we can use it as a starting point for our discussion. The term “lexicalization” has been used to refer to many different phenomena; in Himmelmann (2004:27), five different uses for such word are listed. A typical conception of lexicalization as a historical phenomenon is “adoption into the lexicon” (Brinton & Traugott 2005:18): the English phrase *hand in the cap* became *handicap* through univerbation, a classical example of lexicalization (from phrasal to lexical; Brinton & Traugott 2005:49).

What about the pathway “lexeme > derivational affix”? Is it to be regarded as grammaticalization or lexicalization? The answer pretty much depends on what we locate in the lexicon and in the grammar, respectively.

In short, if we adopt a model of the lexicon by which derivational affixes (specifically, lexical derivation) are part of the lexicon, then the processes we are investigating here are instances of lexicalization; if, otherwise, we believe that derivational affixes are part of the ‘grammar’ (however defined), even if they convey lexical/content meaning, then we are dealing with grammaticalization (Himmelmann 2004:22-23). Thus, as expected, the evolution of lexical items into (commonly accepted) derivational affixes has been understood in the literature sometimes as grammaticalization and sometimes as lexicalization, even with inconsistencies in the treatment (as highlighted both by Himmelmann 2004 and by Brinton & Traugott 2005).

Lehmann (1989), for instance, suggests that the development of Old High German lexeme *haidus* ‘form’ into the Middle High German derivational affix *-heit* (cf. Eng. *-hood*) is an instance of lexicalization; however, in a later work, Lehmann (1995:87) cites Latin *mente* ‘mind (ablative)’ > Romance *-mente* / *-ment* ‘adverb forming suffix’ and Proto-Germanic *līko* > Eng. *-ly* as “grammaticalization of nouns” (Himmelman 2004:24; Brinton & Traugott 2005:64). Blank (2001), just as Lehmann, believes that Eng. *-hood*, which has the same Germanic etymon as Germ. *-heit* (SOED 1993), is a lexicalized item; Ramat (1992) regards the change *haidus* > *-heit* as grammaticalization (quoted in Brinton & Traugott 2005:63-64).

We have seen above (1.2.2) how the ‘idiomatization’ of a lexeme in a complex word and its evolution into an affix(oid) is understood in Booij’s Construction Morphology. As far as the relationship between (his) idiomatization, grammaticalization and lexicalization is concerned, Booij’s treatment is unclear. For instance, he uses the term “productive lexicalization” to refer, for instance, to the phenomenon by which the Maale (a North Omotic language) noun *nayi* ‘child’ conveys the general meaning of ‘agent’ when used in complex words as *bayi nayi* ‘one who brings cattle to the grazing area’, lit. ‘cattle child’, or *waari nayi* ‘goat child’ (one who takes care of goats; Booij 2010:99<sup>25</sup>). In the very same paper, however, Booij quotes many examples from Amerindian languages of “lexical affixes”, i.e. affixes with “a specific, non-grammatical

---

<sup>25</sup> The author quotes the Maale data from Amha (2001:78).

meanings”; he also states that

“[the] rise of derivational morphemes is often qualified as grammaticalization (Aikhenvald 2007: 58), since these morphemes have become affixes. Yet, if situated at the endpoint of grammaticalization, we expect these morphemes to have abstract grammatical properties, whereas a morpheme like *-dom* still has a rather specific meaning. Hence, it seems that there is a cline for such bound morphemes ranging from a more lexical to a more grammatical meaning.”

The key point, as mentioned at the beginning of the preceding section (1.3), to decide whether the genesis of derivational affixes (conveying lexical / content meaning) is to be regarded either as lexicalization or as grammaticalization, is to consider how much such processes have in common with grammaticalization and with lexicalization, especially as far as meaning is concerned. See Himmelman (2004:24):

“the real issue appears to be the question of whether the emergence of derivational formatives shares more similarities with prototypical instances of grammaticization or with prototypical instances of lexicalization (...). If this point of view is rigorously applied, it should in principle be possible to claim that for good theoretical and/or empirical reasons derivational formatives are part of the (grammarians’s) lexicon but at the same time their historical development is an instance of grammaticalization rather than lexicalization.”

Not that Himmelman provides a clear answer to the question: “[t]o decide such issue one would need detailed empirical studies on the various stages and processes involved in the emergence of derivational formatives” (2004:28). This is what we intend to do in the present work: carry out a careful and detailed analysis of historical data. For our purposes, apparently, it is not strictly necessary to decide whether derivation is inside or outside the lexicon, before we can decide whether the process involved in its genesis is actually grammaticalization or lexicalization. What is relevant, in our perspective, is the degree of similarity between ‘established’ grammaticalization (and lexicalization) and the evolution of a lexeme into a derivational affix. We shall deal with the semantic aspects of such processes in 1.3.1.2, where we shall comment on the evolution of the Old Chinese lexeme 性 *xìng* ‘nature,

characteristic' into the bound formant *-xìng* 'the quality of [X] / connected with [X]'.

In the light of the above, it will not come as a surprise that, as mentioned in the preceding section, lexical derivation has apparently never been a major issue in grammaticalization studies. Those works on grammaticalization (and lexicalization) which we quoted here mostly provide a superficial treatment of the issue; this can be explained, in our opinion, by the fundamental 'theoretical embarrassment' caused by derivational affixes which convey lexical/concrete meaning. Albeit many have proposed that affixes in lexical derivation are often the product of the grammaticalization of lexemes as compound constituents (or in collocations), such processes of grammaticalization have never been a subject for deeper investigations, to the best of our knowledge, whereas much has been written on the genesis of 'more grammatical' categories as tense, number, etc. The issue of how lexical derivational affixes come into being has been investigated mostly in the frame of morphological research on the borderline between derivation and compounding (see, among others, the essays in Dressler *et al.* 2005). In the literature, one finds a number of excellent studies on the history of present-day derivational affixes; these, however, are normally carried out on one language (or on one language family) only, and have a descriptive (rather than analytical) focus.

Also, lexical derivational affixes have mostly been ignored in comparative research. Let us take as an example the 'World Lexicon of Grammaticalization' (Heine & Kuteva 2002), a cross-linguistic study of recurrent pathways of grammaticalization, with a list of typical source and target meanings of grammaticalization processes (Appendix I and II). Among the target meanings of grammaticalization, only FEMALE and MALE may be regarded as notions (marginally) involved in lexical derivation. Also, deverbal agentive nouns are one of the very few derivational meanings which are cross-linguistically quite common, as remarked by Ricca (2005:32); according to him, they are non-prototypical instances of derivation, and they are close to the borderline with inflection.

We pointed out above (1.2) the reasons for which typology 'prefers' derivation to inflection: basically, the fact that derivation is not organized in paradigms, with 'categories' and 'values', and the seemingly unlimited

number of meanings which may be expressed derivationally, having very few derivational ‘categories’ which may be compared across languages. The same happens for grammaticalization research; as lexical derivation is ‘unfit’ for cross-linguistic comparison, it has been given little consideration in studies aimed at finding regularities (or, even, universals) in pathways of grammaticalization (as Heine & Kuteva 2002, quoted above).

Moreover, as remarked by Bauer (2002:38-9), it is not easy to find reliable data on derivation for many languages:

“brief grammatical descriptions inevitably give brief descriptions of derivation; some grammarians consider derivational morphology as something of a side issue in grammatical description (particularly if they are attempting to provide a concise description), and thus give it little attention; it is frequently unclear to the reader of a description (possibly because the categories do not easily apply to the language in question) what is inflection and what is derivation; writers of descriptions (particularly descriptions of lesser-known languages) may not have all the information to answer questions which can be answered for other languages - accordingly descriptions are not strictly comparable”.

This is true also for Mandarin: as we shall see (2.2), we find very different descriptions of derivation in different works on Modern Chinese morphology. Also, as we have shown above (1.2.2), even in the ‘familiar’ Indo-European languages of Europe the classification of a phenomenon as derivation or as compounding (or as inflection) is not always uncontroversial.

Moreover, the fact that derivation is not obligatory, differently from inflection, makes it even more difficult to analyse the processes of grammaticalization of derivational affixes with the parameters which are usually found in the literature. In the next section, we shall highlight how Lehmann’s (1995) well-known ‘Parameters of Grammaticalization’ appear as inadequate for the analysis of the genesis of lexical derivation.<sup>26</sup>

---

<sup>26</sup> A critical revision of Lehmann’s parameters has been carried out also by Bisang (2008). However, as we shall see below (1.3.2), his focus is on grammaticalization processes in general as they occur in languages of East and Mainland South-East Asia, rather than on lexical derivation, being thus very different from our analysis.



### 1.3.1.1 The ‘Parameters of Grammaticalization’

In Lehmann (1995:121 ff.) six parameters aimed at assessing the degree of grammaticalization of a linguistic sign are proposed. Such parameters of grammaticalization are based on the notion of ‘autonomy’: “the more freedom with which a sign is used, the more autonomous it is. Therefore the autonomy of the sign is converse to its grammaticality, and grammaticalization detracts from its autonomy”. Autonomy (and its reverse, grammaticality) is a gradual property; the degree of autonomy may be assessed on the basis of three major parameters, namely ‘weight’, ‘cohesion’ and ‘variability’, which are manifested along two dimensions, the paradigmatic one and the syntagmatic one.

The six parameters are presented as such in Lehmann (1995:123):

Table 1.1. The parameters of grammaticalization (Lehmann 1995)

	<b>Paradigmatic</b>	<b>Syntagmatic</b>
<b>Weight</b>	integrity	structural scope
<b>Cohesion</b>	paradigmaticity	bondedness
<b>Variability</b>	paradigmatic variability	syntagmatic variability

According to Lehmann, loss of ‘weight’ corresponds to an increased degree of grammaticalization. At the paradigmatic level, weight is ‘integrity’, i.e. the possession of “a certain substance which allows it [a sign] to maintain its identity, its distinctness from other signs, and grants it a certain prominence in contrast to other signs in the syntagm» (Lehmann 1995:126). The notion of integrity is a complex one, and may be applied both to phonology and to semantics. As far as phonology is concerned, the loss of phonological substance results in a loss of integrity; as seen above (1.2.2), affixoids apparently fail to qualify as affixes since they have no difference in the phonological form with the corresponding free morph. As far as semantics is concerned, grammaticalization is said to involve ‘desemanticization’, a notion which has two interpretations, according to

Lehmann, i.e. either a loss of (related) meanings, or the evolution from a concrete meaning into a more abstract one.

The example of the first kind of desemanticization provided by Lehmann is that of the Latin preposition *dē*, which lost its delative sense ('movement from the surface of something') in the evolution towards Romance languages, as in French *de*, conveying "the sheer notion of a relation between two entities" (Lehmann 1995:128). The second kind of desemanticization is seen, as said above, as the evolution of a concrete meaning into an abstract one (Lehmann 1995:129):

since the initial meaning is richer, more specific, it is also more palpable, more accessible to the imagination (...) and, in this sense, more concrete; whereas the meaning of strongly grammaticalized signs, such as 'of', 'will' or 'and', do not yield mental images, cannot be illustrated and are, in this sense, more abstract.

This is a fundamental point behind the notion of grammaticalization: as we shall see (2.1.1), also in traditional Chinese philology the idea that 'empty' words (i.e. grammatical morphemes) originated from 'full' words (i.e. words with lexical content) was present since the XIV century (Heine, Claudi & Hünemeyer 1991). Going back to the notion of abstract (*vs.* concrete) meaning, we may say that it may be applied easily to transpositions, i.e. processes which have the sole function of altering the word-class of the base. When one deals with examples such as Polish *-ówka* 'type of vodka made from NOUN' (quoted in 1.2), it is less clear whether the criterion of abstractness makes sense. Nevertheless, if one can identify the lexeme from which the affix originates, abstractness may be seen as a gradual property; in other words, even when the kind of meaning conveyed by a would-be derivational affix is quite concrete, one can still compare it with the meaning of the original lexeme and see if it is less concrete.

However, it appears that Lehmann's 'abstract' has a meaning close to that of 'relational'; relational meaning is typically part of the domain of inflection, rather than to that of derivation (specifically, of lexical derivation). This is a key issue in our research and we shall discuss it at length in the next section.

At the syntagmatic level, the parameter of weight is reflected in the ‘structural scope’ of a sign, i.e. “the structural size of the construction which it helps to form” (Lehmann 1995:143). One of the examples of structural scope reduction proposed by Lehmann is the grammaticalization of main verbs into auxiliary verbs as It. *avere* ‘to have’, which operate at clause level as main verbs and at phrase level as auxiliaries. The parameter of structural scope reduction may be applied to the analysis of derivational affixes as well. For instance, the Mandarin morpheme 者 *zhě* in Classical Chinese was used as a demonstrative particle (‘one who Vs’, ‘one who is ADJ’), among other functions; in the modern language, it may combine with nouns, verbs, adjectives and, also, phrases, always forming nouns<sup>27</sup>:

- (22) a. 黠武主义者  
*dúwǔ zhǔyìzhě*  
 militaristic-ideology-*zhe* ‘militarist’
- b. 参加者  
*cānjiāzhě*  
 participate.in-*zhe* ‘participant’
- c. 爱国者  
*àiguózhě*  
 patriotic-*zhe* ‘patriot’
- (23) 破坏社会治安者 (ex. from Dong X. 2004:85)  
*pòhuài shèhuì zhì'ān zhě*  
 destroy society public order *zhe* ‘disturber of public order’

---

<sup>27</sup> Incidentally, we shall remark that the fact that 者 *zhe* may combine with nouns, adjectives and verbs is a violation of a seemingly established principle as the ‘Unitary Base Hypothesis’ (Aronoff 1976) or, rather, its modified version (Scalise 1984). As we shall see (3.2.5), such freedom of combination has been documented in many recent works (e.g. Montermini 2001, Plag 2004).

Judging from the examples above, 者 *zhě* has not fully grammaticalized into a suffix, since its scope may still include a phrase; however, we believe that examples like that in (23) are residual instances of its historical usage as a particle<sup>28</sup>, as we shall argue in 3.2.5. When applying the criterion of scope reduction, therefore, one should carefully distinguish synchronically productive word formation patterns from ‘vestiges’ of some previous stage of the language, possibly limited to some specific register and/or to some diamesic variety (as, say, formal writing).

The criterion of loss of syntagmatic weight, i.e. scope reduction, partially overlaps with a well-known principle in morphological theory, i.e. the ‘Lexical Integrity Hypothesis’, prohibiting the interaction of syntax and morphology (see Lieber & Scalise 2006 for a brief history of this hypothesis). That is to say, according to the various versions of the Lexical Integrity Hypothesis, an example like that in (21) could not be a product of morphology, since an affix could not attach to a syntactic constituent (a phrase). However, a few examples of “phrasal derivation” are attested e.g. in English (Lieber & Scalise 2006):

(24) a. self-sufficient-ish

b. New Years Day-ish

(25) post digestive disorder complications

In (24a-b), the derivational suffix ‘-ish’ is attached to a (seemingly) phrasal constituent<sup>29</sup>; in (25), the scope of the prefix ‘post-’ includes the phrase ‘digestive disorder’. Examples like those presented here may also be interpreted as a further confirmation of the validity of the parameter of

<sup>28</sup> See Yuan Y. (1997) and Dong X. (2004:85-89) for a diachronic and synchronic description of the functions of 者 *-zhě* (see also Hong B. 2005).

<sup>29</sup> The morpheme ‘-ish’ could also be analysed as a degrammaticalized sign: “(...) for some speakers *ish* has become a free morpheme with roughly the meaning ‘approximately’” (Spencer 2005, qtd. in Lieber & Scalise 2006). With a cursory Google search, we found examples like *seveneightish* (shoe size), which might support Spencer’s analysis, although such usage seem not so frequent.

structural scope reduction; since lexical derivation is ‘less grammaticalized’ than, say, inflection, i.e. it is further from the prototypical grammatical categories, derivational affixes may occasionally broaden their scope. Such an issue is particularly challenging for lexicalist approaches to word formation.

Coming to the question of the distinction between compound constituents and derivational affixes, the parameter of scope reduction does not prove to be useful. Compare exx. (26) and (27):

(26) 來自中國的人

*lái zì Zhōngguó de rén*  
 come from China STR PTC person  
 ‘person who comes from China’

(27) 中國人

*Zhōngguó rén*  
 China-person ‘(a) Chinese’

In (26), 人 *rén* ‘person’ is used as a word, and it can have a phrasal modifier as 來自中國 *lái zì Zhōngguó*, with the insertion of the marker of modification 的 *de*. In (27) there is no such marker, and 人 *rén* is modified by the word 中國 *Zhōngguó* ‘China’; compare the phrase 中國的人 *Zhōngguó de rén* ‘person from China’ (and compare ex. 14b). Therefore, the scope of 人 *rén* in (27) is more limited than in (26); this can tell us whether we are dealing with syntax or morphology, but it gives us no hint as to whether 人 *rén* is to be treated as a lexical item or as a derivational morpheme.

The second major parameter of Lehmann’s, ‘cohesion’, is translated into ‘paradigmaticity’ (at the paradigmatic level) and ‘bondedness’ (at the syntagmatic level). The paradigmaticity of a sign is “the formal and semantic integration both of a paradigm as a whole and of a single subcategory into the paradigm of its generic category”, whereas ‘bondedness’ is “the intimacy with which it [the sign] is connected with another sign to which it bears a syntagmatic relation” (Lehmann 1995:132,

147).

The criterion of paradigmaticity is not relevant since, as we have repeatedly remarked above, lexical derivation may not be arranged into paradigms. This parameter, therefore, is of little significance as far as derivation is concerned. The property of bondedness is a gradual one, going from simple juxtaposition to merger; any increase in bondedness is termed ‘coalescence’. The usual path is for a juxtaposed element to lose accent, becoming thus a clitic, which later may become a bound morph and, eventually, the boundary between this sign and the base is lost, leading to the loss of identity of the original morpheme.

A clitic stage in the pathway from free morph to affix is deemed necessary also by Hopper & Traugott (2003:142):

While there is not always evidence of a clitic pre-stage in the grammaticalization of affixes out of autonomous lexical words, the fixing or “freezing” and loss of lexical autonomy involved in the process presupposes a clitic stage. In the example of French *-ment*, Spanish *-mente* which we discussed above, and in other examples of derivational affixes such as English *-hood*, *-ly*, etc. out of full nouns, it may be assumed that at one stage the eventual affix was attracted to what came to be its future stem and came to form an accentual unit with it. (...) [i]t is the frequent syntactic collocation of a particular word class, such as noun, with a particular type of clitic, such as an adposition, that most typically leads to morphologization.

So, clitics are supposed to play a key role in processes of morphologization. A different position is held by Lehmann (1995:149-50), who provides the example of the Latin coordination marker *que*, “which is appended to the first word of the second conjunct (as in *cum in ramo sedebat caseumque devorare in animo habebat* [‘as it sat on a tree, it had the intention of eating some cheese’])”; therefore, the clitic is not always hosted by the constituent with which it has a grammatical relationship.

The criterion of bondedness has a limited significance for our research, since in word formation, as we have seen, there is no difference between compound constituents and would-be affixes; the only exception is represented by the few cases of tone neutralization in compounding (see footnote 16), which however is not a rule in Chinese word formation.

The third major parameter put forth by Lehmann is ‘variability’, i.e.

“the freedom with which the language user chooses a sign” (1995:137); which yields the criteria of ‘paradigmatic variability’ and of ‘syntagmatic variability’.

At the paradigmatic level, ‘variability’ means that a sign may be substituted by another element in the same paradigm, ‘intraparadigmatic variability’, or else that sign is omitted, and the category is left unmarked; this is termed ‘transparadigmatic variability’. In some languages, such as e.g. Burmese, there may be a degree of flexibility in assigning classifiers to certain nouns; this is an instance of intraparadigmatic variability. A reduction in transparadigmatic variability, on the other hand, corresponds roughly to the ‘obligatorification’ of a category (i.e. the marker for that category cannot be omitted). Once more, both parameters are suitable for the analysis of the genesis of typical grammatical categories, i.e. inflectional categories, which may be in competition with some other form in a paradigm and which should be obligatorily expressed. This does not apply to lexical derivation, especially as far as obligatoriness is concerned.

At the syntagmatic level, variability is understood as such (Lehmann 1995:158):

The syntagmatic variability of a sign is the ease with which it can be shifted around in its context. In the case of a grammaticalized sign, this concerns mainly its positional mutability with respect to those constituents with which it enters into construction. Syntagmatic variability decreases with increasing grammaticalization.

So, if an adverb is grammaticalized into a case affix, the adverb increasingly loses freedom of position in the sentence; whereas an adverb possibly can be located in different places in a clause, it is not so for a case affix, which is normally put either on the left side or on the right side of a word.

We quoted above (1.2.2) the case of Eng. ‘-type’, which can be analysed as an affixoid since it has a distribution different from that of the lexeme ‘type’ (*a type of* vs. *blues-type*, ex. 16). Lehmann also highlights that a different position of a grammaticalized sign from its lexical ‘forefather’ is caused both by the effect of coalescence, which causes the grammatical element to be adjacent to its ‘lexical support’, and also by processes of analogy with functionally equivalent constructions

(1995:159-160). However, processes of evolution as that of *type*, by which a word used in a syntactic structure becomes an affix, are not that frequent in Chinese. As hinted above, the trend for Mandarin (at least, for nominal suffixes) seems to be the semantic evolution of a head constituent in compounding which, in Mandarin attributive compounding, is normally the right-hand constituent (cf. Packard 2000:39, Ceccagno & Scalise 2006:255). Thus, the real issue is the distinction between the two domains, namely compounding and derivation. However, the criterion of syntagmatic variability may be useful in the analysis of languages as Italian, where attributive compounds are normally left headed (and the adjective normally follows the head noun in syntax), whereas typically lexical derivational morphemes are suffixes.

To sum up, in this section we have proposed a review of Lehmann's 'Parameters of Grammaticalization', and we have tried to show that they have been designed with inflectional morphology in mind, and some of them are inadequate for the analysis of processes of morphologization of 'our' derivational affixes. In the next section, we shall discuss the semantic correlates of the morphologization of affixes, to assess whether they fit in the notion of 'desemanticization', as it is conceived in the literature on grammaticalization.

### 1.3.1.2 Grammaticalization and 'bleaching'

In the research on grammaticalization, it is commonly assumed that meaning 'fades away' in the evolution towards grammar. In the preceding section, we quoted Lehmann's term 'desemanticization'; a common term used to denote the notion of the 'weakening' in meaning is 'bleaching', a notion reminiscent of von der Gabelentz's *verbleichen* and Meillet's *affaiblissement* (Hopper & Traugott 2003:94). This correlate of grammaticalization has been given different names in the literature, and definitions vary as well<sup>30</sup>; we shall be using the term 'bleaching' in what follows as a convenience term, to indicate generally any conception of the weakening in meaning in grammaticalization.

What most approaches have in common is that non-grammatical, lexical

---

<sup>30</sup> See Campbell (2001:118-9) for a discussion of different conceptions of desemanticization.



meaning is seen as concrete, material, whereas grammatical meaning is seen as abstract, relational (Heine, Claudi and Hünemeyer 1991:41-45); in traditional Chinese philology as well it was argued that “empty words” (虛字 *xūzì*), i.e. function morphemes, originate from “full words” (實字 *shízi*), lexical morphemes (Heine, Claudi and Hünemeyer 1991:5, Xing 2003:3-4). As a starting point for further discussion of the different conceptions of ‘bleaching’, we shall first give one Chinese example of the evolution of a lexeme into a (possible) derivational affix, namely 性 *xìng* ‘nature, character, disposition’ > ‘the quality of [X] / connected with [X]’.

As mentioned in 1.1.4 (ex. 4a-b), the morpheme 性 *xìng* ‘nature, character, disposition’ is found in a number of complex words in Mandarin, such as 重要性 *zhòngyǎo xìng* ‘importance’ (important-*xìng*); in such usage, it has been regarded by some as derivation (cf. e.g. Chen R. 1986, Luo J. 2004). In Old Chinese, 性 *xìng* was a lexeme, a free form, the meaning of which included ‘inherent property’, ‘immutable nature’, ‘life’, ‘temperament’; these are the meanings listed in a dictionary of the Classical language (GHYDCD 2000, my translation):

- a. ‘quality, intrinsic properties or characteristics of sthg.’ (性质，指事物所具有的本性、特点);
- b. ‘Indicates inherent properties of the human being’ (指人的本性);
- c. ‘(Buddhism) The opposite of 相 *xiàng* [physiognomy]. The inherent, inner non-modifiable properties of things, such as heat for fire, or dampness for water’ (（佛）与“相”相对。指事物内在的和不可改变的本质，如火的热性、水的湿性);
- d. ‘Biological life, vitality’ (生命、生机);
- e. ‘Disposition, temperament’ (性情、脾气).

According to Luo J.'s account (2004:91-93<sup>31</sup>), between the fourth and the third century BCE, 性 *xìng* was normally used only as a free form, and we have attestations of it as a constituent in complex words from the first century BCE (in the 史記 *Shǐjì* 'Records of the Grand Historian').

At the stage of Middle Chinese, 性 *xìng* was used in complex words, mostly as the right-hand constituent, acting as the head (10 out of 12 bimorphemic words in the 世說新語 *Shì Shuō Xīn Yǔ* 'New Tales of the World', fifth century CE), a tendency which is even stronger in a tenth century text as the *Dūnhuáng Biànwén* (敦煌變文). In the above mentioned works, the non-head constituent may be not only a noun, but also a verb or an adjective (e.g. 定性 *dìngxìng* 'quiet mind', 慈悲性 *cíbēixìng* 'benevolence, pity'; Luo J. 2004:92), which was not common in previous texts; it therefore appears that the combinatory possibilities for 性 *xìng* complex words have increased.

In the Early Modern Chinese (13th-19th cent.) texts analysed by Luo J., 性 *xìng* is more often used as a constituent in a complex word than as a free morph; Luo J. also claims that, during this period, the meaning of 性 *xìng* becomes 'emptier', i.e. more general (2004:92), although he does not make explicit what is meant by 'emptier'. However, the examples provided by Luo J., actually still seem to bear a rather concrete meaning, i.e. 'nature, disposition', which is not fundamentally different from its lexical meaning (see the list above). We found some more Early Mandarin examples in the *Academia Sinica* tagged corpus, as e.g. 急性 *jíxìng* '(of) impatient disposition', from the adjective 急 *jí* 'impatient, urgent', used as an attribute for people (e.g. in 水滸傳 *Shuǐhǔzhuàn* 'Water Margin', 14th cent.); here the meaning conveyed by 性 *xìng* is clearly still that of 'disposition, temperament'. The word 急性 *jíxìng* has survived into Modern Mandarin, but it acquired a new meaning, namely 'acute' (associated mainly with diseases; cf. Chen 1986:89); thus, in such case the complex word does not denote a stable characteristic (as one's

---

<sup>31</sup> See the source for the complete list of the texts included in Luo's sample.

disposition), but a changeable property<sup>32</sup>. We believe that this could be interpreted as meaning generalization.

What about the modern language, then? As said above, 性 *xìng* is no longer a free form, and is thus virtually never used in isolation. When used as the right-hand head constituent in a complex word, it may combine with nouns, verbs, adjectives and also with adverbs (as e.g. 经常性 *jīngchángxìng* ‘regularity’, from 经常 *jīngcháng* ‘regular’). It appears that, at least since the Middle Chinese period, there has been a word formation template such as

(28)  $[[x]N/A/V [xìng]N]N$  ‘the nature or spirit of [X] N/A/V / connected with [X]N/A/V’

The meaning conveyed by 性 *xìng* in such schema is not fundamentally different from that which it could convey as a lexeme: see e.g. a word as 佛性 *fóxìng* ‘nature of the Buddha’ (i.e. ‘the awareness of all living creatures’; HYDCD 1993).

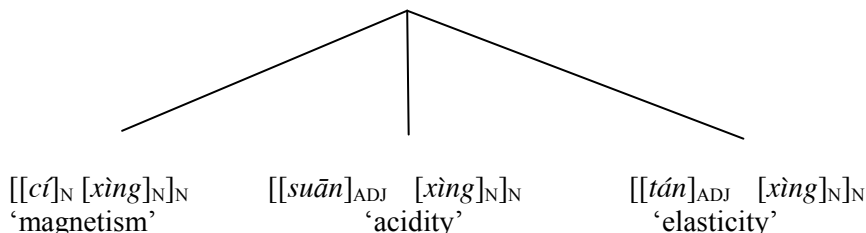
However, in Early Modern texts, a word as 忍性 *rěnxìng* ‘endurance, tolerance’ (lit. ‘endure-*xìng*’) is attested<sup>33</sup>; we believe that such a word is the instantiation of the constructional idiom quoted above, namely  $[[x]_X [xìng]N]N$  ‘the property of [X] / connected with [X]’. Complex words as 磁性 *cíxìng* ‘magnetism’, 酸性 *suānxìng* ‘acidity’, 弹性 *tánxìng* ‘elasticity’ are connected to such schema:

---

<sup>32</sup> *Contra* Luo J. (2004:94; my translation): “In all of its meanings, [-性 -*xìng*] always designates an inherent property’, and this conditions the choice of what abstract word or phrases may be the ‘X’ [i.e. the ‘base’] before ‘*xìng*’”.

<sup>33</sup> One attestation of 忍性 *rěnxìng* (seemingly) meaning ‘endurance’ was found in “The Story of a Marital Fate to Awaken the World” (醒世姻緣傳 *Xǐngshì Yīnyuàn Zhuàn*, XVII cent., from the *Academia Sinica* corpus).

(29)  $[[x]_X [xìng]_N]_N$  ‘the property of [X] / connected with [X]’



In the words in (29), 性 *xìng* does not appear to convey the meaning ‘nature’ or ‘spirit’, but rather it has become a morpheme forming abstract nouns (cf. Chen 1986:89). If we take, for instance, the Modern Chinese word 必然性 *bìránxìng* ‘inevitability, certainty’, we may note that the non-head 必然 *bìrán* ‘inevitable, certain’ is an adjective, but it is also a noun (‘necessity’) in itself; thus, the function of -性 *-xìng* here is that of carrying a word class (and the [+abstract] feature), not affecting the meaning of the whole word in any other way.

The historical evolution of -性 *-xìng* may be sketched as such:

(30) ‘the nature or spirit of [X]<sub>N</sub>’ > ‘the nature or spirit of [X]<sub>N/A/V</sub>’ > ‘the property of [X] / connected with [X]’ / ‘suffix forming abstract nouns’

It must be pointed out that -性 *-xìng* cannot be always interpreted as a mere indicator of noun class:

(31) 爆发 *bàofā* ‘to burst out’ → 爆发性 *bàofāxìng* ‘explosiveness’

If the function of -性 *-xìng* were just that of building an abstract noun, 爆发性 *bàofāxìng* could well mean ‘explosion’ rather than ‘explosiveness’. We believe that the notion of ‘property’ is still present in the constructional idiom underlying -性 *-xìng* complex words (‘the property of [X] / connected with [X]’); this explains why such complex

words cannot be associated, for instance, with an event reading.

Having sketched the diachronic evolution of the lexeme 性 *xìng* into the word-formation element -性 *-xìng*, let us discuss the different conceptions of bleaching which we find in the literature. We shall go back to -性 *-xìng* in **1.3.2.1**, where its evolution will be compared to that of Ger. *-heit* / *-keit*, having an analogous function in word formation.

As said at the beginning of this section, the process of grammaticalization is most often understood as entailing some shift from a ‘concrete’ meaning into an ‘abstract’ meaning; this idea is found also in traditional Chinese philology. Hopper & Traugott (2003:94) invoke the notion of “pragmatic enrichment”, which should occur in the early stages of grammaticalization, with some “redistribution” or “shift” in meaning; however (my italics),

“[t]here is no doubt that, over time, meanings tend to become weakened during the process of grammaticalization. (...) As grammaticalized forms become increasingly syntacticized or morphologized they unquestionably cease over time to carry *significant semantic or pragmatic meaning*.”

However, in the kind of derivational phenomena considered here, i.e. lexical derivation, something akin to lexical/content meaning rather than “purely” grammatical meaning is involved and, this, we cannot expect some sort of *absolute* abstraction of meaning (cf. the Polish affix *-ówka*, meaning ‘type of vodka made from NOUN’, **1.2**); rather, we must envisage some notion of ‘relative’ abstraction. Our proposal is that when a lexeme develops a new meaning, available when used in word formation, with a fixed position and with stable selectional properties, if (and only if) the meaning conveyed by such constituent is ‘more abstract’ than when it is (or was) used as a lexeme, then we are dealing with a process of grammaticalization and a new derivational affix is born. This holds even if there is no formal differentiation with the original lexeme.

How do we understand (relative) ‘abstraction’, then? The term ‘abstraction’ has already been used in the analysis of the evolution from a concrete meaning to an abstract one by Heine, Claudi and Hünemeyer (1991:43-45); in their work, such term has several possible readings. Those which seem to be most interesting in our perspective are

*generalizing abstraction* and *isolating abstraction*: the former is defined as “reducing the number of distinguishing features of a concept to its most ‘central characteristics’ or ‘nucleus’”, the latter “separates one particular property or feature that is not necessarily the ‘core’ or ‘nucleus characteristic’ of that concept” (Heine, Claudi and Hünemeyer 1991:43). The process of *generalizing abstraction* involves taxonomic reasoning: a lexeme is taken to a higher taxonomical level (hyperonymy: *cork-oak* → *oak* → *tree* → *plant*); *isolating abstraction*, on the other hand, corresponds to the identification of the whole lexeme with one of its features.

Let us try to apply such notions to one of the Dutch example seen above (1.2.2), *boer* ‘farmer’ > ‘seller of [X]<sub>N</sub>’. We shall use the metalanguage of *Lexical Semantics* (cf. Lieber 2003) for a tentative representation of the “body” of the lexeme *boer*, i.e. its encyclopaedic features:

- (32) <worker>  
 <agriculture>  
 <runs a farm / works in a farm>  
 <**sells agricultural products**>

In its usage as an ‘affixoid’, i.e. as a bound constituent in the constructional idiom  $[[x]_N [boer]_N]_N$  ‘seller of [X]<sub>N</sub>’, one of the encyclopaedic feature defining the corresponding lexeme is ‘isolated’, that is the fact that the sells agricultural products (for a living). The isolated features is not a core semantic trait for the definition of *boer*, the fact that a farmer works in a farm certainly being more relevant for the definition of *boer*’s intensional meaning<sup>34</sup>.

The process of meaning abstraction has gone further for *-boer*, since it is used also in reference to a seller of non-agricultural products, such as coal (*kolenboer*) or even broadband cable services (*kabelboer*, fn. 19; and compare *pornoboer*!); the appropriate rendering of the meaning that *-boer* conveys is ‘seller of [X]<sub>N</sub>’, as seen above. A similar process of abstraction of meaning might be operating also in the Maale case quoted above (1.3.1)

---

<sup>34</sup> A dictionary definition of *boer* is “someone whose trade consists in agriculture and/or cattle-breeding” (“iem. wiens bedrijf bestaat uit landbouw en/of veetelt”; VD 2005, my translation).

from Booij (2010:99), *nayi* ‘child’ > ‘agent (in the domain of cattle herding)’, as e.g. in *waari nayi* ‘one who takes care of goats’. Since cattle herding is typically an activity for children in the society of Maale speakers, it is likely that some encyclopaedic feature like <herds cattle> is present in the “body” of *nayi* ‘child’; such feature has been isolated in a constructional idiom. More data is needed, however, to support this analysis.

The evolution of 性 *xìng* ‘nature, character, disposition’ > -性 *-xìng* ‘the property of [X] / connected with [X]’, outlined above, may be analysed as an instance of generalizing abstraction. From the point of view of semantics, a noun indicating inherent and everlasting properties of people or things, or the disposition of a person, evolved into a nominal suffix, indicating just any property. The generalization in meaning goes together with an increase in combinatory possibilities: if the lexeme 性 *xìng* originally could combine (essentially) with nouns, it then broadens the range of ‘bases’ to include also adjectiveness and verbs. At this early stage of the process, 性 *xìng* is still used to indicate properties inherent to men or things. Given such an “environment”, 性 *xìng* further reduces (i.e. generalizes) its intensional meaning: 性 *xìng* complex words begin to be used to indicate just any property, not only inherent ones, and may be associated with all sorts of referents. The further increase in the combinatory possibilities of 性 *xìng* is proved by the fact that, in Modern Mandarin, morphemes belonging to any major word class may combine with -性 *-xìng* (albeit with differences in productivity), building abstract nouns (still conveying the meaning ‘property’).

So far, so good. But, according to Heine, Claudi and Hünemeyer (1991:43-45), neither generalizing nor isolating abstraction are involved in grammaticalization, the relevant process being *metaphorical* abstraction: a given sign undergoing grammaticalization may become endowed with a new meaning, metaphorically connected with the original one. Heine, Claudi and Hünemeyer propose the example of the Ewe<sup>35</sup> noun *vi* ‘child’, which is reportedly in the course of grammaticalization into a

---

<sup>35</sup> A Niger-Congo language of Ghana.

suffix and thus acquired several meanings, as e.g. INEXPERIENCED and UNSUCCESSFUL (Heine, Claudi and Hünemeyer 1991:79 ff.):

(33) *núfǎlǎ* ‘teacher → *núfǎlǎ-ví* ‘inexperienced teacher’

This suggests a metaphorical change, from the category of PERSON (‘child’) to that of QUALITY (‘inexperienced, unsuccessful’). This amounts to saying that an affix may develop a new meaning, entirely different (albeit related) from that of the lexeme from which it originates; this is why, according to Heine, Claudi and Hünemeyer, the ‘traditional’ conception of bleaching is not appropriate for the characterization of grammaticalization processes, just because it “implies that its output is necessarily part of its input; that is, what happens in the course of grammaticalization is that concepts are merely reduced in their intensional content while their extension is increased” (1991:43). We disagree with their stance, and we believe that the facts that the meaning shifts involved in grammaticalization may include metaphor and also metonymy and that, ultimately, they will lead to abstraction of meaning are not in principle incompatible.

Our position on the interplay between metonymy and metaphor in processes of grammaticalization is not far from Hopper & Traugott’s. They also believe that processes as metaphor and metonymy are involved in grammaticalization, and yet they do not see this as a reason for setting aside the notion of bleaching which, as mentioned above, they see as a necessary correlate of grammaticalization. They also explicitly state that “in grammaticalization (...) the meaning will always be derivable from the original lexical meaning by either metaphorical or conceptual metonymic inferencing. Therefore, meaning changes in grammaticalization are never arbitrary” (Hopper & Traugott 2003:94-5).

Some explanation on how metaphor and metonymy are involved in the genesis of derivational affixes are due; however, we believe that the understanding of the above mentioned issue largely depends on what notions of metaphor and metonymy we assume. Going back to the Dutch example *boer* ‘farmer’ > *-boer* ‘seller of [X]N’, one could well say that metonymy is involved, as in isolating abstraction there is a conceptual association by contiguity (‘a farmer makes a living by selling agricultural



products' > 'a farmer is a seller'; cf. the features represented in 32).

Let us now review Heine, Claudi and Hünemeyer's treatment of the above mentioned Ewe *vi* 'child' > *vi* 'suffix'. We said above that one of the meanings which the suffix can convey is that of INEXPERIENCED; another meaning is that of MEMBER, "within a political, sociocultural, or geographically defined community" (Heine, Claudi and Hünemeyer 1991:85):

- (34) a. *Eβe* 'Ewe' → *Eβe-vi* 'an Ewe'  
 b. *du(me)* 'village' → *dume-vi* 'a native of a village'  
 c. *Tógó* 'Togo' → *Tógó-vi* 'a native of Togo, a Togolese'

Heine, Claudi and Hünemeyer suggest that the semantic structure (the body?) of 'child' consists of two basic components, namely YOUNG and DESCENDANT-OF; different "channels of conceptual expansion" are involved in the various meanings which the suffix may convey (1991:86). A metaphor from PERSON to QUALITY underlies meanings conveyed by the affix as e.g. SMALL, INSIGNIFICANT, TYPICAL BEHAVIOUR; the changes from one meaning category to another is not abrupt, but follows intermediate steps (1991:89; compare figure 3.2, p. 87). The intermediate steps "are contiguous, or metonymous, [but] they nevertheless show a relation to one another that can be described as being "weakly metaphorical" in nature" (1991:89).

What about MEMBER? According to Heine, Claudi and Hünemeyer's analysis (1991:84), "[t]here is another development that has the effect that the feature YOUNG, which forms one of the two main components of the noun *vi* 'child', is "bleached out" – with the result that the second component, DESCENDANT-OF, is generalized". Examples as those in (34) are said to be "the result of an analogical (metaphorical) transfer of the kind parents:child to community:individual (*ivi*). Another step in the chain of evolution of *-vi* which directly follows MEMBER is that termed TYPICAL BEHAVIOUR: "[t]he implicature that someone who is a member of a certain group exhibits behaviour that is representative of that group appears to have invited another conceptual interpretation, namely

that *-ví* also denotes a person who adheres to the TYPICAL BEHAVIOUR of that group” (1991:85-6):

- (35) a. *amedzró* ‘foreigner, alien’ → *amredzó-ví* ‘somebody who behaves like an alien’
- b. *amedáhe* ‘poor person’ → *amedáhe-ví* ‘a deplorable person, somebody who suffers because he or she is poor and therefore deserves pity and attention’
- c. *ameyibɔɔ* ‘black person’ → *ameyibɔɔ-ví* ‘somebody who shows a typical African behaviour, adheres to African values’

In a word as *amredzó-ví* (35a), the notion DESCENDANT-OF is no longer relevant, since the noun can refer to someone who acts as a foreigner, no matter whether he or she is or is not actually one. There are also some *-ví* derived words which have no significant difference in meaning with the underived noun, as e.g. *ha'metɔ* vs. *ha'metɔ'-ví*, both meaning ‘member of a club/society’; here, “the meaning of *-ví* has been bleached out” (Heine, Claudi and Hünemeyer 1991:86).

To sum up, the process of grammaticalization of the polysemous suffix *-ví* in Ewe begins when one of the two meaning “components” (virtually equivalent to encyclopaedic features, in our opinion), YOUNG, is isolated and follows a line of semantic evolution, and another component, DESCENDANT-OF, follows another line, which leads to meanings such as e.g. MEMBER (cf. 34) and TYPICAL BEHAVIOUR (cf. 35). It seems to us that processes just described are not fundamentally different from the isolating abstraction which we invoked for Du. *-boer*, as far as the early stages of the evolution are concerned. Moreover, a development as DESCENDANT-OF > MEMBER > TYPICAL BEHAVIOUR (> null?) appears to go in the direction of further generalization (cf. what was said above about *amredzó-ví*, 35a). Also, note that such a development in meaning happens in word formation, both for *-boer* and *-ví*; the semantic connection with the lexeme *vi* ‘child’ is apparently lost once the process of grammaticalization has begun, judging from the agrammaticality of the

example below (from Heine, Claudi and Hünemeyer 1991:89; italics in the source, glosses altered):

- (36) \*βu'kulá-*ví* nyé *vi*'  
 driver-*ví* be child  
 'A driver who has not yet acquired a driving licence is a child.'

The word *βu'kulá-*ví** 'somebody who knows how to drive but has not yet acquired a driving licence' is a derivative of *βu'kulá* 'driver'. The meaning conveyed by *-ví* is NOT YET PASSED AN EXAM (by the metaphor YOUNG > INEXPERIENCED > NOT YET PASSED AN EXAM<sup>36</sup>); one cannot equate *-ví* and *vi*', since a *βu'kulá-*ví** is not really a child.

Having illustrated Heine, Claudi & Hünemeyer's notion of metaphorical abstraction with an example in the domain of derivation, we want to stress the fact that their "metaphor/metonymy" approach, in our opinion, does not involve anything significantly different from isolating and generalizing abstraction. Their analysis of the shift *vi*' 'child' > *-ví* involves all sorts of abstraction: isolating (cf. the 'split' between YOUNG and DESCENDANT-OF), generalizing (cf. the discussion of ex. 35a) and metaphorical. As one can see in the evolution of Du. *boer* > *-boer*, the kind of semantic shifts by contiguity which occur in isolating abstraction may be understood as metonymy. Metaphor may be invoked to account for semantic shifts in the evolution of 性 *xìng* 'inherent nature (of people and things) / immutable inner properties of things' > -性 *-xìng* 'the property of [X] / connected with [X]': we may analyse this as metaphoric extension from 'inherent nature' to '(any) property' ('importance', 'gravity', 'regularity'), with a generalizing effect.

As said before, this depends much on the conception of "metaphor" and "metonymy" which we assume. In Hopper & Traugott's treatment of Latin

---

<sup>36</sup> "The transition from YOUNG to NOT YET PASSED AN EXAM does not qualify as being metaphorical since the latter feature is typically associated with young people. If, however, this feature is applied to an adult, as in the case with nouns like *βu'kulá-*ví** (...), then a metaphorical relation emerges between a child and an adult having a characteristic associated with children" (Heine, Claudi & Hünemeyer 1991:89).

*mente* ‘mind (ablative)’ > French *-ment* ‘adverb forming suffix’ (2003:140-1; cf. the quotation from Lehmann 1995 above, this section), which according to them is an instance of grammaticalization (specifically, morphologization), they point out that in Latin you had phrases like *clara mente* ‘with a clear mind’, whereas in Modern French *-ment* is “no longer restricted to psychological senses, but is a general adverb formative”, as demonstrated by examples like *doucement* ‘softly’. This evolution could also be interpreted as metaphor, where ‘mind’ is extended to ‘way, manner’ (‘with a certain frame of mind’ > ‘in a certain way’).

Moreover, we want to point out that Du. *-boer*, Ch. 性 *-xìng* and Ewe *-ví* all evolved in a specific environment, i.e. in a fixed position inside a complex word (with selectional restrictions). In a strict (synchronic) descriptive perspective, we would just have to posit two constructional idioms underlying the words in (34) and (35):

(37)  $[[x]N \text{ } ví]N$  ‘member of the community  $[X]N$ ’

(38)  $[[x]N \text{ } ví]N$  ‘somehow who behaves as  $[X]N$ ’

These two schemas would seem to be fundamentally distinct because of the difference in meaning. If, however, we adopt a “dynamic” perspective, the connection appears quite evident; compare the (39a) and (39b) (Heine, Claudi and Hünemeyer 1991:85):

(39) a. *Tógó-tó* ‘an inhabitant of Congo’

b. *Tógó-ví* ‘a native of Togo, a Togolese’ (cf. 34c)

According to Heine, Claudi & Hünemeyer’s account, *-tó* derived words may indicate membership “not necessarily by birth”, but *-ví* should, in principle, indicate ‘membership by birth’; a word as *Tógó-ví*, however is not only used in reference to a native of Togo, but may also refer to a non-native whose behaviour is that of the typical good Togolese, bearing representative characteristics as ‘calm’ and ‘peaceful’. The word formation schemas (37) and (38) are, therefore, connected; the range of potential “bases” for *-ví* complex words has increased, including not only

“communities” to which one may belong only by birth, and the number of possible (even if unattested) derivatives has increased, accordingly.

However, the problem of how to accommodate such polysemy in a CM approach is not resolved, and cases as that of *-vi* are not rare in the World’s languages; templates are constructions, combining form, meaning and function, and thus a change in meaning would have to correspond to a new template (which is what we proposed in 37 and 38). We nevertheless know that, in such cases, the different meanings which a polysemous affix may convey are connected (cf. the quotation from Hopper & Traugott 2003:94-5): will we, then, posit only one overarching template which may encompass all the uses of polysemous affixes, or shall we rather keep the individual schemas? We shall postpone the discussion of such issue to CHAPTER 3 (3.2.2).

