

Bound Roots in Mandarin Chinese and Comparison With European “Semi-Words”¹

Laura Pirani
*University of Verona*²

Compounding processes are morphologically productive in modern Mandarin Chinese. The process called “root compounding” (Sproat and Shih, 1996) is responsible for the creation of a high percentage new two-syllable words in the modern language. Bound root words look like compounds but the constituting parts are bound morphemes or “bound roots”. They have an intermediate status between affixes and words since they are bound like affixes but they behave like compounds constituents. Similarly, various European languages involve bound roots called “semi-words”. Root compounds in Mandarin and European semi-words are very similar: they have the same morphological structure and they share a similar historical and lexical development. The constituting elements are all bound and they behave more as content words with a lexical meaning. Historically, they derive from old free forms (*i.e.*, Mandarin bound roots from Old Chinese monosyllables, European semi-words from Greek and Latin words) and they are still productively used to form words that are part of a more technical and modern vocabulary. The analysis of Mandarin and European root compounding process shows that bound roots are not affixes but lexical words (*i.e.*, nouns) for three reasons. First, because of their lexical meaning. Secondly, they can attach to other bound roots. And finally, they can be both first and second word constituent, as in compounds (cf. Scalise, 1984; Sproat and Shih, 1996; Packard, 1998). Root compounding is a sign of the presence of morphology in Chinese and of a morphological trigger similar to typologically and genetically non related languages.

0. Introduction

This paper examines a word-formation process in Mandarin Chinese that involves one or two bound roots and compares it to a similar morphological process in European languages. This word-formation process in Mandarin is exemplified by the forms

¹ My special thanks to Rint Sybesma (Leiden University) for helping me in writing and revising this paper. Many thanks also to Denis Delfitto (University of Verona), Lisa Cheng (Leiden University), Stefan Rabanus (University of Verona) and Martin Everaert (Utrecht University) for their comments and suggestions to my work.

² Laura Pirani works now at *Q-go*, Diemen, The Netherlands.

presented in (1) and is claimed to be morphologically productive in the current language (cf. Sproat and Shih, 1996 and Packard, 1998; 2000).

In the following examples as in all this work, Chinese examples are written with characters followed by *pinyin*, then the meaning of the two (or more) words that form the compound and the compound resulting meaning in glosses. Characters are often between square brackets with their grammatical category and bound roots are often marked in bold.

- (1) a. [[电]_N [子]_N]_N (*diàn zǐ*) *electricity particle* ‘electron’
 b. [[磁]_N [带]_N]_N (*cí dài*) *magnetism tape* ‘(magnetic) tape’
 c. [[电]_N [压]_N]_N (*diàn yā*) *electricity press, weigh down* ‘voltage’
 d. [[脑]_N [力]_N]_N (*nǎo lì*) *brain power* ‘intelligence’

Complex words containing bound roots like those in (1) are dubbed “root compounds” by Sproat and Shih (1996) and “bound root words” by Packard (2000).

A similar process is morphologically productive also in Indo-European languages: above all, Romance and Germanic languages combine one or two bound roots and form the so-called “neo-classical compounds” (cf. Ralli (2005)) where the neo-classical bound root(s) is/are dubbed from their Latin or Greek origin. Several examples are presented in (2): the examples (a.- c.) are from a Romance language, Italian and those in (d.³ - f.) are from a Germanic language, English. In these examples neo-classical roots are marked in bold and labelled as nouns since they refer to the category ‘noun’ and they were nouns in the old language.

(2) Italian⁴

- a. [[**Semi**]_N[conduttore]_N]_N ‘semi-conductor’
 b. [[**Macro**]_N[economia]_N]_N ‘macro-economy’
 c. [[**Teo**]_N [**log**]_N ia]_N⁵ ‘theology’

³ Examples d. is taken from V. Waszink (2006). See this study for a recent and insightful discussion on different types of neo-classical compounds.

⁴ The Italian words presented here exist also in English.

⁵ [-ia] is a suffix, thus it is not part of the neo-classical bound root *-log-*. The same is for the suffix *-y* in example (2f.).

English

- d. **[[hydro]_N[gel]_N]_N** *hydr-o-gel*⁶
- e. **[[micro]_N[wave]_N]_N** *micr-o-wave*
- f. **[[neuro]_N[log]_Ny]_N** *neur-o-log-y*

The examples in (2a., b., d. and e.) present words made up by one bound root at the left hand side of the compound; the compounds in (2c.) and (2f.) are composed by two bound roots, both of Ancient Greek origin.