To sum up, the aim of the discussion in this section was that of showing that the semantic processes involved in the evolution of a lexeme into a derivational affix are not fundamentally different from the familiar mechanisms of grammaticalization, as generalization, metonymy and metaphor. However, since the derivational phenomena considered here involve lexical/content meaning, the mechanisms of metonymy and, especially, metaphor, may operate differently from grammaticalization involving “pure” grammatical/relational categories (see Hopper & Traugott 2003:81 ff. for some examples).

Having made clear our position on the relationship between lexical derivation and grammaticalization, let us now deal with the issue of how derivation is to be related with lexicalization.

### **1.3.1.3 Grammaticalization and lexicalization**

In 1.3.1, we have briefly illustrated how the evolution of a lexeme into a derivational affix is seen by some authors as grammaticalization and by some authors as lexicalization; sometimes, it is actually the same people who regard the very same phenomenon as grammaticalization and as lexicalization in different works (see e.g. the quotations from Lehmann). We accepted Himmelmann’s (2004) suggestion that the crucial point is whether the genesis of (lexical) derivational affixes is closer to prototypical grammaticalization or to prototypical lexicalization; in 1.3.1.2

and 1.3.1.3, we have shown how the semantic (and formal, albeit with differences) processes involved in the creation of derivational affixes are not fundamentally different from those involved in more typical instances of grammaticalization. We must now have a look at the ‘other side’ of the issue, namely the relationship between derivation and lexicalization.

Himmelman (2004) proposes that processes of grammaticalization and lexicalization may be distinguished according to three parameters:

- (i) host-class formation;
- (ii) change of syntactic context;
- (iii) change of semantic-pragmatic context.

“Host-class” refers to “the class of elements the gram is in construction with”; “syntactic context” refers to “the larger syntactic context in which the construction at hand is used”, similarly to “semantic-pragmatic context” (Himmelman 2004:32-3). In grammaticalization, which is understood here as a process involving constructions, rather than individual morphemes, host-class, syntactic context and semantic-pragmatic context are all expanded. To give an example, when demonstratives grammaticalize into articles, they do so typically when appearing before a noun; so the process regards the item in a construction (DEM NOUN → ART NOUN; Himmelman 2004:31). The shift from demonstrative to articles typically entails expansion of the host class (articles may then appear with proper names and unique entities), expansion of the syntactic context (the construction with the article may start to appear obligatorily e.g. in adpositional expressions), and expansion of semantic-pragmatic context (articles may have “associative anaphoric uses”, as “a wedding – the bride”, whereas demonstratives cannot; 2004:32-3).

According to Himmelman, in grammaticalization those three levels of expansion typically co-occur; however, it is not clear whether all of them need to be present for a process to qualify as grammaticalization. He believes that expansion of the semantic-pragmatic usage contexts is “the core defining feature of grammaticization processes” (2004:33). It is important to stress the fact that, in Himmelman’s understanding, semantic-pragmatic context expansion is neutral as to the mechanisms

involved, “whether grammaticization involves a loss of meaning or rather a transfer of meaning, whether it involves metonymy or metaphor or both, etc.” (Himmelmann 2004:39, endnote 9); the only relevant feature is that “a given construction is used in a larger set of contexts than it was used before”.

As to the relationship between lexicalization and grammaticalization, Himmelmann holds the view that they are not the opposite of one another, highlighting the similarities between the two processes: for instance, on the formal level, erosion and fusion are correlates of both processes (2004:38). As far as the semantic-pragmatic context is concerned, in lexicalization both expansion and narrowing may occur; the meaning changes are non-directional, whereas grammaticalization necessarily involves an expansion, as said above.

What is the “actual point of opposition” between grammaticalization and lexicalization, then? According to Himmelmann (2004:37-8), this is “lexical generality”:

“[i]n lexicalization a specific string of items is conventionalized. In grammaticization the process of conventionalization applies to an expression pattern consisting of at least one fixed item (the grammaticizing element which becomes the increasingly general construction marker) and a growing class of items which enter into this construction.”

If we apply Himmelmann’s parameters to the instances of genesis of derivational affixes illustrated above, it clearly appears that they resemble more grammaticalization than lexicalization. Let us have another look at the case of 性 *-xìng* ‘the quality of [X] / connected with [X]’.

First and foremost, in lexical derivation we are dealing with patterns, and not with “a specific string of items”, as is intrinsic in the notion of constructional idiom: for 性 *-xìng*, there is an increase in generality of the grammaticalizing morpheme, which can combine with an increasingly bigger set of “variables” (i.e. the [X] slot in the template); the increase may involve word classes (from nouns only to any major word class, for 性 *-xìng*) and/or semantic features (from nouns qualifying an inherent nature to any word indicating any property).

As far as the syntactic and semantic-pragmatic contexts are concerned,

it appears that in the evolution from the lexeme 性 *xìng* into an affix there has been an expansion by metaphorical extension (from ‘inherent nature’ to ‘(any) property’; compare Himmelmann’s analysis of Ger. *großer Wurf* ‘big throw’ > ‘great success or achievement’); the number of different contexts in which Modern Mandarin Chinese -性 *-xìng* complex words may be used accordingly increased.

To sum up, in this section we have shown how, according to different treatments of grammaticalization and lexicalization, the evolution of lexemes (compound constituents) into derivational affixes conveying lexical/content meaning resembles more closely typical instances of grammaticalization, rather than lexicalization.

In what precedes, our analysis focussed on the semantic aspects of grammaticalization; in the next section, we will take a closer look at the characteristic of grammaticalization processes in the languages of the area to which Mandarin belongs, especially as far as the formal correlates are concerned.

### 1.3.2 Grammaticalization in East and South-East Asian Languages

The languages of East and mainland South-East Asia, the area to which Mandarin Chinese belongs, are characterized by a number of common features due to prolonged contact, such as e.g. lack of inflection, lexical tone, classifier constructions (for an overview, see Goddard 2005). Some of the features of the languages from this area are especially relevant for our research since, as shown by Bisang (1996, 1998, 2001, 2004, 2008), they influence the way in which grammaticalization works. In this section, we shall illustrate the key features of grammaticalization in East and South-East Asian languages.

As mentioned in 1.3.1.1, there seems to be general agreement on the point that the change from a more concrete to a more abstract meaning generally involves an increase in ‘morphosyntactic integration’ (i.e. reduction in autonomy, cfr. Lehmann 1995, Bisang 2004:109). Bybee, Perkins & Pagliuca (1994:20) even suggest that there is a necessary link between semantic and phonetic reduction: “(...) the development of grammatical material is characterized by the dynamic coevolution of meaning and form”. The idea that grammaticalization somehow entails



formal evolution is also inherent in the notion of *cline of grammaticalization* (as defined in Hopper & Traugott 2003:6); an example of cline is that in (40) (ivi, p. 7):

(40) content item > grammatical word > clitic > inflectional affix

Clines may be significantly different from one another, but “[g]enerally, they involve a unidirectional progression in bondedness, that is, in the degree of cohesion of adjacent forms that goes from loosest (“periphrasis”) to tightest (“morphology”)” (Hopper & Traugott 2003:7).

However, Bisang observes that, for East and South-East Asian languages, grammaticalization typically does not involve “coevolution of form and meaning”, lacking thus what seemed to be almost a universal of grammaticalization. This is motivated, according to him, by some features of the languages in the area:

- a. lack of obligatory (grammatical) categories;
- b. weak correlation between lexicon and morphosyntax;
- c. predominance of pragmatic inference;
- d. existence of rigid syntactic (word order) patterns.

The absence of obligatory categories is subsumed by Bisang under the label of *indeterminateness* (2004:111-112), allowing the omission of arguments (*pro-drop* character) and of grammatical categories as number, tense and aspect for verbs, and number and referentiality for nouns. In most languages of East and South-East Asia (including Mandarin), there are no paradigms, no grammatical values of a category which must be obligatorily expressed (in the relevant syntactic context).

“Weak correlation between lexicon and morphosyntax” means that there is some degree of freedom in the usage of a lexical item, as far as its word class is concerned. The very same word may be placed, in different contexts, in the syntactic slot of a noun or of a verb (a phenomenon usually referred to as 詞類活用 *cílèi huóyòng* ‘flexible use of lexical categories’ in Chinese linguistics; Jiang S. 2005:225-229, see also ZHANG Bo. 1994). For instance, Ch. 工作 *gōngzuò* may be used either as a noun (‘work, job’) or as a verb (‘to work’):

- (41) 你的工作不錯啊  
*nǐ de gōngzuò bùcuò à*  
 you DET job not-bad INTERJ ‘your job is not bad, ah’
- (42) 她工作得很好  
*tā gōngzuò de hěn hǎo*  
 3SG work DEG very good ‘she works very well’

Such freedom was much greater in earlier stages of the language; see an example from Old Chinese (Bisang 2004:114, my emphasis and glosses, characters added):

- (43) 公若曰爾欲吳王我乎  
*Gōng Ruò yuē ěr yù Wú wáng wǒ hū*  
 Gong Ruo say 2SG want Wu king 1SG Q  
 ‘Gong Ruo said: “do you want to deal with me as the King of Wu was dealt with”?’<sup>37</sup>

In (43), the proper name 吳王 *Wúwáng* ‘king Wu’ is used in the verb slot of a transitive predicate, between agent and patient (Bisang 2004:113). In such construction, a proper name is understood as (simplifying) ‘consider / treat s.o. as [proper name]’; since the king Wu referred to was murdered, the inferrable meaning of the sentence is ‘do you want to kill me?’.

The first two characteristics illustrated above, indeterminateness and weak correlation between lexicon and morphosyntax, are closely connected with two more characteristics of the languages of the area, namely the predominance of pragmatic inference and the existence of rigid syntactic (word order) patterns. Pragmatic inference plays an essential role in those languages which have no overtly expressed obligatory grammatical categories; rigid syntactic patterns guide the interpretation and the consequent processes (reanalysis; Bisang 2008).

---

<sup>37</sup> Example from the 左傳 *Zuǒ Zhuàn* (*Commentary of Zuo*), which is supposed to have been written earlier than 389 BCE.

How do the properties illustrated here influence the processes of grammaticalization in East and South-East Asian languages? One important peculiarity of the area is that grammaticalization processes do not follow (unidirectional) clines (see above); here a grammaticalized item does not show a gradual evolution from a ‘less grammatical’ to a ‘more grammatical’ function, but rather keeps different interpretations which may be “recovered” by means of pragmatic inference, which operates at all the stages of grammaticalization (and not only in the early stages, as suggested by Hopper & Traugott 2003; see Bisang 2008:21-22): “[o]ne can see the step from lexical item to grammaticalized item but it is often hard to clearly distinguish between more and less grammaticalized items” (Bisang 2008:23). This means that “one and the same marker may express different grammatical concepts in different situations or in different constructions” (Bisang 2008:16). This is made possible in the first place by the weak correlation between lexicon and morphosyntax, by which the same lexeme may occur in different syntactic environments (Bisang 2004:116-117).

The polysemy of grammaticalized items is tightly connected with another feature of East and South-East Asian languages, namely the lack of coevolution of form and meaning; such lack is caused by the indeterminateness of those languages, as pointed out by Bisang (1996:535):

“[i]n a language in which almost every grammatical category almost always can be inferred from the context, i.e., in a language where there is almost no obligatory grammatical category, even a highly grammaticalized linguistic item shows a higher degree of informative value than in a language showing a lower degree of indeterminateness. This higher degree of informative value is reflected by the fundamental phonological stability of a linguistic sign even in a context of high grammaticalization.”

An example of ‘East Asian’ grammaticalization is the Mandarin lexeme 在 *zài*, which may act as a verb, meaning ‘be (at)’ (44), as an adposition (45) or as a progressive marker (46; adapted from Bisang 2004:117, glosses altered, characters added):

- (44) 她在圖書館  
*tā zài túshūguǎn*  
 3SG.F be.at library ‘she is at the library’
- (45) 他在醫院死了  
*tā zài yīyuàn sǐ-le*  
 3SG.M at hospital die-PFV ‘he died at the hospital’
- (46) 他在穿皮鞋 (qtd. from Li & Thompson 1981:221)  
*tā zài chuān píxié*  
 3SG.M PROG put.on leather-shoe ‘he is putting on leather shoes’

Each of the different ‘identities’ of 在 *zài* is recoverable through pragmatic inference, and there are no differences in shape (no coevolution of form and meaning)<sup>38</sup>.

However, the examples of grammaticalization in East and South-East Asian languages which may be found in the literature generally involve ‘typical’ grammatical categories, as tense, aspect, definiteness, and so on. What about lexical derivation? Many, if not most, of the Mandarin word formation elements which we shall take into consideration seemingly fall into the category of ‘class nouns’ (which will be discussed again in 3.2.1), defined by Bisang (1996:525) as “generic terms on a rather high level of abstraction from which more concrete nouns can be derived by further determination (cf. e.g. Engl. tree → apple tree)”. Examples of class nouns include 學 *xué* ‘scientific discipline’ (語言學 *yǔyánxué* ‘linguistics’), 人 *rén* ‘person’ (寄件人 *jìjiàn rén* ‘sender’; cf. exx. 13-14), 論 *lùn* ‘thesis’ (進化論 *jìn huà lùn* ‘evolutionary theory’), 性 *xìng* ‘nature, character’ (不定性 *bùdìng xìng* ‘uncertainty, indeterminacy’; cf. exx. 4a-c

<sup>38</sup> Ansaldo and Lim (2004:346-347) observe that when 在 *zài* is used as an adposition (as in 45), it actually bears a weakened stress, which “may be realized as a lower tone and interpreted as such”. We are not sure whether this may be interpreted as phonetic erosion induced by grammaticalization, or just as a consequence of general trends in sentence-level prosody (for an overview, see Shen X. 1990).

and sect. 1.3.1.2), 主義 *zhǔyì* ‘-ism’ (社會主義 *shèhuìzhǔyì* ‘socialism’; Bisang 2001).

Bisang believes that class nouns are grammaticalized items, originating from nouns (1996:533, 546-547). However, he sees compounding (his ‘modification’) and derivation as points along a continuum which cannot be unequivocally distinguished: “Since these two processes are mutually related by a continuum of grammaticalization/lexicalization, a clear-cut distinction is not possible”. He further suggests that “[t]he derivational morphemes are suffixes because they are further lexicalized/grammaticalized from items in the position of class nouns”; he thus concedes that some items, as 主義 *zhǔyì* ‘-ism’, “can be described as derivational affixes”, but he fails to provide criteria for derivational status (Bisang 2001).

It appears to us the set of class nouns, as conceived by Bisang, is heterogeneous collection of word formation elements, which share some properties if looked at synchronically, namely the fact that they appear in the head position of a complex word (even a noun phrase, according to Bisang 1996), and that they have a rather abstract meaning, even though such degree of abstractness has not been clearly defined by Bisang.

However, in a constructionist perspective, class nouns may be seen as a slot in a construction, i.e. an environment in which grammaticalization not only *may* take place, but, also, is somehow facilitated. This is the idea behind the notion of ‘attractor positions’ (Bisang 1996:523-528, his italics):

“[f]rom the paradigmatic perspective, slots which *attract* linguistic items in order to grammaticalize them. In this sense, they operate as a kind of *melting pot* or as a kind of *catalyst* for linguistic items to be grammaticalized into different types of grammatical functions. If, for example, an element falls into the domain of the attractor position for TAM [Tense, Aspect and Mood] it will be grammaticalized into a TAM marker. In their paradigmatic function, attractor positions promote metaphoric processes.”

Attractor positions are defined according to their position with respect to the head noun or the main verb. Bisang represents them in ‘maximum patterns’, in which no element is obligatory but the head. These are the

maximum patterns for nouns (Bisang 1996:525, his italics):

- (47) a.                    CL/Q                N            CN        RELN    CONJN  
       b.            CONJN    RELN    CN        N                    CL/Q    (DEMA)<sup>39</sup>

The two patterns are specular, and they differ as to the relative position of head and attributes in a given language: (47a) is the head-attributes pattern, as in Mandarin; (47b) is the opposite pattern, found e.g. in Thai.

Bisang, in the framework of *Construction Grammar* (cf. 1.2.2), assumes that constructions bear meaning, and they can both be the ‘frame’ within which grammaticalization occurs and the product of such a process (Bisang 1998:13-14): “[i]n pragmatics, constructions often provide the basic patterns for processes of reanalysis and analogy”; even a single word may be regarded as a construction. In a construction, “certain positions can attract further items into a new function by the mechanism of analogy”, and these are attractor positions (Bisang 1998:16); it is important to stress that attractor positions operate only as elements in a construction, “i.e. within a framework where several potential grammatical concepts co-operate with each other and with the semantics of the main verb or the head noun, respectively” (Bisang 1996:528).

As seen in 1.2.2, in the framework of Construction Morphology, word formation schemas, i.e. the constructions which operate in word formation, are both ‘produced’ by the language user as they encounter a certain number of words of a certain type and, also, they are employed by the user to build new words (cf. the quotation from Booij 2009:207); just as maximum patterns, they are both the product of conventionalisation (not necessarily grammaticalisation, needless to say) and the *loci* where new items may be ‘attracted’ and develop a new function. A schema as

---

39. Bisang’s glosses are:

CL/Q = classifier / quantifier

N = noun

CN = class noun

RELN = relational noun (expressing locative case)

CONJN = conjunctive noun (to join clauses)

DEMA = demonstrative adverbial.

(48)  $[[a]_N [b]_{Ni}]_N$  ‘ $N_i$  with relation R to X’ (cf. 18)

Which underlies all right-headed complex words with a nominal head in a given language (in our case, Mandarin, but also Dutch or English, among others), acts as a pattern just as Bisang’s maximum pattern, the difference being that here only one relation and one attractor position is present. We believe that a development such as that of Du. *boer* ‘farmer’ > *-boer* ‘seller of  $[X]_N$ ’ is not different from what Bisang analyses as the grammaticalization of a class noun into a derivational suffix (as 主義 *zhǔyì* ‘-ism’); a constructional idiom, in CM terms, represents the conventionalisation of an item analogous to a class noun, i.e. indicating a rather general notion. The notion of abstractness/generality in meaning of an item, here, is not understood in an absolute (and vague) sense as by Bisang, but, as stated before (1.3.1.2), in a relative sense, as ‘being more abstract/general than the original (lexical) meaning. In 3.2.1 we shall discuss further maximum patterns in relation to their role in processes of areal convergence.

Thus, to sum up, the notions of maximum patterns and attractor positions, together with the lack of coevolution of form and meaning, are the most relevant aspects of grammaticalization in East and South-East Asian languages for our research. In the next section, we shall provide a concrete example of the differences in processes of grammaticalization in the “familiar” Indo-European languages and in Chinese, using two derivational affixes with a very similar story and analogous functions, namely Ger. *-heit*, and Ch. -性 *xìng* ‘the property of  $[X]$  / connected with  $[X]$ ’, which was illustrated above (1.3.1.2).

### 1.3.2.1 Comparing ‘Western’ and ‘Eastern’ Grammaticalization Phenomena

In the Indo-European languages of Europe, many fully grammaticalized signs may be traced back to ‘words’, to free morphs, as Fr. *-ment* ‘adverb forming suffix’ < Latin *-mente* ‘mind’ (cf. 1.3.1). In some other cases, a lexemic ‘forefather’ of an affix is not attested, and one can only trace the origin of an affix back to another affix, usually from the mother language; this is often the case for lexical derivational affixes in Romance languages, which

inherited many of their affixes from Latin, at which stage they were already fully morphologised, as It. *-aio* ‘dealer in [X]<sub>N</sub>’<sup>40</sup> (*giornalaio* ‘newsagent’) < Latin *-arius* (*piscarius* ‘fishmonger’; cf. Tekavčić 1980:28 ff.; cf. also Magni 2008). We do not want to imply that *all* derivational affixes originate from lexemes; rather, we might say that, in cases such as It. *-aio* < Lat. *-arius*, we cannot in principle exclude the possibility that the suffix was connected with an unattested lexical morpheme, possibly from some earlier, undocumented stage of the language.

For some derivational morphemes, actually, a non-lexemic origin may be easily demonstrated, as e.g. the English suffix *-ness*:

“[f]orming nouns expressing a state or condition, especially from adjectives and (originally past) participles, as *bitterness*, *conceitedness*, *darkness*, *hardness* (...), also occasionally from adverbs, such as *everydayness*, *nowness*, etc., and in other nonce uses. Also in extended senses ‘an instance of a state or condition’, as a *kindness* etc., ‘something in a state or condition’, as *foulness* etc., and in a few other exceptional uses, as *witness*” (SOED 1993, my expansions).

The suffix *-ness*, thus, is not the product of the grammaticalization of a lexeme, but, rather, originates from a verbal affix, formed by the (former) consonantal ending of the past participle form of strong verbs and the suffix for weak verbs: Old Eng. *-nes(s)*; compare Old High Ger. *-nessi*, *-nassi*, *-nissi*, modern Ger. *-nis*. It is not uncommon for verbal inflection to develop class-changing (transpositional) derivational functions, as shown in (9): Ger. [*singen*]V > [*singende*]V ‘sing-PTCP.PRS’ > [*singende*]ADJ.

In Western Germanic languages, among others, we have several instances of lexical derivational suffixes with a clear lexemic origin, as Eng. *-hood*, defined in the Shorter Oxford English Dictionary (SOED 1993; my expansions) as: “originally a Germanic name meaning ‘person, sex, condition, rank, quality’. Forming nouns of condition or quality or indicating a collection or group from nouns and adjectives, as *childhood*, *falsehood*, *sisterhood*”. The Modern suffix is the evolution of Old English *-hād*; compare the parallel Old High German form *-heit* and Old Saxon *-hēd*. In Modern German, the cognate suffix is *-heit* (*/-keit*), forming words as *Freundlichkeit* ‘friendliness’;

<sup>40</sup> Incidentally, we shall remark that not all Italian *-aio* derived words denote ‘dealers’, as sometimes *-aio* conveys a locative meaning (*letamaio* ‘dung-heap’; cf. Magni 2008).



according to the Wahrig dictionary, the original lexeme *heit /heid* could mean “Person, Stand, Rang, Wesen, Beschaffenheit, Art, Geschlecht” (WDW 2000:614<sup>41</sup>). Other cognate forms are the Gothic lexeme *haidus* ‘kind, manner’ (SOED; in WDW, ‘Art und Weise’) and Old Norse *heidr* ‘honour, worth’.

The Mandarin suffix which appears as functionally close to Ger. *-heit* and Eng. *-hood* (and, incidentally, *-ness*) is 性 *-xìng* ‘the property of [X] / connected with [X]’, the product of the grammaticalization of the Old Chinese lexeme 性 *xìng*, which had, among others, the meanings ‘nature, characteristic’; such development has been already described in 1.3.1.2. In short, a polysemic lexeme as 性 *xìng*, which could convey meaning as ‘inherent property’, ‘immutable nature’, ‘life’, ‘temperament’, was originally used as a free form; it then developed a usage as the right-hand constituent in complex words, and its meaning underwent generalization, from ‘inherent property’, ‘temperament’, to ‘the nature or spirit of [X]N/A/V’, and then to ‘the property of [X] / connected with [X]’, and it can be also argued that, in many cases, it may nowadays be interpreted as a mere indicator of noun class (cf. Chen 1986:89). Such generalization in meaning is accompanied by a broadening of the range of possible ‘bases’, from just nouns to any major word class. We also argued that the semantic shift from ‘inherent nature (of people and things) / immutable inner properties of things’ to ‘the property of [X] / connected with [X]’ may be analysed as metaphoric extension, from ‘inherent nature’ to ‘(any) property’ (重要性 *zhòngyàoxìng* ‘importance’).

What about the shift from Old High German *heit* to Modern German *-heit*? Such evolution has been analysed in detail in Lightfoot (2005), who uses such case as a *bona fide* example of grammaticalization. The reconstructed West Germanic noun *\*haid*, which was associated with the meanings ‘way, nature, appearance, property, characteristic, person, position, rank, honor, sex’, is found again in Old High German as *heit /heid*, both as a free form and as a bound word constituent (and compare the Gothic and Old Norse forms quoted above).

In Old High German texts, the tendency for *heit /heid* when used as a

---

<sup>41</sup> According to the Wahrig Dictionary, the Old High German lexeme may be connected with the Indo-European root *\*kai-* ‘shining, bright’ (“scheinen(d), leuchten(d)”).

free form was to express “person related meanings”, as in *dhrim heidim* ‘three persons’ (= ‘Trinity’; ex. from the Old High German *Isidor*, end of the 8th cent. CE; Lightfoot 2005:594). As the head constituent in complex words, *heit* / *heid* could convey a meaning akin to that of 性 *xìng*: compare *manaheit* and 人性 *rénxìng* (attested in the *Mencius*, 4th cent. BCE), both ‘human nature’. In Otfrid von Weissenburg (ca. 870 CE), *heit* is attested again both as a full noun (meaning ‘person, luminary’) and as a bound form in a word as *zágaheit* ‘timidity, hesitation, cowardice’, from *zag* ‘powerless, bad’; according to Lightfoot’s analysis, here *-heit* is already derivational in nature, because it is “largely devoid of meaning or simply signalling something like ‘quality of’ (...)” and “the semantic basis lies with the first element, namely *zag* ‘powerless, bad’” (2005:596). He also remarks that “[o]ne could posit the separate, analytical, nominal meanings of the two components in the latter example (i.e. ‘powerless’ plus ‘nature’ to render ‘powerless nature’), but that is perhaps lacking some of the sense of the holistic meanings like ‘timidity’ and so on” (*ibidem*).

Lightfoot also mentions the phrase *uuízent héit* ‘knowing person’ and the complex word *uuízentheit* ‘knowledge, consciousness’ (from Notker’s translation and commentary of Boethius, ca. 1000 CE). There is a clear difference in meaning between the phrase, in which *heit* is a free lexeme, and the complex word, in which *-heit* is a bound morpheme; a plural is possible for the former (*uuízent héite* ‘knowing people’) but not for the latter, which “seems to be good evidence for a semantic transition of the analytic interpretation moving toward the holistic, derivational one” (Lightfoot 2005:594-595). We discussed above the Early Mandarin word 急性 *jíxìng* ‘(of) impatient disposition’, from the adjective 急 *jí* ‘impatient, urgent’ (used as an attribute for people), where the meaning conveyed by 性 *xìng* is still one of its lexical meanings, namely ‘disposition, temperament’; a word 急性 *jíxìng* is also attested in Modern Mandarin, meaning ‘acute’ (associated mainly with diseases; cf. Chen 1986:89), and thus does not denote a stable characteristic (as one’s disposition); The lexemic nature of 性 *xìng* in 急性 *jíxìng* ‘(of) impatient disposition’ is also proved by the possibility of adding the

“dummy affix” (Lin 2001:82) -子 *-zi* to it: 急性子 *jíxìngzi* ‘of impatient disposition’, ‘impetuous person’ (Chen R. 1986:89). We believe that the distinction between 急性 *jíxìng* ‘(of) impatient disposition’ and 急性 *jíxìng* ‘acute’ is analogous in nature to that between *uuízent héit* ‘knowing person’ and *uuízentheit* ‘knowledge, consciousness’, namely between the same form used as a lexical item and as an affix. Lightfoot’s arguments for the derivational status of *uuízentheit* ‘knowledge, consciousness’ are summarised below (adapted from Lightfoot 2005:595):

- a. *-heit* bears a “generalized” meaning, ‘characteristic of, quality of’ or “virtually solely functioning as a nominalizer” (cf. 30)
- b. the “semantic basis” is in the left-hand constituent *uuízent* ‘knowing, conscious’
- c. the constituent mostly occurs as bound
- d. *-heit* is “in a systemic relation to other derivational suffixes”, i.e. it competes e.g. with *-tuom* (cognate with Eng. *-dom*)
- e. *-heit* is phonologically short

The semantic arguments in favour of the affixhood of *-heit* in words as *zágaheit* ‘timidity’ or *uuízentheit* ‘knowledge, consciousness’ (a. and b.) are akin to those which we suggested in our discussion of grammaticalization and bleaching (1.3.1.2). The difference lies in the fact that here, again, reference is made to a notion of ‘general’ which is apparently absolute, rather than relative as we suggested: we want to stress once more the point that notions as ‘generality’ and ‘abstractness’ may not be defined precisely but, rather, are to be taken as relative, as ‘more’ or ‘less’ than another (earlier) stage. Point c. and e. are not really relevant for Mandarin, since many (if not most) lexical morphemes are bound as well, and bondedness is not a sufficient criterion for affixhood (albeit a

necessary one)<sup>42</sup>; as to the size of *-heit*, in Mandarin nearly all would-be affixes are monosyllabic, just like lexical morphemes. The competition with other morphemes with an analogous function is a rather interesting point, albeit this does not happen for *-性* *-xìng* which, to the best of our knowledge, has no real ‘competitors’.

It clearly appears, in our opinion, that the processes behind the grammaticalization of Old High German *heit* / *heid* into the suffix *heit*, which then survived into Middle and Modern German, are not significantly different from those described before for the development of *性* *xìng* into Mandarin *-性* *xìng*: a polysemic lexeme undergoes a generalisation in meaning, and becomes a repeated marker in a ‘word family’, i.e. a set of complex words sharing a common constituent (in CM terms, a set of words which are the instantiation of the same constructional idiom; cf. exx. 19-20, above). Such shift happens in a specific environment, namely in a fixed position in complex words with a certain type of ‘non-head’ constituents.

At the stage of Middle High German, *heit* as a full noun becomes very rare, and “[t]he norm in this era is for the abstract suffix *-heit* to mean ‘characteristic of, condition of, manner of’ as in *rîchheit* ‘the condition of being rich, wealthy’ (...)” (Lightfoot 2005:598); as said above (1.3.1.2), in a text sample of Early Modern Chinese (13th-19th cent.) analysed by Luo J. (2004:92), *性* *xìng* is more often used as a constituent in a complex word than as a free morph, whereas in Old Chinese it was typically a free form. It is also worth remarking that for some Middle High German *-heit* complex words the interpretation may be ambiguous between the lexical meaning and the ‘derivational’ meaning, as for *hübescheit* (from *hübesch* ‘well educated and mannered’), which is reported to bear the meaning ‘well educated and mannered nature’ (which Lightfoot terms “part plus

---

<sup>42</sup> We highlighted above the parallel between *uuizent heit* ‘knowing person’ and *急性 jíxìng* ‘(of) impatient disposition’, on the one side, and *uuizentheit* ‘knowledge, consciousness’ and *急性 jíxìng* ‘acute’ on the other side, basing on the fact that in the former, *heit* and *性* *xìng* are used in their lexical ‘identity’, as nouns, whereas in the latter they are used in their affixal identity, as proved by the difference in meaning. Note that in Old High German the distinction was also one between unbound vs. bound usage, whereas in Chinese, not surprisingly, in both cases *性* *xìng* is bound.

part” meaning), but also ‘beauty’ (Lightfoot 2005:598); this means that

“the transformation from part plus part into holistic meaning must have been an ongoing, contextually based process, since *-heit* was already often felt to signal derivational meaning in the Old High German era, yet it still could go through the same development for a form like *hübescheit*, which first came about only in the Middle High period. That is, in some cases the meaning ‘well educated nature’ likely gave way to the simple, holistic meaning of ‘beauty’ (...), and thus the cycle of compounding and suffixing could reoccur with the various fusions at various times. Ultimately, the morphosyntactic and semantic nature of *-heit* is determined by the context in which it is found”.

This is also visible for *-性 -xìng*, as shown with the 急性 *jíxìng* ‘(of) impatient disposition’ / ‘acute’ example above; compare also 忍性 *rěnxìng*, meaning either ‘to restrain one’s temper’ or ‘endurance’ (cf. footnote 31). Moreover, since the 9th century, *-heit* began to combine also with adjectives (as Middle High Ger. *wisheit* ‘wisdom’) and nominalised infinitives (*Unwissenheit* ‘carelessness’), and, generally speaking, with a broader range of nouns, including living beings (*tierheit* ‘animal kingdom’<sup>43</sup>). An increase in the combinatory possibilities occurred also in the history of *-性 -xìng*; we argued that this is a consequence of its generalization, i.e. the reduction in its intensional meaning (from ‘inherent property’ to just any property), and this coincides with Lightfoot’s analysis of *-heit*.

It is also interesting to compare the fate of *heit* and 性 *xìng* as free nouns. In Modern Standard German, *heit* is no longer used as a free form, possibly because perceived as too archaic; it is in fact preserved in Bavarian, which is regarded as an archaic dialect (Lightfoot 2005:601). In Modern Mandarin too, as already said above, 性 *xìng* is never a free form. However, the original lexemic meaning of *heit* and *xìng* may still be seen in complex words, many of which have been preserved from earlier stages of the languages at issue, as e.g. Ger. *Gottheit* ‘divinity, godly nature’ (<

<sup>43</sup> Interestingly, in Modern German *Tierheit* can mean ‘animality, bestiality’. The shift in reference from ‘animal kingdom’ to ‘animality, bestiality’, i.e. the condition of being an animal reminds us of Ch. 忍性 *rěnxìng*, ‘to restrain one’s temper’ and ‘endurance’.

Middle High Ger. *goteheit*), *Christenheit* ‘Christianity’ (< *chrīstanheit*) and Ch. 本性 *běnxìng* ‘inherent quality’, 任性 *rènxìng* ‘stubborn, headstrong’. Moreover, whereas in modern German *heit* is not found on the left side of a word, in Mandarin it may well be used in any position, bearing one of its lexemic meanings: 性急 *xìngjí* ‘impatient, short tempered’, 性能 *xìngnéng* ‘natural capacity, function’.

As to the phonological shape of *-heit* and *-性 -xìng*, there seems to have been no reduction, and the differences between the modern forms and those of the preceding stages of the language are due to regular sound change (as the shift from [ei] to [ai] in *-heit*): compare the distinction between the Modern English word *doom* and the affix *-dom* (see above, **1.2.2**). However, since the lexeme *heit* disappeared from the language, *-heit* as a bound form is normally classified as a suffix; in a language as Mandarin, where many lexical morphemes are bound, and grammaticalization without phonological alteration is the rule rather than the exception, we rely mostly on semantic criteria to identify derivational affixes.

In this section, we have shown how the processes which lead to the evolution of a lexeme into a derivational affix may be very similar in genetically unrelated and typologically distant languages, as German and Chinese. Incidentally, in the case of Ger. *-heit* no significant phonological reduction has occurred in the process of grammaticalization, which is what we expected for Chinese (as a language of the East Asian area, cf. **1.3.2**).

It might be worthwhile to explore the connection between meaning abstraction and phonological reduction, i.e. if the kind of meaning expressed in lexical derivation, which can be more ‘concrete’ than ‘pure’ grammatical meaning (tense, number, etc.), prevents reductions in shape<sup>44</sup>. This, however, is far beyond the aims of the present work. Having discussed in detail the nature of the processes of development of lexemes into affixes in a language as Chinese, we are now in a position to

---

<sup>44</sup> Incidentally, we shall remark that phonological reduction for grammaticalized items is actually attested in Mandarin for aspect markers as 了 *-le* (perfective; cf. fn. 16) and 著 *-zhe* (progressive), which convey ‘typical’ grammatical categories.

reconsider the notion of affixoid which, as we shall see in CHAPTER 2, is a central one in the research on Chinese word formation<sup>45</sup>.

### 1.3.2.2 Affixoids reconsidered

The notion of ‘affixoid’ has been introduced in the literature on word formation, seemingly, to label those items which somehow lie between compounding and derivation, possessing hybrid properties, “lexical elements caught up in such a transition of status from the constituent of a compound to a derivational morpheme” (Olsen 2000:902), as seen before (1.2.2). Examples of affixoids include, among others, Eng. ‘-ware’ (hardware), Ger. *-arm* ‘low in [X]N’ (*fettarm* ‘low-fat’), etc. The definition of affixoid we used as a starting point in our research is that of Construction Morphology, according to which affixoids are morphemes which occur both as free lexemes and as constituents in complex words (in a fixed position), but their meaning is “specific and more restricted” in word formation (Booij 2005:114), as Du. *-boer* ‘seller of [X]N’ vs. the lexeme *boer* ‘farmer’.

Another problematic category for the distinction between compounding and derivation is that of neoclassical constituents (also ‘semi-words’), bound lexical morphemes of Greek or Latin origin which have no corresponding free form, as Eng. ‘anthropo-’, ‘-logy’, etc. Whereas the notion of a bound morpheme conveying lexical meaning is somehow ‘anomalous’ for a language as English, this is not unusual for Mandarin, the lexicon of which contains many bound lexical roots, as 校 *xiào* ‘school’ or 軍 *jūn* ‘army’, which are never used in isolation (cf. the compound 軍校 *jūnxiào* ‘military academy’).

Affixes and affixoids, as seen above, share an important characteristic, namely that they both ‘exist’ only in word formation (in CM terms, as part of schemas); an affixoid may convey a certain meaning only as part of a complex word. The fundamental difference between affixes and affixoids is that the former do not have a corresponding (homophonous) lexeme in

---

<sup>45</sup> *Contra* Ten Hacken (2000:356): “(...) the idea of introducing one or more intermediate classes between derivation and compounding seems to be restricted to an episode in German linguistics of the 1970s and 1980s” (and cf. the authors quoted in 1.2.2).

the same synchronic stage of the language; thus, an affixoid is ‘promoted’ to affixal status if the connection with the lexeme is lost, either because of sound change (as Eng. *doom* vs. the affix *-dom*), or because the lexeme becomes obsolete, as in the case of Old and Middle High Ger. *heit* / *heid* (1.3.2.1).

Is a distinction based on the above mentioned criteria meaningful? Firstly, the fact that a lexeme falls out of use is, basically, an incident of history; let us compare the evolution of Du. *boer* ‘farmer’ with that of Old / Middle High Ger. *heit*. In both cases, two lexemes have acquired a specific meaning as the right-hand constituent in complex words; such meaning is ‘abstract’ enough to instantiate a productive pattern with derivation-like characteristics (cf. 1.2.2; see also Lightfoot’s criteria in 1.3.2.1). The difference between *-boer* ‘seller of [X]N’ and *-heit* (/ *-keit*) ‘characteristic / quality of [X]’, thus, is just that *boer* is still a free lexeme in the Dutch lexicon; from such a perspective, if *boer* had been ousted from the language by a competing lexeme (which is what happened to *heit*; Lightfoot 2005), we would now, most likely, consider *-boer* as a derivational suffix. The process behind the genesis of *-boer*, however, would be exactly the same. This means that ‘affixoid’ as it is defined in CM is not an intermediate category between that of derivational affix and that of compound constituent; it looks more like a *subcategory* of derivational affixes, including those elements which may still be connected to a lexeme synchronically. As far as the process of grammaticalization / morphologization is concerned, we believe that there is no difference between an affix and an affixoid.

What about the connection between sound change and affixal status? This makes sense only if one assumes a model of grammaticalization by which such a process is inevitably correlated with some degree of change on the formal level (as e.g. Bybee, Perkins & Pagliuca 1994; see above, 1.3.2, and cf. also Lehmann’s parameters, 1.3.1.1). In the latter case, we should say that only after the affixoid has undergone some sort of sound change it can be regarded as a ‘true’ derivational affix. However, since in Mandarin grammaticalization of a sign without sound change is the norm rather than the exception, such criterion proves to be inadequate.

A residual problem is that of word class. In CM, affixes do not bear a word class, which belongs to the construction as a whole (see the schema



in 21), whereas affixoids do. However, as seen above (e.g. in 19), even for affixoids the part of speech tag is attached to the construction; the difference is that, in this case, the tag is identical to that of the affixoid, since affixoids act as heads. In our opinion, this is a minor problem, especially if one regards derivational suffixes as the (categorical) heads in complex words (see e.g. Bisetto & Scalise 2007).

To sum up, we believe that we may do without the notion of affixoid, at a theoretical (rather than descriptive) level. This is because, on the one side, the processes of grammaticalization / morphologization of lexemes into derivational elements occur irrespective of the fact that the lexemes has or has not lost its connection with the newborn affix and, on the other side, the very notion of “coevolution of form and meaning” may not be applied as such to Chinese, in which even highly grammaticalized signs may retain their phonological shape (and, also, other lexical usages). The category of semiwords appears to be meaningful for the Indo-European languages of Europe, as English or Italian, in which there is a strong tendency to associate ‘lexical’ with ‘free’ and ‘grammatical’ with ‘bound’; in Mandarin Chinese, where a large number of lexical morphemes are never (or nearly never) used in isolation, there seems to be no point in positing another subclass of morphemes other than the ‘traditional’ distinctions lexical *vs.* grammatical and free *vs.* bound. Such issue will be reprised in **2.2.2**.

However, the notion of ‘affixoid’ has been employed in many works on Chinese morphology, as we shall see in the next chapter. We shall argue that this is due to the application of ‘Western’ categories and criteria in the analysis of word formation in Mandarin.



## **CHAPTER 2**

### **PREVIOUS TREATMENTS OF LEXICAL DERIVATION IN CHINESE LINGUISTICS**

In this chapter we shall provide an overview of the work that has been done to date on derivation and, more synthetically, related issues in Chinese linguistics. Firstly, we shall discuss how fundamental notions in the study of morphology as ‘morpheme’, ‘affix’ and ‘derivation’ have been applied to Chinese (especially Mandarin); we believe that such introduction is needed, since the different applications of those notions which one meets in the literature are, more often than not, due to different understandings of Chinese word formation phenomena. In the second part of the chapter, we shall focus on how lexical derivation has been dealt with in the research on Chinese morphology, especially in recent years, to draw some insights for our own analysis.

#### **2.1 Some Key Notions of Morphology in Chinese Linguistics**

##### **2.1.1 The Notion of ‘Morpheme’ and Related Issues**

In the Chinese linguistic literature before the twentieth century there were no notions equivalent to ‘root’ or ‘affix’, partly because of the characteristics of Classical Chinese as a written language. Thus, with the introduction of such (and other) notions at the turn of the century, new words had to be coined to denote them. This is far from being a purely terminological question; the introduction of the fundamental notions of word formation made the birth of modern Chinese linguistics possible and, also, gave rise to a debate as to whether ‘Western’ linguistic categories should be applied to Chinese, since they were conceived with Indo-European languages in mind.

As pointed out in the preceding chapter (1.1.3), ‘morpheme’ and ‘word’ have a tendency to overlap in Chinese, as in all isolating languages; in this section, we shall briefly discuss some conceptual and terminological issues related to the notion of ‘morpheme’ and its classification, drawing

mainly from Pan, Ye & Han (2004).

The first systematic description of (Classical) Chinese grammar was the well-known 馬氏文通 *Mǎshì Wéntōng* ('Mr Ma's grammar') by Ma Jianzhong, published in 1898 (Casacchia 2006, Gu 2006)<sup>1</sup>, in which 'Western' categories as 'prefixed (form)' (前加 *qiánjiā*) and 'suffixed (form)' (后附 *hòufù*) have been introduced to China. In the earlier literature, only alterations of syllables/morphemes were taken into consideration, rather than the combination of them (Pan, Ye & Han 2004:14-16). In Ma's grammar, however, there was still no clear separation between the notions of 'character' (字 *zì*) and '(syntactic) word' (詞 *cí*); since, as seen above (1.1.4; see also 1.1.3), (nearly!) all characters correspond to a syllable, and there is a strong tendency towards a 1:1 correspondence between units of speech (syllables) and units of meaning (morphemes), it comes as no surprise that in the Chinese tradition there was no notion of 'word' distinct from that of 'character' ( $\approx$  morpheme). Such association is better motivated for the Classical language, in which there were many more monosyllabic 'words', i.e. free forms, than in Modern Mandarin (Feng S. 1998); however, even today, as Chao Y. puts it (1968:136),

“[w]hatever conception of the syntactic word we shall find scientifically justifiable to define, it plays no part in the Chinaman of the street's conception of the subunits of the Chinese language. Thus, if one wishes to ask what the syntactic word *xianzai*<sup>2</sup> 'now' means, one would say: “現在”这两个字是什么意思? “*Xianzai*” *zhe liangge zi shi shenme yisi?* ‘What is the meaning of these two *zi* “*xianzai*”?’” (Chao Y. 1968:138).

---

<sup>1</sup> The first grammar of Modern Chinese was Li Jinxi's 新著國語文法 *Xīnzhù Guóyǔ Wénfǎ* ('A New Grammar of the National Language'), published in 1921 (Casacchia 2006:361).

<sup>2</sup> Chao's italics. The romanization of Chinese has been altered to pinyin for the sake of consistency. The seemingly unusual mixture of traditional and simplified Chinese characters is explained with the fact that, back in 1968, the official list of simplified characters was slightly different from today.

In the scientific community, however, the notions of ‘character’ and ‘word’ were separated at least since Zhang Shizhao’s 1907 work 中等國文典 *Zhōngděng Guówén Diǎn*, who suggested that even though a character may correspond to a word, this does not entail that any word corresponds to one character (qtd. in Pan, Ye & Han 2004:97); it was Li Jinxi (cf. fn. 1) who then defined the ‘Chinese’ word, the 詞 *cí*, *grosso modo* as it is defined in the Western linguistic tradition<sup>3</sup>.

The next step would be the introduction of the terms translating the notion of morpheme by Chen Wangdao in 1940, namely 辭素 *císù* ‘(constituent) element of diction’ and 語素 *yǔsù*, ‘(constituent) element of language / expression’. Chen Wangdao somehow still keeps apart the notion of bound morpheme (as element of diction) and free morpheme (as element of language, of expression), and is thus still different from the notion of “smallest meaningful unit of language structure”; Chen Wenbin (1955) later proposed 詞素 *císù* ‘element of the word’. However, since ‘morphemes’ are not necessarily part of words, but rather, as in the case of free morphemes, may constitute phrases as sentences, 詞素 *císù* does not appear to be an appropriate rendition of the notion of morpheme (cf. Zong S. 1997, Piccinini 2005). It was Zhu Dexi which, having been influenced by structuralist ideas, suggested that the correct translation of ‘morpheme’ should be the above mentioned 語素 *yǔsù* ‘element of language / expression’; Lü Shuxiang later made the 語素 *yǔsù* as the basic unit of grammatical analysis (1979, qtd. in Pan, Ye & Han 2004:99-100; cf. Zong S. 1997). Moreover, according to Lü S., whereas in ‘Western languages’ (西方語言 *xīfāng yǔyán*) ‘word’ and ‘sentence’ are the most important units, in Chinese, “for historical reasons”, ‘morpheme’ and ‘phrase’ are not less relevant than the word.

The issue of how to define ‘free’ and ‘bound’ in Chinese has also been the object of much debate. Zhang Sho. (1957) applies the traditional Chinese distinction between ‘full’ (實 *shí*) and ‘empty’ (虛 *xū*)

---

<sup>3</sup> Needless to say, the definition of the ‘word’ is still a matter of debate in the scientific community also in the West. For an overview on the issue, see Dixon (2002) and Ramat (2005), among others.

morphemes, i.e. between (roughly speaking) lexical and grammatical / function morphemes, and further distinguishes between ‘independent’ (独用 *dúyòng*) and ‘non-independent’ (非独用 *fēidúyòng*) morphemes; here ‘independent’ means ‘being able to stand alone’, thus coinciding with the notion of (syntactic) word. ‘Empty’ morphemes may be prefixed (前加 *qiánjiā*) or suffixed (後加 *hòujiā*). The relevant innovation introduced by Zhang S. is that ‘free’ vs. ‘bound’ and ‘lexical’ vs. ‘grammatical’ are conceived as independent distinctions, whereas previous authors as Chen Wangdao and Lu Zhiwei (1964) associate ‘free’ with ‘lexical’ and ‘bound’ with ‘grammatical’, possibly because of the influence of Western linguistic thought. This is especially relevant since Mandarin, as said before, has a very large number of bound lexical morphemes (cf. 1.3.2.2). In *Mandarin Primer* (1952), Chao Y. proposes that the morphemes of Chinese may be either always bound (粘著 *niánzhuó*) or free (自由 *zìyóu*), but free morphemes are free only in some contexts, and bound in others (cf. also Chao Y. 1968:144). In his later work *A Grammar of Spoken Chinese* (1968:143-146), Chao Y. classifies morphemes according to their ‘combinability’: morphemes may be bound on the left side, on the right side, alternatively on both sides or always on both sides; he also divides free and bound morphemes into further subcategories according to the ‘ease’ with which they combine, as we shall see below (2.2.2).

In Yin F. (1984; qtd. in Pan, Ye & Han 2004:104-106 and 116-117) it is proposed, *contra* Chao Y., that there are no ‘true’ bound morphemes, but, rather, each morpheme has a ‘degree of dependence’ (隸屬程度 *lìshǔ chéngdù*), ranging from ‘fully independent’ (完全獨立 *wánquán dúlì*) to ‘conditionally independent’ (條件獨立 *tiáojiàn dúlì*). Such degree of dependence may be tested with ‘discriminants’ (判別式 *pànbíeshì*): for instance, a nominal morpheme is free if it can appear in a ‘numeral-classifier-noun’ construction, as 一把刀 *yì bǎ dāo* ‘one CLF knife’; a morpheme as 壁 *bì* ‘wall’ is normally used only as a constituent in complex words (牆壁 *qiángbì* ‘wall’), but may also be found in a phrase as 碰了一次壁 *pèng-le yí cì bì* ‘hit-PFV one CLF wall’, ‘hit a wall once’, is thus classified as ‘conditionally independent’.

The distinction between bound and free morphemes, as seen above, may prove to be rather problematic, since the status of a morpheme may also depend on the (micro-)syntactic context (cf. Zhou & Marslen-Wilson 1994), differently from what happens in the Indo-European languages of Europe, in which free or bound status are normally stable characteristics; the intermediate category of ‘semi-free morpheme’ (半自由語素 *bànzìyǒu yǔsù*) has been proposed to label those morphemes with hybrid properties (first proposed in Lü S. 1962). However, different authors have different understandings of such notion. In Zhang Z. (1981, qtd. in Pan, Ye & Han 2004: 115-116), semi-free morphemes are those bound morphemes as 語 *yǔ* ‘expression, language’ which may combine with other morphemes without restrictions (e.g. 英語 *yīngyǔ* ‘English (language)’), whereas ‘non-free’ (不自由 *bù zìyǒu*) morphemes have a fixed position, as 第 *dì* ‘marker of ordinal numbers’, which is always prefixed (第七 *dìqī* ‘seventh’).

A recent proposal is that of Dong X. (2004:45ff.), who defines ‘semi-free morphemes’ as those morphemes which “may not constitute a word in isolation, and must combine with other morphemes to do so, but, under certain conditions, may appear in the syntactic slot of a word” (2004:46, my translation). Among her semi-free morphemes are 刊 *kān* ‘periodical, publication’ (e.g. 期刊 *qīkān* ‘periodical’) or 館 *guǎn* ‘building’ (圖書館 *túshūguǎn* ‘library’), lexical bound morphemes which may occupy the syntactic slot of a free form, if preceded by certain monosyllabic bound elements, mostly demonstrative in nature: 本刊 *běnkān* ‘this periodical’, 此館 *cǐguǎn* ‘this building’. According to Dong X., verbal morphemes as well may be semi-free, as 知 *zhī* ‘to know’ and 感 *gǎn* ‘to feel’, which may be used freely if precede by monosyllabic adverbs, as in 已知 *yǐzhī* ‘known’, 不感 *bùgǎn* ‘not feel’. The motivation for the ‘ambiguous’ nature of such morphemes is prosodic: semi-free morphemes are (syntactically) free, but also ‘prosodically deficient’ (韻律不足 *yùnlǜ bù zú*) and, thus require another syllable/morpheme to build a minimal prosodic word (on the prosodic word in Chinese, see below, 3.2.1.1), just as clitics require a host (Dong X. 2004:56-57). Dong X. believes that the semi-free morphemes of Chinese share many similarities with what Anderson (1992) terms ‘phrasal affixes’, i.e.

clitics; the difference, according to her, lies in the fact that Chinese semi-free morphemes show no phonological fusion with their host and, generally, bear a lexical rather than grammatical meaning (Dong X. 2004:56, fn. 1).

In Dong X.'s model, to sum up, Mandarin morphology has a threefold distinction: free, bound and semi-free; morphemes may be also classified, according to their semantics, into lexical ('full') and grammatical ('empty'), as shown below (elaborated from Dong X. 2004:90-91):

Table 2.1. A classification of Chinese morphemes (Dong X. 2004)

	<b>Lexical</b>	<b>Grammatical</b>
<b>Free</b>	May constitute a 'full' word in isolation	May constitute an 'empty' word in isolation
<b>Bound</b>	Bound lexical roots, may not constitute a word in isolation	Affixes
<b>Semi-free</b>	Bound lexical roots which may occupy the slot for a free form in certain prosodic contexts	May act as affixes and as 'empty' words or clitics, in certain contexts

This model allows for elements which operate both in word formation and in syntax, i.e. semi-free morphemes, as *-者* *-zhě*, which, as seen in **1.3.1.1** (exx. 22-23), may form (agentive) nouns combining with nouns, verbs, adjectives, and also with phrases. Such kind of items, which look like 'anomalous' affixes with a seemingly double nature (morphological and syntactical), will be discussed in detail in the next chapter (3.2.4).

Packard (2000:67 ff) has a rather different stance on the distinction between free and bound morphemes in Mandarin. According to him, "[t]he characteristics "bound" and "free" are clearly definable in Chinese, despite claims to the contrary"; the fact that a given morpheme may be



bound or free, depending on the context, is explained as such:

“a morpheme that possesses varied characteristics or identities potentially has separate (though undoubtedly related) entries in the mental lexicon for each identity. (...) a given characteristic applies unambiguously for any given morpheme as a function of its lexical identity, and the native speaker knows clearly what the usage is for any given entry in the mental lexicon”

So, for instance, a morpheme as 木 *mù* ‘wood, tree’ is normally bound (木匠 *mùjiāng* ‘carpenter’); however, when it means ‘numb’, it may occupy a syntactic slot, as in

- (1) 我舌头木了  
*wǒ shétou mù-le*  
 1SG tongue numb-PFV  
 ‘my tongue is numb’ (ex. from Packard 2000:68)

In yet other cases, ‘bound’ or ‘free’ identity may depend on the style of the text (classical *vs.* modern, spoken *vs.* written, etc.).

Packard’s classification is based on the combination of two properties of a morpheme: on the one side, free *vs.* bound status and, on the other side, having functional *vs.* lexical meaning. The latter is virtually analogous to the traditional distinction of ‘empty’ and ‘full’ signs seen above; however, Packard takes such distinction as not discrete, “and indeed there is a continuum upon which the concepts of content and function rest, with most function morphemes finding their historical origins in the grammaticalization of content morphemes”. The combination of the two distinctions yields a four-way classification of morphemes, which is presented in table 2.2 (exx. from Packard 2000:74, table 12:

Table 2.2. A classification of Chinese morphemes (Packard 2000)

	<b>Lexical</b>	<b>Grammatical</b>
<b>Free</b>	‘Root words’: 冰 <i>bīng</i> ‘ice’	Function words: 和 <i>hé</i> ‘and, with’, 的 <i>de</i> ‘determiner’
<b>Bound</b>	Bound roots: 房 <i>fáng</i> ‘house’	Affixes -者 <i>-zhě</i> ‘agentive suffix’, -过 <i>-guo</i> ‘experiential past’

If we compare Packard’s proposal to Dong X.’s, we may realise that the only major difference is that in the former there are no intermediate (sub-)classes between ‘free’ and ‘bound’; whenever a morphemes acts both as a bound and as a free form, this means, according to Packard, that we are dealing with separate identities of a form.

However, Packard (2000) has the same treatment for morphemes as 木 *mù* ‘wood, tree’ / ‘numb’ as for other morphemes which, we believe, cannot be lumped together, as those bound morphemes which appear as free in different styles or different varieties of Mandarin (as 刊 *kān* ‘periodical, publication’, quoted above). Moreover, Packard does not seem to take into consideration ‘phrasal affixes’, a phenomenon which is not found only in Mandarin. In 2.2.2, we shall discuss Packard’s definition of affixes, which is an especially relevant aspect for the purposes of our research.

The brief presentation of some issues related to the application of the ‘conventional’ notions of morphological analysis to Modern Chinese which we have given here will be instrumental in understanding the difficulties with the notion of ‘affix’, the topic of the next section, and the consequences that this had on the treatment of Chinese word formation in the literature.

### 2.1.2 ‘Root’ vs. ‘Affix’, ‘Derivation’ vs. ‘Compounding’

The terminology related to the notion of ‘affix’ has been introduced into Chinese as early as 1931 by Qu Qiubai, who suggested that in Chinese ‘roots’ (字根 *zìgēn*), ‘prefixes’ (字頭 *zìtóu*) and ‘suffixes’ (字尾 *zìwěi*) are used to build new lexical items, just as French speakers use the same elements of the Latin language to build words (Qu Q. 1957a, qtd. in Pan, Ye & Han 2004:65). The similarities between semi-words and Chinese morphemes which are ‘creatively’ used in word formation (see 1.3.2.2), thus, had been noticed relatively early by Chinese linguists (and see Packard 2000:77). In the work by Qu Q. quoted here, one finds such ‘suffixes’ as -家 *-jiā* ‘professional / expert of [X]<sub>N</sub>’ (政治家 *zhèngzhìjiā* ‘politician’) or -主義 *-zhǔyì* ‘-ism’, which are termed 新式的字尾 *xīnshì de zìwěi* ‘new-style suffixes’; such category (partly) overlaps with Bisang’s class nouns (1.3.2) and, as we shall see, it is most likely the first attempt in a long series to coin an *ad hoc* label for the Chinese formants of the kind of lexical derivational affixes.

The necessity for a new term to label Chinese (possible) affixes is explained, as suggested earlier (1.3.2.2), with the difficulties arising from the application to Mandarin of a category as that of derivational affix, which has been conceived with inflectional (or, more generally, synthetic) languages in mind. We have already seen how an intermediate category between ‘affix’ and ‘root’ has been proposed also for ambiguous word formation elements in the languages of Europe (1.2.2); however, the issue is much more relevant for Mandarin, and the separation of compounding and derivation in such language poses a serious challenge.

The ‘theoretical embarrassment’ caused by lexical derivation in Mandarin Chinese is especially evident in the huge differences in the treatment of such phenomenon by different authors. Pan, Ye and Han (2004:77 ff) surveyed the literature on Modern Chinese word formation and analysed a sample of 14 representative works on the topic (see the source for the list), ranging from 1932 to 1982, in which they found more than 400 different morphemes which have been classified as either

‘affixes’ or ‘affixoids’<sup>4</sup>. After deleting from their count the cases of reduplication, which are not relevant here, 340 morphemes have been considered; among those, only 16 are reported to have been quoted as examples of affix(oid)s in the majority of the works sampled, whereas as many as 223 morphemes have been labelled as affix(oid)s only once (Pan, Ye & Han 2004:81). These 16 ‘qualified’ formants will be introduced in the next section (table 2.3); here we shall just remark that the criteria for the definition of affixes (and related notions) which may be found in the Chinese linguistic literature are extremely varied and only partly overlap with those put forth in ‘Western’ linguistics. As to lexical derivation, a major problem which every linguist faced is that of the distinction between an affix and a compound constituent, especially if the item at issue has a high frequency of use as a (bound) word formation constituent. Even though this is not, obviously, the *only* problem, it appears to be the biggest obstacle in the development of a sound treatment of affixation for Mandarin with cross-linguistic consistency. Such issue will be the topic of par. 2.2.

In the Chinese linguistic literature, the term commonly used nowadays to indicate the notion of ‘root’ is 詞根 *cígēn*, lit. ‘word root’; such term is significantly different from Qu Q.’s 字根 *zìgēn* quoted above, which still made reference to the notion of ‘character’ rather than to that of ‘word’. There have been (at least) two different understandings of 詞根 *cígēn*. According to Zhu D. (1982, qtd. in Pan, Ye & Han 2004:94), in Chinese morphology the notion of ‘root’ (詞根 *cígēn*) is opposed to that of ‘affix’ (詞綴 *cízhù*); in his analysis, Mandarin has only 12 affixes, and all the rest are roots. However, Lü S.’s position appears as more appropriate for the analysis of Chinese, since he takes into account the peculiarities of the language, compared to the languages of Europe (1979:94, qtd. in Pan, Ye & Han 2004:94; my translation):

“Word formation in Western languages is based on derivation, and roots are opposed to affixes; Chinese morphology is based on compounding and roots are

---

<sup>4</sup> Here we are using ‘affixoid’ as a convenience term to indicate any label that has been used in Chinese linguistics for borderline items, as e.g. 類詞綴 *lèicízhù* or 準詞綴 *zhǔncízhù*, which shall be introduced in what follows.

opposed to ‘root words’ [根詞 *gēncí*], which, namely, may be used in isolation and may also form complex words”.

Packard’s notions of ‘root word’ and ‘bound root’ (see table 2.2) seem in line with Lü S.’s position, and they seem to overlap with the notions of ‘word’ and ‘root’ in the general linguistic literature: ‘root words’ resemble (syntactic) words, since they may both be used as they are or combine with other roots to form a compound word; ‘bound roots’ resemble ‘our’ roots, which cannot be used in isolation. In some inflectional languages, ‘full’ lexical morphemes are typically bound, since they require the proper inflectional ending to be used in an actual sentence: for instance, the Italian word *amica* ‘(girl)friend’ is made up of the root *amic-*, bearing the lexical meaning ‘friend’, and the feminine singular ending *-a*. In other languages, as e.g. English, morphological processes may be based on free forms, fully-fledged words: the base form ‘dog’ may be inflected for plural (‘dogs’), or a diminutive may be attached to it (‘doggie’). In Mandarin, normally, a bound root may be used only in combination with other bound roots or (root) words, always lexical in nature, rather than grammatical (but cf. the discussion of semi-free morphemes in 2.1.1).

The problem, again, lies in the definition of the affix and in its distinction from bound roots or, more specifically, from those bound roots which are very frequent in word formation, especially if they appear in a fixed position with a consistent meaning. See, for instance, the words in (2a-c), all containing the bound lexical morpheme 工 *gōng* ‘work’:

- (2) a. 木工  
       *mùgōng*       ‘carpenter’  
       wood-work
- b. 電工  
       *diàngōng*     ‘electrician’  
       electricity-work

- c. 工德  
*gōngdé* ‘work ethics’  
 work-ethics

In the words (2a-b), 工 *gōng* is the right-hand constituent, and bears the meaning ‘worker (related to [X]<sub>N</sub>)’. In (2c), however, 工 *gōng* is found on the left side of the word, and bears the meaning ‘work’. Does this mean that the morpheme 工 *gōng* has two separate identities, and the words in (2a-b) are the product of a process of word formation akin to derivation? As we shall see, this issue has been much discussed in the literature.

The introduction of the notion of ‘affix’ in itself has generated a heated debate; as pointed out by Ma Q. (1995:101), “for Chinese, there are disagreements in understanding for any unit of analysis at all levels, but for none of these the terminological divergences are such as for the affix” (my translation). The earliest Chinese term for ‘suffix’, 語尾 *yǔwěi*, lit. ‘expression tail’ was introduced by Hu S. (1930, qtd. in Pan, Ye & Han 2004:67); such term, alongside with 語頭 *yǔtóu* ‘prefix’, lit. ‘expression head’ was in use until the forties.

Nowadays, the term which seems to have gained common acceptance for ‘affix’ is 詞綴 *cízhù* ‘word affix’, which is, in principle at least, consistent with the notion of ‘affix’ in the general literature (an affix should be attached to a word); however, in 1979 Lü S. still advocated for the term 語綴 *yǔzhù*, ‘expression affix’, since, as he suggested, such elements may combine not only to a root / word, but also to a phrase (qtd. in Pan, Ye & Han 2004:68). As said in the preceding section, items as 者 *-zhě*, which may bear the same meaning when attached to a phrase and to a root / word, posed a problem to many linguists; we shall see below (3.2.5) that such phenomena may be found also in a language as English.

For the purposes of the present study, what seems most interesting is how Chinese linguists analysed and labelled ‘hybrid’ items, i.e. our affixoids. We mentioned at the beginning of this section Qu Q.’s class of 新式的字尾 *xīnshì de zìwěi* ‘new-style suffixes’, as 家 *-jiā* ‘professional / expert of [X]<sub>N</sub>’ or 主義 *-zhǔyì* ‘-ism’, akin to Bisang’s ‘class nouns’;

these are among the items of Mandarin word formation which are most often quoted as examples of affix(oid)s in the Chinese linguistic literature. In a later work (Qu Q. 1957b, qtd. in Pan, Ye & Han 2004:66), Qu Q. proposes a partition of affixes into ‘semantic affixes’ (意義上的字尾 *yìyì shàng de zìwěi*) and ‘grammatical affixes’ (文法上的字尾 *wénfǎ shàng de zìwěi*). The former category coincides with Qu Q.’s own ‘new-style suffixes’, and he regards them as both affixes and ‘words’ / ‘lexical items’ (字眼 *zìyǎn*); ‘semantic affixes’, according to this author, are also roots (字根 *zìgēn*).

In Wang L. (1951 and 1985), the term 記號 *jìhào* ‘marker(s)’ indicates morphemes with grammatical meaning, which are opposed to 新增記號 *xīnzēng jìhào*, lit. ‘newly-added markers’; the latter, according to Wang L., are equivalent to affixes (詞尾 *cíwěi*) since they ‘correspond to Western suffixes’ (Wang L. 1951:304, qtd. in Pan, Ye & Han 2004:67, my translation). Wang L. believes that an often-recognised affix as the ‘dummy’ nominalizer 子 *-zi* (1.3.2.1) is too ‘empty’ to be regarded as a true affix, since affixes in ‘Western’ languages have a full meaning; items as 性 *xìng* ‘the property of [X]’ or 化 *-huà* ‘-ise, -ify’ (e.g. in 現代化 *xiàndàihuà* ‘modernise’) are ‘full’ enough to qualify as suffixes and, moreover, they also correspond to analogous ‘European’ affixes. It clearly appears, thus, that Wang L. had a peculiar interpretation of ‘Western’ models of derivation, and applied them in an even more peculiar way to the analysis of Mandarin. He regards as affixes those elements which translate as affixes in English or French, obviously an inadequate criterion; also, he sets as a requirement for ‘affixhood’ conveying full meaning whereas derivational meaning, although not grammatical, strictly speaking, is often the product of the semantic generalisation of some preexistent lexical item. Such Eurocentric approach, we believe, hampered the development of the research in the field of derivation and compounding and, as we shall see, similar positions have been held by other Chinese linguists.

The common terms used nowadays to designate the ‘affixoid’ are 類詞綴 *lèicízhùì*, lit. ‘simil-affix’ or 準詞綴 *zhǔncízhùì*, lit. ‘quasi-affix’ (Ma

Q. 1995); Lü S. (1979), consistently with his understanding of the affix (see above), uses the term 類語綴 *lèiyǔzhuì*, with the ‘(linguistic) expression’ rather than the ‘word’ as the base of affixation. The main characteristic of affixoids, according to Lü S., is that they are not yet fully devoid of meaning (1979, qtd. in Ma Q. 1995:103; cf. Guo F. 1982:250); his examples may also be subsumed under the label of class nouns, as e.g. 館 *guǎn* ‘building’ (圖書館 *túshūguǎn* ‘library’; Lü S. 1953, qtd. in Pan, Ye & Han 2004:67).

In all of the approaches outlined here, starting from Qu Q. in the thirties and up to Ma Q. in the nineties, ‘affixoids’ are understood roughly as Bisang’s class nouns, “generic terms on a rather high level of abstraction” (1.3.2); it is assumed that ‘affixoids’, however they are termed, still bear lexical meaning, whereas a ‘true’ affix should be ‘empty’. This is not incompatible with the definition of lexical derivation, in which no typical grammatical meaning is involved; however, it does not solve the problem of the distinction between derivation and compounding and, above all, the category of ‘affixoid’, defined as such, appears as inconsistent, both intralinguistically (what is the exact dividing line between a root and an affixoid?) and cross-linguistically (since the notion of affixoid was not conceived with Mandarin in mind). Let us see now how the phenomena of derivation and compounding have been understood in Chinese linguistics.

The terms commonly used in Mandarin to translate ‘derived word’ and ‘compound’ are, respectively, 派生詞 *pàishēngcí* and 複合詞 *fùhécí*, which may be found in a work from the fifties as Cao B. (1952, qtd. in Pan, Ye & Han 2004:77); according to Cao B., Chinese words can be either simple (簡單詞 *jiǎndāncí*), derived or compounded. However, since, again, the notion of derivation was perceived by some as a ‘foreign’ notion, its reception has not been straightforward.

Pan W. (1990:99), for instance, believes that the notion of ‘derivation’ is suitable for English, a language in which derived words, roots and stems form a ‘word family’ (詞族 *cízú*), built around a stem. However, in Chinese, according to him, affixes do not have such a function, and affixation is rather phonologically motivated; items as -化 *-huà* ‘-ise, -ify’ and -者 *-zhě*, termed by Pan W. 新興詞綴 *xīnxīng cízhuì*, lit. ‘new and developing affixes’ are regarded as ‘versatile constituent of



compounding’ (see below, **2.2.1**) and, thus, outside the domain of derivation. So, Pan W. suggests that the term 附加法 *fùjiāfǎ* ‘affixation’ rather than derivation is used, and he insists on the point that Mandarin affixes are not heads. The idea of the word families with a common root as the product of derivation (as 派生法 *pàishēngfǎ*) is found also earlier, in Zhang Shi. (1956:49 and 46-47; qtd. in Pan, Ye & Han 2004:94), who gives the root 力 *lì* ‘force’ as an example, from which you can ‘derive’ 力量 *lìliàng* ‘force’ (lit. ‘quantity of force’), 電力 *diànlì* ‘electricity’, 生產力 *shēngchǎnlì* ‘productivity’; it does not matter, thus, whether the constituent is prefixed or suffixed, and whether it conveys some ‘special’ (however defined) meaning. Needless to say, such an understanding of ‘derivation’ is quite far from that which is found in the literature on word formation.

The application of the notion of ‘compound’ to Chinese has also been far from painless. However, even a partial review of such issue would require much space and, here, we shall limit ourselves to the essential, focusing on borderline issue between compounding and derivation<sup>5</sup>. We learned from the general linguistic literature that the input of compounding are lexemes, but there has been much debate on the exact nature of such constituents, and different terms are used to define them, as seen before (**1.2.2**). Since the compound is indeed defined by its constituent parts, the identification of the latter is the key to the delimitation of the category; as the reader will expect, the positions in the literature are quite varied. Here we shall just propose a comparison between two representative works, and we shall go back to the issue in **2.2.2**.

In his influential grammar of Mandarin Chinese, Chao Y. (1968) suggests that, in order to fall into the category of compounds, a Chinese word must contain at least a bound (lexical) root. This is because, according to Chao Y., it is hard to draw a neat border between words formed by other words (i.e. free forms) and phrases, given the characteristics of Mandarin; thus, if a word contains a bound root, it is

---

<sup>5</sup> On the notions of ‘word’, ‘compound’ and ‘phrase’ applied to Chinese, see, among others, Duanmu S. (1998), Dai J. (1998), Packard (2000) and Feng S. (2001).

necessarily built in the morphological component of the language, rather than in syntax. Packard (2000) holds the opposite view, sticking to the ‘traditional’ definition of the compound as a word made up of other words (as e.g. in Fabb 1998:66). Thus, in his opinion, the ‘true’ compounds of Chinese are those words which consist of ‘root words’, free lexical morphemes (see table 2.2), whereas a word containing one (or more) bound morpheme is a ‘bound root word’, as 電腦 *diànnǎo* ‘computer’, made of the bound lexical morphemes 電 *diàn* ‘electricity’ and 腦 *nǎo* ‘brain’ (ex. from Packard 2000:81). He further suggests that the opinion according to which most Mandarin complex words are compounds is explained by the incorrect equivalence between the Chinese word 複合詞 *fùhécí* ‘compounded word’, an early label for two-syllable words, and the English word ‘compound’ (Packard 2000:78).

As we may see, the application of the ‘Western’ notions of derivation and compounding to a language as Mandarin Chinese, with significantly different characteristics of the lexicon and morphology, produced strong divergences in the treatment of Chinese word formation. Generally speaking, most contemporary scholars (especially the Chinese-speaking ones) would classify as compounds all multimorphemic expressions which qualify as ‘words’ (see Dai J. 1998 and Duanmu S. 1998), but lack the properties of phrases (see e.g. Lin H. 2001 and Dong X. 2002, 2004); those authors which recognise derivation as a separate phenomena, needless to say, will likely regard some of those multimorphemic words as derived.

Following this brief presentation of terminological issues, let us now turn to a crucial question for our research, namely the status of lexical derivation in Chinese linguistics.

## 2.2 Lexical Derivation in Chinese Linguistics

### 2.2.1 Overview

In the preceding section, we introduced the topic of the reception of the ‘Western’ notion of derivation and compounding, which gave rise to a heated debate in Chinese linguistics, and we anticipated some significant

data about the strong disagreement which one finds in the literature on Mandarin word formation. We mentioned the data from Pan, Ye and Han (2004:77 ff) which, in a sample of 14 representative works on Chinese word formation ranging from 1932 to 1982, found 340 different morphemes which have been classified as ‘affixes’ or ‘affixoids’; only 16 among those have been quoted in the majority of the works sampled, whereas 223 morphemes have been labelled as affix(oid)s only once (Pan, Ye & Han 2004:81). These 16 ‘qualified’ formants are shown in the table below.

Table 2.3. Affixes and affixoids in Pan, Ye & Han (2004)

Nr.	Morpheme	Gloss	Examples <sup>6</sup>
1	-巴 <i>-bā</i>	Noun forming affix for body parts or people	嘴巴 <i>zuǐbā</i> / <i>zuǐba</i> ‘mouth’
2	-度 <i>-dù</i>	‘Degree’	硬度 <i>yìngdù</i> ‘hardness’
3	-兒 <i>-r</i>	Noun forming affix (with prosodic function)	花兒 <i>huār</i> ‘flower’
4	反- <i>fǎn-</i>	‘Anti-, counter-’	反革命 <i>fǎngémìng</i> ‘counter- revolutionary’
5	-化 <i>-huà</i>	‘-ise, -ify’	工業化 <i>gōngyèhuà</i> ‘industrialise’
6	-家 <i>-jiā</i>	Expert, artist’	作家 <i>zuòjiā</i> ‘writer’

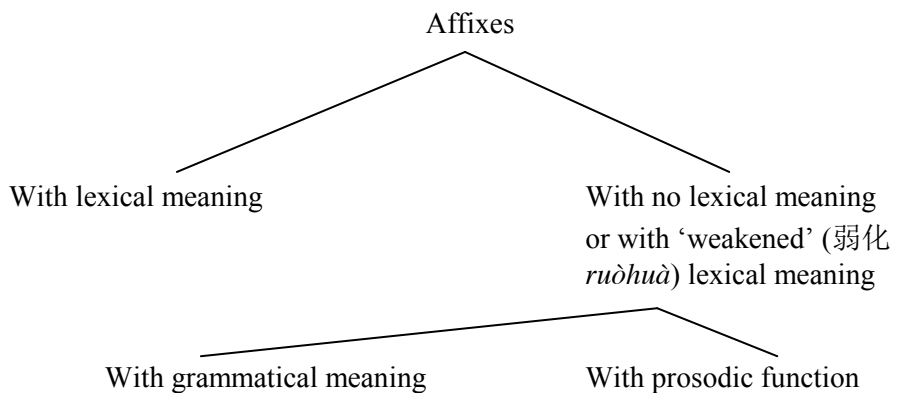
<sup>6</sup> Examples from Pan, Ye & Han (2004, appendix I).

7	老- <i>lǎo-</i>	‘Old (showing respect)’	老師 <i>lǎoshī</i> ‘master, teacher’
8	-了 <i>-le</i>	Perfective aspect marker	走了 <i>zǒu-le</i> ‘left’
9	-們 <i>-men</i>	Collective / plural marker	我們 <i>wǒmen</i> ‘we’
10	-然 <i>-rán</i>	Adverbial suffix	忽然 <i>hūrán</i> ‘suddenly’
11	-頭 <i>-tōu</i>	Noun forming affix (with prosodic function)	石頭 <i>shítōu</i> ‘stone’
12	-性 <i>-xìng</i>	‘The quality of [X]’	革命性 <i>gémìngxìng</i> ‘revolutionary quality’
13	-員 <i>-yuán</i>	‘Member, staff’	教員 <i>jiàoyuán</i> ‘teacher’
14	-者 <i>-zhě</i>	Agentive suffix	作者 <i>zuòzhě</i> ‘author, writer’
15	-著 <i>-zhe</i>	Progressive aspect marker	吃著 <i>chīzhe</i> ‘eating’
16	-子 <i>-zi</i>	Noun forming affix (with prosodic function)	桌子 <i>zhuōzi</i> ‘table’

As mentioned before, the criteria for the identification of affixes (or affixoids) in the Chinese linguistic literature are very varied, including e.g. having a fixed position, morphological cohesion, productivity (however defined) and combinability with a large number of bases. Most proposals are either very restrictive, i.e. they provide criteria according to which only an extremely small number of morphemes would be affixes, or very ‘lenient’, allowing for a potentially enormous number of morphemes to

qualify as affixes or affixoids, with no independently motivated distinction between affix(oid) and compound constituents, as we shall see in greater detail in the next section. For instance, according to Packard's definition, Mandarin would only have a very limited number of affixes (2000:174; see below, 2.2.2); if we accept such view, derivation is, at best, an embryonic phenomenon, and the focus should be on compounding or, generally speaking, the combination of lexical roots (this view is shared e.g. by Dong X. 2004, as we shall see; see also Xiao T. 1984, qtd. in Pan, Ye & Han 2004:93). Items with a relatively 'strong' lexical meaning, as 化 *-huà* '–ise, –ify' and 性 *-xìng* 'the property of [X]' are defined as 'affixoids' or as 'new and developing affixes', 'versatile constituent of compounding' (Pan W. 1990; see above, 2.1.2); as pointed out in the preceding chapter (1.3.2.2), we believe that the distinction between 'affix' and 'affixoid' is not significant at the theoretical level, especially in a language as Mandarin, and we may well do without it. Some other linguists have a broader notion of 'affix'; according to Yip P. (2000:60), even 人 *rén* 'person' (see exx. 13-14 and 26-27, CHAPTER 1), may be regarded as a suffix; such a vague and broad definition of the affix has, in our opinion, no cross-linguistic validity (see the discussion in the next section).

In the works from the sample of Pan, Ye and Han, the partition of affixes which seems to gather the most consensus may be summarised as such (2004:86):



In Chinese linguistics, affixation is taken as a ‘macro-category’, and the analysis of word formation phenomena is based on formal criteria. In our linguistic tradition, word formation processes as inflection and derivation have been defined not only according to their formal devices of expression, but also taking into consideration their syntactic, functional, structural and semantic features; in the Chinese tradition, it seems to us that to ‘prove’ the existence of morphology for the Chinese language many authors have tried to identify a class of morphemes which could appear as similar to affixes in ‘Western’ languages, from a semantic and formal point of view. Moreover, Chinese linguists have focused on the grammatical *vs.* lexical distinction, rather than on the inflection *vs.* derivation distinction, arguably because of the virtual lack of inflection in the language.

The affixes with ‘prosodic function’ (i.e. which provide a ‘prosodic support’ for the building of a word; see Feng 1998, 2001), namely –兒 *-r*, –頭 *-tou* and –子 *-zi*, form a peculiar group, since they have neither lexical nor grammatical meaning and, thus, they appear as markedly different from the ‘typical’ affixes of Indo-European languages (according e.g. to Pan W. 1990, Dong X. 2004); however, they are the only Mandarin affixes which underwent some phonological reduction (namely, loss of tone and, for –兒 *-r*, also segmental erosion), just as affixes in the languages of Europe (Dong Z. 2003). If phonological reduction is taken as a criterion for affixhood, then we would probably have to exclude all of ‘our’ lexical derivational affixes since, as seen before (1.3.2, 1.3.2.1; see the discussion of –性 *-xìng* ‘the quality of [X]’), they have no formal differences with the corresponding lexeme. Moreover, such a criterion would be inadequate for a language belonging to the East Asian area, where grammaticalization without changes in the shape of a sign is the norm (1.3.2). On the other hand, those Mandarin affixes which actually show some reduction, i.e. the above-mentioned affixes with prosodic function, are too ‘empty’ and, thus, far from the typical lexical derivational affix<sup>7</sup>; moreover, they seem to have limited (or no) productivity in the contemporary language and, thus, they are

---

<sup>7</sup> However, we must remark that both –子 *-zi* and –頭 *-tou* bear nominal word class; whatever the category of the ‘base’ morpheme, the resulting word is always a noun (e.g. 想頭 *xiǎngtōu* ‘idea’, lit. ‘think–*tou*’).

of limited significance for a synchronic analysis of Chinese word formation. We thus chose not to consider 兒 *-r*, 頭 *-tou* and 子 *-zi* in our work, also because much has been written on such topic (see e.g. Pirani 2007).

The kind of morphemes which we take into consideration in the present work is represented, in Pan, Ye and Han's sample, by items as 度 *-dù* 'degree', 化 *-huà* 'ise, -ify', 家 *-jiā* 'expert, artist', 性 *-xìng* 'the quality of [X]', 員 *-yuán* 'member, staff', 者 *-zhě* agentive suffix and, marginally, 反 *-fǎn* 'anti-, counter-' and 老 *-lǎo* 'old'; such formants appear as close to lexical derivation, both from a functional and from a semantic point of view. In the next chapter, we shall introduce our sample (3.1.2), for which we shall provide a treatment akin to that for 性 *-xìng* (1.3.2.1), drawing on historical data to assess the nature of such processes as derivation or compounding (or none of the two). We shall also discuss synchronic criteria for affixhood; in the next section section, some representative treatments from the recent literature shall be illustrated.

### 2.2.2 Recent Works

The number of works which deal with topics related to affixation, derivation and compounding in the recent history of Chinese linguistics (namely, in the XXth and XXIst centuries) is very big; due to lack of space, here we shall just quote some representative and relevant works from the recent years. We shall start our review with Chao Y.'s *A Grammar of Spoken Chinese* (1968; Chinese translation 1979) which, although not among the newest, has been a very influential work in the field, as mentioned before.

In Chao Y.'s system, affixes proper should have no lexical meaning, but only grammatical meaning, as the nominal suffixes 頭 *-tou* and 子 *-zi*, or 們 *-men* 'collective / plural marker' (2.2.1, table 2.3); also, 'empty' grammatical formants should have a high token frequency and it should be possible to list them esautively, whereas 'full' lexical morphemes should have a lower frequency. A notion which has been echoed in many works on Chinese morphology (e.g. Pan W. 1990, see the preceding section) is

that of ‘versatile’ morpheme (Chao Y. 1968; in Chao Y. 1979, 結合面寬的語素 *jiéhémiàn kuān de yǔsù*), including items as 者 *-zhě* and 人 *-rén* ‘person’ quoted above, or 師 *shī* ‘master, expert’ (工程師 *gōngchéngshī* ‘engineer’) and 士 *-shì* ‘scholar, person trained in a certain field’ (傳教士 *chuánjiàoshī* ‘missionary, preacher’); again, these may all be regarded as class nouns, in Bisang’s sense. Chao Y.’s position is rather ‘conservative’, and he believes that ‘versatile’ morphemes are compound constituents; however, he also posits a category of ‘new and developing affixes’ (新興語綴 *xīnxīng yǔzhuì* in Chao Y. 1979<sup>8</sup>), including items which we quoted here as 化 *-huà* ‘-ise, -ify’, 性 *-xìng* ‘the property of [X]’ and 論 *-lùn* ‘thesis’ (進化論 *jìnhuàlùn* ‘evolutionary theory’); the last two morphemes are also considered by Bisang (2001) as examples of class nouns (1.3.2). What is, thus, the difference between ‘versatile’ constituents and ‘new and developing affixes’? According to Chao Y., the latter correspond to affixes in foreign languages and, thus, have been imported into Mandarin.

Such a classification does not appear to employ consistent criteria, since affixes proper are identified according to functional and semantic criteria, whereas ‘new and developing affixes’ are identified according to an essentially historical criterion, namely that of the (supposed) origin; such a distinction is not even reminiscent of that between inflection and derivation, since among affixes proper we find items as 子 *-zǐ*, which may act as a nominalizer and, thus, is certainly not inflectional. Moreover, a separation between items as 師 *-shī* ‘master, expert’ and items as 論 *-lùn* ‘thesis’, which both appear in the rightmost position in a number of complex words, bearing a consistent meaning, only because of a different origin of the pattern (autochthonous *vs.* foreign), seems to be of no significance for the understanding of how the morphology of Chinese works.

A distinction between ‘proper’ affixes and ‘new and developing affixes’ is found also in Guo L. (1983). Besides this, there are two aspects of Guo

---

<sup>8</sup> Note that Chao Y. (1979) still uses 語綴 *yǔzhuì* ‘expression affix’, whereas Pan W. (1990) preferred the expression 新興詞綴 *xīnxīng cízhuì*, based on 詞 *cí* ‘word’ (2.1.2).



L.'s model which are especially interesting, in our perspective. Firstly, he proposed that a morpheme may possess different 'identities', connected with different items in the 'mental lexicon', a notion which we find also in Packard (2000, 2.1.1), and, thus, a morph may convey lexical meaning in its 'lexemic' usage, and, also, convey grammatical (functional) meaning; this is in line with the special properties of grammaticalization of Mandarin, a language in which the same item may have 'less grammaticalized' and 'more grammaticalized' identities without differences in shape (1.3.2). Thus, an item as 多 *duō* 'many, much, multi-, poly-' is both a lexeme and a 'new and developing prefix' (新興的前綴 *xīnxīng de qiánzhuì*), and the difference in status may be seen in the different distribution; whereas 多- *duō-* as an adjective normally requires the adverb 很 *hěn* when modifying a noun, as in the phrase 很多人 *hěn duō rén* 'many people', 多- *duō-* as a prefix indeed does modify directly a noun, as in the words 多神教 *duōshénjiào* 'polytheism' and 多音節 *duōyīnjié* 'polysyllabic'. As seen in the preceding chapter (exx. 15-16), the criterion of the difference in distribution has been used also for the justification of affix(oid)s in languages as English.

Secondly, he takes into consideration meaning shift as a diagnostic for affixhood: he regards as 'typical suffixes' (典型的後綴 *diǎnxíng de hòuzhuì*), besides -子 *-zi*, -兒 *-r* and -頭 *-tōu*, also -者 *-zhě* and -家 *-jiā* 'expert, artist'; the latter, which bears lexical meaning, is regarded as a suffix by Guo L. because in many new words which contain such item it does not mean 'someone who is engaged in some specialistic career' (從事某種專業事業的人 *cóngshì mǒuzhǒng zhuānmén shìyè de rén*), but, rather, it is a general human suffix, as in 空想家 *kōngxiǎngjiā* 'dreamer, visionary' or 陰謀家 *yīnmóujiā* 'schemer, conspirator' (1983:254-256). Thus, the 'emptying' of meaning and the productivity of the pattern, together with the function of assigning a stable word class (in this case, noun), are the proof of the fact that -家 *-jiā* is a 'genuine' suffix.

To sum up, the criteria put forth by Guo L. for delimiting the category of 'typical' affixes and that of 'new' affixes show interesting analogies with those

which we outlined in CHAPTER 1 for Mandarin Chinese: ‘emptying’ (i.e. generalisation) of meaning, assigning a stable word class, distributional differences with a corresponding lexeme. Once again, however, the distinction between ‘proper’ and ‘new’ affixes is somehow artificial, and the semantic criteria for it are unclear.

A particularly interesting analysis is that of Ma Q. (1995), who makes an explicit distinction between processes which may or may not build a new lexeme (詞位 *cíwèi*). In Ma Q.’s model, new lexemes may be built by compounding and by affixation, and thus his ‘affixation’ is virtually synonymous with ‘derivation’; if the adding of a morpheme does not produce a new lexeme, then such morpheme, conveying grammatical meaning, is an ‘auxiliary particle’ (助詞 *zhùcí*) and not an affix, as the aspect markers –了 *-le* and –著 *-zhe* (Ma Q. 1995:110-111). In a sense, thus, he bases his treatment on the ‘familiar’ categories of compounding, derivation and inflection, defined according to semantic, functional and formal criteria, consistently with the general literature. He does not explicitly talk about inflection, and this is not surprising, given the typological features of Chinese already discussed; however, he sets a borderline between ‘affixation’ (i.e. derivation) as a word-formation device and the ‘auxiliary particles’ which contribute only (typical) grammatical meaning. He avoids the term ‘affix’ for items as aspect markers since they may attach both to words and to phrases, and thus resemble more clitics than affixes.

According to Ma Q., compound constituents and (derivational) affixes may be clearly distinguished, without resorting to criteria as ‘autochthonous vs. foreign’ (cf. Chao Y. 1968, above); he actually insists on the point that, even when there is a correspondence between a Mandarin word and a foreign derived word (as, say 機械化 *jīxièhuà* ‘mechanise’; Wang L. 1980:311), the formant at issue is ‘domestically made’ (國產 *guóchǎn*; Ma Q. 1995:107), and its development is independent from the foreign model (see below, 3.2.3). The distinction between roots (lexical morphemes) and affixes is based, firstly, on distributional criteria, as in other models discussed before; affixes have a fixed position, i.e. they are either prefixes or suffixes, and, also, they may not constitute a word in isolation. An interesting point is that such features are understood by Ma Q. as ‘relative’ (相對 *xiāngduì*): thus, a morpheme may be used freely in syntax, but has a fixed position as a bound word constituent;

otherwise, a polysemic item may be bound and used in a fixed position in one of its meanings (義項 *yìxiàng*), but used freely in others (Ma Q. 1995:114-115). In line with the approaches of Guo L. (1983) and Packard (2000), a morph is thus allowed to have more than one ‘identity’ which are independent of one another, albeit connected.

The category of affix is then further divided into ‘true affixes’ (真詞綴 *zhēn cízhù*) and ‘quasi-affixes’ (準詞綴 *zhǔncízhù*). ‘True’ affixes are completely devoid of meaning, and they provide ‘prosodic support’ to a word, as e.g. the often-quoted 子 *-zi*, 頭 *-tōu*, and 兒 *-r* (for a list, see Ma Q. 1995:116); they are always bound and, also, they underwent some phonological reduction, mostly loss of tone. ‘Quasi-affixes’ are morphemes with either concrete (實在 *shízhài*) or abstract (抽象 *chōuxiàng*) meaning, with a (relative) fixed position, which cannot be words; they do not undergo phonological reduction as true affixes, but they share with the latter ‘categorical meaning’ (範疇義 *fànchóuyì*), i.e. word class and, possibly, semantic category (as class nouns; Ma Q. 1995:121). Thus, by definition, a quasi-affix corresponds to a root in the lexicon; Ma Q. further argues that the ‘affixal’ meaning of a given item must be an ‘extension’ (引申 *yīnshēn*) of its ‘basic meaning’ (基本意義 *jīběn yìyì*)<sup>9</sup>. Let us now provide an example of Ma Q.’s treatment of roots and quasi-affixes.

The bound morpheme 語 *yǔ* ‘language, expression’ may be used to form about any glottonym in the right-hand (head) position, as in 日語 *rìyǔ* ‘Japanese (language)’; he may also be found on the left side of words, acting as a modifier in terms as 語法 *yǔfǎ* ‘grammar’ (‘language-law’) or 語義 *yǔyì* ‘semantics’ (‘language-meaning’), bearing its basic, ‘core’ meaning of ‘language’. Such an item does not appear in a fixed position,

---

<sup>9</sup> By ‘basic meaning’, Ma Q. (1995:119-120) means the first or the first two meanings which are listed in a dictionary for a character / morpheme. Although such a method is far from perfection, it is generally true that the meanings for an entry are ordered in a chronological fashion, from oldest to newest, and thus it is likely that the extensions in meaning will not be found in the first place; note, however, that chronological order in the listing of meaning is not necessarily always respected (see De Mauro 2005:80-82).

even in a relative sense, and it is productively used in its basic meaning, rather than in an extended meaning and, thus, is not an affix. According to Ma Q.'s own estimate (1995:113; see the source for a description of the sample considered), one may find at least 1277 'versatile' morphemes (Chao Y.'s 結合面寬的語素 *jiéhémiàn kuān de yǔsù*, introduced above; see also the discussion of Yip P. 2000, 2.2.1 and below); among those, (quasi-)affixes must be identified by applying his criteria.

A versatile morpheme which qualifies as a (quasi-)suffix is 學 *xué* 'study, learning, branch of learning'. Ma Q. (1995:114) compares complex words as 法學 *fǎxué* 'science of law' and 社會學 *shèhuìxué* 'sociology', on one side, and 學制 *xuézhì* 'educational system', 博學 *bóxué* 'erudition' (lit., 'plentiful-learning') or 村學 *cūnxué* 'village school'; whereas in the words belonging to the first group 學 *xué* is used with a fixed meaning ('branch of learning'), always in suffixal position, in the words from the second group it appears in different positions, conveying disparate (albeit connected) meanings, as 'study', 'learning' and 'school'. Thus, 學 *xué* is a good example of a morpheme with relative fixed position and meaning and, also, relative bound status, since it may be a free lexeme in its verbal meaning ('to study'). Also, its 'affixal' meaning is not one of its 'core' meanings (fourth meaning in CCD 2002; see footnote 9); we shall go back to the analysis of 學 *xué* in 3.2.1. Ma Q.'s model (as in Ma Q. 1995) includes many interesting aspects, and provides a sound treatment of lexical derivation, consistent with many proposals which may be found in the general literature (see 1.2.2); however, the distinction between 'true' affixes and quasi-affixes deserves further discussion, as we shall see.

Two very different proposals, which may be deemed as representative of the most 'radical' approaches, are those by Packard (2000) and Yip P. (2000), which we already mentioned before (2.2.1, 2.1.2, 2.2.1). Packard believes that Mandarin has both 'grammatical affixes', as aspect markers (–了 *-le*, –著 *-zhe*), and 'word-forming affixes', as 可– *kě-* '–able' (可吃 *kěchī* 'edible') and –度 *-dù* 'degree'; most of his 15 examples are the same as those summarised in table 2.3. He associates 'grammatical affixes' with inflection, and 'word-forming affixes' with derivation,

although he says that they are not equivalent (Packard 2000:77); he believes that a word-forming affix should possess at least two of the following characteristics (2000:70):

- a. may change the word-class of the base
- b. apply selectively
- c. have a relatively variable and unpredictable meaning
- d. attach to free or bound morphemes

If we look at the properties of derivation summarised in **1.2.2**, the only ‘anomalous’ feature is c., since, according to Scalise, Bissetto & Guevara (2005), the meaning of regular derived words is *not* dependent from the context; rather, it is so for compound words (see the ‘dog bed’ example). Therefore, this seems to be evidence in favour of compound constituent status, rather than derivational affix status. What about the distinction between word-forming affixes and compound constituents, then?