For all the languages whose examples are illustrated in (1) and (2) (*i.e.*, Mandarin, Italian and English), the most frequent structure encloses two bound roots. Mandarin, Italian and English are also similar as in most words made up by one bound and one free root, the bound root can be both the constituent to the left or to the right. Let's consider some examples in (3) see the different positions of bound roots (in bold).

(3) Mandarin

- a. 住房 (*zhù fang*) *to live house* 'residence'
- b. 国家 (*guó jiā*) *country family* 'nation'

Italian

- c. **Discoteca** 'disco'
- d. **Telecomunicazione** 'telecommunication'

English

- e. **Discography**
- f. **Cardio**fitness

In most Mandarin root-compounds and in most Italian and English neo-classical compounds, the same bound root can be both at the left and at the right hand side. Let's see one example of a productively used bound root in Mandarin (see (4a.) and (4b.)), Italian (see (4c.) and (4d.)) and English (see e. and f.).

⁶ The linking elements *-o-* and *-i-* are used in both neo-classical noun formation and neo-classical adjective formation. These linking elements mark ancient Greek (*-o-*) and Latin (*-i-*) origin. In English and German, the linking element *-o-* is much more frequent than *-i-*. In Romance languages, they are used with the same frequency.

Some neo-classical compounds do not contain a linking element. The absence of a linking element often derives from the fact that *o* and *i* can be deleted before a vowel.

(4) Mandarin: considering bound root 子(*zi*)

- | | | | |
|----|----|--------------------|---------------------------------------|
| a. | 独子 | (<i>dú zǐ</i>) | <i>only child, son</i> ‘only child’ |
| b. | 子宫 | (<i>zǐ gōng</i>) | <i>child, son womb, uterus</i> ‘womb’ |

Italian: considering the bound root *-term-*

- c. Termografo ‘termograph’
- d. Isotermo ‘isoterm’

English: consider bound root *-graph-*

- e. Graphology
- f. Chronograph

However, bound roots distribution does not seem to be free: cases of bound roots used exclusively at the left or at the right hand of root-compounds can be found both in Mandarin, in Italian and in English. In Mandarin, their position is the left-hand side, in Italian and English is the right-hand side. Because of this restriction, these bound roots seem to be more grammatical than lexical elements. Let’s see some examples in (5).

(5) Mandarin

- | | | | |
|----|-----|---------------------------|----------------------------------|
| a. | 飞机 | (<i>fēi jī</i>) | <i>to fly machine</i> ‘airplane’ |
| b. | 工程师 | (<i>gōng chéng shī</i>) | <i>project master</i> ‘engineer’ |
| c. | 住房 | (<i>zhù fang</i>) | <i>to live house</i> ‘residence’ |

Italian

- d. **Semicerchio** ‘semicircle’
- e. **Neodottore** *new doctor* ‘a University student that has recently graduated’

English

- f. **Hypo**allergenic
- g. **Psycho**-socio-analysis

Mandarin bound roots *ji*, *shi* and *fang* in (5a. - c.) mean ‘machine’ ‘master’ and ‘room, house’ respectively and even if on the surface they seem similar to the bound roots presented in (2), they are morphologically different since their behaviour is closer to that of affixes rather than lexical elements. As happens to affixes, they are found either at the left or at the right hand side, they most frequently attach to free roots, they are not used in lexically restricted contexts. In contrast, bound roots in (2) are both at the left and the

right hand side, attach most frequently to other bound roots and are mostly used for the creation of terms concerning science and technology.

Likewise, the Italian and English bound roots *neo-*, *semi-*, *hypo-* and *pshyco-* originate from Greek like the other bound roots in (2) but nowadays their behaviour is more similar to that of prefixes: they maintain their original lexical meaning and they are exclusively at the left hand side of a word, they tend not to select bound roots and they are used in diverse (and not only scientific) contexts. In contrast with the examples of bound roots presented in (2), in most of the cases in (5), the prefix tends to select a free (and not to a bound) root. With this respect, recent studies point out that these bound roots forming compounds in European languages (cf. Ralli, 2005) have to be treated either as affixes or as roots, not both.

However, looking at the examples above, it seems that several bound morphemes derived from Ancient languages behave more like affixes while others behave more like roots (especially like nouns). Moreover, it seems a typologically different language, Mandarin Chinese shows a similar usage with respect to bound roots.

This paper will propose a new approach to the study of bound roots in typologically non related languages: all these bound roots (or “semi-words” as they are called in Italian) exist in different languages and they are nowadays productively used. Many of them are roots, many are affixes. The criteria that permit to individuate and distinguish the two types are their historical development, their trigger, their distribution, their morphological selectional properties: roots are relicts of ancient languages, thus have the same structure of the language they originate from (cf. Ralli, 2005), are used for the creation of science terminology, they exclusively select one-syllable words and they are found both as right and left constituents and combine with other bound roots (as for lexical elements). The latter type contains bound roots with the same origin but they grammaticalized and are used in multiple and more numerous contexts as affixes: as all affixes do, they are positionally restricted, they do not have to create a specific vocabulary, they tend to attach to free words and they can select both one- and two-syllable words (cf. (5b.) and (5g.)). The first type of bound roots is involved in a compounding process, the second type in a derivational process.

The questions that arise are:

- i. What criteria in particular are responsible for the distinction of bound roots into those that seem more compositional from those more derivational?
- ii. What are the similarities and the differences between Mandarin Chinese semi-words and European semi-words?

This paper will approach each of these questions in the paragraphs that follow.

1. Functions and Properties of Mandarin Chinese Bound Roots

Bound roots in Mandarin Chinese have been analysed by Sproat and Shih (1996) and by Packard (2000). Sproat and Shih claim that in Mandarin Chinese several nominal roots used in derived words productively form “root compounds”; these roots are bound morphemes that semantically represent “kind” (Carlson (1977)). More precisely, they emphasize that the common Mandarin word related to a root would typically denote the canonical instance of that kind (1996: 58-59); so for instance, the root 蚁(*yǐ*) denotes the ant *kind*, whereas the normal Mandarin word for ‘ant’ 蚂蚁(*mǎ yǐ*) denotes the ‘canonical’ ant and can in general be used to denote most (sub)kinds of ants. One can represent the relationship between 蚁(*yǐ*) and 蚂蚁(*mǎ yǐ*) in terms of an IS-A hierarchy and 蚂蚁(*mǎ yǐ*) is a daughter node; the thick edge to the node around 蚂蚁(*mǎ yǐ*) represents the fact that this node denotes the canonical instance of the kind. Other words in which 蚁(*yǐ*) is the head would also be representable as bearing an IS-A relation to the 蚁(*yǐ*) node, such as 军蚁(*jūn yǐ*) ‘army ant’ or 工蚁(*gōng yǐ*) ‘worker ant’. Other cases, such as 白蚁(*báiyǐ*) (white ant) ‘termite’, bear a less clear IS-A relation to 蚁(*yǐ*) - termites being of course not ants, except perhaps in folk taxonomy. However, this kind of extension is common in compounding - cf., English *polecat*, *cuttlefish*, *sea lion*. They assume that monomorphemic nouns, such as 鸡(*jī*) ‘chicken’, which must be considered to be both roots and words, serve both the functions of kind and canonical instance.

In addition to Sproat and Shih, Packard (2000) investigates the productivity and the use of some bound morphemes that frequently form disyllabic constructions in modern Chinese. More precisely, Sproat and Shih on one side and Packard on the other introduce for Chinese monosyllabic bound morphemes respectively the term of “*non word roots*” and “*bound roots*” and they propose they form “*root compounds*” or “*bound root words*” (generally disyllables with either two bound roots or a bound and a free root). Chinese *bound root words* should be kept distinguished from compound-words: in fact, according to Packard’s (2000) definition, a “compound form” in Chinese has only free roots (generally two, *i.e.*, disyllabic compounds are the most frequent).

In contrast with bound root denotation of kind and disyllabic bound root words denotation of the canonical instance of that kind, compound-words do not denote neither the kind nor the canonical instance of that kind⁷. Moreover, differently from other bound morphemes (*e.g.*, inflectional affixes such as the aspectual particle 了(*le*)), Chinese *bound roots* have a lexical meaning⁸. Examples of productive Chinese *bound roots* are presented in (6).

⁷ Compound-words are called “full-compounds” in Sproat and Shih (1996) opposite to “root-compounds”.