Packard believes that word-forming affixes, even though they do not convey typical grammatical meaning, are grammatical in nature, and although “the distinction between ‘grammatical’ and ‘lexical’ exists on a continuum rather than dichotomous scale (...), it is nonetheless possible to draw a distinction between the two” (Packard 2000:71). Thus, between two semantically close items as 者 *-zhě* ‘one who is/does X’ and 員 *-yuán* ‘person whose job/position is X’ (保健員 *bǎojiànyuán* ‘health worker’; our example), only the former may be classified as a word-forming affix, because of the ‘generality’ and ‘abstractness’ of its meaning, “since the meaning of *-zhě* entails that of *-yuán*” and, also, “words formed with *-yuán* tend to have meanings which are more ‘fixed’ and ‘lexicalized’ (Packard 2000:71-73, his glosses). Moreover, word-forming affixes tend to be more productive than bound roots, if the lower number of items listed in a (reverse) dictionary is indicative of higher productivity, since “the vast number of forms which may take *-zhě* as an ending precludes their being exhaustively listed”. Packard also suggests that the addition of word-forming affixes, as opposed to bound

roots, “involves a grammatical change that, in general, is on a par with that caused by a logical operator (...) or a change in thematic role”, as ‘agent’, ‘patient’, ‘having/conferring the property of’, etc. (2000:73).

Even if such an approach may seem appealing, many problems arise in its application to concrete cases. The notions of ‘generality’ and ‘abstractness’ on which the distinction between affixes and roots rests, are vaguely defined by Packard: even if we admitted that 者 *-zhě* has a more general meaning than 員 *-yuán* ‘person whose job/position is X’, where would we set the threshold between ‘lexical’ and ‘grammatical’? What is the point in comparing only two items, among the hundreds of morphemes which have affix-like properties? As to productivity, Packard’s quantitative measure is an indirect one, and it is synchronic in nature; other methods could give different results, especially if diachronic data are taken into consideration. More generally, while productivity might be a prerequisite for the development of a lexical item into a derivational affix (see 1.2.2), it is unclear whether it is also a good criterion to prove that a morpheme is ‘more grammatical’ than another; this probably depends also on *which* measure of productivity we choose to adopt (see Plag 1999, Bauer 2001b). We believe, nevertheless, that it is of doubtful significance to compare, again, only two formants, for the same reasons stated above. The idea that derivation has a limited set of functions is not new: Beard (1998), for instance, suggests that derivational meanings could be based on case functions (subject, object, place, etc.). However, there seems to be a rather strong consensus in the literature on the point that derivational meanings form an open, potentially unlimited set, differently from inflection (see 1.2). Let us now compare Packard’s analysis with Yip P.’s (2000).

The notion of ‘affix’ is characterised by Yip P. “by adopting the two helpful criteria of *desemanticization* and *versatility*” (2000:59, his emphases). He argues that some morphemes (his ‘mononyms’) are “categorical props” or “mere word class indicators” which do not affect the meaning of the morpheme(s) they attach to or, else, “affect them in a most general way (e.g. 记者 *jìzhě* ‘journalist’ in which *zhě* suggests ‘doer’)”; such morphemes are to be regarded, according to Yip P., as “*sub-morphemes* or canonical forms”, under the label of affixes, and

words as 記者 *jìzhě* ‘journalist’ are derived words.

Here, once again, reference is made to some notion of ‘desemanticization’ (compare e.g. Packard’s ‘generality’ and ‘abstractness’); we believe, however, that –者 *-zhě* is not a good example of a ‘desemanticized’ affix, since it has always been used in a similar function, namely as a pronoun substitute or nominaliser since the Old Chinese period, as in 耕者 *gēngzhě* ‘ploughman’ (Pulleyblank 1991:66-67), although, admittedly, the modern usage is much more restricted and derivation-like than its classical usage, as we shall see in 3.2.5. As for ‘versatility’, a notion which is drawn from Chao Y. (1968; Yip P. 2000:59, fn. 7), it resembles closely some conception of productivity, and it is thus based on synchronic data on attested words; Yip P. does not provide any further details on ‘versatility’, and we may thus infer that he means, roughly, ‘appearing in a high number of words’, ‘having a high type frequency’, as in Chao Y. (1968). The parameters of desemanticization and versatility, thus, appear as vague, and they do not seem to have any significance as such, especially in a cross-linguistic perspective, for a distinction between derivational affixes and compound constituents. Yip P. (2000:59) himself admits that it is not yet possible “to draw absolute and unmistakable distinctions between derivation and compounding”; he lists a large number of affixes, such as 36 suffixes for human nouns, including the above mentioned 人 *rén* (2.2.1) and 客 *kè* ‘guest’ (旅客 *lǚkè* ‘tourist’), but he does not provide a detailed reason for the inclusion of all of them. An item as 匠 *jiàng* ‘artisan’, for instance, does not seem to us to be particularly ‘desemanticized’, since it expresses its basic lexical meaning in complex words (as 石匠 *shíjiàng* ‘stonemason’).

In Sun Y. (2000), again, ‘affixes’ are morphemes which form new lexemes and, thus, are defined just as derivation. She does not take ‘desemanticization’ as a criterion for affixhood; rather, she believes that conveying lexical meaning, connected with the ‘lexical’ meaning of the corresponding lexeme, is a specific characteristic of Mandarin affixes. She also maintains that such characteristic actually favoured the development of affixation in Chinese, and that ‘completely emptied’ morphemes may

not easily develop in the language. She also disagrees with the *communis opinio* that it is necessary to separate ‘proper’ affixes and ‘affixoids’, since the latter is a notion which was created because of the failure of Mandarin affixes to comply with criteria for affixhood designed for inflectional languages; once realised that affixation (derivation) in Mandarin is a developing phenomenon, and many items possess hybrid properties as they are not yet fully grammaticalised, an intermediate category between ‘affix’ and ‘lexeme’ / ‘root’ is not required. She thus suggests that the ‘modern’ model of affixation for Mandarin is that of items as 者 *-zhě* ‘agentive suffix’, having a ‘rather strong’ lexical meaning, akin to Bisang’s class nouns, once again; such model has not yet fully developed. The innovative aspect of Sun Y.’s model is the abandoning of the ‘desemanticization’ model and of the consequent distinction between affixes and affixoids; it is again unclear, however, what semantic criteria she employs to identify her affixes. We shall go back to Sun Y.’s proposal in **3.2.1.1**.

In **2.1.1**, we provided an outline of Dong X.’s understanding of the ‘morpheme’ in Mandarin. According to Dong X. (2004), grammatical morphemes may be free, bound or semi-free (table 2.1); bound grammatical morphemes correspond to affixes, as the nominal suffixes 子 *-zi*, 兒 *-er* and 頭 *-tōu*, or 化 *-hua* ‘-ise, -ify’ and 性 *-xing* ‘the property of [X]’. However, she believes that it is unlikely that Mandarin derivational morphology may develop further, because of the characteristics of the Chinese lexicon. Mandarin Chinese, differently from a language as English, possesses the ‘character’ (漢字 *Hànzi*), which represents the ‘junction’ of phonology and grammar (Dong X. 2002:103-106), with a stable relationship between character, syllable and meaning, making it difficult for such unit of writing to lose its semantic value. Many complex words of Modern Mandarin were, originally, phrases; lexicalisation, however, in most cases does not involve the loss of the meanings of the constituents and the boundaries between morphs do not become blurred, and the same morphemes may combine with yet other morphemes to form words. In English, according to Dong X., the ‘junction’ of phonology and grammar is in the word, rather than in the morpheme, and there is no regular correspondence between units of



meaning (i.e. morphemes) and units of phonology / prosody (as the syllable in Chinese). The relative instability in the sound shape makes it easier for meaning to ‘become blurred’ (變得模糊 *biàn de móhu*) or even be lost, turning compound constituents into derivational affixes; loss of meaning and of morpheme boundaries may result in loss of motivation for a compound, which is then perceived as a simple word (see the often-quoted English example *lord* < Old Eng. *hlaforð* < *hlaflweard* ‘one who guards the loaves’). Whereas the dominating pattern of compounding in English is the combination of words, in Mandarin it is the combination of bound roots. Such bound roots may be used quite freely, in word formation, in different positions inside a word and, thus, never grammaticalise into affixes. Such view is in line with Wu F.’s position, who states that Mandarin ‘grammatical words’ (語法詞 *yǔfǎcí*) and ‘clitics’ (附著詞 *fùzhùcí*) do not evolve into inflectional markers, but, rather, combine with a neighbouring word and become morphemes inside that word (2005:25); thus, items as 可- *kě-* ‘-able’ (可吃 *kěchī* ‘edible’, quoted above) and -者 *-zhě* ‘agentive suffix’ (see below, 3.2.5 and 3.2.6) are just morphemes inside a words and, thus, they are lexicalised rather than grammaticalised.

Therefore, the distinction between affixes, affixoids and bound lexical roots is not very relevant for Mandarin (Dong X. 2004:41); any regular word formation pattern with a fixed constituent and a variable slot (with a part of speech and semantic constraints), in which the semantic relationship between the constituents is stable, and the resulting words have a predictable meaning, may be regarded as a ‘word formation rule’ (詞法模式 *cífǎ móshì*). Thus, the morpheme 人 *rén* is the fixed constituent in two different ‘word-formation rules’, ‘toponym + 人 *rén*’ (廣州人 *Guǎngzhōurén* ‘Cantonese’) and ‘ethnonym + 人 *rén*’ (藏族人 *Zàngzúrén* ‘ethnic Tibetan’; compare exx. 12-14, CHAPTER 1). Such a proposal appears as very similar to Booij’s Construction Morphology (1.2.2) in that the emphasis is on the patterns (schemas, in Booij’s terms) rather than on the individual processes; however, such an analysis may be convenient in a strictly synchronic and idiolinguistic descriptive

perspective, but it does not tell us anything about the diachronic processes, lumping together patterns with a different history, and it lacks cross-linguistic breadth, stressing the peculiarities of Mandarin rather than its commonalities with the World's languages.

Let us now turn to a summary of the main points discussed here.

### 2.2.3 Summary

To sum up, in the Chinese literature the most frequently proposed criteria for the identification of affixes appear to be the 'emptying' of meaning, having a stable position, a stable meaning and productivity; the definition of such notions, however, is not the same for all authors. The very concepts of 'affix' and 'derivation' are not understood in the same way by all linguists. Moreover, the criteria provided are often vaguely defined, especially as far as desemanticisation is concerned. No author provides a semantic 'threshold' for affixhood, and we believe that this is because this would make no sense at all; any consideration about a shift in meaning of a sign, be it loss, blur or anything else, can only be based on a comparison with the historical meanings and functions of the same sign, and not with some other (cf. Packard's analysis of 者 *-zhě* and 員 *-yuán* above). As to the 'affixoid', we already proposed that such label is significative only at a descriptive level and it is not relevant for Mandarin, since the tendency for the language is that of having grammaticalised signs which are formally identical to their corresponding lexemes.

Having illustrated the main issues concerning the phenomenon of '(lexical) derivation' and related notions in Mandarin Chinese, we may now propose the analysis of our language data. In the next chapter, we shall provide a diachronic and synchronic analysis of some Chinese morphemes which, given their 'hybrid' properties, are good examples of items at the borderline between derivation and compounding.

## **CHAPTER 3**

### **DERIVATION OR COMPOUNDING? THE MANDARIN CASE**

This chapter constitutes the core of our research, as we shall analyse individual cases of Mandarin morphemes at the borderline between derivation and compounding, possessing hybrid properties. Since the number of items which could be included in our study is extremely big, we selected a small sample of morphemes, divided into five groups, which we regard as representative of the most interesting and relevant derivation-like phenomena in Chinese word formation.

#### **3.1 The Boundary between Derivation and Compounding in Modern Mandarin Chinese**

##### **3.1.1 Methodological Issues**

As seen in the preceding chapter, in the Chinese linguistic literature there seems to be no consensus on whether (affixal) derivation is a productive word formation phenomenon in Modern Mandarin, and even those authors which support the view that Mandarin has derivative morphology do not agree on how the ‘affix’ should be defined, as opposed to a (frequently used) compound constituent. The notion of ‘affixoid’ has been put forth to describe such borderline items; however, even this solution has generated controversy, as different authors assign different items to such class.

In what follows, we shall carry out our analysis of historical and synchronic data on a sample of morphemes which are ‘good candidates’ for derivational status. Our sample is divided into five classes of morphemes, identified according to different motivations; after a general presentation of each group, we shall deal in detail with one or two representative formants per set, for which more data could be found. As mentioned, the criteria according to which the five classes have been identified are not homogeneous, as they include distribution, (supposed)

origin, period of development, etc.; such dishomogeneity is functional for our illustration, as the grouping will help us to gain a better understanding of the phenomena at issue here.

In the next section, we shall thus introduce our sample, providing the motivations behind each class and a few examples of representative morphemes for each group; the remainder of the chapter is devoted to the analysis of the individual classes, with the aim of gaining a better understanding of derivation, both in Mandarin and as a general phenomenon of word formation.

### 3.1.2 Our Sample

As mentioned in the preceding section, the reasons behind the classes of our sample are different for each of them; here, we shall describe and explain them.

The items in the first group in our sample coincide roughly with Bisang's class nouns (see 1.3.2), and are often identified as 'affixoids' or 'versatile morphemes' in the literature (2.1.2, 2.2.2). They are nominal formants with a rather generic meaning, and they are employed in a fixed position with a certain meaning in word formation; examples of such morphemes are -學 *-xué* 'branch of learning', 家 *jiā* 'expert of [X]<sub>N</sub>' or -員 *-yuán* 'member, staff', all seen in CHAPTER 2. Such a class of items has a potentially enormous size, as said before: Ma Q. (1995) enumerates as many as 1277 'versatile' morphemes from his sample (2.2.2) and, given the vagueness of the definition of the category, we will not try to define it clearly. As we shall see, the class nouns of Mandarin mostly convey their basic, lexemic meaning in word formation and, thus, resemble more compound constituents than derivational affixes. However, through the analysis of the historical evolution of individual items, one may find out that for some of them the diachronic processes which operate are those of grammaticalisation, and that their derivational status is supported by historical evidence, as seen for -性 *-xìng* 'the property of [X]' (1.3.2.1).

The second class of formants is not very dissimilar from that of class nouns, as it includes nominal morphemes which bear a general meaning, as -吧 *-bā* 'bar' (話吧 *huàbā* 'call shop') or -米 *-mǐ* 'fan' (球迷 *qiú mí*

‘ball game fan’). These formants have been grouped separately not because of some radical difference from a (strict) linguistic point of view, but rather because of the treatment they received in the literature; an item as 吧 *-bā* ‘bar’ is often regarded as a paradigmatic example of a ‘new tendency’ towards the creation of affixes in Mandarin (see e.g. Wu Y. 2000, Fan L. 2002 Dong Z. 2003, among others). As we shall show below (3.2.2), the mechanism of analogy and the high productivity of such formants played a key role in their perception as suffix(oid)s; also, the fact that they became productive in the Mandarin lexicon (relatively) recently enables us to provide a more detailed picture of their development and, thus, they deserve a separate treatment from class nouns.

In class three, we grouped the affixes which are ‘commonly’ accepted as affixes in the literature (see table 2.3), as 化 *huà* ‘-ise, -ify’ and 性 *-xìng* ‘the property of [X]’; since the latter has been already dealt with extensively (in 1.3.1.2), we shall focus on the former item. Such formants are usually regarded as affixes even in the most ‘conservative’ works (as Packard 2000; see 2.2.2); both because they have a stable word class and, also, because they are semantically and functionally analogous to ‘Western’ (i.e. Standard Average European) affixes. After having determined that they actually qualify as derivational affixes, we shall test the hypothesis that the contact with SAE languages had an influence on their grammaticalisation, focussing on the ongoing developments for 化 *-huà*, evaluating the interaction between external influx and language-internal mechanisms.

The fourth class we decided to take into consideration includes the often quoted ‘agentive suffix’ 者 *-zhě* and 式 *-shì* ‘model, style’ (日式 *rìshì* ‘Japanese-style’), which are two rather peculiar items, compared to other ‘candidate’ affixes. What sets them apart is that they seemingly may combine both with lexical and with phrasal elements, as shown above for 者 *-zhě* (see exx. 22-23, CHAPTER 1); also, their pathways of development differ from those of class nouns (group 1), as we shall see (3.2.5). Both 者 *-zhě* and 式 *-shì* possess properties both of bound word formation elements and of syntactic items (Zhang Yi. 2002a, 2002b); for such reason, as seen before (2.1.1), they are regarded as ‘semi-free

morphemes' in Dong X's model (2004). We shall test the hypothesis that they are items in the course of grammaticalization, which would explain their hybrid properties.

The fifth group is easily motivated, since it includes prefixed items, as 非- *fēi*- 'non' (非暴力 *fēibàoli* 'nonviolence'), 反- *fǎn*- 'counter, anti' (反間諜 *fǎnjiàndié* 'counterespionage'; see table 2.3) and 超- *chāo*- 'ultra-, super-' (超聲波 *chāoshēngbō* 'ultrasonic wave')<sup>1</sup>. Such items mostly correspond to prefixes of SAE languages and, just as 'our' prefixes, they apparently do not change the category of the word/root they attach to; we shall evaluate whether this is actually true and if language contact played a role in the history of some of these formants. In order to gain a thorough understanding of the productive prefixed morphemes of Mandarin, we shall compare the kind of items quoted above with morphemes as 可- *kě*- '-able' (可達 *kědá* 'accessible') and 難 *nán* 'difficult, unpleasant' (難寫 *nánxiě* 'difficult to write'); such morphemes seem to act as (categorical) heads, since they apparently determine the adjectival class of the whole complex word (Ceccagno & Scalise 2006:252).

The groups briefly illustrated here are summarised in table 3.1:

Table 3.1. Our sample

Cl.	Morpheme(s) considered	Other examples	Remarks
1	-學 <i>-xué</i> 'branch of learning'	-家 <i>-jiā</i> 'expert, artist', -員 <i>-yuán</i> 'member, staff'	Akin to class nouns
2	-吧 <i>-bā</i> 'bar'	-迷 <i>-mí</i> 'fan', -族 <i>-zú</i> 'tribe, group'	Patterns with a comparatively short history
3	-化 <i>-huà</i> '-ise, -ify'	-性 <i>-xìng</i> 'property of [X]'	Commonly accepted as suffixes, correspond to SAE affixes

<sup>1</sup> See Wang F. (1998) and Yip P. (2000) for a list of possible prefixes in Mandarin.

4	-者 <i>-zhě</i> ‘agentive suffix’ -式 <i>-shì</i> ‘model, style’		Items which attach both to word/roots and to phrases, ‘semi-free morphemes’
5	非 <i>-fēi-</i> ‘non’ 可 <i>-kě-</i> ‘-able’	反 <i>-fǎn-</i> ‘Anti, counter’ 超 <i>-chāo-</i> ‘ultra-, super-’ 難 <i>-nán-</i> ‘difficult, unpleasant’	Prefixed morphemes, partly correspond to SAE prefixes

Let us now turn to the analysis of each group, taking into consideration the relevant morphological and lexicographic issues.

### 3.2 Analysis of the Sample

#### 3.2.1 *Class Nouns in East and South-East Asian Languages*

In the preceding chapter, we labelled as class nouns a large group of formants of Mandarin Chinese; class nouns are defined by Bisang (1996:525) as “generic terms on a rather high level of abstraction from which more concrete nouns can be derived by further determination (cf. e.g. Engl. tree → apple tree)” (1.3.2). Such a notion may be found also in the Chinese literature: items with the above mentioned characteristics are termed 大類名 *dàlèimíng*, lit. ‘name(s) of a major type’ in Cheng X. (1992c, 1992d). We already mentioned that Bisang believes that class nouns are a product of the grammaticalization of nouns (1996:533, 546-547) and some of them can be seen as derivational affixes; nevertheless, since he regards compounding and derivation as categories with blurred boundaries, he does not provide criteria for a clear-cut distinction between compound constituents and derivational elements (Bisang 2001). Thus, according to Bisang, although we cannot (always) equate class nouns to derivational affixes, it is anyway true that the former are just not the same as other nouns in the lexicon of a language; however, he fails to provide a rigorous definition of such category of word

formation elements. The same vagueness is a problematic aspect also in the definition of all those notions akin to class nouns in Chinese linguistics; among those, the proposal which appears as most interesting is that of Ma Q. (1995), which, as seen before (2.2.2) defines ‘quasi-affixes’ (a subclass of affixes, rather than an intermediate category between ‘root’ and ‘affix’) not only as bound items which indicate a (broad) semantic category and appear in a fixed position, but, also, he further specifies that their ‘affixal’ meaning must be an extension of their basic (lexical) meaning. Such a characterisation of affixes is in line with our ‘diachronic’ approach (1.3.2.1).

Class nouns are not only a Chinese phenomenon; they are a common sight in East and South-East Asian languages, as e.g. Thai *khon* ‘person’ → *khon-khâi* ‘sick person’ (Bisang 1996:546). We have seen before (1.3.2) that Bisang sees class nouns as one of the ‘attractor positions’, “slots which *attract* linguistic items in order to grammaticalize them” by the mechanism of analogy, in a ‘maximum pattern’. Maximum patterns are constructions, and they can be a frame for processes of grammaticalization and, also, a product of those processes; even a single word may be regarded as a construction<sup>2</sup> (Bisang 1998:13-14). We thus remarked that in the framework of Construction Morphology the constructions of word formation, termed word formation schemas, are both ‘produced’ by language users as they encounter a certain number of words of a certain type and, also, they are employed to build new words (Booij 2009:207). Thus, they appear as analogous to Bisang’s constructions (maximum patterns), and they contain slots which may ‘attract’ new items: a constructional idiom is the product of the conventionalisation of an item indicating a rather general notion, analogous to a class noun (see the Dutch example *boer* ‘farmer’ > *-boer* seller of [X]<sub>N</sub>, 1.2.2).

In a language contact situation as that of East and South-East Asia, the inferences which are born out of the communicative needs between

---

<sup>2</sup> In constructionist approaches, generally speaking, “[a]ny linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency” (Goldberg 2006:5); in such perspective, individual words and even morphemes may be analysed as constructions.



speakers with different native languages may increase the necessity of mechanisms as reanalysis, metonymy and metaphor (see 1.2.2, 1.3.2.1), which are at work in processes of grammaticalization; Bisang (1998:17) even suggests that “the existence of linguistic areas (*Sprachbünde*) may be due to the cross-linguistic spreading of the above mechanisms of grammaticalization”. According to Bisang, this happens with the diffusion of constructions, the function of which is triggering the above mentioned mechanisms. A full discussion of the (possible) relationship between language contact and the mechanisms of grammaticalization is far beyond the scope of our research; in what follows, we shall limit ourselves to a specific aspect of language contact in the area to which Chinese belongs, namely contact through writing.

The aspect of language which is normally seen as primary is, needless to say, speaking: as we all know, the majority of the World’s languages do not possess a writing systems, and their transmission is only oral. Accordingly, the study of language contact is focussed on the spoken interactions between language users; however, the influence of languages on one another may also take place as a result of contact by writing, “a form of indirect language contact in which the great majority of the people involved do not interact at all. Consequently, contact is limited to a small minority of individuals on the basis of political, economic, religious and cultural relations” (Bisang 2001:189). In the development of Mandarin word formation patterns as those based on class nouns, it is very likely that foreign models played an influential role, interacting with tendencies inherent in the Chinese language: the main ‘medium’ for the transmission of such influx are, predictably, translations of European works (English, but also German, French, etc.), involving the ‘reproduction’ of many words which were not part of the Chinese lexicon (on the reception of Western notions through translated works see, among others, Masini 1993 and the contributions in Lackner, Amelung & Kurtz 2001). To give but a couple of examples, words as 化學 *huàxué* ‘chemistry’ and 光學 *guāngxué* ‘optics’ were introduced into the Chinese lexicon by Western translators during the second half of the XIXth century (Masini 1993:81 and appendix II). The impact of translations on the Chinese lexicon, especially during the XIXth and the early XXth century was tremendous; as we shall see, whereas European languages could have only an indirect

influence on Chinese (i.e. through loan translations), the fact that many Japanese neologisms were actually written with Chinese characters provided the conditions for a stronger impact on the Chinese lexicon (Wang L. 1980:519 ff., Masini 1993:iii).

How were these words rendered in Chinese? The Chinese language is characterised, as already mentioned (1.1.4), by a strong tendency towards the overlapping of units of speech (prosody), units of meaning and units of writing, i.e. towards the correspondence of syllable, morpheme and character. Morphemes which are made of more than one syllable / character are quite uncommon (see ex. 6, CHAPTER ONE), and in most words there are as many morphemes as there are syllables; “there is a rather strong (...) principle in Chinese and Vietnamese which makes sure that each syllable must have its meaning and which somehow seems to be connected with the fact that the smallest meaningful element has to be the syllable” (Bisang 2001:192; see also Sun J. 2005). Hence, polysyllabic unanalysable words have never been welcome in the Chinese lexicon (see Sapir 1921, qtd. in Bisang 2001); the borrowing of the phonological form of foreign words often involves the creation of such unsegmentable items, as 奧林匹克 *àolínpíkè* ‘olympic’, and is thus a dispreferred strategy. As Bisang (2001:191) puts it, “[w]hen Chinese came in contact with Sanskrit and later with languages such as English, French, German and Russian it formed new words by using its own lexicon and by extending rules of word formation which already existed in the language”. A classification of the strategies for the enrichment of the Chinese lexicon has been put forth by Masini (1993:128 ss.):

- a. ‘phonemic loans’, i.e. the reproduction of the phonological form of a foreign word, as 鴉片 *yāpiàn* ‘opium’
- b. ‘hybrids’, i.e. the combination of a phonemic loan with a native morpheme, usually a class nouns, as 基督教 *jīdūjiào* ‘christianity’, the sum of the loan *jīdū* ‘Christ’ and the morpheme *jiào* ‘teaching, religion’)

- c. ‘loan translations’ (or ‘syntactic loans’), words or phrases coined in China, based on the structure of a foreign term, as 鐵路 *tiělù* ‘railway’ < Ger. *Eisenbahn*
- d. ‘semantic loans’, “terms which existed in the traditional lexicon, but assumed a new meaning on the basis of a foreign model” (1993:129), as 新聞 *xīnwén* ‘news’ < ‘recently heard facts’
- e. ‘graphic loans’, i.e. the adoption of Japanese words or, rather, the borrowing of their written forms and of their meaning, while the reading is Chinese, as 電話 *diànhuà* ‘telephone’ (lit. ‘electric words’) < Jap. *denwa*
- f. ‘autochthonous neologisms’, words which were coined without resorting to a foreign model word, as 飛機 *fēijī* ‘airplane’, lit. ‘fly-machine’

The category of ‘graphic loans’ may be further divided into the subcategories of ‘original loans’, i.e. autochthonous Japanese words (made of Chinese characters/morphemes) and words taken by the Japanese from classical Chinese texts, and ‘return loans’, Chinese words which had fallen out of use and were ‘recovered’ by Japanese authors. For instance, 銀行 *yínháng* ‘bank’ (Jap. *ginkō*) is an original loan from Japanese, whereas 世界 *shìjiè* ‘world’ (Jap. *sekai*) was originally a classical Chinese word used in Buddhist texts to translate the Indian notion of *loka*, i.e. the ‘cosmos’, “understood as time (*shi* 世) and space (*jie* 界)” (Masini 1993:147).

Although the topic of the lexical acquisitions in the history of Chinese is clearly not a central one in our research, we shall employ historical data on hybrids, loan translations and graphic loans to illustrate some diachronic tendencies in the development of Chinese word formation. Here some general trends will be illustrated, whereas in the next section we shall focus on the interaction of Chinese and Japanese in lexical developments connected with the core issue of our study, namely the genesis of derivational formants.

As said above, the relative ‘impermeability’ to phonemic loans has made necessary to resort to other autochthonous material for the building of neologisms, also ‘revitalising’ word formation patterns which were already attested in the language (but not frequently used). We may thus hypothesize that the creation of loan translations, hybrids and autochthonous neologisms provided a stimulus for the usage of Chinese word formants and, also, favoured the ‘conventionalisation’ of morphemes of the kind of class nouns; in a constructionist perspective, such environment has the conditions for the grammaticalization of a lexeme (a class noun) into a derivational affix, as seen for -性 *-xìng* ‘the property of [X]’ (1.3.1.2; see also the analogous case of Ger. *-heit*). As to word formation patterns based on a class nouns, Bisang (2001) remarks that the modifier-modified order in the noun phrase has been the standard throughout all the history of Chinese, and, for instance, words containing a class noun as 家 *jiā* ‘expert of [X]<sub>N</sub>’ (3.1.2) are attested since the stage of Middle Chinese (as e.g. 詩家 *shījiā* ‘poet’, 8th century); such patterns were among the ‘resources’ which translators could use to build neologisms translating foreign notions; also, new items could be ‘attracted’ into the position of class nouns (Bisang 2001:200):

“the processes of Chinese word formation are basically language internal, i.e., they are not copied from another language, but their use and their diffusion within the Chinese lexicon is triggered by written contact with Standard Average European Languages. (...) Translators from prestigious European languages somehow had to imitate the textual structure of the original text and therefore also copied such European techniques into their own language. The extent to which word formation processes can be observed in Chinese and Vietnamese can be seen as an example of imitating the structure of the original text although the formal inventory of how word formation is realized in these languages is rather autochthonous”

Here are a few more examples of ‘Western’ terms which were adopted in the Chinese lexicon as loan translations or as autochthonous neologisms, having a class nouns as their head (exx. adapted from Bisang 2001:200):

(1)	機械論	政治家	社會主義
	<i>jīxièlùn</i>	<i>zhèngzhìjiā</i>	<i>shèhuìzhǔyì</i>
	machine- <i>lùn</i>	politics- <i>jiā</i>	society- <i>zhǔyì</i>
	‘mechanics’	‘politician’	‘socialism’

Items as 論 *lùn* ‘theory’, 家 *jiā* ‘expert of [X]<sub>N</sub>’ and 主義 *zhǔyì* ‘doctrine, -ism’, introduced before, which have either always been present as such in the Chinese lexicon or have been adopted by analogy with Japanese neologisms, as seems to be the case for -主義 *-zhǔyì* (Masini 1993:220), are thus a Chinese ‘imitation’ of foreign structures with autochthonous word formants. So, it seems that rather than copying the class nouns by themselves, the lexicon adopted word formation patterns, constructional idioms (compare Dong X.’s ‘word-formation rules’, 2.2.2) which, rather than reproducing the morphological structure of the foreign word, as in calques (Eng. school bus > It. *scuolabus*), reproduce its ‘semantic structure’; see the following examples, containing the class nouns -學 *-xué*:

(2)	動物學	電學	經濟學
	<i>dòngwùxué</i>	<i>diànxué</i>	<i>jīngjìxué</i>
	animal- <i>xué</i>	electricity- <i>xué</i>	economy- <i>xué</i>
	‘zoology’	‘electricity (science)’	‘economics’

The class noun -學 *-xué* is often regarded as equivalent to the neoclassic constituent Eng. *-logy* (and to the correspondent forms in the other SAE languages; see e.g. Wang F. 1998:72). However, just by looking at the examples in (2) it appears as evident that such Chinese item conveys the meaning of ‘branch of science’, independently from whether the corresponding ‘European’ word contains the constituent *-logy* (*-logie*, *-logia*, etc.; note that 經濟學 *jīngjìxué* ‘economics’ is an original graphic loan form Japanese), and it has no productive competitors in the Modern language. This pattern has been employed (either in Chinese or in Japanese) whenever the name for a branch of science needed to be created, even when there was no foreign model word, as for 電學 *diànxué*, an

autochthonous neologism (see Masini 1993, appendix II), and has thus become a ‘cover marker’ for a semantic category. The same may be said of other class nouns, as 論 *lùn* ‘theory’, which is found in terms as 多元論 *duōyuánlùn* ‘pluralism’ and 分子論 *fēnzǐlùn* ‘molechular theory’, which in English are built according to different models; if such terms were the mere transposition of foreign words, one would expect a formant as 主義 *zhǔyì* ‘-ism’ to appear in words as the above mentioned 多元論 *duōyuánlùn* ‘pluralism’ or in 二元論 *èryuánlùn* ‘dualism’ (compare 個人主義 *gèrénzhǔyì* ‘individualism’).

In some words, -學 *-xué* is semantically ‘redundant’, as e.g. 物理 *wùlǐxué* ‘physics’, as 物理 *wùlǐ* means ‘physics’ by itself (ex. from Ma Q. 1995), and such redundancy is possible only if the base has two syllables (/ morphemes); Sun Y. (2000) remarks that monosyllabic ‘affixes’ (i.e. the vast majority of ‘affixes’) may freely attach to polysyllabic words, creating prosodically ‘heavy’ words which should be dispreferred (see Feng 1998, 2001). In the next section, we shall elaborate on the relationship between prosody and word formation in the history of the Chinese language, and we shall assess the influence that contact by writing had on the development and diffusion of word formation pattern based on a class nouns and, more generally, on Chinese word formation.

### 3.2.1.1 Some Remarks on the Role of Japanese on the Development of Chinese Word Formation

As mentioned earlier (1.1.3), the dominant model in the Modern Chinese lexicon is the multimorphemic word: according to the figures in Xing J. (2006), about 80% of Mandarin words are made of more than one morpheme. As to the number of syllables, most complex words are made of two syllables: according to the figures in Shi Y. (2002:70), above 80% of Modern Chinese words are disyllabic. Given the fact that the vast majority of Chinese morphemes are monosyllabic, we may say that the ‘preferred’ model for Mandarin is the word made of two syllables, each representing one morpheme, as the examples below (see also exx. 1-3, CHAPTER 1):

(3)	公款	軍事	豐富
	<i>gōngkuǎn</i>	<i>jūnshì</i>	<i>fēngfù</i>
	public-money	military-affair	plentiful-abundant
	‘public money’	‘military affairs’	‘rich, abundant’

Words as 公款 *gōngkuǎn* and 軍事 *jūnshì* have the same structure of English compounds as *atomic bomb*, *ghost writer* or *blackboard*; they are of the modifier-modified type. The ‘motivation’ for a word as 豐富 *fēngfù*, however, is less apparent, since the meaning of the whole compound is not fundamentally different from the meaning of its parts if considered individually, i.e. its constituents are (near-)synonymous<sup>3</sup>; as we shall see, the building of many ‘anti-economic’ compounds as 豐富 *fēngfù* is a consequence of the spreading of the ‘disyllabic model’ in the Chinese lexicon (Feng S. 1998).

The Old Chinese lexicon, especially before the Han Dynasty (206 BCE – 220 CE), was prevalently monosyllabic; only about 20% of the words were made of two syllables (in the written language) before 200 BCE (Shi Y. 2002:72). From the point of view of syllable structure, the tendency throughout the recorded history of Chinese has been towards simplification; the evolution of the Chinese syllable is summarised in table 3.2:

Table 3.2. The evolution of syllable structure in the history of Chinese<sup>4</sup>

Stage	Minimum syllable size	Maximum syllable size	Coda consonants
Old Chinese (ca. 1000 BCE)	CVC	CCCMVCCC	At least ten different consonants
Middle Chinese (ca. 800 CE)	CV	{C, S} V {C, S}	[m], [n], [ŋ], [p], [t], [k]
Modern Chinese	V	{C, S} V C	[n], [ŋ]

<sup>3</sup> Wälchli (2005) terms such compounds as ‘synonymic co-compounds’.

<sup>4</sup> Here ‘C’ stands for ‘consonant’, ‘V’ for ‘vowel’, ‘M’ for ‘medial’ and ‘S’ for ‘semivowel’.

In Modern Mandarin Chinese, as said before (1.1.2), the minimal syllable is made of a single vowel (e.g. 餓 *è* ‘hungry’), with only two possible codas ([n] and [ŋ]) and only 21 possible initial consonants, never combined. Thus, from the point of view of phonology, the syllable has been progressively deprived of consonant clusters and of most final consonants, and the number of homophonous syllables has accordingly increased; this is often given as an explanation for the dysyllabification of the lexicon, since adding a syllable reduces the risk for homophony (see e.g. Packard 2000, Lin H. 2001; see also Lüdtke 1985<sup>5</sup>).

In prosodic terms, the simplification of syllable structure which occurred in the history of Chinese has resulted in a loss of ‘weight’ (Feng S. 1998:225; see also Feng S. 1996, 1997). In the framework of ‘Prosodic Morphology’ (McCarthy & Prince 1993, 1998) the smallest independent unit of prosody is the ‘prosodic word’, realised by the ‘foot’; it is given as a rule that the foot must be binary (in terms of ramification), either under syllabic or moraic analysis<sup>6</sup>. Let us compare the minimal syllable in Old and Middle Chinese, analysed according to the parameters of Prosodic Morphology described above (adapted from Feng S. 1998:288; ‘σ’ stands for ‘syllable’, ‘μ’ for ‘mora’):

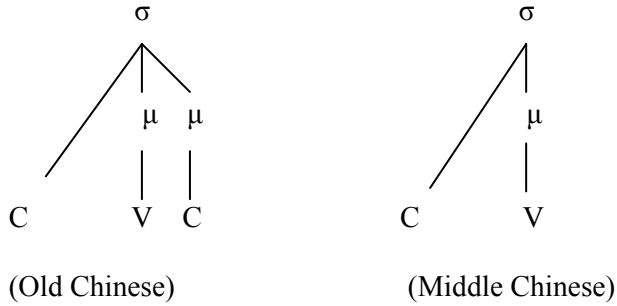
---

<sup>5</sup> According to Lüdtke “given a statistically relevant number of speech events (...), the stochastic process of both slur and prolixity will inevitably result in phonetological shrinking and semantactic accretion” (1985:356); when a sign becomes “intolerably short”, it is either substituted by a longer one (e.g. Latin *multum* ‘much / many’ > Old French *mout* > [mu], later substituted by *beaucoup*; Lüdtke 1985:361), or new material is added (‘accretion’). Such processes are recursive, and when the added material merges with the original sign, shrinking may occur and, consequently, again accretion.

<sup>6</sup> We shall not discuss here the arguments for the ‘binarity requirement’; the reader is referred to McCarthy and Prince (1993, 1998) and to Feng (1998:227 ff and 2001:165, fn 1).



Figure 3.1 The minimal syllable in Old Chinese (ca. 1000 BCE) and Middle Chinese (ca. 800 CE)



Since a foot and, consequently, a minimal prosodic word must be able to rest upon a binary structure, the monomoraic minimal syllable of Middle Chinese (around 800 CE) does not qualify as a prosodic unit; hence, a minimal prosodic word should be built on two syllables. In Modern Chinese, the standard foot is made of two syllables; a ‘degenerate foot’ (退化音步 *tùihuà yīnbù*) is made just of one, whereas three syllables may constitute a ‘superfoot’ (超音步 *chāoyīnbù*; Feng S. 2001), but this is subject to conditions. A degenerate foot may not constitute a prosodic word, whereas a superfoot may constitute a (super-)prosodic word, but not as freely as a standard foot, as we shall see in greater detail later (Feng S. 1997).

The consequences of such simplification was that many monosyllabic words, i.e. free lexical items, had to form a prosodic unit with another morpheme to constitute a well-formed prosodic word, a minimal unit of prosody. Such “two-word prosodic combinations”, if frequent, might become idiomatised and appear in a fixed order; this often led to lexicalisation of the combination into a compound (Feng S. 1998; see also Dong 2002:37-40). The birth of a large number of complex words, leading to the definition of Modern Mandarin as a ‘language of compounds’ (see 1.1.3), thus, might be partly explained with the interaction between prosody and word formation. Also, the building of (apparently) anti-economical compounds such as those made of (near-)synonymous constituents (see ex. 3 above) may be easily explained: when the

combination of morphemes into disyllabic units became necessary for prosodic reasons, in the early stages it was ‘easier’ to combine (near-)synonyms, since no significant change in meaning was involved, or words which somehow designate a set, as 衣 *yī* ‘shirt + 裳 *shāng* ‘skirt’ = ‘clothes’ (Feng S. 1998:223). The last example, incidentally, is a good example of idiomatization of a ‘two-word prosodic combination’: at the beginning the constituents could appear in both orders (衣裳 or 裳衣); later on, the order 衣裳 *yīshang* became fixed, and the word was lexicalised as ‘clothes’ (Feng S. 1998).

Having briefly introduced some facts on the interaction between phonology, prosody and word formation in the history of Chinese, let us now have a closer look at how word formation patterns based on class nouns are created and diffused; specifically, we shall analyse data on the morpheme -學 *-xué* ‘branch of learning’, focussing on the role which contact with Japanese possibly had in the development of such pattern.

Some connection between language contact and the development of certain tendencies in Mandarin word formation has been already suggested in the literature, e.g. by Wang L. (1989) and Masini (1993). The former proposed that “disyllabification in Chinese was mainly caused by two factors. The first is phonological simplification; the second is the acquisition of foreign words” (1989:165, my translation). Masini (1993:122) states that

“I believe that loans from western languages further encouraged this move toward polysyllabism. The autochthonous neologisms that originated in the XIXth century were all polysyllables, except for the terms used for chemical elements (...), characters of Japanese origin (...), phonemic loans used to indicate unit of weight and measure (...) and some characters of dialectal origin.”

Such ‘move’ toward polysyllabism is not only quantitative: quoting data from Wang L. (1985), Masini (1993:123) remarks that until the XIXth century between 70% and 80% of disyllabic ‘words’ were built according to the ‘associative structure’, i.e. they were made of coordinated items,

often (near-)synonymous (see above)<sup>7</sup>; from the XIXth century on, “disyllabic neologisms of foreign origin generally assumed a determining-determined structure”.

What about the distinction between compounds and derivatives? Masini lists around 300 disyllabic neologisms from the period between 1840 and 1898 (Masini 1993, appendix II); among those 60 are reported to contain a prefix or a suffix, according to his definition. Masini defines derivational affixes on the basis of their ‘paradigmaticity’ (1993:123-124):

“What changes is the semantic relationship of the word with the whole class of words sharing either the same suffix or prefix (...). In Chinese, affixes are word-morphemes that were originally free, but subsequently degraded into a sort of semantic mark repeated in a class of words. If placed in front of the base, they are prefixes; if placed after, they are suffixes.”

The fact that a morpheme which was once free becomes bound is a condition for affixhood, but surely not a sufficient one since, as remarked earlier, many lexical morphemes of Mandarin are actually bound. As to the relationship among words sharing a common constituent, this is akin to the notion of paradigmatic relations in Construction Morphology (1.1.2); in certain cases, ‘paradigmatic’ word formation is virtually identical to analogy, as in the Dutch example below (from Booij 2007:37):

(4) <i>boter-briefje</i>	→	<i>margarine-briefje</i>
butter-letter		margarine-letter
‘marriage contract’		‘cohabitation contract’

In *margarine-briefje*, an idiosyncratic interpretation of *margarine* is possible because of the existence of the model word *boter-briefje*. However, as remarked in 1.1.2, paradigmatic word formation cannot be equated with simple analogy, since in many cases there is no model word, as for Du. *-baron* ‘baron’ > ‘rich dealer in [x]<sub>N</sub>’ (*avfal-baron* ‘rich dealer in trash’; Booij 2007:37).

It is worth remarking that Masini (1993:124) does make a distinction

---

<sup>7</sup> On the development of coordinative compounding in the history of Chinese, compare the data in Cheng X. (1992c, 1992b, 1992a e 1992d).

between ‘proper’ affixes and affixoids (“affix-like formatives”). The former “belonged essentially to two closed, highly restricted classes”; they are the ‘traditional’ prefixes 老- *lǎo-* ‘old’ and 小- *xiǎo-* (usually before a surname; see CHAPTER 1, fn. 15), etc., and ‘dummy’ suffixes as 子- *-zi*, 頭- *-tōu*, etc. Affix-like formatives, which were “added” during the history of Chinese to the ‘traditional’ affixes, are not yet fully “de-lexicalized”, and they form an open class. However, Masini suggests that such a distinction is relevant only in a rigorous synchronic perspective, and he carries out a diachronic analysis of the phenomenon of ‘affix-like’ word formation; we cannot but agree with his position (see 1.3.2.2). Masini remarks that in the tendency towards the ‘reduction’ of lexical items into ‘affix-like formatives’ was at work in Chinese even before the XIXth century, albeit not very developed. For instance, a ‘prefixoid’ as 洋- *yang-* ‘foreign’ was found in words as 洋煙 *yángyān* ‘opium’ (lit. ‘foreign smoke’) and 洋布 *yángbù* ‘calico’; in such case, “*yang* can no longer be considered simply an adjective, since the overall meaning of the word to which it gave origin was not the sum of the various meanings of the elements composing it” (Masini 1993:125). This, however, does not seem to be a valid point, since the conventionalisation of an idiosyncratic meaning (e.g. ‘foreign cloth’ = ‘calico’) is a feature of lexicalization, rather than of grammaticalization. This (infelicitous) example aside, Masini’s suffix-like formatives correspond mostly to class nouns. He suggests that the development and diffusion of word-formation patterns based on class nouns was already visible in the period considered in his research, i.e. the XIXth century; such tendency would become much stronger since the XXth century, due to the influence of Japanese in the introduction of many suffix-like formatives and, what’s more, the influence of Japanese was not only limited to the ‘exporting’ and diffusion of word formation patterns, but, also, to the size of such words, stimulating the creation of trisyllabic words containing an affix-like formative (Masini 1993:126). Let us now analyse some of the data put forth by Masini on trisyllabic neologisms, focussing on 學- *-xué* ‘branch of learning’.

In Masini’s sample, 學- *-xué* ‘branch of learning’ and 機- *-jī* ‘machine’

(通風機 *tōngfēngjī* ‘ventilator’) are the most ‘productive’ suffix-like formatives (i.e. those which are contained in more neologisms); they were both already part of the ‘inventory’ of Chinese class nouns before the XIXth century. The meaning ‘branch of learning’ for the lexeme 學 *xué* has been available at least since the Middle Chinese period (GHYDCD 2000); earlier complex words as 史學 *shǐxué* ‘the science of history’ (attested in the 晉書 *Jìn Shū* ‘Book of Jin’, VIIIth century) and 算學 *suànxué* ‘arithmetic’ (新唐書 *Xīn Táng Shū* ‘New Book of Tang’, XIXth century) were generally composed of two syllables / morphemes. However, “[t]he development of affix-like formatives is closely linked to their association with disyllabic bases. From a morphological point of view, this will be a major source of neologisms during the XXth century” (Masini 1993:125). Thus, the enlargement of the size of complex words containing a class noun was not only a formal change but, according to Masini, it also provided a ‘frame’ for word formation patterns which favoured the development of suffix-like formatives and, thus, had an influence on Mandarin word formation. While it is true that it has always been possible to juxtapose two lexical morphemes in a modifier-modified relation, if a suffix-like formative as -學 *-xué* was attached to a disyllabic word, the grammatical morpheme 之 *zhī* (compare Modern Mandarin 的 *de*) was often added to overtly mark the relationship between the constituents. For instance, ‘geometry’ as a subject of study was commonly indicated as 幾何之學 *jǐhé zhī xué* in texts before the XIXth century, whereas in the contemporary language the marker has been dropped (幾何學 *jǐhéxué*). In the older form 學 *xué* cannot be regarded as an affix, given the presence of the syntactic overt marker 之 *zhī*, but in the newer word 學 *xué* is located in a syntagmatic environment which is ‘compatible’ with affixal status, a construction (Bisang’s maximum pattern) in which 學 *xué* is in the slot (attractor position) for class nouns (see above, 3.2.1).