⁸ Cf. also Packard (1998:16, 17): “Also, unlike “typical” affixing languages, Chinese has a large class of morphemes (which we may call ‘bound roots’) that possess certain affixal properties

- (6) a. 头 (*tóu*) ‘head, hair’
 e.g., 头脑 (*tóu nǎo*) *head brain* ‘mind’
- b. 国 (*guó*) ‘country’
 e.g., 国家 (*guó jiā*) *country family* ‘nation’
- c. 师 (*shī*) ‘teacher, master’
 e.g. 工程师 (*gōng chéng shī*) *project master* ‘engineer’
- d. 机 (*jī*) ‘machine’
 e.g. 飞机 (*fēi jī*) *to fly machine* ‘airplane’

However, even though bound roots in (6) can be all treated like compound constituents because of a lexical meaning, I propose that Mandarin bound roots are not all part of the same phenomenon: many bound roots are currently used as right and left constituent in compounds and by attaching to other bound morphemes they change their lexical meaning; in addition, they are used in a more technical and scientific environment. Bound roots showing this behaviour are those in (6a.) and (6b.).

In contrast, there are cases of bound roots like those in (6c.) and (6d.) that tend to have a more fixed position in compounds, to select free morphemes; they tend not to change their meaning according to the context.

To support this hypothesis, the example (7) shows the distribution, the selectional properties and the change of meaning that a bound root like 头(*tóu*) ‘head, hair’ (in (6a.) can undergo:

- (7) a. 头路 (*tóu lù*) *tóu class* ‘first class (good etc.)’
 b. 城头 (*chéng tóu*) *city wall tóu* ‘top of the city wall’

The two examples in (7) show the possibility of finding the bound root 头(*tóu*) both as constituent to the left (see a.) and to the right (see b.); in addition, both in (7a.) and in (7b.), 头(*tóu*) attaches to a bound morpheme in order to create a new independent word. Finally, in the words in (7), 头(*tóu*) is not used with its main (and most frequent) meaning of ‘head, hair’: in a. it means ‘first’, in b. it means ‘top’. Thus, the bound root 头(*tóu*) shows a specific distribution (it is not positionally restricted, it can be both on the left and on the right), it has the selectional property [+bound] as it tends to select bound morphemes and it can undergo changes of meaning.

Opposite to 头(*tóu*), the example (8) presents the different behaviour of the bound root 师 (*shī*) (already seen in (6c.)).

(namely, they are bound and productive in forming words), but encode lexical rather than grammatical information, and furthermore may occur as the component at the right or at the left hand side of a word”.

- (8) a. 老师 (*lǎo shī*) *old teacher, master* ‘teacher’
 b. 工程师 (*gōng chéng shī*) *project teacher, master* ‘engineer’

In compound-words like those in (8), the bound root 师 (*shī*) is the right hand side constituent and select free morphemes (like adjectives, verbs or two-syllable nouns). It always has the meaning of ‘teacher, master’.

All considered, Mandarin Chinese bound roots have to be distinguished into two groups: the first group includes bound roots like 头 (*tóu*), they are less positionally restricted and select bound morphemes; the second group includes bound roots like 师 (*shī*), they are positionally restricted and select free morphemes. I will analyse them in detail in the next paragraph.

2. Two Types of Bound Roots in Mandarin

As illustrated in the previous paragraph, Mandarin bound roots fall into two distinct classes. I propose this distinction on the grounds of three criteria:

- a) The selection by a semi-word of bound/free morphemes;
- b) Their position of a semi-word in a compound;
- c) The change of meaning of a semi-word according to the context.

According to these three morpho-semantic criteria, semi-words can be divided into:

- I. What I call “semi-nouns”: a conspicuous groups of semi-words that select bound morphemes, can be both right and left hand side word constituent, change their meaning according to the context. They are used to create a scientific and technical vocabulary. These semi-words are lexical and not grammatical elements (even if they are bound morphemes) as they behave like nouns. An example of semi-nouns, 子 (*zǐ*) ‘child, son’ is presented in (9).