In Masini’s sample, the first trisyllabic -學 *-xué* complex word (i.e. without the marker 之 *zhī*) is 植物學 *zhìwùxué* ‘botany’ (1859); throughout the XIXth century, 學 *xué* is used both as a free word and as a

bound suffixed constituent. For instance, 天文之學 *tiānwén zhī xué* and 天文學 *tiānwénxué* for ‘astronomy’ are both attested, and the same goes for 動物之學 *dòngwù zhī xué* and 動物學 *dòngwùxué* ‘zoology’ (Masini 1993:126); Masini, as mentioned before, proposed that Japanese acted as a model for the creation of such trisyllabic neologisms. It must be remarked again that the pattern  $[[X]_N [學]_N]_N$  ‘the study of  $[X]_N$ ’ already existed in Chinese, although the non-head, mostly, had to be monosyllabic, and Japanese also imported Chinese disyllabic words as 化學 *huàxué* ‘chemistry’ (giapp. *kagaku*) in the XIXth century; moreover, as Masini points out, “at the end of the XIXth century, graphic loans from Japanese included no compound words composed of suffix-like formatives not already used in Chinese” (1993:140). Anyway, the trisyllabic 學 *-xué* complex word which were coined in Japan and adopted in Chinese during this period are at least fourteen, many more than the disyllabic 學 *-xué* words which were imported in Japan (Masini 1993:149-150); although trisyllabic forms met with some resistance, as shown above with 天文之學 *tiānwén zhī xué* ‘astronomy’ and 動物之學 *dòngwù zhī xué* ‘zoology’, “they later paved the way for the acceptance of other trisyllabic compounds, first those of Japanese origin and then the autochthonous inventions created by analogy with the imported version” (Masini 1993:149-150).

To sum up, the trend in the XIXth century seems to have been towards the acceptance and diffusion of more asyndetic modifier-modified complex words, also with a disyllabic modifier; the modified constituent was always a monosyllabic class noun, as 學 *-xué* ‘branch of learning’ and 機 *-jī* ‘machine’. It has been hypothesized that contact by writing with Japanese played a role in such developments, providing the model for the coinage of trisyllabic complex words of the kind described above; the Japanese words imported in Chinese during the XIXth (and early XXth) century were mostly neologisms created (or revitalised) to translate words from European languages<sup>8</sup> and, thus, Standard Average European

<sup>8</sup> “Following the formation of the Meiji government [明治 *Meiji*, 1868-1912], Japan embarked on a remarkable endeavor to adopt everything possible from Western civilization.

languages had an indirect influence on the ‘modernization’ of the Chinese lexicon. During the XXth century, as contact with Japanese became closer, the number of trisyllabic words containing a class noun grew accordingly, and the Japanese language, according to Masini, “further stimulated the introduction of many other affix-like formatives (especially suffix-like) which significantly enriched the Chinese lexicon” (1993:126), as 性 *-xìng* ‘the nature of [X]’, the diffusion of which has apparently been influenced by Japanese neologisms as 可能性 *kanōsei* and 重要性 *jūyōsei*, the ‘translation’ of Eng. ‘possibility’ and ‘importance’, accepted in Chinese (*kěnéngxìng* and *zhòngyàoxìng*; Wang L. 1980:230; see below, 3.2.3)<sup>9</sup>. Moreover, the acceptance of trisyllabic modifier-modified words contributed to the ‘revitalization’ of old Chinese formatives as the already mentioned 家 *-jiā* ‘expert of [X]<sub>N</sub>’ or 者 *-zhě* ‘agentive suffix’ (Masini 1993:150, fn. 48; see Dong X. 2002:301-303 for diachronic examples of complex words in which 家 *-jiā* conveys its ‘affixal’ meaning). Given the situation outlined here, we may also hypothesize that the development of hybrids, i.e. the combination of a ‘phonemic loan’ with an autochthonous element, often a class noun (see above, 3.2.1) is connected with the general trends in the ‘restructuring’ of Mandarin word formation, namely, with the diffusion of ‘suffix-like formatives’; as Masini points out, a marked growth of ‘hybridization’, just as that of ‘suffixation’, will occur only “once closer contacts were established with the Japanese language”, i.e. at the beginning of the XXth century (1993:140).

As to the diffusion of the ‘disyllabic modifier + 學 *xué*’ pattern, we shall propose here a comparison between quantitative data drawn from Masini’s sample, which may be regarded as indicative of the situation in the XIXth century (Masini 1993, appendix II), and from a reverse lexicon of Modern Mandarin (NXCD 2005). In the first line of table 3.3 we provide the number of 學 *-xué* complex words having a ‘base’ of, respectively, one, two or

---

(...) In every field, with every borrowing, came new words and new terminology” (Takashima 1998:4).

<sup>9</sup> The fact that Japanese acted as a stimulus for the diffusion of 性 *-xìng* in its ‘affixal’ function must not obscure that the process of grammaticalization of the lexeme 性 *xìng* seems to have been already under way before, as shown before (1.3.1.2 and 1.3.2.1).

more than two syllables in Masini's sample and their 'relative' weight, i.e. the percentage of the total<sup>10</sup>; in the second line, the same figures from NXCD (2005) will be provided.

Table 3.3. Complex nouns with 學 *xué* as the head in Masini (1993) and in NXCD (2005)

Work		Number of syllables of the modifier			Total
		1	2	>2	
Masini (1993)	Words	20	18	0	38
	%	52.6	47.4	0	100
NXCD (2005)	Words	20	73	41	134
	%	14.9	54.5	30.6	100

This is, admittedly, an improper comparison, because of the different nature of the two sets of data taken into consideration (a sample of texts in the first case, a lexicographic work in the second case); nevertheless, we believe that such figures are 'usable' as an indicator of a certain tendency. Let us comment briefly on the data in table 3.3.

As to the data in Masini, we may notice that there is a slight prevalence of complex words with a monosyllabic modifier, and no word bigger than three syllables. However, the count is based on Masini's own labelling of 'suffix-like' usage for -學 *-xué* and we thus considered also words as 科學 *kēxué* 'science', 文學 *wénxué* 'literature', 藝學 *yìxué* 'technology' and 述學 *shùxué*, either 'art' or 'technology', in which it is highly doubtful that -學 *-xué* actually conveys the meaning 'branch of science'; moreover, the latter two terms for 'technology' seem to have fallen out of use (they are not included in a broad dictionary as CCD 2002). We may remark that as many as 13 (65%) disyllabic words are autochthonous neologisms, two of them are 'semantic loans' and only four are original loans from Japanese (法學 *fǎxué* 'science of law' is a return loan). The proportions are reversed for neologisms with a disyllabic modifier: 11

<sup>10</sup> Figures rounded to the first decimal place.



words, i.e. almost two thirds of this group, are original graphic loans from Japanese; of the remaining seven words, six are classified as autochthonous neologisms, but one of those, 地質學 *dìzhìxué* ‘geology’, might also be a Japanese loanword according to some<sup>11</sup>, and terms as 動重學 *dòngzhòngxué* ‘dynamic mechanics’ and 靜重學 *jìngzhòngxué* ‘static mechanics’ have been later replaced by 動力學 *dònglìxué* e 靜力學 *jìnglìxué*; 幾何學 *jǐhéxué* ‘geometry’, as seen above, is just the expansion of 幾何 *jǐhé*, a word which had existed in Chinese for at least two hundred years<sup>12</sup>. It appears, thus, that Japanese was actually a stimulus for the creation of trisyllabic words.

In the sample drawn from NXCD (2005), which should give us a picture (albeit a somewhat simplified one) of the contemporary language, it clearly appears that the ‘disyllabic modifier + 學 *xué*’ pattern is the dominant one, accounting for 54.5% of the total number of 學 *xué* complex words. The prevalence of such model is even more evident if we take into consideration the fact that the vast majority of words in the NXCD sample with a modifier larger than two syllables (37 out of 41) may be regarded as recursively formed on the base of an existing disyllabic word, as e.g. 地理學 *dìlǐxué* ‘geography’ → 經濟地理學 *jīngjìdìlǐxué* ‘economic geography’; also, all the words with a trisyllabic base (except 形而上學 *xíng’érshàngxué* ‘metaphysics’) seem to be built by adding a prefix-like element to an existing disyllabic modifier, as e.g. 微生物學 *wēishēngwùxué* ‘microbiology’ (lit. ‘micro-organism-*xué*’). Moreover, the words with a monosyllabic modifier in the dictionary sample were mostly coined not later than the XIXth century. In short, the contemporary data from NXCD suggests that the ‘monosyllabic modifier + 學 *xué*’ pattern

<sup>11</sup> Masini (1993:84): «Scholars of Modern Chinese lexicon have tended to put all the emphasis on this second phase, the introduction of Japanese neologisms into Chinese, being convinced that these terms were original loans from Japanese. Instead, a good many of these terms had actually reached Japan from China and then returned to China several decades later.»

<sup>12</sup> Back in 1607, Matteo Ricci and Xū Guāngqǐ (徐光啓) translated the title of Euclid’s “Elements of Geometry” as 幾何原本 *jǐhé yuánběn* (Masini 1993, appendix II).

virtually ceased to be productive after the XIXth century.

Our data, thus, provide further support to Masini's analysis of the evolution of Chinese word formation and, especially, of the role of Japanese in the diffusion of the 'disyllabic modifier + class noun' word formation pattern. In the next section, we shall elaborate on the prosodic correlates in such developments of the Chinese lexicon.

### 3.2.1.2 Concluding Remarks on Class Nouns

In the preceding section, we presented and discussed some diachronic data on -學 *-xué* 'branch of learning', chosen as a representative example of class nouns; we highlighted the fact that contact by writing with Japanese, a language which during the XIXth and early XXth century was significantly enriched with 'technical' vocabulary translating Western notions, has led to the diffusion in Chinese of patterns of word formation built around a suffixed class noun with a disyllabic (or larger) modifier. With the acceptance and diffusion of such word templates, Mandarin has apparently become more 'receptive' of new class nouns which were grammaticalised / conventionalised in Japanese and, also, has 'revitalised' attested suffix-like formatives, as the above mentioned -家 *-jiā* 'expert of [X]<sub>N</sub>'. The influence of Standard Average European, thus, was mostly an indirect one, at least in the XIXth century (but we shall go back to this point in 3.2.3).

From the prosodic point of view, the affirmation of word formation schemas by which trisyllabic complex words are built is in line with the evolution of the 'prosodic word' in Chinese, which we briefly outlined in the preceding section. We said that three syllables constitute a 'super-foot' in Mandarin, and such a combination may realise a prosodic word (i.e. a minimal independent unit of prosody) only under certain conditions; specifically,  $[2_N + 1_N]_N$  structures ('2' stands for 'disyllabic', '1' for 'monosyllabic') are allowed, whereas  $[1_N + 2_N]_N$  structures are not, generally speaking, since the 'natural foot' in word formation is 'dextrorse', i.e. it builds (disyllabic) feet from left to right (Feng 1997, 2001; exx. from Duanmu S. 2000, '/' stands for 'foot boundary').

- (5) 动物园 → /动物/园  
*dòngwùyuán* /dòngwù/ *yuán*  
 animal-garden ‘zoo’
- (6) 野生动物 → /野生/动物/  
*yěshēng dòngwù* /yěshēng/dòngwù/  
 wild animal ‘wild animal’
- (7) 北京动物园 → /北京/动物园/  
*Běijīng dòngwùyuán* /Běijīng/dòngwù/ *yuán*  
 Peking zoo ‘Peking zoo’<sup>13</sup>
- (8) \*鞋工厂  
*xiégōngchǎng*  
 shoe-factory ‘shoe factory’  
 (ex. from Feng S. 2001:172)

As shown in (5),  $[2_N + 1_N]_N$  complex words do not violate the rules of foot building in Mandarin word formation; a word form as (8), however, is unacceptable, since the first constituent is a monosyllabic noun, 鞋 *xié* ‘shoe’, followed by a disyllabic noun, having thus a  $[1_N + 2_N]_N$  structure. Needless to say, such models allows for prosodic rules to ‘look into’ the morphological structure of a words; otherwise, a word as (8) could be grouped as  $[xié[gōng/chǎng]]$  and be prosodically well-formed. The ‘dextrorse rule’ of foot building is best understood as ‘cyclic’, meaning that the grouping of syllables into prosodic units is done step by step, starting from the smallest grammatical unit (Duanmu S. 2000; Chomsky, Halle & Lukoff 1956); see the examples below (Feng S. 2001:172):

<sup>13</sup> Five-syllable words do not always respect the ‘dextrorse rule’ in foot building; for a discussion of counterexamples, see Duanmu S. (2000).

- (9) a. 金店  
*jīndiàn*  
 gold/metal-shop  
 ‘jeweller’s shop’
- b. 五金店  
*wǔjīndiàn*  
 five-gold/metal-shop  
 ‘hardware shop’
- c. \*金商店  
*jīn shāngdiàn*  
 gold/metal-shop  
 ‘jeweller’s shop’

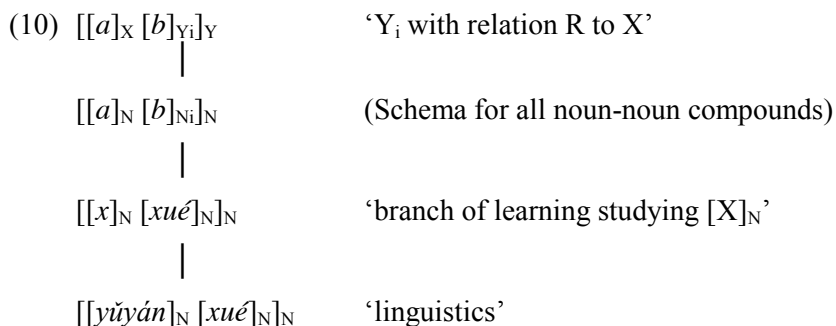
The word in (9a) is disyllabic, and thus corresponds to a standard foot; the word in (9b) has a structure analogous to that in (5), and thus satisfies the conditions for the building of a super-foot. The fact that (9c) is unacceptable, just as (8), is explained with the inconsistency between prosodic and morphological units (*\*[jīn[shāng/diàn]]*)<sup>14</sup>.

The word formation schemas based on class nouns which became common in the XIXth and, especially, XXth century are built according to the  $[2_N + 1_N]_N$  structure or, better,  $[2_X + 1_N]_N$ , since words with a verbal modifier as 打火機  $[[dǎhuǒ_V] + jī_N]$  ‘lighter’ are also attested. According to the works by Feng S. on the interaction of prosody and morphosyntax throughout the history of the Chinese language (Feng S. 1996, 1997, 1998 and 2001, among others), the most significant changes in the structure of the prosodic word occurred much earlier than the period considered here, as mentioned in the preceding section. Taking into consideration Masini’s data on the ‘resistance’ of the Chinese lexicon to

<sup>14</sup> However, words as 金項鏈 *jīnxiàngliàn* ‘golden necklace’ are actually acceptable, despite having a  $[1_N + 2_N]_N$  structure. According to Feng S. (2001:172), this is because the first morpheme, 金 *jīn*, is used as an adjective and, thus, a ‘syntactic word’ (句法詞 *jùfǎcí*) is built; in syntax, foot building goes from right to left (*[jīn/[xiàngliàn]/]*); we shall go back to this point below (3.2.5).

the acceptance of trisyllabic complex words (see the preceding section), we believe that we may hypothesize that the influence of Japanese has somehow ‘forced’ the diffusion of word formation models which, although prosodically acceptable, were not as ‘preferred’ as disyllabic words. Feng S. (2001; see also Feng S. 1998), however, does not consider data from the ‘transitional period’ between 1840 and 1919, i.e. between Early Mandarin and Modern Mandarin Chinese (see 1.1.1).

As to the nature of class nouns, we have already made clear that they cannot be equated to derivational affixes, in our opinion, especially since their definition is quite vague; however, we do believe that word formation schemas based on class noun provide a good syntagmatic environment for the grammaticalization of a lexeme into a derivational formant (cf. the notion of ‘attractor position’ from Bisang 1996, discussed in 1.3.2 and 3.2.1). A class noun in itself is not fundamentally different from any other constituent of compounding, as part of a schema in which it conveys a certain meaning; so, an individual word as 語言學 *yǔyánxué* ‘linguistics’, containing the class noun -學 *-xué* ‘branch of learning’ is the instantiation of a word formation schema which is, in turn, connected to more general schemas (compare the schema in 18, CHAPTER 1):



If we want to determine whether -學 *-xué* is to be regarded just as a class noun or as a grammaticalised item, a derivational suffix, we are faced with a tough challenge, since the usage of the lexeme 學 *xué* as a marker for subjects of study has started very early, with attestations dating

from as early as the VIIIth century (at the latest) and is thus usually regarded as one of the ‘historical’ meanings of the lexeme. The fact that at some point it was no longer possible to use 學 *xué* with the meaning ‘branch of learning’ as a free form, but only as a bound constituent located to the right of a complex word is a necessary (but not sufficient) condition for affixhood (see Lehmann’s parameters of grammaticalization, **1.3.1.1**); in this respect, we may compare 學 *xué* to Du. *-boer* ‘seller of [X]<sub>N</sub>’, which cannot be used with such meaning as a word, but has a corresponding lexeme meaning ‘farmer’ (**1.2.2**), just as 學 *xué* may be used in Modern Chinese as a free verbal form, meaning ‘to study’. Also, both *-學 -xué* and *-boer* convey these particular meanings, namely ‘branch of learning studying [X]<sub>N</sub>’ and ‘seller of [X]<sub>N</sub>’, only when they are in their bound usage as right-hand constituents. As to semantics, Ma Q. (1995) suggests that *-學 -xué* may be regarded as an affix, since the meaning ‘branch of learning’ is not one of its ‘core’ meanings, i.e. it is not among the first two listed in a dictionary (see the discussion in **2.2.2**); actually, in a general reference dictionary as the 漢語大詞典 *Hànyǔ Dàcídiǎn* (HYDCD 1993), ‘branch of learning’ is listed as the seventh meaning (out of twelve) for 學 *xué*. This method is not always reliable, since the ordering of meanings in a dictionary usually follows a chronological order, but it is not always so (see fn. 9); in this case, the earliest attestations of 學 *xué* as a lexeme meaning ‘branch of learning’ apparently date to the VIIth century, and its usage as a bound word constituent with such meaning may be located around the VIIIth century, whereas, as a lexeme meaning ‘to study’, 學 *xué* is found virtually throughout all the recorded history of Chinese (e.g. in the 詩經 *Shījīng* ‘Book of Songs’; HYDCD 1993). So, it appears that the meaning ‘branch of learning’ was developed comparatively late in the history of 學 *xué*, and, thus, it may be interpreted as an extension of its primary meaning, ‘to study’; given the flexibility of Chinese in using verbs as nouns and vice versa, which was even stronger in Old Chinese (see exx. 41-43, **1.3.2**; see also Bisang 2008), we may well hypothesize that 學 *xué* was used as a noun, indicating the action of studying, and thus got conventionalised as ‘the study (of [X]<sub>N</sub>)’, which we nowadays understand as

‘branch of learning studying  $[X]_N$ ’. The shift in meaning, in such case, appears to us as metonymic in nature, and thus we believe that, from a strict semantic perspective, the class noun -學  $-xué$  would qualify as a lexical derivational affix. However, the  $[[x]_N [xué]_N]_N$  ‘branch of learning studying  $[X]_N$ ’ pattern, despite the early start, was ambiguous between syntax and morphology even in the XIXth century, as seen with the 天文之學  $tiānwén zhī xué$  vs. 天文學  $tiānwénxué$  ‘astronomy’ examples above (3.2.1), most likely for prosodic reasons. Hence, it might be claimed that the semantic evolution outlined above actually occurred *before* 學  $xué$  became a bound form, and thus such development is (at least, partly) independent from the constuctional idiom. In the discussion of -性  $-xìng$  ‘the quality of  $[X]_N$ ’ (1.3.1.2), we remarked that even in Early Modern Chinese texts 性  $xìng$  was still used, sometimes, as a free morpheme, as a word, although its grammaticalization was already under way; the difference lies in the fact that the ‘signs’ of its semantic and functional evolution are visible in complex words, rather than in its usage as a free form, without the presence of any marker of modification (as 之  $zhī$ ). In short, the interpretation of the evolution 學  $xué$  ‘to study, study’ > -學  $-xué$  ‘branch of learning studying  $[X]_N$ ’ as grammaticalization of a derivational affix might be controversial.

From the perspective of language contact, it has been suggested that there has been an ‘exchange’ of word formation schemas between China and Japan, in both directions. Schemas of the kind of  $[[a]_X [CN]_N]_N$  ‘member  $a$  of category CN’, where ‘CN’ stands for class noun, are attested in the Chinese language at least since the IVth century BCE, as e.g.  $[[x]_X [fū]_N]_N$  ‘person engaged in manual labour of the  $x$  kind’ (漁夫  $yǔfū$  ‘fisherman’; Cheng X. 1992c:101). From the data analysed here, it seems that many such schemas were imported in Japanese as:

(11)  $[[a]_X [CN]_N]_N$  ‘member  $a$  of category CN’ (CN is monosyllabic)

And, later, were ‘re-imported’ in China as

(12)  $[[a]_X [CN]_N]_N$  ‘member  $a$  of category CN’ (CN is mono- or disyllabic)

Needless to say, the ‘monosyllabic requirement’ in (11) is to be taken as a (strong) tendency, rather than as a rule. As seen in the preceding section, the simplification in the structure of the minimal syllable caused the ‘enlargement’ of the minimal independent unit of prosody (the prosodic word), and such disyllabic ‘frame’ has led to the lexicalization of many phrases of two syllables. It may be suggested that the trisyllabic ‘super-foot’ of Modern Chinese, has become the prosodic frame which favoured the creation of a large number of class nouns, items which may eventually grammaticalise into lexical derivational affixes. Such hypothesis appears to be plausible, from a quantitative point of view: in a language as Chinese, an increase in the number of syllables of the modifier of a schema is virtually equivalent to an increase in the number of *morphemes*; by getting rid of the ‘monosyllabic requirement’ for the modifier, it became possible to build words of greater complexity (i.e. with a more complex semantic structure). Hence, the birth of formatives with derivational features might be both an indirect consequence of the greater freedom in building of complex words, and a direct consequence of the ‘creative’ use of morphemes of Chinese origin in Japan (especially in the XXth century, as e.g. -化 *-huà* ‘-ise, -ify’; see below, 3.2.3) and in China; in the literature (Guo L. 1983; Shen M. 1986, 1995), it has been already pointed out that ‘new’ affixes (however defined; see 2.2.2) ‘prefer’ disyllabic modifiers, and this may be interpreted as part of the tendencies in word formation outlined here.

In the following section, we shall deal with the issue of ‘newly coined’ formatives with derivation-like features, and we shall see how productivity and analogy play an essential role in the ‘conventionalization’ of such items.

### 3.2.2 Some ‘Newly-coined affixes’: on Analogy and Productivity

In the recent literature, many authors have pointed out that Mandarin word formation has shown a ‘new tendency towards affixation’ (新的詞綴化傾向 *xīnde cízhùihuà qīngxiàng*, Shen M. 1986; see also Wu Y. 2000, Fan L. 2002 Dong Z. 2003). The suffixed items which are usually quoted as examples of such tendency, mostly, may be analysed as class nouns, since they indicate (general) semantic categories and form ‘paradigms’, as



-吧 *-bā* ‘bar’, -秀 *-xiù* ‘show’ and -迷 *-mí* ‘fan’:

(13) 網吧	真人秀	棋迷
<i>wǎngbā</i>	<i>zhēnrénxiù</i>	<i>qímí</i>
net- <i>bā</i>	real-person- <i>xiù</i>	chess+ <i>mí</i>
‘internet café’	‘reality show’	‘chess enthusiast’

Prefixed items in this (broadly defined!) group are mostly monosyllabic adjectival morphemes which correspond to lexical morphemes but have a different distribution, as seen before for 多- *duō-* ‘multi-, poly-’ (2.2.2); moreover, they underwent some ‘generalization’ / ‘broadening’ (泛化 *fànhuà*) in meaning, as e.g. 高 *gāo* ‘tall, high’, which indicates ‘high degree of’ in complex words as 高蛋白 *gāodànbái* ‘high protein’ (Shen M. 1995:36). We postpone the discussion of prefixed items to 3.2.6; here we shall rather focus on the suffix-like elements introduced above and, specifically, on -吧 *-bā* ‘bar’.

The bound formative 吧 *bā* ‘bar’, quite productive in recent years, is a loan from English which, however, entered Mandarin in a rather peculiar way. The loan 吧 *bā* ‘bar’ is first found in the hybrid word 酒吧 *jiǔbā*, lit. ‘alcohol-bar’ (but compare fn. 18, below), defined as ‘bar; counter at which alcoholic beverages are served in a Western-style restaurant or hotel’ (CCD 2002, my translation); after the acceptance of such hybrid, a number of words have been created by analogy, as 水吧 *shuǐbā* ‘water bar’ (a place where mostly soft drinks are served) or 咖啡吧 *kāfēibā* ‘coffee shop’. In the creation of new forms by analogy, the original meaning of 吧 *bā* gets ‘blurred’, as in the following examples<sup>15</sup>:

(14) 陶吧	布吧	串吧
<i>táobā</i>	<i>bùbā</i>	<i>chuàn bā</i>
pottery- <i>bā</i>	cloth- <i>bā</i>	skewer- <i>bā</i>
‘pottery bar’	‘cloth bar’	‘skewer bar’

<sup>15</sup> The first two examples and the related comments are quoted from Fan L. (2002:136).

The English translations of the Mandarin examples in (14) obviously deserve further explanation. A 陶吧 *táobā* ‘pottery bar’ is a pottery workshop where customers may create their own products, and a 布吧 *bùbā* ‘cloth bar’ is textile workshop in which customers can take part in the manufacturing process; a 串吧 *chuànbā* ‘skewer bar’ is a restaurant specializing in skewers.

In such words, the meaning of 吧 *bā* is somehow generalised, and it “broadly indicates an entertainment place with a particular function or supplied with some special equipment”, as per the definition of the Xinhua dictionary of neologisms (XH 2003, my translation; see also Dong Z. 2003:8). In the example below<sup>16</sup>, 吧 *bā* is used as a sort of ‘hyperonym’ for ‘entertainment place’, actually encompassing places which do not have much in common with the ‘original’ 酒吧 *jiǔbā* ‘(alcohol) bar’:

- (15) 影吧里留个影, 布吧里染快布,  
*yǐngbā li liú-ge yǐng, bùbā li rǎn kuài bù*  
 picture- *bā* LOC leave CLF picture cloth-*bā* LOC dye CLF cloth  
 玻璃吧做花瓶: 上海人“吧”里玩个够  
*bōlibā zuò huāpíng Shànghǎirén bā li wán-ge gòu*  
 glass-*bā* make flower-vase Shanghai-person *bā* LOC have.fun CLF enough

‘Take a picture at a picture bar, dye a piece of cloth at the cloth bar, make a flower vase at the glass bar: Shanghai people will have enough fun at “bars”’

Note that when 吧 *bā* is used as a hyperonym in the passage, it is marked with double quotation marks, to indicate that such usage is ‘unorthodox’. This reminds us of Eng. ‘ism’ when used as an hyperonym for any belief or ideology the name of which ends in -ism; in this case, however, the ‘lexicalised’ version of the suffix has been accepted into the lexicon and, thus, it may no longer be regarded as anomalous (see Ramat

<sup>16</sup> Example from the newspaper 北京晚报 *Běijīng Wǎnbào*, 13/3/1999 (qtd. in XHXYD 2003, my translation).

1992).

Another semi-bound usage of 吧 *bā* as a hyperonym is the verb 泡吧 *pàobā* ‘waste time, wander in “bars”’; such verb is a ‘separable word’ (離合詞 *lǐhécí*), i.e. a verb-complement structure the constituents of which may be separated by some items, as quantifiers or aspect markers, as in the example below<sup>17</sup>:

- (16) 每星期泡一次吧4,000/月  
*měi xīngqī pào yí cì bā sìqiān yuè*  
 every week dawdle one CLF *bā* 4.000 month  
 ‘If you go to bars once a week, that’ll make 4.000 [JPY] a month’

Such usage of 吧 *bā* could be termed as ‘conditionally independent’ (see the discussion of Yin F. 1984, **2.1.1**), since it may appear in a ‘numeral-classifier-noun’ construction. Going back to semantics, it is important to stress that 吧 *bā* in 泡吧 *pàobā* may refer to various kind of ‘bars’, as e.g. an internet café (網吧 *wǎngbā*, ex. 14; XH 2003).

As to the meaning shift that 吧 *bā* has undergone, it appears that the definition of the Xinhua dictionary quoted above does not cover all the instances of 吧 *bā* complex words, as meaning abstraction has gone further for such formative. In table 3.4 (adapted from Arcodia 2011), we shall present a small convenience sample of 30 words containing 吧 *bā* as the right-hand constituent, taken from the literature (Wu 2000, Fan 2002, Dong 2003) and from websites. For each word, a gloss, an explanation of its meaning and the source are provided; since most of these words are not listed in dictionaries, sometimes the definition has been understood by the author from the context in which the word was found.

---

<sup>17</sup> Example extracted from a Chinese forum; the thread is about the cost of life in Tokyo (<http://bbs.metroer.com/t-113336-1-1.html?%26amp%3Bascdesc%3DDESC>).

Table 3.4. Sample of words containing 吧 *bā* as the right-hand constituent (from Arcodia 2011).

Word	Gloss	Meaning <sup>18</sup>	Source
玻璃吧 <i>bōlibā</i>	glass-bar	A glass workshop where customers may create their own products.	Fan (2002:136); Beijing Wanbao, 13/3/1999 (quoted in XH 2003)
布吧 <i>bùbā</i>	cloth-bar	A textile workshop where customers may create their own products	Fan (2002:136)
茶吧 <i>chábā</i>	tea-bar	A Western style tea house	www.nciku.com
串吧 <i>chuànbā</i>	skewer-bar	A restaurant specializing in skewers	Own observation
創意吧 <i>chuàngyìbā</i>	creative-bar	A company offering strategic consulting to other enterprises for their survival or for their development, providing ideas, solutions and decision-making	Wu (2000:77)
迪吧 <i>dībā</i>	disco-bar	Disco bar	Own observation
果吧 <i>guóbā</i>	fruit-bar	A place selling fruit products	Own observation

<sup>18</sup> A question mark in brackets is added when the meaning of the word has been inferred by the author from the context in which the term itself is found.

話吧 <i>huàbā</i>	talk-bar	A call shop	www.nciku.com
畫吧 <i>huàbā</i>	painting-bar	A painting school, some- designed as a café	<a href="http://sh.msn.100du.com/sh/home.php?sid=1180956709506">http://sh.msn.100 du.com/sh/home. php?sid=11809567 09506</a>
花卉吧 <i>huāhuìbā</i>	flowering.plant -bar	A place where flowers are used for healthcare	Wu (2000:76)
懷舊吧 <i>huáijiùbā</i>	yearn.for.the. past-bar	A (virtual) meeting place dedicated to a rock band, to a videogame, etc. for the nostalgics (?)	<a href="http://tieba.baidu.com/?kw=toto%BB%B3%BE%C9">http://tieba.baidu. com/?kw=toto% BB%B3%BE%C9</a>
健身吧 <i>jiànshēnbā</i>	keep.fit-bar	A place equipped with fitness machinery	Wu (2000:74)
酒吧 <i>jiǔbā</i>	alcohol-bar	A place selling alcoholic beverages in a Western- style restaurant or hotel.	CCD (2002)
嚼吧 <i>juébā</i>	chew-bar	Room offering free chewing gum to office workers complaining of high pressure in high-end office complexes	www.shanghaidaily. com/buzzword (24/08/2008)
咖啡吧 <i>kāfēibā</i>	coffee-bar	café, coffee bar	Own observation
烤吧 <i>kǎobā</i>	roast-bar	A restaurant specializing in barbecue food	Wu (2000:74)

啤吧 <i>píbā</i>	beer-bar	Bar specializing in beer (?)	<a href="http://www.flickr.com/photos/u-suke/2416618474/in/set-72157604481600832/">http://www.flickr.com/photos/u-suke/2416618474/in/set-72157604481600832/</a>
球吧 <i>qiúbā</i>	ball-bar	A site offering information on ball games	<a href="http://www.qiuba.net">www.qiuba.net</a>
水吧 <i>shuǐbā</i>	water-bar	A place for relaxation that exclusively sells various types of drinks and drinking water	<a href="http://www.nciku.com">www.nciku.com</a> ; XH (2003)
陶吧 <i>táobā</i>	ball-bar	A pottery workshop where customers may create their own products	Fan (2002:136)
貼吧 <i>tīebā</i>	paste-bar	Online 'bar' to publish fans' posts related to their idols	<a href="http://www.shanghaidaily.com/buzzword">www.shanghaidaily.com/buzzword</a> (06/08/2007)
痛快吧 <i>tòngkuài bā</i>	unconstrained -bar	A shop where a customer pays for venting his or her tension, anger or frustration by violently punching or smashing goods	<a href="http://www.shanghaidaily.com/buzzword">www.shanghaidaily.com/buzzword</a> (30/09/2007)
外語吧 <i>wàiyǔbā</i>	foreign language -bar	A place for learning foreign languages	Wu (2000:76)
網吧 <i>wǎngbā</i>	net-bar	Internet café	XH (2003)
文化吧 <i>wénhuàbā</i>	culture-bar	A place (e.g. in a factory) where cultural activities are held	<a href="http://www.nssh.gov.cn/2008/11-21/162450.html">www.nssh.gov.cn/2008/11-21/162450.html</a>

香水吧 <i>xiāngshuǐbā</i>	perfume-bar	A place where the customers' health is restored using perfumes and music	Wu (2000:76)
眼吧 <i>yǎnbā</i>	eye-bar	An optometry clinic where a computer-manipulated environment claimed to be beneficial for eye health is created to help ease eye stress and disorders	www.shanghaidaily.com/buzzword (21/12/2008)
氧吧 <i>yǎngbā</i>	oxygen-bar	A leisure or business place which provides oxygen therapy equipment for people to inhale oxygen	Wu (2000:74)
影吧 <i>yǐngbā</i>	picture-bar	A public place where films and TV shows may be watched for a charge	baike.baidu.com/view/1136323.htm
遊戲吧 <i>yóuxìbā</i>	game-bar	An amusement arcade	Wu (2000:74)

Even in a limited sample as that presented in the table above, we have a remarkable number of different meanings conveyed by -吧 *-bā*. The dictionary definition of -吧 *-bā* as “an entertainment place with a particular function or supplied with some special equipment”, quoted above, may apply to instances such as 茶吧 *chábā* ‘Western-style tea house’, 烤吧 *kǎobā* ‘restaurant specializing in barbecue food’ or 痛快吧 *tòngkuai bā*, ‘shop where a customer pays for venting his or her tension, anger or frustration by violently punching or smashing goods’; however, we also have “bars” which do not seem to be connected with

entertainment, as 創意吧 *chuàngyìbā* ‘creative bar’, which is a kind of enterprise in the field of business consulting, or 話吧 *huàbā*, a call shop.

Such polysemy is not easy to accommodate, in a Construction Morphology approach, as pointed out in 1.3.1.2, since in word formation templates form, meaning and function are associated, and any shift in meaning would entail the birth of a new schema, theoretically; diachronic data tell us that the different meanings that a polysemous item conveys are connected, and thus we might posit a single schema which may encompass all of the uses of -吧 *-bā*.

In Cheng L. (2004:54), the ‘radial network’ (輻射網絡 *fúshè wǎngluò*) model for the analysis of polysemous word formatives is proposed. In such model, the ‘typical meaning’ of a sign is the centre, which forms a which forms a meaning network with its peripheral meanings. Such proposal is not new, since it was used by Jurafsky (1996) to represent the pathways of semantic evolution for the category of diminutives in morphology, and the ‘network’ itself is based on the notion of ‘radial category’ (Lakoff 1987). Below is our representation of the radial network for 吧 *bā*:



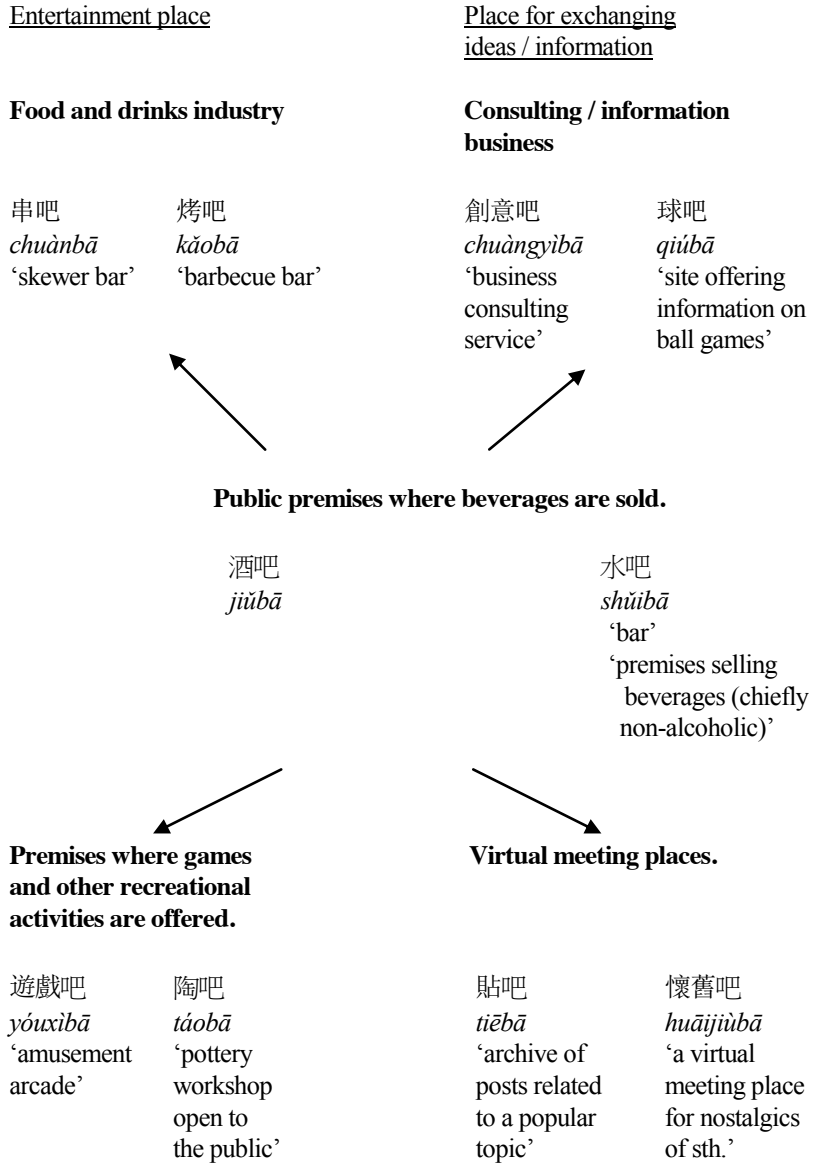


Figure 3.2 Radial network for 吧 bā

In the representation of the radial network proposed in figure 3.2, words containing -吧 *-bā* as the right-hand constituent are divided into two groups, sharing meaning similarities, defined in the header of the column. Each group is further subdivided into two sets, historically “derived” from the central meaning (here, that of ‘bar’); the original lexical meaning, thus, lies at the centre of the scheme. Through such a modality of representation, both the relationships between the ‘centre’ of the scheme and the ‘peripheral’ meanings and the connections among the different ‘sub-families’ are visible. although we are using synchronic data, the diachronic dimension is involved as well, since the central lexical meaning of the lexeme is also the oldest, even though for -吧 *-bā* complex words the time distance between the various usages is very small, being 酒吧 *jiǔbā* ‘(alcohol) bar’ a word with a comparatively short history<sup>19</sup> (cf. the schema in Jurafsky 1996:542).

The radial network model is a convenient method from the descriptive point of view, but it has a serious limitation, in that it does not make explicit the possible relations between *all* of the meanings of -吧 *-bā* by the mechanisms of abstraction, as described in CHAPTER 1 for Ewe *ví* ‘child’ > *ví* ‘human derivational affix’ (1.3.1.2). A polysemy as that of Ch. -吧 *-bā* or Ewe *-ví*, which can convey meanings as disparate as YOUNG, TYPICAL BEHAVIOUR and NOT YET PASSED AN EXAM, among others, may be analysed in two ways. One could say that each line of development corresponds, synchronically, to a different suffix, and all of those suffixes are homophonous in the present stage of the language (such possibility is suggested by Heine, Claudi & Hünemeyer 1991:95). Otherwise, we may propose that the different ‘channels’ of semantic evolution be subsumed under one overarching schema and, thus, we are

---

<sup>19</sup> The word 酒吧 *jiǔbā* is attested in Wang Meng’s “別依阿華 (*bié yī Ā-huá*)” (‘don’t rely on A-Hua), published in 1981. The version 酒吧間 *jiǔbājiān*, with the added character 間 *jiān* ‘room, space’ was apparently attested as early as the 1930s, e.g. in Ding Ling’s “詩人亞洛夫 (*Shīrén Yàluòfū*)” (‘The poet Jarov’ HYDCD 1993), but we believe that this is of no significance for our analysis, since the various  $[[x]_X [bā]_N]_N$  words considered here could not be created by analogy with 酒吧間 *jiǔbājiān*, given the different structure.

dealing with one polysemous affix, rather than with several related affixes. Thus, let us have a closer look at the meaning shifts undergone by -吧 *-bā*, to evaluate which of the two possible analyses outlined above best fits the data.

Using the metalanguage of ‘Lexical Semantics’ (Lieber 2003), we shall try to give a (tentative) representation of the ‘body’ of the lexical meaning of 吧 *bā* (as in 酒吧 *jiǔbā*), i.e. a list of its salient encyclopaedic features:

- (17) <public premises>  
 <selling drinks>  
 <entertainment place>  
 <place where recreational activities may be held>  
 <meeting / socialization place>  
 <free access>  
 <economical activity>

In terms of ‘isolating abstraction’, the semantic features which are isolated in the various instances of -吧 *-bā* are difficult to identify; For instance, <place where recreational activities may be held> would fit for many -吧 *-bā* complex words, but not e.g. for 創意吧 *chuàngyìbā* ‘business consulting service’ or 話吧 *huàbā* ‘call shop’. Commonalities in meaning are easy to find inside subsets, but not across them; also, the four partitions presented here might even be insufficient, and we might want to add e.g. a fifth subset, ‘healthcare / fitness’, including 健身吧 *jiànshēnbā* ‘premises equipped with fitness machinery’ and 眼吧 *yǎnbā* ‘an optometry clinic where a computer-manipulated environment claimed to be beneficial for eye health is created to help ease eye stress and disorders’, among others. The CM notion of ‘paradigmatic relations’ would have to be applied within each subset (subparadigms), and we would have a word formation schema for each subgroup (see Dong X.’s ‘word formation rules’, 2.2.2).

Such data, however, might be interpreted in a very different way. The fact that it is very hard to accommodate all the ‘extensions’ basing on the

encyclopaedic features of the ‘lexical’ meaning of -吧 *-bā* may be indicative of the fact that generalizing abstraction has occurred, leading to an increase in the extensional meaning of the sign; as pointed out before (1.3.1.2), this is a characteristic correlate of grammaticalization. Very often, the meaning of a derivational affix is very general and, thus, it is in no way strange that it may designate a huge variety of referents (see Ewe -vī). It appears that by the mechanism of generalizing abstraction, the morpheme -吧 *-bā* (as part of the word 酒吧 *jiǔbā*) has been taken to a higher taxonomical level; such evolution may be sketched as such:

(18)  $[[x]_N [bā]_N ]_N$  ‘premises selling drinks or food belonging to category  $[X]_N$ ’ > ‘place (actual or virtual) where a service related to  $[X]_{N/ADJ}$  is offered or where information related to  $[X]_N$  may be exchanged or where  $[X]_V$  may be done’

The process starts with analogy, by which *hapax legomena* are built which, later, through meaning extension, may become a subclass of word with a common element, i.e. a paradigm (in the sense of CM), which is what apparently happened to -吧 *-bā* (Fan L. 2002:137). As mentioned above, words were first formed by analogy with 酒吧 *jiǔbā* ‘(alcohol-)bar’; following the creation of a number of neologisms, a constuctional schema is ‘born’. We illustrated before (1.3.1.2) the evolution of -性 *-xìng* ‘the property of  $[X]$  / connected with  $[X]$ ’; just as -性 *-xìng*, -吧 *-bā* too gets associated with word classes other than nonus, namely verbs (嚼吧 *juébā*, lit. ‘chew-bar’, 健身吧 *jiànshēnbā*, ‘keep.fit-bar’) and adjectives (痛快吧 *tòngkuai bā*, lit. ‘unconstrained-bar’). At the same time, the semantic area of -吧 *-bā* complex words broadens, and newly-coined ‘bars’ include places which offer drinks and food, but also other services (香水吧 *xiāngshuǐbā*, ‘A place where the customers health is restored using perfumes and music’) and all sorts of gathering places, whether for playing games (遊戲吧 *yóuxìbā* ‘amusement arcade’) or for exchanging information on a topic (貼吧 *tiēbā* ‘post bar’, i.e. ‘webpage where fans publish posts related to their

idols'), or even a consulting business (創意吧 *chuàngyìbā* 'business consulting service'). From the semantic point of view, we believe that metaphor is at work, by which the 'core' meaning of 吧 *bā* is extended to include any place which can be associated with the defining features of a 'bar', be it actual or virtual; thus, the word formation schema in (18) may generalise over all of the complex words.

However, the meaning of -吧 *-bā* is not general enough to define it as a pure locative, and the semantic connections with its 'core' meaning 'bar' (from 酒吧 *jiǔbā*) are still visible, especially when -吧 *-bā* competes with some other formative. For instance, to indicate a public place where people go to drink tea, the lexeme 茶 *chá* 'tea' may be associated with -吧 *-bā*, but also with 館 *guǎn* 'building' and, actually, both 茶吧 *chábā* and 茶館 *chágǔǎn* are attested, with no blocking effect. However, 茶吧 *chábā* and 茶館 *chágǔǎn* denote two quite different referents: whereas the former refers to a Western-style tea house (see table 3.4), the latter refers to a traditional Chinese tea house and, thus, it appears that -吧 *-bā* has a semantico-pragmatic 'flavour' of modernity and 'westernity'. Another comparison which may be suggested is that between 文化宮 *wénhuàgōng*, lit. 'culture palace', a "large scale, well-equipped cultural palace, usu. having a cinema, a lecture hall, a library, etc.", 文化館 *wénhuàguǎn*, lit. 'culture building', a "cultural centre[,] establishment where cultural work for the masses is carried out and where people take part in cultural and recreational activities" (CCD 2002) and 文化吧 *wénhuàbā*, which we understood as a place (e.g. in a factory) where cultural activities are held (see table 3.4); in such case, the distinction is one of size/scale, and a 'bar' is associated with the idea of a comparatively small place (smaller e.g. than a 'palace'). Thus, 'modern / Western' and 'of limited size/scale' are both features which -吧 *-bā* complex words inherited from the original 酒吧 *jiǔbā* '(alcohol) bar', and are part of its pragmatico-semantic colouring.

In our opinion, of the two possible analyses presented here, namely either having a number of related schemas or positing only one

overarching schema for -吧 *-bā* complex words, we believe that the latter fits best the data and, also, is more consistent with the characteristics of (lexical) derivational formatives, cross-linguistically. The fact that the schema proposed in (18) is very general is actually not surprising, as derivational affixes are often the product of ‘meaning abstraction’; -吧 *-bā* apparently followed a pathway of metaphorical extension, a common fact in grammaticalization, with increased lexical generality and contextual expansion (see 1.3.1.2, 1.3.1.3). Despite the generality of the schema, the connections with the original meaning of ‘bar’ are still present, as shown by the semantico-pragmatic features described above.

One more clarification is needed. The fact that we support such analysis for -吧 *-bā* does not entail that we believe that the same may be said about just any polysemic affix, as Ewe *-ví*, as each individual case is independent from every other and should be considered by itself; actually, in the history of languages we have many examples of items which, due to different developments, at some point become ‘estranged’ (as e.g. Old Eng. *dom* / *-dom* > Modern Eng. *doom* and *-dom*; see above, 1.2.2).

In 2.2.2, we quoted (from Ma Q. 1995) the bound morpheme 語 *yǔ* ‘language, expression’ as an example of an item which may not be regarded as an affix, despite its ‘versatility’ in word formation, as it is found both on the left side and on the right side of complex words, bearing its basic meaning. We should point out that 吧 *bā* as well may be found as the left-hand constituent in complex words (exx. from Arcodia 2011):

(19) 吧女	吧台	吧員
<i>bānǚ</i>	<i>bātái</i>	<i>bāyuán</i>
bar-woman	bar-counter	bar-personnel
‘barmaid’	‘bar counter’	‘bartender’

However, in such words 吧 *bā* is used in its ‘core’ meaning, i.e. that of ‘bar’; it thus appears that we have two related homophonous items, namely 吧 *bā* ‘bar’, a bound lexical morpheme, and -吧 *-bā*, a derivational suffix.

One residual issue is the possible role that English had in the

‘morphologization’ of 吧 *bā*. It is true that a neologism as 網吧 *wǎngbā* might have been coined as a calque of Eng. ‘internet café’; however, the semantic development of -吧 *-bā* as word formative appears as completely autonomous. We may compare -吧 *-bā* to -領 *-lǐng* ‘collar’, which just as Eng. ‘collar’ is used to indicate certain categories of workers, as ‘white collar’. According to Fan L. (2002:137), Ch. 白領 *báilǐng* ‘white collar’ was the model after which 藍領 *lánlǐng* ‘blue collar’ was created by analogy; other neologisms which were created following this pattern are 粉領 *fěnlǐng* ‘pink collar’ (‘woman working in the service industry, as a secretary, etc.’), 黑領 *hēilǐng* ‘black collar’ (‘person engaged in menial manual labour’) and 金領 *jīnlǐng* ‘golden collar’ (‘top-level executive’), among others. In English as well many ‘-collar’ words were created in the XXth and XXIst century, as ‘green-collar’ (‘person working in environmental care’), ‘scarlet-collar’ (‘woman working in internet pornography’) and ‘open-collar’ (‘person working from home, mostly using the Internet’). Even for -領 *-lǐng* ‘collar’, it appears that English provided the model (i.e. the constructional idiom), but the subsequent developments were autonomous for each of the two languages; we have both words which are structurally identical but convey different meanings in the two languages, as 黑領 *hēilǐng*, ‘person engaged in menial manual labour’, corresponding to Eng. ‘black collar’, which however indicates a person operating in the black market, and Mandarin words which have no equivalent in the other languages, as 金領 *jīnlǐng* ‘golden collar’ (‘top-level executive’). Thus, -吧 *-bā* and -領 *-lǐng* are not fundamentally different, as to the role which English had in their evolution. The word formatives -吧 *-bā* ‘bar’ and -領 *-lǐng* ‘collar’ are a good illustration of how an item of the Chinese lexicon may develop a specific usage inside a word formation schema, with some extent of influence from a foreign language (English, in both cases). The fact that such items can spread (relatively) easily is not surprising, given the fact that Mandarin is very receptive towards class nouns; needless to say, not all of such items are to be regarded as derivational affixes.

In the next section, we shall deal with another suffix-like formative of Mandarin which is often regarded as a product of ('Japanese-mediated') European influence, namely -化 *-huà* '-ise, -ify'.

### 3.2.3 'Chinese' vs. 'European' Affixes?

In the literature on Chinese word formation, there are two groups of items which are commonly regarded as affixes, even in many 'conservative' works. The most representative morphemes in the first group are the 'dummy' nominal suffixes -兒 *-r*, -頭 *-tou* and -子 *-zi*, the main function of which is to provide a 'prosodic support' for the building of a word (see Feng 1998, 2001). As said before (2.2.1), they form a peculiar group, since they have neither lexical nor grammatical meaning; they sometimes perform nominalization (as for 想頭 *xiǎngtōu* 'idea', from the verb 想 *xiǎng* 'to think') but, very often, they attach to nominal morphemes and, thus, are semantically and functionally redundant. Hence, they appear as markedly different from the 'typical' affixes of Indo-European languages (Pan W. 1990, Dong X. 2004). The other group of 'commonly accepted affixes' includes, items as -化 *huà* '-ise, -ify' and -性 *-xìng* 'the property of [X]'; they are normally regarded as affixes because they have a stable word class and, also, because they are semantically and functionally analogous to 'Western' (i.e. Standard Average European) affixes (3.1.2). Needless to say, the analogy with English or French affixes is hardly a criterion for 'affixhood'; we shall thus look into the history of these formatives, to evaluate whether their evolution belongs in the domain of grammaticalization, and to assess the possible influence of SAE languages in their development. Since -性 *-xìng* 'the property of [X]' was dealt with extensively in CHAPTER 1 (1.3.1.2, 1.3.2.1), we shall focus here on -化 *huà* '-ise, -ify'.

Some remarks are needed, however, as to the origin of -性 *-xìng* 'the nature of [X]'. We mentioned before (3.2.1.1) that, according to Wang L. (1980:230), the introduction of -性 *-xìng* into Mandarin as the (rough) equivalent of Eng. *-ty*, *-ce* [sic!] and *-ness* has occurred through the 'mediation' of Japanese loanwords as 可能性 *kanōsei* and 重要性 *jūyōsei*,



the ‘translation’ of Eng. ‘possibility’ and ‘importance’, accepted in Chinese as *kěnéngxìng* and *zhòngyàoxìng* (see also Wang L. 1989); later on, *-xìng* was used in China as a redundant marker, as in 必要性 *bìyàoxìng* ‘necessity’, which was introduced from Japan as just 必要 *bìyào* (Jap. *hitsuyō*) ‘necessity’, and then *-性 -xìng* was added to it<sup>20</sup>. In Masini (1993:140, 150), both *-性 -xìng* and *-化 huà* are classified as ‘suffix-like formatives’ of Japanese origin<sup>21</sup>; the only neologisms containing *-性 -xìng* in his sample (covering the period from 1840 to 1898) are 炭性 *tànxìng* ‘carbon’ and 銀性 *yínxìng* ‘silvering’ (uncertain translation). However, in 炭性 *tànxìng* ‘carbon’ the formative *-性 -xìng* is clearly not used in its ‘affixal’ meaning, since it indicates a concrete referent, and the same is true for 銀性 *yínxìng* ‘silvering’ (if the translation is correct). According to Chen R. (1986:89-90), the evolution of the lexeme 性 *xìng* into an affix started before the XXth century and was already ‘mature’ in the years following the “May Fourth Movement” of 1919<sup>22</sup>; our data confirm that the evolution of 性 *xìng* into the suffix *-性 -xìng* ‘the nature of [X] / connected with [X]’ was well under way before the XXth century. We believe, thus, that it is misleading to ‘credit’ the Japanese with the introduction of *-性 -xìng* to China; rather, it appears that the acceptance of many *-性 -xìng* complex word from Japanese

<sup>20</sup> Note that the form 必要性 *hitsuyōsei* for ‘necessity’ is attested in Modern Japanese; it is unclear, however, whether such lexeme was ‘imported’ from China as such or whether it is an autochthonous, (partly) independent creation.

<sup>21</sup> “In the second half of the XIXth century, most trisyllabic Japanese loans had the suffix-like formative [學] *xue*. Later, other such formatives proved to be very productive, and contributed to rendering Modern Chinese more receptive to trisyllabic terms than it had been. (...) Particularly *hua* 化, a suffix-like formative used to form adjectives and nouns, and *xìng* 性, to form attributes and nouns” (Masini 1993:150, fn. 48).

<sup>22</sup> An anti-imperialist, nationalistic movement growing out of student protests on the Fourth of May 1919 in Beijing. Following the May Fourth Movement, Classical Chinese was substituted by Vernacular Chinese (白話 *báihuà*) as the standard written language. According to Wang L. (1980:586; my translation), “the twenty or thirty years following the May Fourth Movement are the period of the great change in the Chinese lexicon”.

provided a stimulus for the diffusion of a pattern of word formation which was already attested in the Chinese language (see words as 忍性 *rěnxìng* ‘endurance, tolerance’, attested as early as the XVIIth century; CHAPTER 1, fn. 33), possibly encouraging the building of trisyllabic neologisms, as seen for -學 *-xué* ‘branch of learning’ (3.2.1.1, 3.2.1.2; see also fn. 20). The ‘historical’ lexemic meanings of 性 *xìng* ‘nature, character, disposition’ are still visible in Modern Mandarin words as 黨性 *dǎngxìng* ‘party spirit / character’ and others; Luo J. (2004:94) suggests that these are all ‘projections’ of the diachronic evolution of such formative. This is a common fact when there is a ‘split’ between a lexeme and a corresponding (bound) formative, as for Du. *boer* / *-boer*, since the former may still be used with the original meaning ‘farmer’; the difference is that in the case of 性 *xìng* / -性 *-xìng*, both the lexeme and the corresponding affix are bound. Let us now turn to the analysis of the data on -化 *huà* ‘-ise, -ify’.

The morpheme -化 *-huà*, functionally equivalent to Eng. ‘-ise’ and ‘-ify’, is also often interpreted as a ‘Western’ suffix which entered the Chinese lexicon via Japanese; according to Wang L. (1980:311) this happened, again, in the years following the May Fourth Movement, through the acceptance of ‘Japan-made’ neologisms as 工業化 *gōngyèhuà* ‘industrialise’, 機械化 *jīxièhuà* ‘mechanise’, 現代化 *xiàndàihuà* ‘modernise’ (see above, 2.1.2). However, Wang L. concedes that after the acceptance of such model, more -化 *-huà* complex words were created by analogy in China, as 形象化 *xíngxiànghuà* ‘symbolise’ and 規律化 *guīlǜhuà* ‘regularise’; Ma Q. (1995:107), as said before, believes that even formatives as -化 *-huà* are ‘domestically made’, and, although they got diffused following an external influence, their development is independent from the foreign model. To assess such claims, let us now present and discuss some diachronic data on the development of 化 *huà*.

Under the heading 化 *huà*, no less than 25 distinct meanings are listed in the “Comprehensive Chinese Dictionary” (漢語大詞典 *Hànyǔ Dà Cídiǎn*, HYDCD 1993; 9 in the “Great Dictionary of Classical Chinese”

GHYDCD 2000<sup>23</sup>). However, the meanings which are relevant for the evolution of 化 *huà* into the bound formative ‘-ise, -ify’ are only those connected with the idea of ‘change’, namely:

- a. ‘Change human habits’ (转移人心风俗 *zhuǎnyí rénxīn fēngsú*)
- b. ‘Change’ (变化 *biànhuà*; 改变 *gǎibiàn*)
- c. ‘Change in quality’ (质变 *zhìbiàn*)

Thus, 化 *huà* was used in the since Old Chinese and until Early Mandarin as a verb, whereas in the Modern language is found only in compound words and in ‘classical’ idioms as 化干戈為玉帛 *huà gāngē wéi yùbó*, lit. ‘to transform weapons in gems and silk (i.e. friendship)’, corresponding to Eng. ‘burying the hatchet’. An early usage of -化 *-huà* as a ‘suffix’ is found in Huang Yuanyong’s ‘My Confession’ (懺悔錄 *chànhuīlù*, early XXth century; qtd. in HYDCD 1993), still written in Classical Chinese (1.1.1):

- (20) 安東者，號稱吾國土地，而完全日化者也。  
*Āndōng zhě hào chēng wúguó tǔdì ér*  
 Andong EMPH be.known.as my-country territory and.yet  
*wánquán rìhuà zhě yě*  
 completely Japan-*huà* NMLZ PART  
 ‘Andong, it is known as a part of my country, and yet it has been completely japanized’

Since Huang Yuanyong died in 1915, well before the beginning of the May Fourth Movement, we may say that -化 *-huà* complex words built after the word formation schema  $[[X]_X [huà]]_V$  ‘to cause s.o. / sthg. to become X’ were already attested in the language even before that period.

We searched for words with the (generic) ‘X-*huà*’ structure in the *Academia Sinica* corpus of Early Mandarin (13th-19th cent.) and we had 57 hits, 48 of which are tagged as verbs, all disyllabic; below are some

<sup>23</sup> We excluded from the count the use of 化 *huà* as a surname.

examples :

(21) 溶化	坐化	軟化
<i>rónghuà</i>	<i>zuòhuà</i>	<i>ruǎnhuà</i>
dissolve- <i>huà</i>	sit- <i>huà</i>	soft- <i>huà</i>
‘to melt’	‘die while sitting’	‘soften’

In 溶化 *rónghuà*, 化 *huà* bears the meaning ‘to melt’, and in 坐化 *zuòhuà* it means ‘die’, both (broadly) connected with the idea of ‘change’, but not synonymous with it, needless to say. The word 軟化 *ruǎnhuà* ‘soften’, attested in the seventeenth century vernacular novel “The Story of a Marital Fate to Awaken the World” (醒世姻緣傳 *Xǐngshì Yīnyuàn Zhuàn*), seems to have been formed according to the ‘to X-ise’ schema introduced above; in the passage where it was found, however, it is used as an intransitive verb. Note that the word is still part of the Modern Mandarin lexicon, but nowadays it can be used both transitively and intransitively, just as the correspondent English ‘-en’, ‘-ise’ and ‘-ify’ verbs. Thus, -化 *-huà* ‘-ise, -ify’ complex words actually existed well before the XXth century, and it is possible that more types were attested, but were not found due to the limited nature of our sample. Nevertheless, our data seem to support the view that the pattern became productive only recently.

As to the word class and other features of -化 *-huà* ‘-ise, -ify’ complex words, such as transitivity, Zhang Yu. (2002:50-52) divided -化 *-huà* words from the modern lexicon into four groups:

a. X 化<sub>1</sub>; may be followed by an object, but may not be modified by a degree adverb, as shown below:

丑化 →	丑化现实生活 →	*最丑化
<i>chǒuhuà</i>	<i>chǒuhuà xiànréalsí shēnghuó</i>	<i>zuì chǒuhuà</i>
‘uglify’	‘uglify real life’	‘most uglified’

b. X 化<sub>2</sub>; may not be followed by an object, may mostly be used in a passive sentence, normally may not be modified by a degree adverb:

僵化 →	*僵化了思想 →	思想被僵化了
<i>jiānghuà</i>	<i>jiānghuà le sīxiǎng</i>	<i>sīxiǎng bèi jiānghuà le</i>
‘become rigid’	‘made thinking rigid’	‘thinking was made rigid’

c. X 化<sub>3</sub>; may be modified by a degree adverb and may not normally act as predicates, if they are not modified by such an adverb, suffer from restrictions in passivization

很女性化 →	*他女性化
<i>hén nǚxìnghuà</i>	<i>tā nǚxìnghuà</i>
‘(very) feminised’	‘he is feminised’

d. X 化<sub>4</sub>; may act as predicate in a simple sentence only if preceded by adverbs as 已經 *yǐjīng* ‘already’ or 剛剛 *gānggāng* ‘a moment ago’, may not be modified by degree adverbs:

已经工业化了 →	*国家工业化了
<i>yǐjīng gōngyèhuà le</i>	<i>guójiā gōngyèhuà le</i>
‘already industrialised’	‘the country has industrialised’

Such differences are analysed by Zhang Yu. (2002:52) as functional in nature. He proposes a ‘weakening ranking’ of the verbal value of -化 *-huà* complex words, in which each of the four groups corresponds to a word (sub-)class:

X 化 <sub>1</sub> >	X 化 <sub>2</sub> >	X 化 <sub>3</sub> >	X 化 <sub>4</sub>
V (transitive)	V (intransitive)	ADJ	N

This does not mean that such words are actually adjectives, nouns, etc.; -化 *-huà* complex words are basically verbs, and may also be used as

nouns (e.g. 四個現代化 *sì gè xiàndàihuà* ‘the Four Modernizations’) just as many other verbs of Mandarin (工作 *gōngzuò* ‘to work’ / ‘job’, etc.). Such ranking, rather, aims at showing that the ‘prototype’ of the -化 *-huà* verb is the transitive one (X 化<sub>1</sub>), and the more we move to the right of the ranking, the more properties of transitive verbs are lost. Such grouping influences the choice of possible ‘bases’: 化<sub>1</sub> only combines with monosyllabic adjectives, X 化<sub>4</sub>, is productively used only with multisyllabic nouns, and 化<sub>3</sub> and 化<sub>4</sub>, predictably, do not show such neat preferences.

According to Zhang Yu. (2002:53), such ranking is a reflection of the pathway of grammaticalization of 化 *huà*, and the ranking is to be interpreted also as a ‘emptying’ (i.e. abstraction) ranking (where 化<sub>4</sub> is the most grammaticalised sign). Not surprisingly, the early example 軟化 *ruǎnhuà* ‘soften’ has a monosyllabic modifier as a base, just as required by 化<sub>1</sub>, i.e. the least grammaticalised -化 *-huà*; note, however, that 軟化 *ruǎnhuà* is used as an intransitive verb in our only example, but this does not mean that a transitive usage was also possible already at that time (it is actually possible in Modern Mandarin; see above). With a cursory search of the *Academia Sinica* corpus of Mandarin Chinese, we found that most of the -化 *-huà* words in the sample are trisyllabic, i.e. they have a disyllabic base; also, for words with a monosyllabic base, it appears that 化 *huà* may convey different meanings (as in ex. 21 above), whereas trisyllabic -化 *-huà* complex words mostly conform to the schema  $[[X]_X [huà]]_V$  ‘to cause s.o. / sthg. to become X’. The history of 化 *huà* shows, again, the interaction between autochthonous word formation patterns and the influence of Japanese neologisms, acting as ‘mediators’ between China and Europe. We may suggest that Japanese has borrowed a word formation pattern, only to give it back to Chinese later, when it had become more ‘ripe’; following the model of trisyllabic -化 *-huà* complex words created in Japan, many such trisyllabic neologisms are created in Chinese (see the examples quoted above), similarly to what happened for -學 *-xué* ‘branch of learning studying  $[X]_N$ ’ (3.2.1.1, 3.2.1.2).

As to the connection between ‘weakening’ of the transitive verbal

character and grammaticalization suggested by Zhang Yu., we may remark that the position in the ranking of individual words is not necessarily connected with their origin: words imported from Japan as the above quoted 工業化 *gōngyèhuà* ‘industrialise’, 機械化 *jīxièhuà* ‘mechanise’, 現代化 *xiàndàihuà* ‘modernise’ (Wang L. 1980:311) belong to different groups (工業化 *gōngyèhuà* belongs to the fourth group, N, and the latter two belong to the third, ADJ); another word of Japanese origin as 歐化 *Ōuhuà* ‘Europeanise’ has the same distribution of the second group (V intransitive), just as the above quoted autochthonous word 形象化 *xíngxiànghuà* ‘symbolise’. Hence, it seems that the process of weakening of the verbal features of -化 *-huà* complex words is language-internal, rather than conditioned by foreign influences or by the ‘age’ of individual words.

We also believe that it is highly doubtful that the loss of verbal features and grammaticalization for -化 *-huà*. The English equivalent *-ise / -ize* is of Greek origin, and entered (Middle) English through Old French (which received it from Late Latin (e.g. *baptizāre* ‘to baptise’ < Greek βαπτίζειν *baptízein*); the same suffix evolved from Latin into It. *-izzare*. The original Ancient Greek suffix, -ίζειν *-ízein* ‘the adoption or imitation of customs / fashions of certain groups’, could combine with proper names and ethnonyms, as in βαρβαρίζειν *barbarízein* ‘to barbarise’, and had only intransitive usage (DELI 1999); with a cursory dictionary search of Eng. *-ise / -ize* words and It. *-izzare* words (in DEM 2000), we found out that nearly all of the modern verbs with such suffix in English and Italian are either transitive or both transitive and intransitive (as ‘to barbarise’, quoted above). Also, the few cases of verbs which are always intransitive sometimes have a different origin, as It. *agonizzare* ‘to be in the throes of death’, from Late Latin *agonizāre* ‘to fight’ (< Greek αγωνίζεσθαι *agonízesthai*); the contemporary meaning is the result of analogy with the word *agonia* ‘agony’. Thus, the evolution from the ‘ancient’ suffix to the modern ‘European’ equivalents went in the direction intransitive > transitive, rather than the other way round. Both the diachronic Chinese data and the comparison with English and Italian tell us that there seems to be no connection between the loss of verbal features and increasing grammaticalization; rather, what all

-化 *-huà* complex words in the four groups proposed by Zhang Yu. (2002) have in common is that they are all verbs conforming to the schema in (22):

(22)  $[[X]_X [huà]]_V$  ‘to cause s.o. / sthg. to become X’ / ‘to become X’

Thus, they seem to allow both transitive and intransitive usage; a larger sample of data is needed, however, to find out whether there are any -化 *-huà* complex words which are actually used only intransitively or only transitively. A different explanation must be proposed for the different distribution of verbs from each of the four groups, but this is beyond the aims of the present research.

To sum up, we believe that -化 *-huà* shows several characteristics of grammaticalised items, in terms of structural scope reduction (from verb phrases to single words) and increased bondedness (from free to bound in a fixed position; 1.3.1.1); it is found in a stable position with a stable meaning, and it has a very different distribution and usage from the original verb, which anyway is no longer used in the modern language (compare Eng. ‘mock’ and ‘-type’. exx. 15-16, CHAPTER 1). As to semantics, its ‘affixal’ meaning is tightly connected with its ‘core’ meaning of ‘change’ but, yet, the abstraction of meaning is visible, since, as pointed out by Zhang Yu. (2002:53), the meaning of ‘change’ for 化 *huà* historically derives from ‘to educate’ (教育 *jiàoyù*), ‘to change through education’ (教化 *jiàohuà*); it appears, thus, that the evolution in meaning has gone in the direction of generalisation, from ‘educate’ / ‘change through education to ‘change (into something else)’ to ‘to cause s.o. / sthg. to become X’ / ‘to become X’. Hence, we believe that -化 *-huà* is best understood as an affix. As to the role of foreign loanwords in the development of 化 *huà* into a productive word formative, the situation appears as analogous to that of 學 *xué* ‘branch of learning’, i.e. a word formation pattern already attested in the history of the Chinese language becomes very productive following the reception of many Japanese graphic loans translating ‘Western’ notions, with an expansion of the size of the modifier (from 1 to 2 or more syllables).

Let us now turn to the discussion of data on -者 *-zhě* ‘agentive suffix’



and -式 *-shì* ‘model, style’, two formatives which appear as rather different from those discussed in what precedes.

### 3.2.4 Between Morphology and Syntax: -者 *-zhě* and -式 *-shì*

The morpheme -者 *-zhě* has been often described in the literature as an agentive suffix (e.g. in Packard 2000, Yip P. 2000, Dong X. 2004). Such formative actually possesses many characteristics of typical derivational affixes, as being found in a fixed position in a ‘word family’, conveying a stable meaning; in the reverse lexicon NXCD (2005), in all of the words listed under the heading 者 *zhě* such morpheme acts as an agentive suffix<sup>24</sup>. Also, the notion of ‘agent’ is often expressed by means of derivational affixes, cross-linguistically (see Heine & Kuteva 2002).

As mentioned in 1.3.1.1, in Classical Chinese 者 *zhě* was a demonstrative, “a pronominal substitute for the head of a noun phrase” and a mark of nominalization (see ex. 20 above), among other functions (Pulleyblank 1995; ex. from the 韓非子 *Hán Fēizǐ*, IIIrd cent. BCE):

(23) 楚人有鬻盾與矛者

*Chǔ rén yǒu yù dùn yǔ máo zhě*

Chu person there.be sell shield and spear DEM

‘In the kingdom of Chu there is a person selling shields and spears’

In the Classical language, when 者 *zhě* directly follows a (mono- or disyllabic) verbal or adjectival lexeme, the resulting structure is analogous to a ‘modern’ word containing a suffix (or, generally speaking, a class noun), as e.g. 來者 *láizhě* ‘the one who comes’, 賢者 *xiánzhě* ‘the one who is virtuous; virtuous person’; in fact, Hong B. (2005:188) analyses 者 *zhě* in Old Chinese primarily as an affix. In Cheng X.’s works on the

<sup>24</sup> The only exception is the idiom (成語 *chéngyǔ*) 犖犖大者 *luòluòdàzhě* ‘major items, salient points’; however, since such idioms are normally built (somehow) following Classical Chinese syntax, the above mentioned exception is of no significance for our research.

history of the Chinese lexicon, ranging from the pre-Qin language (先秦 *Xiān Qín*, before 221 BCE) up to the Five Dynasties era (五代 *Wǔdài*, 907-960 CE; Cheng X. 1992a-d), 者 *zhě* is analysed as a suffix (詞尾 *cíwěi*) which has not yet fully developed (grammaticalised?) in its early attestations, because of the syntactic (rather than morphological) nature of items as 學者 *xuézhě* ‘scholar’ or 使者 *shǐzhě* ‘emissary, messenger’. Such forms are syntactic, according to Cheng X., since more morphemes may be added to them, as in 修學者 *xiūxuézhě*, 出使者 *chūshǐzhě*, without significantly altering their meaning, whereas in Modern Mandarin structures as 學者 *xuézhě* and 使者 *shǐzhě* are lexicalised as such. (Cheng X. 1992c:93-94). Starting from the Three Kingdoms and Northern and Southern dynasties periods (三國兩晉南北朝 *Sānguó Liǎngjìn Nán-Běi Cháo*, 220-589 CE, i.e. the early period of Middle Chinese), according to Cheng X., -者 *-zhě* is to be regarded as a ‘proper’ noun forming suffix (Cheng X., 1992a:68-69).

In **1.3.1.1**, we presented some data on the usage of 者 *zhě* in the contemporary language, showing how such morpheme may form noun, combining with other nouns, verbs, adjectives and phrases (see exx. 22-23, CHAPTER 1). This suggests that ‘more grammaticalised’ and ‘less grammaticalised’ usages coexist in the same synchronic stage of the language, which is one of the areal fetures of the East and South-East Asian *Sprachbund* (**1.3.2**). The relevant parameter of grammaticalization, here, is that of structural scope reduction (**1.3.1.1**), i.e. the ‘size’ of the construction in which 者 *zhě* is a part; if in 使用者 *shǐyòngzhě* ‘user’ 者 *zhě* attaches to a word (the verb 使用 *shǐyòng*), in 不符合條件者 *bùfúhétiáojiànzhě* ‘not qualified’ (lit. ‘not meeting conditions’), being thus ambiguous between a word formative and a function word, a free grammatical morpheme. In Classical Chinese, according to Dong X. (2004:85), 者 *zhě* was a function word (虛詞 *xūcí*), operating in the syntactic domain, which could be added to verbs or verb phrases to perform nominalization; in such structures, 者 *zhě* indicated the subject of the verb / VP (or the referent of the adjective), and it was fully productive (i.e. it could nominalize about any verb / VP), as expected for

an item of syntax, lacking the ‘arbitrary gaps’ which are typical of word formation processes (see Haspelmath 2002). According to Dong X. (2004:86-87), the difference between the usage of 者 *zhě* as a function word and as a word formation element lies in the relationship between such item and those items it combines with; whereas in its ‘classical’ usage 者 *zhě* indicates the subject of the (verbal or adjectival) predication (see Zhu D. 1983, Yuan Y. 1997), in its morphologised usage it has mainly agentive meaning, a semantic category often overlapping with that of grammatical subject. Words as 學者 *xuézhě* ‘scholar’ or 使者 *shǐzhě* ‘emissary, messenger’, quoted above, were preserved in the lexicon because word formation rules are based on theta roles (patient, agent, etc.) rather than on syntactic notions as subject and object (Dong X. 2004, quoting Anderson 1992). Moreover, Dong X. (2004:87) remarks that meaning specialization has occurred for some -者 *-zhě* complex words, as the above quoted 學者 *xuézhě* ‘scholar’, which does not merely indicate any person engaged in study, but, specifically, a somehow prominent person in some field of learning (such meaning is attested at least since the Northern Song period, 960-1127 CE; GHYDCD 2000); this happens because “lexicon and morphology are connected (...) and thus the forms generated by morphology may possess some idiosyncratic meaning, differently from syntactic structures” (Dong X. 2004:87, *contra* Packard 2000:73; see also Dong X. 2002).

So far, so good. However, in Modern Mandarin, we have not only instances of 者 *zhě* attaching to a phrasal constituent, but also ‘normal’ -者 *-zhě* complex words with a lexical ‘base’ which show properties both of words and of phrases (Dong X. 2004:89):

- (24) 这本书的读者  
*zhè běn shū de dúzhě*  
 this CLF book DET read-*zhě*  
 ‘The reader of this book’

- (25) 外国的读者  
*wàiguó de dúzhě*  
 foreign-country DET read-*zhe*  
 ‘(a) foreign reader’
- (26) 计算机的发明者  
*jìsuànjī de fā míngzhě*  
 computer DET invent-*zhe*  
 ‘the inventor of the computer’
- (27) ??外国的发明者  
*wàiguó de fā míngzhě*  
 foreign-country DET invent-*zhe*  
 ‘foreign inventor’

According to Dong X., the reason for the oddity of (27) is that 外國的 *wàiguó de* ‘foreign’ is not the internal argument of 發明 *fā míng* ‘to invent’, whereas in (26) 計算機 *jìsuànjī* ‘computer’ satisfies the valence requirement of such verb; thus, it clearly appears that the argumental structure of the verb inside 發明者 *fā míngzhě* is still relevant and, therefore, such item is syntactic in nature. The analogous examples in (24) and (25), however, are both acceptable, and this is because 讀者 *dúzhě* is fully ‘lexical’ and, thus, the valence of the verb 讀 *dú* ‘to read’ is not ‘visible’ to syntax.

Incidentally, the form which is most ‘lexical’ in the comparison above is the disyllabic one, 讀者 *dúzhě* ‘reader’. Dong X. (2004:87-89) suggests that prosody plays a role in the perception of such forms and in their acceptance as lexical items: trisyllabic (2 + 1) structures are prosodically acceptable (3.2.1.2) and, thus, may be easily perceived as words by the language users. If a trisyllabic -者 *-zhě* complex word contains a disyllabic verb-object compound, as e.g. 納稅者 *nàshuìzhě* ‘taxpayer’ (lit. ‘pay-tax-*zhe*’), it may be accepted as a lexical item (be it a compound or a derived word); if, otherwise, the valency of the verb is not

satisfied inside the compound, as in 發明者 *fāmíngzhě* ‘inventor’, the scope of 者 *zhě* includes everything that preceds, including the possible object (26-27) and, hence, we may see that 者 *zhě* operates at the syntactic level, even though 發明者 *fāmíngzhě* is a well-formed prosodic word. In such cases, the semantic (28) and prosodic (29) structure do not overlap (Dong X. 2004:88; see also *ivi*, fn. 1):

(28) [[计算机的发明] 者]  
[[*jīsuànjī de fāmíng*] *zhě*]

(29) [[计算机] 的 [发明者]]  
[[*jīsuànjī*] *de* [*fāmíngzhě*]]

Such inconsistency is interpreted by Dong X. as an indicator of the ‘hybrid’ status of 者 *zhě* between a clitic particle and a word formation affix.

In Guo L. (1983) it is claimed that 者 *zhě* is most productive with polysyllabic bases, whereas its combination with monosyllabic items suffers from arbitrary gaps; according to Dong X. (2004:88), this means that in complex words with a monosyllabic left-hand constituent 者 *zhě* is more morphologised, since gaps are typical of word formation patterns, whereas syntactic rules should exhibit full productivity, as mentioned above. With a quantitative analysis of -者 *-zhě* complex words in the reverse lexicon of Modern Chinese NXCD (2005), we found out that the majority of items have a disyllabic base (55, accounting for 59.1% of the total), much more than items with a monosyllabic base (19, or 20.45%) and with a base of more than two syllables (also 19); also, no disyllabic base belongs to the adjectival class, and no monosyllabic base is (exclusively) a noun. We may also remark that -者 *-zhě* complex words with a (monosyllabic) adjectival base have mostly been inherited from the Classical language: for instance, 長者 *zhǎngzhě* is attested with the

meaning ‘senior, elder’ in the ‘Mencius’ (孟子 *Mèngzǐ*, IIIrd cent. BCE) and as ‘virtuous man’ in the 史記 *Shǐjì* (Ist cent. BCE).

The data presented here suggest that a formal and semantic evolution of 者 *zhě* into an agentive suffix has occurred, and such development is more clearly visible in disyllabic words. The -者 *-zhě* complex words which we found in the reverse lexicon NXCD seem to conform to one of these three word formation schemas:

(30) [[X]V *zhě*]N ‘agent of verb X, X-er’ (no restrictions on the size of X)

(31) [[X]N *zhě*]N ‘person doing X, having opinions proper of X or possessing the characteristic X (X is never monosyllabic)’

(32) [[X]ADJ *zhě*]N ‘X person (X is always monosyllabic)’

Words as 編者 *biānzhě* ‘editor, compiler’, 記者 *jìzhě* ‘journalist’ or 侵略者 *qīnlüèzhě* ‘invader’ are the product of the schema in (30); items like 自由職業者 *zìyóuzhíyèzhě* ‘professional’, 共產主義者 *gòngchǎnzhuyìzhě* ‘communist’ or 無產者 *wúchǎnzhě* ‘proletarian’ conform to (31); words as 老者 *lǎozhě* ‘old person / people’ (attested in Confucius’ ‘Analects’, Vth cent. BCE) are the product of (32). If we took into consideration only the semantic aspect, the schema in (31), with a little adaptation, would suffice to accommodate all of the -者 *-zhě* complex words considered here (compare the treatment of -吧 *-bā*, 18, 3.2.3); however, the restrictions as to the size of the determiner are different for each schema, and this is something that cannot be ignored, if a construction is a combination of meaning, function and form (Michaelis & Lambrecht 1996, Goldberg 2006).

Let us reconsider the data on -者 *-zhě* complex words. Before we said that items with an adjectival base in our NXCD sample mostly seem to have a rather long history and, also, they are invariably disyllabic (i.e. the base is monosyllabic); however, with a cursory Google search we found items as 聰明者 *cóngmíngzhě* ‘intelligent (person)’, 漂亮者

*piàoliangzhě* ‘beautiful (person)’, 單純者 *dānchúnzhě* ‘simple, naive (person)’. None of those words were found in standard dictionaries of Mandarin and this, according to Packard, is indicative of high productivity, which makes it difficult to list exhaustively all -者 *-zhě* complex words (2000:73; see 2.2.2). We suggest two other possible interpretations for such data: the fact that items with a disyllabic adjectival base are occasionalisms, and thus are not registered by lexicographers; another explanation is that their being fully transparent makes their inclusion in dictionaries superfluous. However, the high number of Google hits for items as 聰明者 *cóngmíngzhě* and 漂亮者 *piàoliangzhě* tells us that they should not be occasionalisms; also, words as 剝削者 *bōxiūzhě* ‘exploiter’ and 失業者 *shīyèzhě* ‘unemployed (person)’ are also wholly transparent, and yet they are listed in dictionaries. In addition, we may remark that full productivity is typical of syntactic rules (and, possibly, of inflectional morphology), rather than of derivation, as pointed out before. Thus, it seems that the restrictions on the size of -者 *-zhě* complex words with an adjectival base are not strict and that the tendency for ‘new’ words is to have a plurisyllabic base, as noted by Guo L. (1983), and this is part of a general trend in Modern Mandarin word formation, as seen for -學 *-xué* ‘branch of learning’, -吧 *-bā* and -化 *-huà* ‘-ise, -ify’ above. Hence, a schema as

- (33) [[X] N/V/ADJ *zhě*]<sub>N</sub> ‘person doing X, having opinions proper of X or possessing the characteristic X’

Covers all the instances of -者 *-zhě* in its ‘affixal’ meaning; restrictions on the size of the variable (X) are to be seen in a diachronic perspective, with monosyllabic bases being more ‘classical’ and plurisyllabic bases being more ‘modern’.

Before closing this section, we shall deal briefly with the case of -者 *-zhě* complex words containing a transitive verb and its object, as 愛國者 *àiguózhě* ‘patriot’, lit. ‘love-country-*zhě*’ (ex. 22c, CHAPTER 1). He Y. (2004) points out that in Mandarin both ‘V-OBJ-*zhě*’ and ‘OBJ-V-*zhě*’

structures are attested; however, if the V and the OBJ are monosyllabic, only ‘OBJ-V-*zhě*’ is possible, with a few exceptions (as 肉食者 *ròushízhě* ‘meat eater’, also ‘high-ranking government official’, lit. ‘meat-eat-*zhě*’). Moreover, the morphological collective marker -們 *-men* may attach only to ‘OBJ-V-*zhě*’ structures (He Y.2004:3):

- (34) a. 謠言製造者們  
*yáoyán-zhìzào-zhě-men*  
 rumour-fabricate-*zhě*-COLL
- b. \*製造謠言者們  
*zhìzào-yáoyán-zhě-men*  
 fabricate-rumour-*zhě*-COLL  
 ‘rumour-mongers’

On the other hand, ‘V-OBJ-*zhě*’ structures may be modified by an adjunct, whereas ‘OBJ-V-*zhě*’ may not (He Y.2004:3):

- (35) a. 積極製造謠言者  
*jījí-zhìzào-yáoyán-zhě*  
 active-fabricate-rumour-*zhě*
- b. \*積極謠言製造者  
*jījí-yáoyán-zhìzào-zhě*  
 active-rumour-fabricate-*zhě*  
 ‘wild rumour-monger’

According to He Y., this is because ‘OBJ-V’ structures are ‘true’ endocentric compounds, and thus ‘OBJ-V-*zhě*’ structures are actually ‘lexical’ compounds, with the same features of other compounds in the language; ‘V-OBJ’ structures are verb phrases, conforming to the syntactic order of verb and object, and thus they can be modified by an adverbial adjunct, as in (35a). In ‘V-OBJ-*zhě*’, the base is a stem which is “looped back” from syntax (Pinker 1999, qtd. in He Y. 2004). When an item is generated in syntax and then ‘looped back’ to word formation as a stem, it



may show ‘inappropriate’ behaviour, such as failure to inflect, as e.g. Eng. ‘cuts package’ → ?\*‘cuts packages’ (Pinker 1999)<sup>25</sup>; this explains the anomalous behaviour of items as 製造謠言者們 *zhìzàoyáo yánzhě* ‘rumour monger’ (34b, 35a).

To sum up, our data apparently confirm that -者 *-zhě* appears to have an ambiguous status, in that it combines with lexical items, but also with phrases and with ‘hybrid’ structures, as seen above (34b, 35a). Here, Lehmann’s parameter of structural scope reduction seems to overlap with a principle of lexicalist morphology, i.e. the ‘Lexical Integrity Hypothesis’ (see Lieber & Scalise 2006 for an overview of the different versions of such principle and for the related criticism); if it is true, generally speaking, that a word formation element should not combine with a syntactic item (‘No-Phrase Constraint’), we have several counterexamples to this principle, as e.g. phrasal compounds (Eng. ‘floor of a birthcage taste’) and, also, we see some degree of variation which is not necessarily analysable as a transitional state towards scope reduction. See the examples below (from Lieber & Scalise 2006:9-12):

(36) a. self-sufficient-ish

b. New Years Day-ish

(37) post digestive disorder complications

In (36a-b), the English derivational suffix ‘-ish’ attaches to a phrasal base; in (37), the prefix ‘post-’ scopes over the phrase ‘digestive disorder’. If for ‘-ish’ a degrammaticalization analysis might be proposed (compare *ism, ex*; Ramat 1992, 2001)<sup>26</sup>, for ‘post-’ it is more likely that the ‘phrasal’ usage is an English innovation, rather than a ‘vestige’ of the distribution of the Latin preposition. In the case of -者 *-zhě*, it seems that residual and

<sup>25</sup> Pinker regards ‘cuts packages’ as ungrammatical; however, with a cursory Google search we found several instances of such form and, thus, we added a question mark to the asterisk.

<sup>26</sup> According to Spencer (2005, qtd. in Lieber & Scalise 2006): “(...) for some speakers *ish* has become a free morpheme with roughly the meaning ‘approximately’”.

innovative uses coexist in Modern Mandarin; however, from the semantic point of view, there seems to be no real difference between ‘syntactic’ and ‘affixal’ uses of -者 *-zhě*. We shall get back later to the parallelism between 者 *zhě* as a function word and clitics in the Indo-European languages of Europe; let us now deal with another ‘semi-free morpheme’ of Mandarin, namely -式 *-shì* ‘style, model’.

According to Dong X. (2004), both -者 *-zhě* and -式 *-shì* ‘style, model’ are semi-free morphemes in Modern Chinese, i.e. as items which operate both as affixes in word formation and as particles / clitics in syntax; however, the two items seem to have undergone a rather different evolution, as we shall see. In the Classical language, -式 *-shì* had several meanings, and three of those seem to be connected with the usage at issue here, namely (GHYDCD 2000):

- a. ‘style, form’ (式样 *shìyàng*, 格式 *géshì*);
- b. ‘example, model’ (榜样 *bǎngyàng*; 模范 *mófàn*);
- c. ‘method, norms’ (法式 *fǎshì*; 规格 *guīgé*).

In a reference dictionary of Modern Chinese (CCD 2002), words as 新式 *xīnshì* ‘new-style’, 西式 *xīshì* ‘Western style’ or 舊式 *jiùshì* are listed under the first among those meanings, ‘style, form’. Zhang Yi. (2002a:189, 2002b:96) proposes a distinction between the usage of -式 *-shì* as a ‘root’ and as an item attaching to ‘affixed words’ (附綴式單詞 *fùzhuìshì dāncí*) and to units larger than a word (超詞單位 *chāocí dānwèi*).

The first group includes words as 模式 *móshì* ‘model, pattern’, in which 模 *mó* is near-synonymous with 式 *shì*, or 等式 *děngshì* ‘equality’ (in mathematics; 等 *děng* here means ‘to be equal’); these, may all be regarded as complex words (mostly, nouns), which form a rather closed set, and generally they either contain a morpheme which is virtually

synonymous with 式 *shì*, or, if the relationship is modificational, 式 *shì* conveys a meaning different from ‘style, form’, as e.g. 開幕式 *kāimùshì*, in which 式 *shì* stands for ‘ceremony, ritual’.

The second group includes items in which 式 *shì* bears the meaning ‘style, model’, as 西式 *xīshì* ‘Western style’ quoted above, sometimes with a phrasal constituent, as in the example below (from Zhang Yi. 2002b:97):

- (38) 并非做文字游戏式的插科打诨  
*bìngfēi zuò wénzì yóuxì-shì de chākēdǎhùn*  
 be.really.not make writing game-*shì* DET wisecrack  
 ‘not really wisecracking (using puns)’

In (38), -式 *-shì* is attached to the phrase 做文字遊戲 *zuò wénzì yóuxì* ‘to make puns’. Only this second group, obviously, is relevant for the purposes of our research.

Zhang Yi. (2002a, 2002b) proposes an analysis of -式 *-shì* structures taking into consideration their distribution and the (sub-)type of modificational relation between the ‘base’ and the formative at issue. As a matter of fact, -式 *-shì* structures are not consistent in terms of word class; they resemble mostly ‘non-predicative adjectives’, i.e. noun modifiers which may not be stand-alone predicates, differently from ‘standard’ (i.e. predicative) Mandarin adjectives. However, -式 *-shì* structures differ from non-predicative adjectives in that they cannot be negated with 非 *fēi*-; also, many among these structures are actually nominalised, and may be preceded by quantifiers. Typically, -式 *-shì* structures act as attributive modifiers (定語 *dìngyǔ*, Zhang Yi. 2002b:98); as to semantics, the modification relationship between the ‘base’ and -式 *-shì* may be either of the ‘descriptive-explicative’ type (39) and of the ‘metaphoric-analogic’ type (40). Each type may be further divided into subtypes ‘a’ (if the noun they modify is concrete) and ‘b’ (if the noun they modify is abstract); below are some examples for each subtype (Zhang Yi. 2002b:101-104,

2002a:200-208):

- (39) a. 西班牙式斗牛士  
*xībānyáshì dòuniúshì*  
 Spain-*shì* bullfighter  
 ‘Spanish-style bullfighter’
- b. 外交式的幽默  
*wàijiāoshì de yōumò*  
 diplomacy-*shì* DET humour  
 ‘diplomatic humour’
- (40) a. 模特儿式的三围  
*mótershì de sānwéi*  
 model-*shì* DET three-measurements.of.a.woman  
 ‘model-like measures’
- b. 奴隶式的盲从  
*núlìshì de mángcóng*  
 slave-*shì* DET blind-follow  
 ‘slavish blind following’

The reason for such distinction, in Zhang Yi.’s analysis (2002b:104), is that these four (sub-)groupings are interpreted as a continuum of abstraction, from (39a) to (40b), in which the pathway of evolution of 式 *shì* from root to ‘metaphoric auxiliary pseudo-particle’ (準比況助詞 *zhǔnbǐkuàng zhùcí*) should be visible:

- (41) Root > pseudo-affix > clitic  
 > metaphoric auxiliary pseudo-particle

Thus, according to Zhang Yi.’s treatment, we are dealing with degrammaticalization (or, better, ‘transcategorization’; Ramat 2001:397), since a pseudo-affix is said to evolve into a clitic and, then, into a pseudo-particle, rather than the other way round. Also, it appears that

Zhang Yi. understands the different kinds of relation between an -式 *-shì* structure and its modifier correspond to a different degree of semantic abstraction. Some remarks are needed on those two points.

We already mentioned that we have instances of ‘recognised’ affixes scoping over a phrase also in English (exx. 36-37). We may suggest that -式 *-shì* is not fundamentally dissimilar from Eng. *-ish*, which seemingly acts as a ‘semi-free morpheme’ (*mutatis mutandis*) in the language; thus, for -式 *-shì*, we are dealing with an instance of apparent degrammaticalization, rather than with the ‘resurfacing’ of some historical usage, differently from -者 *-zhě*. Incidentally, we shall remark that both *-ish* and -式 *-shì* words with a phrasal constituent seem to be mostly occasionalisms. Zhang Yi. (2002b:98) believes that there is a connection between the number of syllables of the expression and its ‘lexicality’, in that disyllabic -式 *-shì* words (i.e. those with a monosyllabic ‘base’) become ‘standard words’ (定性詞 *dìngxìngcí*) more readily, whereas ‘larger’ words have an occasional nature; also, ‘standard words’ have a ‘morpheme + affix’ structures, whereas ‘occasionalisms’ have a ‘word/phrase + clitic’ structure. According to Zhang Yi., given the difficulty in distinguishing between morpheme, words and phrase in Chinese, the distinction between ‘standard words’ and hapax legomena is not always clear and, very often, an item is perceived as belonging to the former or to the latter group because of its size:

- |   |  |   |
|---|--|---|
| (42) 英式<br><i>yīngshì</i><br>‘English-style’  | 法式<br><i>fǎshì</i><br>‘French-style’         | 德式<br><i>déshì</i><br>‘German-style’        |
| (43) 法国式<br><i>fāguóshì</i><br>‘French-style’ | 比利时式<br><i>bǐlìshìshì</i><br>‘Belgian-style’ | 意大利式<br><i>yìdàlishì</i><br>‘Italian-style’ |

The expressions in (42), according to Zhang Yi. (2002b:98), are perceived as ‘lexical words’ (詞彙詞 *cíhuìcí*), whereas those in (43) are

regarded as nonce ‘grammatical’ words (臨時性的語法詞 *línshíxìng de yǔfǎcí*), or phrases, mainly because of the number of syllables: compare 法式 *fǎshì* and 法國式 *fǎguóshì*, both meaning ‘French-style’. Such distinction, however, is not discrete, since there is a continuum between word and phrase, in Zhang Yi.’s model.