- (9) a. 量子 (*liàng zǐ*) *quantity particle* ‘quantum’
 b. 电子 (*diàn zǐ*) *electricity particle* ‘electron’
 c. 子音 (*zǐ yīn*) *small element sound* ‘(phonet.) consonant’
 d. 半子 (*bàn zǐ*) *half son* ‘son in law’

- II. What I call “semi-affixes”: a restricted group of semi-words that select free morphemes and are exclusively found as right hand side constituent; they always keep the same meaning. They are not exclusively used for the creation of a technical and scientific vocabulary. An example of semi-affixes, 机 (*jī*) ‘machine’ is presented in (10).

- (10) a. 飞机 (*fēi jī*) *to fly machine* ‘airplane’
 b. 电话机 (*diàn huà jī*) *telephone, phone machine* ‘telephone (set)’

As shown in (9) and (10), semi-nouns are involved in root compounding; semi-affixes are involved in a derivational process.

These two groups of bound roots (or semi-words) share the origin from Old Chinese free nouns. However, even though they have the same origin, it seems they did not have the same development: many semi-words have probably grammaticalized⁹ and are nowadays used more like affixes than like nouns.

Despite this difference, they share a high frequency in morphologically productive processes of the modern language.

Nevertheless, since compounding is morphologically more productive than derivation, semi-affixes are less productively involved in word-formation processes, thus the morphological process that uses semi-nouns is more productive than the one that uses semi-affixes¹⁰.

The presence and the division of bound roots in two groups do not seem to be peculiar of Mandarin Chinese. In the next part, we will examine a similar situation and behaviour in European languages.

3. European Semi-Words

Semi-words are productively involved in European languages compounding. The use of these semi-words have been analysed by Scalise (1994:269-271) with respect to Italian and by Bauer (1983), Warren (1990), Adams (2001) and Baeskow (2004) with respect to English.

Starting from the analysis of Italian semi-words, Scalise calls them “*semi-parole*” and claims they have an uncertain status between “words” and “affixes”. He proposes “*semi-parole*”’s behaviour to be far from affixation and to move near the “word”’s status: this can be supported by the fact that an Italian semi-word can be found both as first and as second word constituent, as happened in compound-forms. Consider the case of *filo* (‘-phil-’)¹¹.

- (11) a. [[Franco] [filo]] ‘Francophile’
 b. [[Filo] [antropo]] ‘philanthrope’

⁹ In this analysis, the term “grammaticalization” refers to a morphological process of semantic change by which a content word (lexical morpheme) changes into a functional word or grammatical affix. The process of grammaticalization includes a phonological reduction and a semantic bleaching. Cf. A. C. Harris (1990).

¹⁰ The least productivity of semi-affixes is confirmed also by the fact that they are a small group compared to semi-nouns.

¹¹ Examples and considerations are taken from Scalise (1994: 269).

A standard affix cannot have such a free behavior: if it is on the base's right hand side it is a suffix, if it is on the base's left hand side, it is a prefix. According to Scalise (1994), the "*semi-parole*" seem to move near the role of compound constituents: consider the parallel between the use of *filo* in (11) and the use of the noun *finestra* 'window' in the compound-forms *ferma-porta* ('door-stop') vs. *porta-finestra* ('French window'): a noun can be both "Word1" (*i.e.*, constituent to the left) and "Word2" (*i.e.*, constituent to the right) in a compound.

I suggest that Scalise's analysis can be extended to other European languages: let's see several examples in (12): a. and b. are from French, c. and d. from Spanish, e. and f. from Dutch and g. and h. from German. Semi-words are in bold.

- (12) a. **Bibliographie** 'bibliography'
 b. **Téléphone** 'telephone'
 c. **Geografía** 'geography'
 d. **Psicología** 'psychology'
 e. **Bibliotheek** 'library'
 f. **Sociologie** 'sociology'
 g. **Grammophon** 'grammophone'
 h. **Voltmeter** 'voltmeter'

In 12 (a.-i.) the semi-word can be both at the right- and at the left hand side and are used with a lexical meaning as happens to compound constituents. Thus, following Scalise (1994) I argue for treating all the cases of European semi-words presented in (12) as lexical forms used as compound constituents.

However, as Bauer (1983) and Baeskow (2004) notice for English neo-classical compounds, there are cases of "hybrid formations" where the resulting disyllabic word is made up by a native morpheme and a neo-classical bound root. In most cases, the native morpheme is a free form. Example (13) presents the cases of *-logy* and *bio-* as semi-words in current American English.