We believe that the distinction between items as those in (42) and those in (43) is of little theoretical significance, and apparently describes only some sort of conception of the language users, although we do believe that structural size plays a role in the acceptance of an expression as a word; what matters most, in our perspective, is the distinction between expressions with -式 *-shì* based on a root or on a free word, and those with a phrasal base (see ex. 38). As to the supposed occasional nature of the expressions with a phrasal base, we must take into account two general problems of lexicography. Firstly, as pointed out by Gaeta & Ricca (2003:64),

“for commercial and practical reasons, they [dictionaries] do not aim at the comprehensive documentation of productively-formed, transparent forms, but rather cover the more frequent and idiosyncratic terms (...). Moreover, even when aiming at complete coverage, lexicographers often overlook new, regular formations, just because they are regular. This is especially true for those word formation processes whose semantic content is not particularly profiled, such as action nouns, quality nouns, relational adjectives, etc.”

If we consider -式 *-shì* complex ‘words’ as those in 42 and 43, following the pattern ‘nation / country + -式 *-shì*’, it will appear as obvious that a language user could build such a term for any nation, when needed (傣式 *dǎishì* ‘Dai-style’, 維吾爾式 *wéiwú'ěrshì* ‘Uyghur-style’<sup>27</sup>, exx. from the web); hence, it is not convenient to list all of those forms in a dictionary, and any speaker can produce and understand them easily. When -式 *-shì* attaches to a phrasal constituent, we may reasonably suppose that such forms are produced for special pragmatic or rethoric reasons, and “complex words with primarily syntactic and/or pragmatic

<sup>27</sup> Dai people and Uyghurs are two ethnic minorities of the P.R.C.

functions are more likely to escape attention than words that require substantive semantic processing in the mental lexicon” (Renouf & Baayen 1998:188); such considerations hold also for -者 *-zhě*.

We may also remark that items as ‘self-sufficient-ish’ (36a) or ‘seven/eight-ish (years old)’ (with different spellings) are not listed in general dictionaries, but are actually not that uncommon (examples from the web):

- (44) Self sufficient 'ish'.com - The urban guide to almost self sufficiency<sup>28</sup>.
- (45) Last night MJ said why don't we buy a plot of land, live in a caravan and be self sufficient-ish<sup>29</sup>.
- (46) When I was young, about seven/eighthish, I lived in a house where there were three or four loose stone steps<sup>30</sup>.
- (47) i [sic!] think we should start off real early like half seven eightish (...)<sup>31</sup>.

Note that in (47) the adjective ‘half’ is used before ‘seven’, obtaining an unusual combination in English. Here *-ish* seem to bear the meaning ‘approximative’, ‘approximatively’ (see fn. 26); such function is similar to that of ‘mock-’, ‘-type’ (see ex. 15-16, CHAPTER 1) and ‘-shape’ / ‘-shaped’ as bound constituents (Renouf & Baayen 1998:188):

“While the affixes themselves are clearly well-established in journalistic prose, their function seems to be more pragmatic in nature than lexical. The affixes *-type* and *-shaped* give writers and speakers the flexibility to express approximation of class membership in a dense morphological form instead of using syntactic periphrastic constructions. They are markers of more informal styles”

---

<sup>28</sup> From [www.selfsufficientish.com](http://www.selfsufficientish.com).

<sup>29</sup> From [www.goal2010.org/category/lifestyle/be-self-sufficient-ish](http://www.goal2010.org/category/lifestyle/be-self-sufficient-ish).

<sup>30</sup> From [http://www.experienceproject.com/group\\_stories.php?g=10308&s=d&sn=10](http://www.experienceproject.com/group_stories.php?g=10308&s=d&sn=10).

<sup>31</sup> From <http://www.bebo.com/Profile.jsp?MemberId=383091113>.

Such pragmatic phenomena are in line with the tendencies in processes of degrammaticalization pointed out by Ramat (2001:397):

“The causes for the existence of degrammaticalization processes are to be sought in the overall tendency to use labels as economic symbols; in such vivid locutions each meaning-bearing element (e.g. suffix morphemes like *-ade*) can be separately manipulated: the label *ade* is more economic than “fruit juice”, though not as transparent (see *bus* in *schoolbus*, *autobus*, etc., as an hyperonym for “public transportation vehicle”). (...) If we had just unidirectional evolution toward grammar we would expect that languages become more and more grammaticalized, which by all evidence is not the case”

In short, “[s]ymbolism and iconism are the contrasting strategies always at work and always in tension in language and thus in linguistic evolution” (Ramat 1992:557); opposite principles may well explain the existence of converse phenomena, as the degrammaticalization / transcategorization of ‘-ish’ and the grammaticalization / morphologization of ‘-type’, ‘-shape’ / ‘-shaped’, etc.

As to the progressive ‘abstraction’ of meaning suggested by Zhang Yi. (2002a, 2002b) in the four subgroups of -式 *-shì* expressions, we believe that there is no fundamental difference in the degree of meaning abstraction among the groups. The fact that Zhang Yi. associates abstraction with a process of degrammaticalization is rather odd; moreover, Zhang Yi. himself argues that the actual relation between -式 *-shì* structures and the modified constituent may differ according to the individual modified item and to the context (2002b:103):

- (48) 中国式的女革命家  
*Zhōngguóshì de nǚgémingjiā*  
 China-*shì* DET female-revolutionary  
 ‘Chinese-style female revolutionary’

- (49) 中国式的社会主义  
*Zhōngguóshì de shèhuìzhǔyì*  
 China-*shì* DET socialism  
 ‘Chinese-style socialism’



- (50) 中国式的家庭  
*Zhōngguóshì de jiātíng*  
 China-*shì* DET family  
 ‘Chinese-style family’
- (51) 中国式的人海  
*Zhōngguóshì de rénhǎi*  
 China-*shì* DET people-sea  
 ‘Chinese-style sea of people’

We shall not go into the details here, and we shall just remark that Zhang Yi believes that each of the four examples belongs to a different group, even though the word at issue is 中國式 *Zhōngguóshì* ‘Chinese-style’ for all of them.

Judging from the data, it appears that the Old Chinese lexeme 式 *shì* has become ‘productive’ in one of its meaning, namely ‘style, model’, attaching to lexical bases; moreover, it seems that degrammaticalization / transcategorization has occurred, and -式 -*shì* is used in Modern Mandarin also as a function word, resembling a clitic, attaching to phrasal bases, most likely for pragmatic reasons. However, we doubt that -式 -*shì* may be regarded as a derivational suffix, since its ‘affixal’ meaning is not fundamentally different from its ‘core’ historical lexical meaning. From such perspective, it is more appropriate to term -式 -*shì* as a class noun, rather than as an affix, differently from -者 -*zhě*.

As to the label ‘semi-free morphemes’ (as understood by Dong X. 2004), we believe that it has descriptive value, but, from a diachronic point of view, morphemes as 者 *zhě* and 式 *shì* had a rather different evolution. While 者 *zhě* was a demonstrative (among other functions; see above), 式 *shì* was a nominal lexeme which evolved into a suffix-like element, a class nouns; the latter has apparently been ‘transcategorised’ and is used also as a particle. A characteristic of these two formatives which moves them ‘closer’ to syntax is their high productivity and transparency in meaning.

Lastly, we shall move our attention towards the left-hand side of the complex word, i.e. to prefixation.

### 3.2.5 Two Models for Prefixation<sup>32</sup>

In the Mandarin lexicon, there are a number of bound word formatives which attach to the left side of words and, mostly, correspond to ‘European’ prefixes, as 非- *fēi*- ‘non, a-’ (非對稱 *fēiduìchèn* ‘asymmetric’), 前- *qián*- ‘former, pre-, ex-’ (前蘇聯 *qiánsūlián* ‘Former Soviet Union’) or 半- *bàn*- ‘half, semi-’ (半導體 *bàndǎotǐ* ‘semiconductor’), among others (see 3.1.2); just as (most) SAE prefixes, they apparently do not change the lexical class of the word they attach to. However, we also have prefixed items belonging to a small, closed set, as 可- *kě*- ‘can, -able’ (可變 *kěbiàn* ‘variable’) and 難 *nán* ‘difficult, unpleasant’ (難關 *nánguān* ‘difficult to govern’), which always bear adjectival class. Whereas class-maintaining prefixation appears to have been the standard for Chinese morphology since the pre-Qin times (i.e. before 221 BCE), those class-changing elements as 可- *kě*- and 難- *nán*- seem to be atypical and have a puzzling behaviour, especially as far as headedness is concerned, as has been first remarked by Ceccagno & Scalise (2006). Also, as we shall see, there are prefixed items which sometimes are class-maintaining and sometimes are class-changing, like 多- *duō*- ‘multi-’, as in 多音節 *duōyīnjié* ‘polysyllabic’.

In table 2.3 we presented 16 morphemes which have been classified as affixes (or affixoids) in the majority of works on Chinese morphology considered by Pan, Ye & Han (2004); among those, we find two prefixes, 反 *fǎn*- ‘anti-, counter-’ (反帝國主義 *fǎndìguózhǔyì* ‘anti-imperialism’) and 老- *lǎo*- ‘old’, often used before surnames (see fn. 15, CHAPTER 1). Some other commonly cited prefixes are 第 *dì*, used to build cardinal numbers, 小- *xiǎo*- ‘young’, understood as the opposite of 老- *lǎo*- and used before the surnames of people who are younger or about the same

<sup>32</sup> This section is mainly based on a talk delivered at the 6th Conference of the European Association of Chinese Linguistics, “A diachronic outlook on prefix-like elements in Chinese word formation” (Poznań, Poland, 26-28 August 2009).

age as the utterer, and 阿- *ā-*, also added to personal names or kinship terms (阿媽 *āmā*, ‘mommy’; see the data in Xu & Cai 2007, Yang Y. 2007). The fate of prefixes, still, has been slightly better than that of suffixes, since those morphs which have been regarded as “true” suffixes are often those which have no meaning or only categorial meaning, such as the often-quoted ‘dummy’ nominal suffixes -子 *-zi*, -儿 *-er* and -头 *-tou*, also because they have undergone some sort of phonological (suprasegmental, actually) reduction, being now toneless (2.2.1). On the other hand, all of the recognised prefixes carry meaning, and the very same morphs are found with other usages: 反- *fǎn-* has the same shape in 反革命 *fǎngémìng* ‘counterrevolutionary’ as in 相反 *xiāngfǎn* ‘contrary’, although it is obviously not a prefix in the latter. Criteria for the identification of prefixes, generally speaking, seem to be even vaguer than those employed for suffixes (2.2.1, 2.2.2).

This, however, is not only a Chinese problem: the recognition of the existence of prefixes came much later than that of suffixes in the Western linguistic tradition, as illustrated in Montermini (2008:13 ff.). Whereas the labels ‘prefix’ and ‘suffix’ entered the vocabulary of most European languages at the end of the XIX century, for a long time after that prefixing was regarded as a special kind of compounding, and suffixes only were assigned to derivation. This was because, among other reasons, many present-day prefixes in *Standard Average European* languages were actually prepositions or adverbs in Latin and Ancient Greek; also, the phonological form of many prefixes is identical to that of prepositions: this is the case, for instance, of the Italian prefix *con-* ‘con- / com-’ as in *connazionale* ‘compatriot’, having the same shape of the preposition *con* ‘with’. This happens also with Chinese prefixes / prefixoids, as with the 反- *fǎn-* example or with 高- *gāo-* ‘high degree of’ (3.2.2), which still have lexical usages. The fact that prefixes have been regarded as “a doubly ‘marginal’ phenomenon, (...) the special case of a bigger phenomenon” (Montermini 2008:9; my translation), i.e. either derivation or compounding, has actually many more reasons, which have to do both with universal tendencies in word formation and with a Indo-European bias in the mind of many linguists which have dealt with the problem; here

we shall sum up only some major points. First of all, there is a well-known typological tendency to prefer suffixes to prefixes, i.e. the number of suffixes is bigger than that of prefixes in the languages of the world: in the *World Atlas of Language Structures* (Haspelmath *et al.* 2005), out of a sample of 894 languages, only 148 are dominantly prefixing, whereas 496 are dominantly suffixing, even though this count is limited to inflectional morphology (see the table in Montermini 2008:51). The psycholinguistic argument for this preference is that the brain processes words from the beginning to the end and, therefore, the beginning of a word is much more relevant than the end for recognition; the most relevant element, the lexical morph, is then placed first (see Stump 2001:708-10, Montermini 2008:52). Also, since the formulation of the well-known “Righthand Head Rule” by Williams (1981), it has been believed that suffixes only could determine the lexical category of the whole word, even if later it was accepted that sometimes even a prefix could bear a word-class, as in the English *en-* deadjectival / denominal verbs (e.g. *ennoble*). This, however, was deemed to be just an exception to a firm rule: the head, and the lexical category of the word, are borne by the rightmost element. This distinction in terms of categorizing force, which makes prefixes somehow “weaker” than suffixes, suffers from an Indo-European bias and, anyway, is contradicted even by data from SAE languages (see the discussion in Montermini 2008:185 ff.). Class-changing prefixation is well documented in many languages: see the many examples of denominal prefixed verbs in Afro-Asiatic and Austronesian languages (Montermini 2008:211).

Going back to Mandarin, in table 3.5 we shall quote data on the treatment of prefixes in some major works of Chinese word formation, taken and abridged from Xu & Cai (2007:133) and Yang Y. (2007:52; my translations):

Table 3.5. Prefixes, prefixoids and related categories in nine sources (Xu &amp; Cai 2007, Yang Y. 2007).

Source	Typology	Examples
Chao Y. (1968)	‘narrowly-defined’ prefixes	老- <i>lǎo</i> - ‘old’, 第- <i>dì</i> - ‘-th’
	‘new and developing’ prefixes	準- <i>zhǔn</i> - ‘quasi-’ 非- <i>fēi</i> - ‘non-, a-’
	‘versatile’ prefixes	可- <i>kě</i> - ‘-able’ 難- <i>nán</i> - ‘difficult’
Lü S. (1979)	prefixes	老- <i>lǎo</i> - ‘old’
	prefixoids	準- <i>zhǔn</i> - ‘quasi-’
Ren X. (1981)	prefixes	反- <i>fǎn</i> - ‘anti-, counter’
	quasi-prefixes	準- <i>zhǔn</i> - ‘quasi-’
Ma Q. (1995)	prefixes	老- <i>lǎo</i> - ‘old’
	quasi-prefixes	準- <i>zhǔn</i> - ‘quasi-’
Guo L. (1983)	‘new and developing’ prefixes	多- <i>duō</i> - ‘multi-’ 非- <i>fēi</i> - ‘non-, a-’
Zhang Bi. (2002)	prefixes	老- <i>lǎo</i> - ‘old’
	prefixoids	準- <i>zhǔn</i> - ‘quasi-’
XHCD (2002)	prefixes	老- <i>lǎo</i> - ‘old’

Wang H. (2002)	prefixoids	準- <i>zhǔn</i> - ‘quasi-’
XHCD (2005)	prefixes	老- <i>lǎo</i> - ‘old’, 非- <i>fēi</i> - ‘non-, a-’

The total number of prefixes and prefixoids in each work varies considerably, ranging from 22 in Lü S. (1979) to three only in the 2002 edition of the dictionary 現代漢語詞典 *Xiàndài Hànyǔ Cídiǎn* (XHCD 2002). With the exception of Ren X. (1981), all of these works regard only those few forms which we quoted above as 阿- *ā*-, 老- *lǎo*-, 第- *dì*- and 小- *xiǎo*- as “true” prefixes, whereas those forms as 非- *fēi*- ‘non-, a-’ are normally treated as quasi-prefixes or anyway as something non-canonical. Here we see, again, that the notion of ‘emptying’ (虛化 *xūhuà*) is crucial: in its radical version, the kind of semantic bleaching that goes together with grammaticalization of a lexical morpheme into an affix is seen as total loss of lexical meaning (see 2.1.2, 2.2.2). Cross-linguistic evidence and, indeed, common sense, suggest that prefixes (just as suffixes) should express some sort of meaning, although arguably not as rich in intension as the kind of meaning of typical content words (nouns, verbs, adjectives), or, at least, bear a word-class.

Let us first present and analyse data on 非- *fēi*- ‘non-, a-’, a representative member of the first category of prefix-like formatives. As mentioned at the beginning of this section, class-maintaining prefixes (as 有- *yǒu*-, added before names of ethnic groups, and 老- *lǎo*-, ‘old, respectful term’) are attested since pre-Qin Chinese (Cheng X. 1992a, 1992b, 1992c); 非- *fēi*- looks like a clear instance of a categorially transparent prefix, and it may be added mainly to nouns and to non-predicative adjectives (3.2.4):

(52) 非动物	非导体	非官方
<i>fēidòngwù</i>	<i>fēidǎotǐ</i>	<i>fēiguānfāng</i>
<i>fēi</i> -move-thing	<i>fēi</i> -conductor	<i>fēi</i> -official
‘inanimate object’	‘non-conductor’	‘unofficial’

In all of those cases, the base word retains its own word class after prefixation. The morpheme 非 *fēi* as a (broadly defined) negator is attested since the stage of Old Chinese, both as a free morpheme and as a word formation element; it was used both as a verb and as an adverb (Dong X. 2002:260-261, Hong B. 2005:107; see also Guo L. 1983, Shen M. 1986). The subtypes of negation which 非 *fēi* could express in the Classical language are listed below (GHYDCD 2000):

- a. negative copula (不是 *bù shì*);
- b. negator for verbs (不 *bù*)<sup>33</sup>;
- c. negator of existence (无 *wú*).

According to lexicographers (see CCD 2002), the modern ‘prefixal’ usage is connected with its historical function as a negative copula, as in the example below (Confucius’ ‘Analects’, Vth cent. BCE, qtd. in Ōta 1987:276):

(53) 我非生而知之者
<i>wǒ fēi shēng ér zhī zhī zhě</i>
1SG NEG.COP born and know DEM NMLZ
‘I was not born learned’

In Classical Chinese, 非 *fēi* was also used as a word formative, as mentioned above, e.g. in 非常 *fēicháng* ‘extraordinary’ (attested in the 史記 *shǐjì* ‘Records of the Grand Historian’, 1st cent. BCE); however, Cheng

---

33. Nella lingua moderna, 不 *bù* funge da negatore per *unbounded elements*, mentre i *bounded elements* vengono negati da 没 *méi*; nell’accezione b, quindi, “negazione” andrà inteso nel primo senso (Shi Y. 2002:201-202).

X. (1992a-d), in his surveys of Chinese morphology since the pre-Qin times to the Tang dynasty, does not list 非 *fēi* either among prefixes or among grammatical words. The morpheme 非 *fēi* in Classical Chinese had also other usages as the first member of a complex word, expressing different subtypes of negations, but the synchronically productive word formation patterns are (54) and (55) only:

- (54) [*fēi* [X]<sub>N</sub>]<sub>N</sub> ‘non X’ →  
非暴力 *fēibàoli* ‘nonviolence’
- (55) [*fēi* [X]<sub>ADJ</sub>]<sub>ADJ</sub><sup>34</sup> ‘lacking the property X’ →  
非常任 *fēichángren* ‘nonpermanent’;

Also, 非 *fēi* has lost its free status in Modern Mandarin Chinese, although it can still be used “freely” in set expressions like 非...不可 *fēi...bùkě* ‘must, will inevitably’ (on such patterns, see Dong Z. 2006). In the contemporary language, 非- *fēi-* is often regarded as a ‘new and developing prefix’ or as a ‘prefixoid’ / ‘quasi-prefix’ (table 3.5; see also Dong X. 2002:260-261). Shen M. (1986:93; 1995:36) analyses 非- *fēi-* as a developing prefix; according to him, the ‘emptying’ (i.e. abstraction) in meaning of an item is proved by the increase in the number of possible ‘bases’, connected with a semantic ‘generalization’ / ‘broadening’, as seen before for 高 *gāo* in words like 高蛋白 *gāodànbái* ‘high protein’ (3.2.2; see Shen M. 1995:36).

Judging from the available data, it seems to us that Mandarin the word formation schema(s) underlying 非- *fēi-* complex words is the evolution of a pattern which was already existent in Old Chinese (see the 非常 *fēicháng* ‘extraordinary’ example above). In the 非- *fēi-* complex words built according to the schemas represented above (54-55), the function of 非- *fēi-* is that of negator for nouns and non-predicative adjectives (i.e. attributive-only items), which is the same function of the negated copula

---

<sup>34</sup> In this section, we shall employ the part of speech tag ‘ADJ’ improperly to indicate non-predicative adjectives.



(不是 *bù shì*) in Modern Chinese. We searched for lemmas with 非 *fēi* as the first constituent in the *Wenlin* dictionary (WL 2007)<sup>35</sup>; among the items found, 55 are listed as nouns, 49 as attributive forms (i.e. non-predicative adjectives), 3 are listed with more than one word-class identity and 2 contain a phrasal constituent; also, we found a single predicative adjective, 非凡 *fēifān* ‘outstanding, extraordinary’ (already attested in the 後漢書 *Hòu Hàn Shū* ‘Book of the Later Han’, Vth century CE; HYDCD 1993). Besides, a small number of forms (two nouns and two verbs) not conforming to the schemas in (54) and (55) were found; in such forms, 非 *fēi* is not a negator but, rather, conveys the meaning ‘wrong, erroneous’, already attested in Old Chinese (漢書 *Hàn Shū* ‘Book of Han’, IIInd cent. CE), as in 非計 *fēijì* ‘ill-conceived plan’ and 非望 *fēiwàng* ‘wild hope’<sup>36</sup>. We thus disagree with Guo L. (1983) and Shen M. (1986, 1995), since it seems that there has been no increase in the number of possible ‘bases’ for 非 *fēi*; rather, almost all productively formed 非- *fēi*-complex words are based either on a noun or on a non-predicative adjective; from the semantic point of view, no generalization in meaning has occurred, and 非- *fēi*- has not become a generic negator, as proved by the existence of synchronic ‘competitors’ which convey other subtypes of negation: 無 *wú* ‘negator of existence’ (無軌 *wúguǐ* ‘trackless’), 不 *bù* ‘negator for events’ (不銹鋼 *bùxiùgāng* ‘inox, stainless steel’), 未 *wèi* ‘not yet’ (未婚 *wèihūn* ‘unmarried’). Also, it is not so clear whether 非- *fēi*- is actually always class-maintaining, since we have several examples as 非貿易 *fēimàoyì* ‘non-commercial’, which seems to conform to the schema in (56):

<sup>35</sup> We chose the *Wenlin* dictionary because it has a very large number of lemmas (about 196.000) and, also, the lexical category of each lemma is indicated. We excluded from the count idioms, since they are not representative of Modern Chinese word formation, and multi-word expressions in which the 非 *fēi* complex word acted as a modifier, because the word class label is provided only for the whole expression (and not for the modifier only).

<sup>36</sup> For the sake of completeness, we may remark that we found also a verbal form in which 非 *fēi* apparently acts as a negator, 非證 *fēizhèng* ‘disconfirm’ (also, ‘disconfirmation’).

(56) [*fěi* [X]<sub>N</sub>]<sub>ADJ</sub> ‘non X’

This issue, however, deserves further reflection on the status of non-predicative adjectives in Chinese; Zhang Bo. (1994) believes that non-predicative adjectives are to be seen as an intermediate step in the noun-verb continuum; Deng, Wang & Li (1996:238) claim that non-predicative adjectives, over time, tend to become closer to standard (predicative) adjectives. No matter what one’s views on transcategoriality in Chinese are, it is a fact that words such as 非貿易 *fēimàoyì* ‘non-commercial’ are born as non-predicative adjectives.

Another apparently “categorially transparent” prefixed morpheme in Chinese is 零- *líng-* ‘zero’, which most likely entered the lexicon through analogy with some English complex words like *zero risk*, *zero emission* and the like; its ‘prefixed’ usage is clearly distinct from its identity as the numeral ‘zero’. According to Cheng L. (2004)’s analysis, 零 *líng* has been grammaticalized as a prefixoid since it has developed, in its prefixed use, a different meaning from that which it conveys in its core lexical usage:

(57) [*líng* [X]<sub>X</sub>]<sub>X</sub> ‘X starting from zero’ →  
零突破 *língtūpò* ‘zero breakthrough’

However, the status of 零- *líng-* as a categorially transparent prefix seems to be contradicted by examples like that in (58), where a noun is apparently turned into a non-predicative adjective, an attributive form:

(58) [*líng* [X]<sub>N</sub>]<sub>ADJ</sub> ‘lacking X’ → 零風險 *língfēngxiǎn* ‘zero risk’

The situation is parallel to that of 多- *duō-* which, as seen before (2.2.2), has a different behaviour when used as an adjective (59a-b, 60) and as a prefixed constituent (61):

(59) a. 她認識很多外國人  
*tā rènshi hěn duō wàiguórén*  
1SG.F know very many foreigner

- (59) b. \*她認識多外國人  
*tā rènshi duō wàiguórén*  
 1SG.F know many foreigner  
 ‘She knows many foreigners’
- (60) 王朔的著作很多  
*Wáng Shuò de zhùzuò hěnduō*  
 Wang Shuo DET work very many  
 ‘Wang Shuo’s works are numerous’
- (61) 中國是多民族國家  
*Zhōngguó shì duōmínzú guójiā*  
 China COP multiethnic country  
 ‘China is a multiethnic country’

With the exception of a few set phrases, the adjective 多 *duō* ‘much, many’ can modify a noun only if preceded by another modifier (Lü 1980: 111-2; cf. Guo L. 1983), sometimes semantically redundant, as 很 *hěn* ‘very’ in (59a); (59b) is therefore ungrammatical. Also, 多 *duō* can be a predicate and come after the subject, as in (60). In (61), 多 *duō* is conjoined to the noun 民族 *mínzú* ‘nationality’ without being itself modified, and it carries a different semantic value from the corresponding adjective: whereas adding 很多 *hěnduō* to 外國人 *wàiguórén* ‘foreigner’ in (59a) ‘adds’ some meaning to the noun without altering its word class and distributional properties, in (61) the morpheme apparently turns the base noun into a non-predicative adjective, an attributive-only form. So, we have clear distributional and semantic differentiation between the adjectival use and the prefixal use of 多 *duō*. Just as 零-*líng-*, 多-*duō-* does not only mean ‘many’, as in (59a), but rather ‘having many X’:

- (62) [*duō* [X]<sub>N</sub>]<sub>ADJ</sub> ‘having many X’ →  
 多功能 *duōgōngnéng* ‘multi-functional’

For 零- *líng*- ‘zero X’, it is just the polarity which is reversed. This phenomenon is far more common than we had expected: there are several cases of prefixes which are only apparently transparent, but which often turn the base noun into a non-predicative adjective, such as the above mentioned 高- *gāo*- in words like 高蛋白 *gāodànbái* ‘high protein’ (3.2.2; see Shen M. 1995:36):

- |      |                         |                     |
|------|-------------------------|---------------------|
| (63) | 高蛋白食品                   |                     |
|      | <i>gāodànbái shípǐn</i> |                     |
|      | high-protein food       | ‘high protein food’ |

The same goes for the ‘negator of existence’ 無- *wú*- ‘without, -free, -less’; compare the nouns in (64) and the non-predicative adjectives in (65):

- |      |                       |                        |
|------|-----------------------|------------------------|
| (64) | 無機鹽                   | 無底洞                    |
|      | <i>wújīyán</i>        | <i>wúdǐdòng</i>        |
|      | <i>wú</i> -organ-salt | <i>wú</i> -bottom-hole |
|      | ‘inorganic salt’      | ‘bottomless pit’       |
| (65) | 無糖                    | 無條件                    |
|      | <i>wútáng</i>         | <i>wútiáojiàn</i>      |
|      | <i>wú</i> -sugar      | <i>wú</i> -condition   |
|      | ‘sugar-free’          | ‘unconditional’        |

It would seem that 無- *wú*- is sometimes class-maintaining and sometimes class-changing. Our opinion, however, is that the words in (64) are formed through a two-step process, as represented in (66):

- (66)  $[[wú\ Y]_{ADJ}\ [X]_{N}]_N$  ‘X lacking Y’

Whereas the words in (65) seem to be built according to a schema analogous to those for 非- *fēi*- ‘non X’, 零- *líng*- ‘lacking X’, 多 *duō*-

‘having many X’ and 高- *gāo-* ‘having a high degree of X’ (56, 58, 62 and 63). So, basically, it would seem that these are all instances of class-changing prefixation, but some of those prefixes, actually, are sometimes class-changing and sometimes class-maintaining (as 非- *fēi-* ‘non X’ and 高- *gāo-*). In a constructionist perspective, the word class is assigned to the complex word by the construction itself, rather than by the affix; still, it remains to be explained why some complex words created, for instance, according to the schema in (62), are actually nouns, such as e.g. 多媒體 *duōméitǐ* ‘multimedia’ and the item 多功能 *duōgōngnéng* is listed in dictionaries also as a noun (WL 2007).

The second group of prefix-like elements which we mentioned at the beginning of this section is a closed set of ‘fully autochthonous’ prefixed elements, with a comparatively long history, which always assign adjectival class to the word they help to build; they have been analysed by many (see table 3.5; see also Ceccagno & Scalise 2006):

- (67)  $[[kě]_V [X]_V]_{ADJ}$  ‘which may be X-ed’ →  
可吃 *kěchī* ‘edible’
- (68)  $[[hǎo]_{ADJ} [X]_V]_{ADJ}$  ‘easy / pleasant to X’ →  
好些 *hǎoxiě* ‘easy to write’
- (69)  $[[nán]_{ADJ} [X]_V]_{ADJ}$  ‘difficult / unpleasant to X’ →  
難辦 *nánbàn* ‘hard to handle’

These three word-formation patterns apparently form a closed set and, apparently, are no longer productive. The morph that has been analysed most often as a prefix is 可- *kě-*; basically, it may be added to any single syllable verb to form an adjective (Lü S. 1980:243). It seems that words built according to the schemas in (67-69) are attested at least since the Middle Chinese period, as e.g. 可憎 *kězēng* ‘hateful, disgusting’ (世說新語 *Shì shuō xīn yǔ*, Vth century CE; Cheng X. 1992c), 好看 *hǎokàn* ‘good-looking’ (10th century; HYDCD 1993), 難熬 *nán’áo* ‘hard to

endure’ (西廂記諸宮調 *Xīxiāngjì Zhūgōngdiào*, XIIIth century); the lexeme 可憐 *kělián* ‘pitiable’ is actually attested in the *Mencius* (4th cent. BCE), and 難老 *nánlǎo* ‘hard to age’ (> ‘long-lived’) is already found in the *Book of Songs* (詩經 *Shījīng*). However, at the stage of Old Chinese, it is unclear whether such structures were phrases or words, and the issue deserves more investigation; we may anyway be rather sure that such model has existed at least for 1500 years.

Generally speaking, the head of productively formed complex words in Chinese is identified to the right for adjectives (Ceccagno & Basciano 2007). As for 可- *kě-* ‘which may be X-ed’ (67), we are faced again with a case for which, apparently, word class assignment is performed “by rule”, i.e. the part of speech tag belongs to the construction; other complex word with a  $[[Y]_V [X]_V ]_{ADJ}$  structure, as e.g. 知名 *zhīmíng* ‘famous’, have been often treated as exocentric compounds in the literature (see e.g. Ceccagno & Scalise 2006:251). The patterns in (68) and (69) are not usually regarded as instances of prefixation (see table 3.5) and, as far as their word class is concerned, could be analysed as left-headed, contrarily to what is believed about adjectives. Ceccagno & Scalise (2006:252), however, suggested that words built around 好 *hǎo* and 難 *nán* could be analysed as “emerging cases of derivation” and, therefore, the left-hand constituent should be taken as a class-assigning prefix; the semantic contribution of 好 *hǎo* and 難 *nán* to the complex word, we may add, also qualifies them as heads.

A provisional conclusion which may be drawn from the data presented above is that compound adjectives are right-headed in Chinese and those prefixed adjectival morphemes which turn the complex word into an adjective could be regarded as class-changing prefixes. As we shall see below, many among such prefixed words also display peculiar prosodic properties.

We have already dealt earlier (3.2.1.2) with the prosodic structure of disyllabic and trisyllabic complex words in Mandarin. We quoted Feng S. (2001) and Duanmu S. (2000)’s treatment of trisyllabic prosodic words, and we pointed out that  $[2_N + 1_N]_N$  structures are allowed, whereas  $[1_N + 2_N]_N$  structures are not, generally speaking, and this is because the ‘natural

foot' in word formation is formed from left to right. However, a word as 金項鍊 *jīnxiàngliàn* 'golden necklace' (Feng S. 2001:172; see above, fn. 14) is acceptable, because the first nominal morpheme is used as a non-predicative adjective; in such case, a 'syntactic word' (句法詞 *jùfǎcí*) is built, and in syntax foot building goes from right to left, making a structure as [*jīn*/[*xiàngliàn*]/] possible. Thus, 'our' trisyllabic prefixed structures as 多功能 *duōgōngnéng* 'multi-functional' or 高蛋白質 *gāodànbái* 'high protein' and the like are prosodically well-formed only if taken as syntactic objects, contradicting the analysis of such elements as prefixes. Duanmu S. (2000), as said before, stresses the fact that foot building rules must take into account the morphological structure of the words; in 'cyclic' foot building, the grouping of syllables is repeated, cyclically, starting from the smaller grammatical unit (Chomsky, Halle & Lukoff 1956); thus, Duanmu advocates in favour of a morphological treatment for a [ $1_{\text{ADJ}} + 2_{\text{N}}$ ]<sub>N</sub> structure as

- (70) 大房間                    [[/*dàfáng*/]<sub>N</sub> *jiān*]<sub>N</sub>  
       *dàfángjiān*  
       'big room'

Duanmu points out that all [ $1_{\text{ADJ}} + 2_{\text{N}}$ ]<sub>N</sub> compounds should be ill-formed in Chinese, but restrictions on word formation do not apply inside a prosodical foot. A structure such as that in (70), 大房间 *dàfángjiān* 'big room', could be acceptable, according to Duanmu, as the first two syllables are able to constitute a foot (*/dàfáng/*) and thus at the boundary with the other syllable, *jiān*, there is not [ADJ + N], but rather [N + N]. The same analysis could be applied to those models of word formation built around 多- *duō*- 'multi-, having many X' and 高- *gāo*- 'high, having a high degree of X' seen above, or 非- *fēi*- 'non X' and 無- *wú*- 'lacking X', which exhibit the same mismatch between morphological structure and prosodic template:

- (71) 非動物            [[/fēidòng/]<sub>v</sub> wù]<sub>N</sub>            vs.            [fēi [dòngwù]<sub>N</sub>  
           fēidòngwù  
           ‘inanimate object’
- (72) 非導體            [[/fēidǎo/]<sub>v</sub> tǐ]<sub>N</sub>            vs.            [fēi [dǎotǐ]<sub>N</sub>  
           fēidǎotǐ  
           ‘non-conductor’

So, the combination of the semantic and distributional differentiation together with the apparent prosodic anomaly could help the identification of prefixes in Chinese. However, such issue deserves further research.

To sum up, our data has shown that the definition of a class of derivational prefixes in Chinese is indeed challenging and calls for the interplay of semantic, distributional and prosodic criteria. From the prosodical point of view, Chinese prefixed words are anomalous, and the notion of cyclic foot has been employed to account for such anomaly and to set them apart from syntactic structures. The models represented by the schemas in (67-69), albeit very interesting from the point of view of headedness (especially 好- *hǎo*- ‘easy / pleasant to X’ and 難- *nán*- ‘difficult / unpleasant to X’), seem to be isolated cases in Chinese word formation. Patterns as those for 非- *fēi*- ‘non X’ and 零- *líng*- ‘X starting from zero’ / ‘lacking X’ appear to us as more typical for the language; prefixed elements as 有- *yǒu*- (for ethnic groups, e.g. 有苗 *yǒumiào* ‘the Hmong’; Wang L. 1980:217) or the above mentioned 老- *lǎo*- ‘old’, still used before surnames, which are categorially transparent, have been attested since the pre-Qin times (before 221 BCE; see Cheng X. 1992c, 1992a, 1992d). However, we have seen that many of those prefixed items sometimes form nouns and sometimes non-predicative adjectives; as in a constructionist perspective the word class may belong to the construction, rather than to the individual words, we should admit that, actually, two schemas are to be posited for some of these formants, without a clear indication of what the restrictions are, i.e. what kind of items ‘fit’ in the variable slot of the ‘nominal’ schema or of the ‘non-predicative adjective’ schema for the same prefixed item (as e.g. 57 and 58 for 零- *líng*-). However, we may also remark that the ‘boundary’ between the category of



noun and that of non-predicative adjective is easily crossed in Mandarin, and we have both nouns used also as non-predicative adjectives, as 專業 *zhuānyè* ‘specialty, major’ (73, ex. from Li Y. 1996), and non-predicative adjectives used as nouns, as 橢圓型 *tuōyuánxíng* ‘ellipse-shaped, elliptic-type’ (74<sup>37</sup>):

- (73) 專業劇團  
*zhuānyè jùtuán*  
 specialised theatre-group  
 ‘specialised / professional theatrical company’
- (74) 前端尖銳的長橢圓型  
*qiánduān jiānrùi de cháng tuōyuánxíng*  
 fore-end point-sharp DET long ellipse-shape  
 ‘elliptic shape with sharp ends’

Also, generally speaking, most nouns can be used as attributes of another noun, as 電腦 *diànnǎo* ‘computer’ in 電腦世界 *diànnǎo shìjiè* ‘computer world’ (Yip & Rimmington 2004:11), just as in English (e.g. ‘nutmeg’ in ‘nutmeg scent’ or, even, ‘the heady, almost nutmeg scent; Bhat 1994:126). Thus, one could just say that the output of the schemas for 非- *fēi* ‘non X’, 多- *duō* ‘multi-, having many X’ start out as nouns and are later used as non-predicative adjectives, sometimes leading to recategorization (i.e. permanent word class shift). However, as said before, a word such as 非貿易 *fēimàoyì* ‘non-commercial’ is born as a non-predicative adjective and, apparently, is not used as a noun; the same goes for 高蛋白 *gāodànbái* ‘high-protein’ (63). Thus, in short, it appears to us that it is more appropriate to posit separate, albeit related, schemas for prefixes which (productively) form nouns and non-predicative adjectives; in the case of 非- *fēi*-, the three schemas proposed in (54-56) may be reduced to two:

<sup>37</sup> Example from <http://www.hudong.com/wiki/%E5%89%91%E6%A6%95>.

- (75) [*fěi* [X]<sub>N</sub>]<sub>N</sub> ‘non X’ →  
非暴力 *fēibàoli* ‘nonviolence’
- (76) [*fěi* [X]<sub>ADJ/N</sub>]<sub>ADJ</sub> ‘lacking the property X’ →  
非常任 *fēichánggrèn* ‘nonpermanent’;  
非貿易 *fēimàoyì* ‘non-commercial’

In the last section, we shall provide a summary of the main conclusions in this chapter.

### 3.3. Summary and Concluding Remarks

The language data illustrated and analysed in this chapter have shown that the growth of word formation patterns with affix-like features (i.e. based on a bound constituent conveying a stable meaning in a fixed position) is the effect of a ‘synergy’ between language-internal tendencies in Chinese and the indirect influence of ‘European’ languages, the impact of which was mediated by Japanese, through which many Western notions and the related words were introduced in the Chinese language (3.2.1, 3.2.1.1, 3.2.1.2). Many word formation processes with a very long history in China have actually developed in the last two centuries, as 學 *-xué* ‘branch of learning’ and 化 *-huà* ‘-ise, -ify’ (3.2.3); also, word formation patterns which were not in use before the XXth century were introduced, as 吧 *-bā* (3.2.2). Needless to say, not all of the word formation patterns which grew in productivity since the XIXth century, many of which may be subsumed under the category of class nouns (1.3.2), are actually derivational affixes, and each case must be analysed by itself, according to the semantic and distributional criteria set out in CHAPTER 1. Also, we remarked how the development of affixes and word formation patterns with affix-like constituents is tightly connected with the diffusion of trisyllabic words in the Mandarin lexicon.

As to the ‘semi-free morphemes’ 者 *-zhě* ‘person doing X, having opinions proper of X or possessing the characteristic X’ and 式 *-shì* ‘style, pattern’, we argued that the label ‘semi-free’ is synchronically valid,

but overshadows the fact that those items have a very different history; whereas we analysed the former as a derivational suffix, the latter does not appear to have grammaticalised, and we actually suggested that it is a transcategorised item (3.2.4).

Lastly, we examined different ‘models’ for prefixation in Modern Mandarin, namely class-maintaining and class-changing (3.2.5). Whereas items belonging to the ‘closed’ group of class-changing prefixes 可- *kě* ‘which may be X-ed’, 好- *hǎo*- ‘easy / pleasant to X’ and 難- *nán*- ‘difficult / unpleasant to X’, apparently no longer productive, are actually always class-changing, items in the first group, as 非- *fēi* ‘non X’ / ‘lacking the property X’, are ambiguous, as they sometimes they turn the noun they attach to into a non-predicative adjective. We proposed that the output category belongs to the schema, and that there is a different schema for each word class identity.



## CHAPTER 4 CONCLUSION

The main aim of this research was to provide a treatment of lexical derivation in Mandarin Chinese, in order to gain a better understanding of Chinese morphology and of derivation as a cross-linguistic phenomenon. Also, we tackled the issue of whether derivational affixes conveying lexical meaning are to be regarded as grammaticalised or lexicalised items, or neither of those, a question on which there seems to be no agreement in the literature, stating our reasons for an analysis of the genesis of derivational phenomena in the framework of grammaticalization theory. In what follows, we shall briefly summarise the most relevant conclusions reached in this work, and we shall suggest some areas for further research.

### 4.1 Lexical Derivation in Grammaticalization Theory

In CHAPTER 1, we pointed out that derivation has an ‘unstable’ placement in theories of grammar, grammaticalization and the lexicon (see Himmelmann 2004, among others). After reviewing some recent proposals, we highlighted the analogies between ‘typical’ grammaticalization (i.e. the genesis of grammatical markers) and the evolution of lexemes / lexical morphemes into derivational affixes, even when they convey lexical meaning.

From the distributional point of view, when a lexeme is used a bound word constituent, appearing in a fixed position (prefixed, suffixed), it means that an increase in bondedness and, often, a reduction in structural scope has occurred, just as e.g. in the ‘creation’ of, say, bound TAM markers (see 1.3.1.1). From the semantic point of view, which is crucial for the characterization of lexical derivation, it seems that the processes commonly accepted as the basis of grammaticalization, like metaphor, metonymy, abstraction (generalising / isolating) are all involved (to a different extent in each instance) also in the morphologisation of derivational affixes, as shown with the comparison between Ewe *vi* ‘child’ > *-vi* ‘(polysemic) suffix’ and Ch. 性 *xìng* ‘nature, spirit’ > –

性 *-xìng* ‘the quality of [X] / connected with [X]’.

A characteristic of the languages of East and South-East Asia, including Mandarin, is that grammaticalization does not involve the “coevolution of meaning and form”, i.e. it does not (necessarily) entail a reduction in the shape of the sign (1.3.2); through the comparison of the histories of Ger. *-heit* (Eng. *-hood*) and Ch. 性 *-xìng* (1.3.2.1), two items with a very similar meaning, usage and function, we showed that phonological reduction (and blurring of boundaries) does not always occur even in an Indo-European language as German and, anyway, the semantic processes operating in every step of the grammaticalization / morphologization of those two formants are strikingly similar. A possible area for further research could be to find out whether there is a connection between the phonological integrity of a grammaticalised sign and the kind of meaning conveyed, i.e. if the ‘concreteness’ of lexical meaning somehow prevents phonological reduction and blurring, as for Ger. *-heit*; since for Mandarin grammaticalization is normally not associated with such alterations in the shape of an item, data from this language is not relevant.

Also, we did not locate lexical derivation either in the grammar or in the lexicon; following Himmelmann’s (2004) suggestion (1.3.1), our aim was only that of showing the similarities between the “emergence” of derivational formatives and “prototypical” instances of grammaticalization. The collocation of derivation in the architecture of language is far beyond the aims of the present work; nevertheless, we hope that our considerations on grammaticalization, morphologization and lexical derivation may contribute to a better understanding of the nature of derivational phenomena.

## 4.2 ‘Inner’ and ‘Outer’ Forces Driving the Evolution of Chinese Word Formation

The Chinese language has always had morphology throughout its history, despite claims of the contrary (1.1.3, fn. 4); however, whereas sub-syllabic grammatical markers (affixes) have been reconstructed for Old Chinese only (Baxter & Sagart 1998, Sagart 1999), morphological processes involving the agglutination of syllables were found since the early texts and thrived into the present day. Affix-like formatives of Old

and Middle Chinese were mostly evaluative or transpositional (i.e. word-class bearing) in nature; in Modern Mandarin, as we have seen, processes of agglutination of lexical morphemes, either bound or free, have been instrumental in the ‘modernization’ of the Chinese lexicon.

Also, following Feng S. (1997, 1998, 2001), we pointed out that the simplification of the syllable structure in the history of the language has led to the ‘disyllabification’ of the Chinese lexicon, to maintain sufficient prosodic ‘weight’ (3.2.1.1); given the strong tendency towards a 1:1 correspondence between syllables and morphemes in Chinese, two-syllable words were also, mostly, bimorphemic words. The acceptance of a large number of Japanese loanwords (graphic loans; 3.2.1) with a ‘2-syllable determiner + 1-syllable determiner’ structure (as 動物學 *dòngwùxué* ‘zoology’) has apparently stimulated the creation of more trisyllabic words with such structure and, thus, the diffusion of more monosyllabic suffix-like formatives (class nouns), as 性 *-xìng* ‘the quality of [X] / connected with [X]’ cited earlier (see table 3.3); in constructionist terms, the above mentioned structure provided an environment for the ‘conventionalisation’ of class nouns (compare Bisang’s maximum patterns; 1.3.2, 3.2.1) and other suffix-like formatives and, eventually, for the grammaticalization of some of those items into derivational affixes.

Thus, as far as the shape and structure of complex words is concerned, we may say that the historical developments in Chinese word formation which favoured the emergence of many lexical derivational formatives were the effect of both language-internal tendencies and Japanese influence (or, better, ‘Japanese-mediated’ European influence). Very often, the word formation schemas containing lexical morphemes which later became affixes were attested in the Chinese language well before the period of intense contact with Japanese (i.e. before the XIXth century); thus, the ‘material’ from which Modern derivational affixes were created is autochthonous in a sense. The ‘synergy’ between ‘inner’ and ‘outer’ forces outlined here in the domains of word formation and prosodic morphology was a fundamental factor in the shaping of the Modern Chinese lexicon (see Masini 1993).

### 4.3 Lexical Derivation as a Cross-Linguistically Valid Category

The subcategory of derivational phenomena which we termed lexical derivation, i.e. those processes which convey (broadly defined) lexical meaning or, at least, bear a word class, is especially problematic as to the distinction with compounding. Typically, the criteria proposed in the literature for the delimitation of those two phenomena are (not surprisingly) synchronic in nature (1.2, 1.2.1, 1.2.2). From the semantic point of view, both lexical derivation and compounding involve lexical rather than (typical) grammatical meaning; whereas in inflectional morphology grammatical information (as e.g. tense, number, etc.) from a definite set are expressed, the range of meaning which may be expressed through derivation is virtually unlimited (see Bauer 2002). Thus, the fundamental criterion for the distinction between derivation and compounding is somehow formal: the constituents of compounding are lexemes, whereas derivation is (more often than not) expressed by affixation; thus, elements which look like lexemes of the language or have properties of lexemes, but behave as derivational affixes in word formation (i.e. appearing in a fixed position with a stable meaning which is not available in their free usage), as the often-quoted Dutch *-boer* ‘seller of [X]<sub>N</sub>’ are sometimes regarded as ‘affixoids’, i.e. as some hybrid entity. The salient formal and semantic features of inflection, derivation and compounding are summarised in table 4.1:



Table 4.1. Inflection, derivation and compounding: salient formal and semantic features.

	<b>Inflection</b>	<b>Derivation</b>	<b>Compounding</b>
<b>Semantic level</b>	Involves grammatical / relational meaning	Involves grammatical/ relational meaning and, often, lexical meaning (including word-class identity)	The individual constituents convey lexical meaning
<b>Formal level</b>	Conveyed by bound morphs or supra-segmental features	Conveyed by bound morphs, sometimes with a structure similar to that of ordinary lexemes of the language	Combines lexemes or other forms endowed with lexical autonomy

Such differentiation of compound constituents and (lexical) derivational affixes is based on a synchronic feature, namely that of bondedness and positional stability; in the Indo-European languages of Europe, typically, free status is associated with lexical meaning, whereas bound status is associated with grammatical meaning. However, we have seen that even in SAE languages we have a number of bound morphemes seemingly lexical in nature, as the so-called neoclassical constituents (*bio-*, *anthropo-*, etc.); in Mandarin, a language in which a very large portion of lexical morphemes are actually bound, appearing in a fixed position with a stable meaning is not a sufficient criterion for labelling an item as an affix.

In a diachronic perspective, when the lexical ‘forefather’ of a derivational morpheme can be identified (as e.g. Latin *mente* ‘mind’ > Fr. *-ment* ‘adverb-forming suffix’), the criterion of phonological reduction is often invoked, especially in models of grammaticalization in which such phenomenon is inevitably connected with alterations in the shape of a sign, i.e. a “coevolution of form and meaning” (Bybee, Perkins & Pagliuca 1994, Bisang 1996). In a language such as Mandarin, a feature of which is the lack of such ‘coevolution’, with highly grammaticalized signs which retain their shape and are used also as lexical items in other contexts, even such criterion is not particularly meaningful. Thus, we tried to argue that the ‘traditional’ diagnostics of grammaticalization of increased

bondedness and scope reduction (**1.3.1.1**), together with the semantic correlates of abstraction mentioned above (**4.1**), are the criteria on which the distinction between lexical morphemes and derivational affixes should rest. Thus, a cross-linguistically valid definition of (lexical) derivation, in our opinion, may only be based on a diachronic analysis, as far as the distinction between compounding and derivation is concerned.

#### 4.4 Suggested Areas for Further Research

From our survey of the recent literature on morphology, it appears that derivation is somehow regarded as a phenomenon ‘between’ inflection and derivation (see table 4.1); however, the definition of inflection is seemingly less problematic than that of derivation (and compounding). In fact, inflection is often defined in terms of certain grammatical categories, as TAM, number, etc.; for derivation, there is no such limited set of meanings and, also, the ‘quality’ of the data on derivation in less-described languages is often far from ideal, since derivation is sometimes seen as a “side issue” in grammatical description (Bauer 2002; **1.3.1**). This, however, is easily turned into a circular argument: without a solid, cross-linguistically valid definition of derivation we cannot improve our understanding of the phenomenon, but such definition must be refined by looking at data from the greatest possible variety of languages. What’s more, if we are to apply our diachronic criteria, we do not only need reliable data from languages in their contemporary stage of development, but also reliable historical data, and this is often impossible for many languages (especially, those with no established written tradition). We hope that this monograph will encourage researchers to undertake the challenging endeavour of collecting data on derivation-like phenomena from languages from different families, geographical areas and types.

Another issue which we could deal with only superficially, due to limited space, is that of multifunctionality, i.e. “categories which, though semantically distinguishable, are marked morphologically in the same way” (Bauer 2002:42). Items as can be understood either as polysemous or as instances of syncretism of different categories, leading to homonymy; in a Construction Morphology approach, since constructions are pairings of form and meaning (and function), each meaning / function should be

associated, in theory, to a specific word formation schema. For some cases, as e.g. 吧 *-bā* ‘bar’ > ‘place (actual or virtual) where a service related to [X]<sub>N/ADJ</sub> is offered or where information related to [X]<sub>N</sub> may be exchanged or where [X]<sub>V</sub> may be done’, the broad range of meanings which the formative at issue may express have been understood as a consequence of the generalisation in meaning occurring in grammaticalization; the meaning of ‘highly grammaticalized’ derivational formatives may be very general and, thus, have a broad extension. In other cases, such as the prefix 非 *-fēi-* ‘non X’, since the output includes productively formed nouns, but also non-predicative adjectives, formed both from nouns and from other non-predicative adjectives, we opted for a ‘polysemy approach’, positing two separate schemas, namely [*fēi* [X]<sub>N</sub>]<sub>N</sub> ‘non X’ and [*fēi* [X]<sub>ADJ/N</sub>]<sub>ADJ</sub> ‘lacking the property X’ (exx. 75-76, 3.2.5). More data is needed, however, to gain a better understanding of multifunctionality in derivation, in relation to constructional approaches.

Lastly, we hope that further research is done on the relationship between prosody and morphology in the evolution of languages; also, more data are needed on such issue in Chinese itself, broadening the variety of texts sampled to obtain a more complete coverage of the lexicon in different periods of time.



## BIBLIOGRAPHY

- Aikhenvald, A. Y. 2007. Typological Distinctions in Word Formation. in T. Shopen (Ed.). *Language Typology and Syntactic Description (Volume III: Grammatical Categories and the Lexicon)*. Cambridge: Cambridge University Press. 1-65.
- Amha, A. 2001. *The Maale Language*. Leiden: Research School of Asian, African, and Amerindian Studies, University of Leiden.
- Amiot, Dany (2005), Between compounding and derivation – elements of word-formation corresponding to prepositions, in Dressler, Wolfgang et alii (eds.), pp. 183-195.
- Anderson, Stephen (1985), Typological distinctions in word formation, in Shopen, Timothy (ed.), pp. 3-56.
- Anderson, Stephen (1992), *A-morphous morphology*, Cambridge, Cambridge University Press.
- Ansaldo, Umberto and Lim, Lisa (2004), Phonetic absence as syntactic prominence. Grammaticalization in isolating tonal languages, in Fischer, Olga, Norde, Muriel & Perridon, Harry (eds.), pp. 345-361.
- Arcodia, G. F. 2006. La Natura Morfologica dei Composti Coordinativi in Cinese Mandarino. *Studi Italiani di Linguistica Teorica e Applicata* 3: 549-578.
- Arcodia, G. F. 2007. Chinese: a Language of Compound Words? In F. Montermini, G. Boyé and N. Hathout (Eds.). *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*. Somerville, Mass.: Cascadilla Proceedings Project. 79-90.
- Arcodia, G. F. 2008. La Derivazione Lessicale in Cinese Mandarino. Milano: Franco Angeli.
- Arcodia, G. F. 2011. A Construction Morphology Account of Derivation in Mandarin Chinese. *Morphology* 21.1:89-130.
- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- Aronoff, M. and F. Anshen. 1998. Morphology and the Lexicon: Lexicalization and Productivity. In A. Spencer and A. M. Zwicky (Eds.). *Handbook of Morphology*. Oxford: Blackwell. 237-247.
- Banfi, E. 2005. Formazione delle Parole in Cinese Mandarino e Questioni di Deriva Tipologica. In M. Grossmann and A. M. Thornton (Eds.). *La*

- Formazione delle Parole. Atti del XXXVII Congresso Internazionale di Studi della Società di Linguistica Italiana (SLI). Roma: Bulzoni. 53-68.
- Bauer, L. 1998. Is There a Class of Neoclassical Compounds, and if So Is it Productive? *Linguistics* 36:403-22.
- Bauer, L. 2001a. Compounding. In M. Haspelmath, et al. (Eds.). *Language Typology and Language Universals*. Berlin-New York: Mouton de Gruyter. 695-707.
- Bauer, L. 2001b. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, L. 2002. What Can You Do with Derivational Morphology. In S. Bendjaballah, et al. (Eds.). *Morphology 2000. Selected papers from the 9th Morphology Meeting*. Amsterdam-Philadelphia: John Benjamins. 37-48.
- Bauer, L. 2005. The Borderline between Derivation and Compounding. In D. Wolfgang, et al. (Eds.). *Morphology and its Demarcations*. Amsterdam-Philadelphia: John Benjamins. 97-108.
- Bauer, L. 2006. Compound. In K. Brown, et al. (Eds.). *Encyclopedia of Language & Linguistics*. Second Edition. London: Elsevier. 719-726.
- Baxter, W. H. and L. Sagart. 1998. Word Formation in Old Chinese. In J. L. Packard (Ed.). *New approaches to Chinese Word Formation*. Berlin-New York: Mouton de Gruyter. 35-75.
- Beard, R. 1998. Derivation. In A. Spencer and A. M. Zwicky (Eds.). *Handbook of Morphology*. Oxford: Blackwell. 44-65.
- Beutel, H. 2005. Wortschatzbesonderheiten im Chinesischen. In A. D. Cruse, et al. (Eds.). *Lexikologie / Lexicology*. Berlin-New York: Mouton de Gruyter. 1064-1071.
- Bhat, D. N. S. 1994. *The Adjectival Category*. Amsterdam-Philadelphia: John Benjamins.
- Biasco, M., W. Mao and E. Banfi. 2003. *Introduzione allo Studio della Lingua Cinese*. Roma: Carocci.
- Bisang, W. 1996. Areal Typology and Grammaticalization: Processes of Grammaticalization Based on Nouns and Verbs in East and Mainland South East Asian languages. *Studies in Language* 20.3:519-597.
- Bisang, W. 1998. Grammaticalization and Language Contact, Constructions and Positions. In A. Giacalone Ramat and P. J. Hopper (Eds.). *The Limits of Grammaticalization*. Amsterdam-Philadelphia: John

- Benjamins. 13-58.
- Bisang, W. 2001. Syntax/Morphology Asymmetry in Vietnamese - a Consequence of Contact by Writing between Vietnamese, Chinese and Standard Average European Languages. In B. Igla and T. Stolz (Eds.). *Was Ich Noch Sagen Wollte ... A Multilingual Festschrift for Norbert Boretzky*. Berlin: Akademie Verlag. 189-201.
- Bisang, W. 2004. Grammaticalization without Coevolution of Form and Meaning: the Case of Tense-Aspect-Modality in East and Mainland Southeast Asia. In W. Bisang, N. P. Himmelmann and B. Wiemer (Eds.). *What Makes Grammaticalization? A Look from its Fringes and its Components*. Berlin and New York: Mouton de Gruyter. 109-138.
- Bisang, W. 2008. Grammaticalization and the Areal Factor – the Perspective of East and Mainland South East Asian Languages. In M. J. López-Cousa and E. Seoane (Eds.). *Rethinking Grammaticalization*. Amsterdam-Philadelphia: John Benjamins. 15-35.
- Bisetto, A. and S. Scalise. 2002. Trasferimento di Costrutti Sintattici in Morfologia. *Lingue e Linguaggio* 1:237-261.
- Bisetto, A. and S. Scalise. 2005. Classification of Compounds. *Lingue e Linguaggio* 4:319-332.
- Bisetto, A. and S. Scalise. 2007. Selection is a head property. *Acta Linguistica Hungarica* 54.4:361-380.
- Blank, A. 2001. Pathways of lexicalization. In M. Haspelmath, et al. (Eds.). *Language Typology and Language Universals*. Berlin – New York: Mouton de Gruyter. 1596-1608.
- Booij, G. 1993. Against Split Morphology. In G. Booij and J. van Marle (Eds.). *Yearbook of Morphology 1993*. Dordrecht-Boston-London: Kluwer Academic Publishers. 27-50.
- Booij, G. 2005. Compounding and Derivation – Evidence for Construction Morphology. In W. Dressler, et al. (Eds.) *Morphology and its Demarcations*. Amsterdam-Philadelphia. John Benjamins. 109-132.
- Booij, G. 2006. Inflection and Derivation. In K. Brown, et al. (Eds.). *Encyclopedia of Language & Linguistics*. Second Edition. London: Elsevier. 654-661.
- Booij, G. 2007. Construction Morphology and the Lexicon. In F. Montermini, G. Boyé and N. Hathout (Eds.). *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*. Somerville, Mass.:

- Cascadilla Proceedings Project. 33-44.
- Booij, G. 2009. Compounding and Construction Morphology. In R. Lieber and P. Štekauer (Eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 201-216.
- Booij, G. 2010. Compound Construction: Schemas or Analogy? In S. Scalise and I. Vogel (Eds.). *Cross-Disciplinary Issues in Compounding*. Amsterdam - Philadelphia: John Benjamins. 93-108.
- Bybee, J. L. 1985. *Morphology. A Study of the Relation Between Meaning and Form*. Amsterdam-Philadelphia: John Benjamins.
- Bybee, J., P. Revere and W. Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the languages of the world*. Chicago: University of Chicago Press.
- Campbell L. 2001. What's Wrong with Grammaticalization. In L. Campbell (Ed.). *Grammaticalization: A Critical Assessment*. Special issue of *Language Sciences* 23:113-161.
- Cao, B. 1952. *Yufa Chubu (Basic grammar)*. Beijing: Gongren Chubanshe.
- Carstairs-McCarthy, A. 1992. *Current Morphology*. London and New York: Routledge.
- Casacchia G. 2006. Chinese Linguistic Tradition. In K. Brown, et al. (Eds.). *Encyclopedia of Language & Linguistics*. Second Edition. London: Elsevier. 358-362.
- CCD. 2002. *The Contemporary Chinese Dictionary*. Beijing: Foreign Language Teaching and Research Press.
- Ceccagno, A. and B. Basciano. 2009a. *Shuobuchulai: la Formazione delle Parole in Cinese*. Bologna: Serendipità.
- Ceccagno, A. and B. Basciano. 2009b. Sino-Tibetan: Mandarin Chinese. In R. Lieber and P. Štekauer (Eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 478-490.
- Ceccagno, A. and S. Scalise. 2006. Classification, Structure and Headedness of Chinese Compounds. *Lingue e Linguaggio* 5:233-260.
- Chao, Y. 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chao, Y. 1979. *Hanyu Kouyu Yufa [Chinese translation of Chao Y. 1968]*. Beijing: Shangwu Yinshuguan.
- Chen, R. 1986. Tan Cízhuì “Xing” (On the affix “Xing”). *Xinjiang Shifan Daxue Xuebao* 2:87-90.



- Chen, W. 1955. Cír Liánxiě de Yǎnbìan, Bànfǎ hé Wèntí (Development, methods and issues in spelling romanised words). In *Hanyu de Cír he Pinxièfǎ* (Word and Romanization in Chinese). Special issue of *Zhongguo Yuwen*. 144-152.
- Cheng, L. 2004. Yuyan Jiechu, Leitui yu Xingtaihua (Language Contact, Analogy and Morphologization). *Waiyu yu Waiyu Jiaoxue* 4:53-56.
- Cheng, X. 1992a. Shishuo Xīnyu Fuyinci Yanjiu (Research on Polysyllabic Words in ‘New Tales of the World’). In X. Cheng (Ed.), *Wei Jin Nanbeichao Hanyu Yanjiu* (Research on Chinese of the Wei-Jin and Northern and Southern Dynasties period). Jinan: Shandong Jiaoyu Chubanshe. 1-85.
- Cheng, X. 1992b. Lunheng Fuyinci Yanjiu (Research on Polysyllabic Words in the ‘Lunheng’). In X. Cheng (Ed.), *Lianghan Hanyu Yanjiu* (Research on Chinese of the Former and Later Han Dynasties). Jinan: Shandong Jiaoyu Chubanshe. 262-340.
- Cheng, X. 1992c. Xianqin Shuangyinci Yanjiu (Research on Polysyllabic Words in the Pre-Qin period). In X. Cheng (Ed.), *Xianqin Hanyu Yanjiu* (Research on Chinese of the Pre-Qin period). Jinan: Shandong Jiaoyu Chubanshe. 44-112.
- Cheng, X. 1992d. Bianwen Fuyinci Yanjiu (Research on Polysyllabic Words in the Bianwen), In X. Cheng (Ed.), *Suitang Wudai Hanyu Yanjiu* (Research on Chinese of Sui, Tang and Five Dynasties period). Jinan: Shandong Jiaoyu Chubanshe. 1-101.
- Chomsky, N., M. Halle and F. Lukoff. 1956. On Accent and Juncture in English. In M. Halle et al. (Eds.), *For Roman Jakobson*. The Hague: Mouton. 65-80.
- Dai, J. X. L. 1998. Syntactical, Phonological and Morphological Word in Chinese. In J. L. Packard (Ed.), *New approaches to Chinese Word Formation*. Berlin-New York: Mouton de Gruyter. 103-133.
- De Mauro, T. 2005. *La Fabbrica delle Parole*. Torino: Utet.
- DeFrancis, J. 1984. *The Chinese Language: Facts and Fantasy*. Honolulu: University of Hawai’i Press.
- DELI. 1999. *Dizionario Etimologico della Lingua Italiana Cortelazzo-Zolli*. Bologna: Zanichelli.
- DEM. 2000. *Il Dizionario della Lingua Italiana per il Terzo Millennio*. Torino: Paravia.

- Deng X., Q. Wang and J. Li. 1996. A Statistical Study of Special Adjectives. In *Pan-Asiatic Linguistics: Proceedings of the Fourth International symposium on languages and linguistics*, Mahidol University at Salaya. 232-241.
- Dixon, R. M. W. 2002. Word: a Typological Framework. In R. M. W. Dixon and A. Y. Aikhenvald (Eds.). *Word: a Cross-Linguistic Typology*. Cambridge: Cambridge University Press. 1-41.
- Dong, X. 2002. *Cihuihua: Hanyu Shuangyinci de Yansheng he Fazhan* (Lexicalization: genesis and development of disyllabic words in Chinese). Chengdu: Sichuan Renmin Chubanshe.
- Dong, X. 2004. *Hanyu de Ciku yu Cifa* (Chinese Lexicon and Morphology). Beijing: Beijing Daxue Chubanshe.
- Dong, Z. 2003. Guanyu Jige Xinxing "Cizhui" de Tantaos (On Some Newly-Coined "Affixes"). *Tangshan Shifan Xueyuan Xuebao* 1:8-10.
- Dong, Z. 2006. Xiandai Hanyu Fuci "Fei" Yongfa zhijian de Guanxi (The relationship among the uses of the Modern Chinese Adverb "Fei"). *Hebei Beifang Xueyuan Xuebao* 6:32-5.
- Dressler, W. U. 1989. Prototypical Differences between Inflection and Derivation. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42:3-10.
- Dryer, M. S. 2003. Word Order in Sino-Tibetan Languages from a Typological and Geographical Perspective. In G. Thurgood and R. J. LaPolla (Eds.). *The Sino-Tibetan Languages*. London and New York: Routledge. 43-55.
- Duanmu, S. (1998), Wordhood in Chinese. In J. L. Packard (Ed.). *New approaches to Chinese Word Formation*. Berlin-New York: Mouton de Gruyter. 135-195.
- Duanmu, S. 2000. Hanyu de Jiezou (Chinese Prosody). *Dangdai Yuyanxue* 4:203-209.
- Fabb, N. 1998. Compounding. In A. Spencer and A. M. Zwicky (Eds.). *Handbook of Morphology*. Oxford: Blackwell. 66-83.
- Fan, L. 2002. Cong "Ba" Tan Hanyuci de Cizhuihua Qingxiang ("Ba" As a Starting Point for a Discussion of the Tendency towards Affixation of Chinese Lexicon). *Xinjiang daxue xuebao* 30:136-137.
- Feng, S. 1996. Prosodically Constrained Syntactic Changes in Early Archaic Chinese. *Journal of East Asian Linguistics* 5.4:323-371.

- Feng, S. 1997. *Hanyu de Yunlü, Cifa yu Jufa* (Prosody, Morphology and Syntax in Chinese). Beijing: Beijing Daxue Chubanshe.
- Feng, S. 1998. Prosodic Structure and Compound Words in Classical Chinese. In J. L. Packard (Ed.). *New approaches to Chinese Word Formation*. Berlin-New York: Mouton de Gruyter. 196-259.
- Feng, S. 2001. Lun Hanyu “Ci” de Duoweixing (The Multidimensional Properties of “Word” in Chinese). *Dangdai Yuyanxue* 3.3:161-174.
- Gaeta, L. 1998. Some Remarks on Analogy, Reanalysis and Grammaticalization. In A. Giacalone Ramat and P. J. Hopper (Eds.). *The Limits of Grammaticalization*. Amsterdam-Philadelphia: John Benjamins. 89-105.
- Gaeta, L. 2005. Morfologia Derivazionale e Tipologia: Quali Universali Linguistici? In N. Grandi (Ed.). *Morfologia e dintorni*. Milano: Franco Angeli. 12-27.
- Gaeta, L. and D. Ricca. 2003. Frequency and Productivity in Italian Derivation: A Comparison between Corpus-Based and Lexicographical Data. *Italian Journal of Linguistics* 15.1:63-98.
- GHYDCD. 2000. *Guhanyu Dacidian* (Great Dictionary of Classical Chinese). Shanghai: Shanghai Cishu Chubanshe.
- Goddard, C. 2005. *The Languages of East and South Asia: an Introduction*. Oxford: Oxford University Press.
- Goldberg, A. 2006. *Constructions at work: the Nature of Generalization in Language*. Oxford: Oxford University Press.
- Grandi, N. 2001. I Suffissi Valutativi tra Derivazione e Flessione: uno Studio Interlinguistico. *Archivio Glottologico Italiano* 86:129-173.
- Grandi, N. 2006. Considerazioni sulla Definizione e la Classificazione dei Composti. *Annali dell'Università di Ferrara, Sezione di Lettere on-line* 1:31-52 (<http://eprints.unife.it/annali/lettere/grandi.pdf>).
- Guo, L. 1983. *Xiandai Hanyu de Qianzhui he Houzhui* (Prefixes and Suffixes in Modern Chinese). *Zhongguo Yuwen*, 4, pp. 250-256.
- Haspelmath, M. 1996. Word-Class-Changing Inflection and Morphological Theory, In G. Booij and J. van Marle (Eds.). *Yearbook of Morphology 1995*. Dordrecht-Boston-London: Kluwer Academic Publishers. 43-66.
- Haspelmath, M. 2002. *Understanding Morphology*. Oxford: Oxford University Press.

- Haspelmath, M., et al. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- He, Y. 2004. The Words-and-Rules Theory: Evidence from Chinese Morphology. *Taiwan Journal of Linguistics* 2.2:1-26.
- Heine, B. & T. Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, B., U. Claudi and F. Hünemeyer. 1991. *Grammaticalization: a Conceptual Framework*. Chicago: The University of Chicago Press.
- Himmelman, N. P. 2004. Lexicalization and Grammaticalization: Opposite or Orthogonal? In W. Bisang, N. P. Himmelman and B. Wiemer (Eds.). *What Makes Grammaticalization? A Look from its Fringes and its Components*. Berlin and New York: Mouton de Gruyter. 21-42.
- Hong, B. 2005. *Litihua Gudai Hanyu Jiaocheng (Consolidation Course in Classical Chinese)*. Beijing: Gaodeng Jiaoyu Chubanshe.
- Hopper, P. J. and E. C. Traugott. 2003. *Grammaticalization (Second Edition)*. Cambridge: Cambridge University Press.
- Hu, S. 1930[1920]. *Guoyu de Jinhua (The Evolution of Chinese)*. In Hu Shi Eenji (Collection of Writings by Hu Shi). Shanghai: Yadong Tushuguan. 252-276.
- HYDCD. 1993. *Hanyu Dacidian (Comprehensive Chinese Dictionary)*. Shanghai: Hanyu Dacidian Chubanshe.
- INKGJT. 1994. *Iwanami Kokugo jiten (Iwanami Dictionary of the Japanese Language)*. Tokyo: Iwanami Shoten.
- Jiang, S. 2005[1989]. *Guhanyu Cihui Gangyao (Outline of Classical Chinese Lexicon)*. Beijing: Shangwu Yinshuguan.
- Jurafsky, D. 1996. Universals Tendencies in the Semantics of the Diminutive. *Language* 72:533-578.
- Klausenburger, J. 2000. *Grammaticalization: Studies in Latin and Romance Morphosyntax*. Amsterdam-Philadelphia: John Benjamins.
- Kuryłowicz, J. 1936. *Dérivation Lexicale et Dérivation Syntactique*. *Bulletin de la Société Linguistique de Paris* 37:79-92.
- Lackner, M., I. Amelung and J. Kurtz (Eds.). 2001. *New Terms for New Ideas: Western Knowledge and Lexical Change in Late Imperial China*. Leiden-Boston: Brill.
- Lakoff, G. 1987. *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

- Lehmann, C. 1989. Grammatikalisierung und Lexikalisierung. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42:11-19.
- Lehmann, C. 1995[1982]. *Thoughts on Grammaticalization*. Munich: Lincom Europa.
- Li, C. and S. A. Thompson. 1981. *Mandarin Chinese. A Functional Reference Grammar*. Berkeley: University of California Press.
- Li, Yu. 1996. Feiwei Xingrongci de Cilei Diwei (Lexical Category of Non-predicative Adjectives). *Zhongguo Yuwen* 1.
- Lieber, R. 2003. Compound Interpretation: Lexical Semantics, Not Syntax. In G. Booij et al. (Eds.). *Topics in morphology: Selected papers from the third Mediterranean Morphology Meeting*, Barcelona: IULA. 241-253.
- Lieber, R. and S. Scalise. 2006. The Lexical Integrity Hypothesis in a New Theoretical Universe. *Lingue e Linguaggio* 5:1-26.
- Lightfoot, D. J. 2005. Can the lexicalization/grammaticalization distinction be reconciled? *Studies in Language* 29.3:583–615.
- Lin, H. 2001. *A Grammar of Mandarin Chinese*. Munich: Lincom Europa.
- Lu, Z. 1964. *Hanyu de Goucifa (Word Formation in Chinese)*. Beijing: Kexue Chubanshe.
- Lü, S. 1953[1941]. *Zhongguo Wenfa Yaolüe (Outline of Chinese Grammar)*. Shanghai: Shangwu Yinshuguan.
- Lü, S. 1962. Shuo “Ziyou” he “Nianzhuo” (On “Free” and “Bound”). *Zhongguo Yuwen* 1:1-6.
- Lü, S. 1979. *Hanyu Yufa Fenxi Wenti (Issues in the Analysis of Chinese Grammar)*. Beijing: Shangwu Yinshuguan.
- Lü, S. 1980. *Xiandai Hanyu Ba Bai Ci (Eight Hundred Words in Modern Chinese)*. Beijing: Shangwu Yinshuguan.
- Lüdtke, H. 1985. Diachronic Irreversibility in Word-Formation and Semantics. In J. Fisiak (Ed.). *Historical Semantics, Historical Word-Formation*. Berlin–New York: Mouton Publishers. 355-366.
- Luo, J. 2004. “X” Xing Cizu Tanwei (On the “X-xing” Word Family). *Jiangtan Shifan Xueyuan Xuebao* 3:91-94.
- Ma, Q. 1995. Xiandai Hanyu Cizhui de Xingzhi, Fanwei he Fenlei (Nature, scope and classification of Modern Chinese Affixes). *Zhongguo Yuyanxuebao* 6:101-137.

- Magni, E. 2008. Conservazione e Innovazione nella Morfologia Derivazionale dell'Italiano: Analisi Sincronica e Diacronica del Suffisso –aio. In E. Cresti (Ed.), *Prospettive nello Studio del Lessico Italiano*. Firenze: Firenze University Press. 497-506.
- Masini, F. 1993. The Formation of Modern Chinese Lexicon and Its Evolution Towards a National Language: The Period from 1840 to 1898. *Journal of Chinese Linguistics Monograph Series 6*. Berkeley.
- McCarthy, J. J. and A. S. Prince. 1993. *Prosodic Morphology I – Constraint Interaction and Satisfaction*. Ms. University of Massachusetts and Rutgers University.
- McCarthy, J. J. and A. S. Prince. 1998. *Prosodic Morphology*. In A. Spencer and A. M. Zwicky (Eds.), *Handbook of Morphology*. Oxford: Blackwell. 283-305.
- Meillet, A. 1958. *Linguistique Historique et Linguistique Générale*. Paris: Champion.
- Montermini F. 2001. The Unitary Base Hypothesis and the Semantics of Word Formation Rules. In P. Bouillon and K. Kanzaki (Eds.), *First International Workshop on Generative Approaches to the Lexicon*, Geneva: University of Geneva.
- Montermini, F. 2008. *Il Lato Sinistro della Morfologia. La Prefissazione in Italiano e nelle Lingue del Mondo*. Milano: Franco Angeli.
- Naumann, B. and P. M. Vogel. 2000. Derivation. In G. Booij, C. Lehmann and J. Mugdan (Eds.), *Morphologie-Morphology*. Berlin-New York: Mouton de Gruyter. 929-943.
- Norde, M. 2009. *Degrammaticalization*. Oxford: Oxford University Press.
- Norman, J. 1988. *Chinese*. Cambridge: Cambridge University Press.
- NXCD. 2005. *Xiandai Hanyu Nixu Cidian (Reverse Order Modern Chinese Dictionary)*. Yanji: Yanbian Renmin Chubanshe.
- Olsen, S. 2000. Composition. In G. Booij, C. Lehmann and J. Mugdan (Eds.), *Morphologie-Morphology*. Berlin-New York: Mouton de Gruyter. 897-916.
- Ōta, T. 1957. *Chūgokugo Rekishi Bunpō (A Historical Grammar of Modern Chinese)*. Tokyo: Konan Shoten.
- Ōta, T. 1987. *Zhongguoyu Lishi Wenfa [Chinese translation of Ōta T. 1957]* Beijing: Beijing Daxue Chubanshe.
- Packard, J. L. 1998. Introduction. In J. L. Packard (Ed.), *New approaches to*

- Chinese Word Formation. Berlin-New York: Mouton de Gruyter. 3-34.
- Packard, J. L. 2000. *The Morphology of Chinese. A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press.
- Pan, W. 1990. *Hanying Goucifa Duibi Yanjiu (Comparative Study of Chinese and English Word Formation)*. X. Lin (Ed.). *Hanyu Luncong (Collected Essays on the Chinese Language)*. Shanghai: Huadong Shifan Daxue Chubanshe. 98-115.
- Pan, W., P. C. Yip and Y. Han. 2004. *Hanyu de Goucifa Yanjiu (Research on Word Formation in Chinese)*. Shanghai: Huadong Shifan Daxue Chubanshe.
- Piccinini, C. 2005. *Mnemotecnica, 'Morfologia' e Apprendimento dei Caratteri Cinesi: uno Studio Preliminare*. In N. Grandi (Ed.). *Morfologia e dintorni*. Milano: Franco Angeli. 126-140.
- Pinker, S. 1999. *Words and Rules: The Ingredients of Language*. London: Phoenix.
- Pirani, L. 2007. *Morphological Processes in the Mandarin Nominal Domain*. Unpublished Ph.D. thesis, University of Verona.
- Plag, I. 1999. *Morphological Productivity. Structural Constraints in English Derivation*. Berlin-New York: Mouton de Gruyter.
- Plag, I. 2004. *Syntactic Category Information and the Semantics of Derivational Morphological Rules*. *Folia Linguistica* 38.3-4:193-225.
- Plank, F. 1994. *Inflection and Derivation*. In R. E. Asher and J. M. Simpson (Eds.). *Encyclopedia of Language and Linguistics*. Oxford: Pergamon. 1671-1678.
- Pulleyblank, E. G. 1995. *Outline of Classical Chinese Grammar*. Vancouver: UBC Press.
- Pulleyblank, E. G. 2000. *From Archaic Chinese to Mandarin*. In G. Booij, C. Lehmann and J. Mugdan (Eds.). *Morphologie-Morphology*. Berlin-New York: Mouton de Gruyter. 1730-1740.
- Qu, Q. 1957a[1931]. *Pubian de Zhongguohua Ziyang de Yanjiu (Research on Common Words in Chinese)*. In Qu Qiubai Wenji (Collected Essays by Qu Qiubai) Beijing: Renmin Wenxue Chubanshe. 683-691.
- Qu, Q. 1957b[1932]. *Xin Zhongguowen Caoan (New Perspectives on the Chinese Language)*. In Qu Qiubai Wenji (Collected Essays by Qu Qiubai) Beijing: Renmin Wenxue Chubanshe. 705-851.
- Ramat, P. 1990. *Definizione di 'Parola' e Sua Tipologia*. In M. Berretta, P.

- Molinelli and A. Valentini (Eds.). *Parallela 4. Morfologia/Morphologie*, Atti del V Incontro Italo-Austriaco della Società di Linguistica Italiana (Bergamo, 2-4 ottobre 1989). Tübingen: Narr. 3-15.
- Ramat, P. 1992. Thoughts on Degrammaticalization. *Linguistics* 30.3:549-560.
- Ramat, P. 2001. Degrammaticalization or Transcategorization? in C. Schaner-Woller, J. Rennison and F. Neubarth (Eds.) *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler*. Torino: Rosenberg & Sellier. 393-401.
- Ramat, P. 2005. Per una Definizione di 'Parola'. In P. Ramat. *Pagine Linguistiche*. Roma: Laterza. 106-121.
- Ren X. 1981. *Hanyu Zaocifa (Word Formation in Chinese)*. Beijing: Zhongguo Shehui Kexue Chubanshe.
- Renouf, A. and H. Baayen. 1998. Aviating Among the Hapax Legomena: Morphological Grammaticalization in Current British Newspaper English. In A. Renouf (Ed.). *Explorations in Corpus Linguistics*. Amsterdam-Atlanta. Rodopi. 181-189.
- Ricca, D. 2005. Appunti per Una Tipologia degli Universali Morfologici. In N. Grandi (Ed.). *Morfologia e Dintorni*. Milano: Franco Angeli. 28-46.
- Rose, J. H. 1973. Principled limitations on productivity in denominal verbs. *Foundations of Language* 10:50-526.
- Sagart, L. 1999. *The Roots of Old Chinese*. Amsterdam-Philadelphia: John Benjamins Publishing Company.
- Sapir, E. 1921. *Language*. New York: Harcourt, Brace and Company.
- Scalise, S. 1984. *Generative Morphology*. Dordrecht, Holland: Foris Publication.
- Scalise, S. 1988. Inflection and Derivation. *Linguistics* 26.4:561-581.
- Scalise, S. 1994. *Morfologia*. Bologna: Il Mulino.
- Scalise, S., A. Bisetto and E. Guevara. 2005. Selection in Compounding and Derivation. In W. Dressler, et al. (Eds.). *Morphology and its Demarcations*. Amsterdam-Philadelphia. John Benjamins. 133-150.
- Shen, M. 1986. *Hanyu Xinde Cizhuhua Qingxiang (The New Tendency Towards Affixation in Chinese)*. Nanjing Shida Xuebao 4:93-98.
- Shen, M. 1995. *Shilun Xin Cizhuhuade Hanminzuxing (On the 'Chineseness' of the New Tendency Towards Affixation)*. Nanjing Shida Xuebao 1:35-41.



- Shen, S. X. N. 1990. *The Prosody of Mandarin Chinese*. Berkeley: University of California Press.
- Shi, Y. 2002. *The Establishment of Modern Chinese Grammar. The Formation of the Resultative Construction and Its Effects*. Amsterdam-Philadelphia: John Benjamins.
- Shi, Y. and C. N. Li. 2001. *Hanyu Yufahua de Licheng (A History of Grammaticalization in Chinese)*. Beijing: Beijing Daxue Chubanshe.
- SOED. 1993. *The New Shorter Oxford English Dictionary*. Oxford: Clarendon Press.
- Spencer, A. 2005. Word-Formation and Syntax. In P. Štekauer and R. Lieber (Eds.). *Handbook of Word-Formation*. Amsterdam: Springer. 73-97.
- Sproat, Richard & Shih, Chilin (1993), Why Mandarin morphology is not stratum-ordered, in Booij, Geert & van Marle, Jaap (eds.), pp. 185-217.
- Stump G. 2001. Affix position. In M. Haspelmath et al. (Eds.). *Language Typology and Language Universals*. Berlin-New York: Mouton de Gruyter. 708-714.
- Sun, C. 1996. *Word-Order Change and Grammaticalization in the History of Chinese*. Stanford: Stanford University Press.
- Sun, C. 2006. *Chinese: a Linguistic Introduction*. Cambridge: Cambridge University Press.
- Sun, J. 2005. Lun ‘Yi Yin Yi Yi’ (On the ‘One Sound, One Meaning’ Principle). In D. Lin (Ed.). *Yuyanxue Luncong (Collected Essays on Linguistics)*. Beijing: Shangwu Yinshuguan. 48-71.
- Sun, Y. 2000. *Xiandai Hanyu Cizhui Tantaos (A Study of Modern Chinese Affixes)*. Hebei Shifan Daxue Xuebao 23:55-58.
- Takashima, T. 1998. Chinese Characters and the Japanese Language 2. *Japanese Book News* 24:4-5.
- Tekavčić, P. 1980. *Grammatica Storica dell’Italiano*. Volume III. *Lexicon*. Bologna: Il Mulino.
- Ten Hacken, P. 2000. Derivation and Compounding. In G. Booij, C. Lehmann and J. Mugdan (Eds.). *Morphologie-Morphology*. Berlin-New York: Mouton de Gruyter. 349-360.
- Terreni, R. 2005. Composti N+N e Sintassi: i Tipi Economici Lista Nozze e Notizia-Curiosità. In M. Grossmann and A. M. Thornton (Eds.). *La Formazione delle Parole. Atti del XXXVII Congresso Internazionale di Studi della Società di Linguistica Italiana (SLI)*. Roma: Bulzoni.

- 521-546.
- Van Gelderen, Elly (2004), *Grammaticalization as Economy*, Amsterdam-Philadelphia, John Benjamins.
- VD. 2005. *Van Dale Groot Woordenboek van de Nederlandse Taal*. Utrecht-Antwerpen: Van Dale Lexicografie.
- Wälchli, B. 2005. *Co-Compounds and Natural Coordination*. Oxford: Oxford University Press.
- Wang, F. 1998. *La Formazione delle Parole nella Lingua Cinese Contemporanea*. Trieste: Università degli Studi di Trieste, Scuola Superiore di Lingue Moderne per Interpreti e Traduttori.
- Wang, H. 2005. *Shilun Xiandai Hanyu de Leicizhui (On Modern Chinese Affixoids)*. *Yuyan Kexue* 5:3-15.
- Wang, L. 1951[1944-1945]. *Zhongguo Yufa Lilun (Chinese Grammatical Theory)*. Shanghai: Shangwu Yinshuguan.
- Wang, L. 1980[1957]. *Hanyu Shigao (Draft History of the Chinese Language)*. Beijing: Zhonghua Shuju.
- Wang, L. 1985[1943-1944]. *Zhongguo Xiandai Yufa (Modern Chinese Grammar)*. Shanghai: Shangwu Yinshuguan.
- Wang, L. 1989. *Hanyu Yufashi (A History of Chinese Grammar)*, Pechino, Shangwu Yinshuguan.
- WDW. 2000. *Wahrig Deutsches Wörterbuch*. München: Bertelsmann Lexikon Institut.
- Williams, E. 1981. On the Notions “Lexically Related” and “Head of a Word”. *Linguistic Inquiry* 12.2:245-274.
- WL. 2007[1997]. *Wenlin Software for Learning Chinese (version 3.4)*. Eureka (CA): Wenlin Institute [based on: ABC Chinese-English Dictionary. Honolulu: University of Hawai’i press].
- Wu, F. 2005. *Hanyu Yufahua Yanjiu de Dangqian Ketu (Current Issues in the Research on Grammaticalization in Chinese)*. *Yuyan Kexue* 4.2.
- Wu, Y. 2000. *Ba – Xin Chansheng de Leihouzhui (Ba – A Newly Coined Suffixoid)*. *Jiangxi Shifan Daxue Xuebao* 4:73-77.
- XH. 2003. *Xinhua Xinciyu Cidian (Xinhua Dictionary of Neologisms)*. Beijing: Shangwu Yinshuguan.
- XHCD. 2002. *Xiandai Hanyu Cidian*. Beijing: Foreign Language Teaching and Research Press.
- XHCD. 2005. *Xiandai Hanyu Cidian*. Beijing: Foreign Language Teaching

- and Research Press
- Xiao, T. 1984. Tan Hanyu Yanjiuzhong de ‘Cigen’ Shuo (On ‘Roots’ in Ge Research on Chinese). *Xinyang Shifan Xueyuan Xuebao* 1:34-39.
- Xing, J. Z. 2006. Teaching and Learning Chinese as a Foreign Language: A Pedagogical Grammar. Hong Kong: Hong Kong University Press.
- Xu, Z. and L. Cai. 2007. Hanyu Qianzhui yu Wailai Yusu de Qianzhui Tiqu (Chinese Prefixes and the Absorption of Foreign Morphemes as Prefixes). *Huaibei Meitan Shifan Xueyuan Xuebao* 6:133-136.
- Yang, Y. 2007. Xinxing Sanyinjie Zhunqianzhuici de Xingzhi, Gongneng yu Fazhan (Nature, Function and Development of Newly-developed Three-syllable Pseudo-prefix Words). *Yuyan Kexue* 1:48-57.
- Yin, F. 1984. Hanyu Yusu de Dingliang Yanjiu (Quantitative Research on Chinese Morphemes). *Zhongguo Yuwen* 5:338-347.
- Yip, P. C. 2000. The Chinese Lexicon. A Comprehensive Survey. London and New York: Routledge.
- Yip, P. C. and D. Rimmington. 2004. Chinese: a Comprehensive Grammar. London and New York: Routledge.
- Yuan, Y. 1997. “Zhe” de Yufa Gongneng jiqi Lishi Yanbian (The Historical Evolution of the Grammatical Functions of “Zhe”). *Zhongguo Shehui Kexue* 3:160-178.
- Zhang, B. 2002. *Xinbian Xiandai Hanyu* (A New Chinese Grammar). Shanghai: Fudan Daxue Chubanshe.
- Zhang, Bo. 1994. Cilei Huoyong de Gongneng Jieshi (Functional Explanation for Transcategoriality). *Zhongguo Yuwen* 5:340-346.
- Zhang, Shi. 1956. *Xiaoxue Cihui Jiben Jiaoyu Zhishi Jianghua* (Introduction to the Philological Lexicon for Basic Education). Hangzhou: Zhejiang Renmin Chubanshe.
- Zhang, Sho. 1957. Lue Lun Hanyu Goucifa (A Short Discussion on Word Formation in Chinese). *Zhongguo Yuwen* 6:1-8.
- Zhang, Yi. 2002a. Zhuci yu Xiangguan Geshi (Auxiliary Particles and Related Patterns). Hefei: Anhui Jiaoyu Chubanshe.
- Zhang, Yi. 2002b. Shuo “X Shi” (On “X-shi”). *Shanghai Shifan Daxue Xuebao* 3:96-105.
- Zhang, Yu. 2002. “Hua” Wei Dongci Gongneng Ruohua de Dengji Xulie (Hierarchy of Functional Weakening of *Hua*-Suffixed Verbs). *Zhongguo Yuwen* 1:50-54.

- Zhang, Z. 1981. Tan Hanyu de Yusu (On the Chinese Morpheme). *Yuyan Jiaoxue yu Yanjiu* 4:4-18.
- Zhou X. and W. Marslen-Wilson. 1994. Words, Morphemes and Syllables in the Chinese Mental Lexicon. *Language and Cognitive Processes* 9: 393-422.
- Zhu, D. 1982. *Yufa Jiangyi* (On Grammar). Beijing: Shangwu Yinshuguan.
- Zhu, D. 1983. *Zizhí hé Zhuǎnzhǐ* (Self-Reference and Reference Shift). *Fangyan* 1:48-55.
- Zong, S. 1997. “Yusu” Shuo, “Cisu” Shuo Liyou Pingxi – Jianlun Hanyu Yusu de Fenlei (A Critical Analysis of the terms “Yusu” and “Cisu” – a Discussion on the Classification of Morphemes in Chinese). *Jinan Xuebao* 4. 134-142.

URL of web corpora:

- Academia Sinica Tagged Corpus of Old Chinese:  
<http://db1x.sinica.edu.tw/cgi-bin/kiwi/akiwi/akiwi.sh>
- Academia Sinica Tagged Corpus of Early Mandarin Chinese:  
<http://db1x.sinica.edu.tw/cgi-bin/kiwi/pkiwi/pkiwi.sh>
- Academia Sinica Balanced Corpus of Modern Chinese:  
<http://db1x.sinica.edu.tw/kiwi/mkiwi/>
- Corpus of the Centre for Chinese Linguistics, PKU (Classical and Modern Chinese): [http://ccl.pku.edu.cn/YuLiao\\_Contents.Asp](http://ccl.pku.edu.cn/YuLiao_Contents.Asp)

## 漢語的派生構詞法

---

作者：馬振國

發行人：張富恭

出版者：文鶴出版有限公司 地址：106 台北市金山南路二段 200 號 8 樓

電話：(02) 2393-4497 傳真：(02) 2394-3810

---

總公司：106 台北市金山南路二段 200 號 8 樓

TEL: (02) 2393-4497 FAX: (02) 2394-6822 / 2394-3810

各區業務及門市

北區：106 台北市金山南路二段 200 號 8 樓

TEL: (02) 2393-4497 FAX: (02) 2394-6822

中區：407 台中市台中港路二段 60-8 號 5 樓之 6

TEL: (04) 2317-0216 FAX: (04) 2314-0002

南區：802 高雄市同慶路 88 號 2 樓

TEL: (07) 2270-888 FAX: (07) 2270-801

網路書店：<http://www.crane.com.tw>

網路書店客服信箱：[crane.book@msa.hinet.net](mailto:crane.book@msa.hinet.net)

法律顧問：蘇家宏律師 TEL: (02) 2365-1288

---

劃撥帳號：01079261 戶名：文鶴出版有限公司

出版日期：2012 年 03 月 初版一刷

定價：450 元

ISBN：978-986-147-523-3 書號：002-0910

---

※著作權所有・翻印必究・非經本公司同意不得轉載、  
仿製或局部使用於雜誌、網路或其他書籍、媒體上。  
※若有缺頁，裝訂錯誤，煩請寄回更換。

### *Lexical Derivation in Mandarin Chinese*

Written by Giorgio Francesco Arcodia

Published by Crane Publishing Co., Ltd.

8F, No. 200, Sec. 2, Jinshan S. Rd., Da-An District,  
Taipei City 10643, Taiwan, R.O.C.

Tel: +886-2-2393-4497 Fax: +886-2-23946822

Crane Website: <http://www.crane.com.tw>

© 2012 Crane Publishing Co., Ltd. All rights reserved.