- (13) a. **Hamburgerology**
 b. **Biofeedback**

Considering Baeskow's analysis and the English examples in (13), I propose that Italian shows "hybrid formations" *à la Baeskow* as well: there are neo-classical bound roots that attach to native free morphemes. Let's see the cases of *-teca*, *micro-* and *hypo-*.

- (14) a. [[Panino]_{-bound}[teca]_{+bound}]-bound 'the place where sandwiches (*panini*) are sold'
 b. [[Micro]_{+bound} [elettronica]_{-bound}]-bound 'microelectronics'
 c. [[Hypo]_{+bound} [tensione]_{-bound}]-bound 'hypotension'

The forms in (13) and (14) are different from those in (11) and (12) which are made up by two bound roots. With respect to constructions like those in (13), Baeskow claims that formations with a neo-classical root and a native free morpheme are considered as violating the Level ordering, causing a discrepancy in the hierarchy of MinLex. She proposes that elements such as *micro-* and *tele-* could be “reinterpreted” as class II prefixes, without being recategorized as affixes and can thus combine with free native constituents (be they simple or complex at level 2). Moreover, Warren (1990) and Adams (2001) state that there are certain Greek and Latin formatives that have become or are in the process of becoming affixes in English.

I agree with this proposal for English and adopt it for Italian: I propose that forms like *bio-* or *micro-* are affixes, thus grammatical and not lexical items also because they are positionally restricted and they have developed a meaning which is the same in all contexts. They can be used with native free morphemes as they are second class affixes. They are very productive and semantically transparent. They seem to grammaticalize¹² from the initial status of lexical items.

Opposite, we have bound roots that can be used both as right- and as left-constituent (like *-graph-*) and have the same meaning of the original neo-classical form and are only used with other bound roots. Their meaning can slightly change according to the contexts. For all these reasons, they can be morphologically considered nouns, thus lexical items.

I will discuss these two types of European semi-words in next paragraph.

4. Two Types of Semi-Words in European Languages

As illustrated in the previous paragraph and as proposed for Mandarin bound roots, semi-words in English and Italian can be divided into two groups according to three morpho-semantic aspects:

- a) The selection by a semi-word of bound/free morphemes;
- b) Their position of a semi-word in a compound;
- c) The change of meaning of a semi-word according to the context.

According to these three morpho-semantic criteria, semi-words can be divided into:

1. What I call “semi-nouns”: a large class of semi-words attaching to bound morphemes, both right and left headed; they change their meaning according to the context and they are used to create a scientific and technical vocabulary. Even if they are bound morphemes, these semi-words are lexical and not grammatical since they behave like nouns. Example of semi-nouns are *-fon-* ‘-phon-’ in Italian and *-graph-* from English. They are illustrated in (15).

¹² Cf. footnote 9.

(15) Italian

- a. [[Fono][log]_{+bound}ia] ‘phonology’
- b. [Micro]_{+bound} [fono] ‘microphone’

English

- c. [[epi]_{+bound} [graph]y] epigraphy
- d. [[grapho][log]_{+bound}y] graphology

2. What I call “semi-affixes”: a smaller class that includes semi-words selecting free morphemes; they are exclusively used as right (prefixes) or at the left hand side constituent (suffixes) and they always keep the same meaning. They differ from semi-nouns as they are exclusively used for the creation of a technical and scientific vocabulary. Examples of semi-affixes from Italian and English are presented in (16). Example (16b.) from Italian exists in English as well. The same for Italian with respect to the English example in (16c.). However, this is not valid for a. and d.: the word that corresponds to *-teca* in Italian does not exist in English¹³ and *-logia* ‘logy’ in Italian is used as a semi-noun¹⁴.

(16) Italian

- a. *-teca* (‘collection’):
it is used in words like [[panino]teca] ‘the place where to buy a sandwich’ (‘*panino*’). The word *panino* is used as a free form. *-teca* is nowadays used as a suffix.
- b. *Neo-* (‘new’):
it is used in words like [neo[laureato]] ‘a student (*laureato*) who has just taken his/her degree’. The word *laureato* is a free form. *Neo-* is nowadays used as a prefix.

English

- c. *Micro-* (‘very small’):
it is used in words like [micro[electronics]], [micro[economics]], [micro[cosm]] [micro[system]] where the words *electronics*, *economics*, *cosmos* and *system* are free forms. *Micro-* is nowadays a prefix.

¹³ This semi-word exists in other Germanic languages like Dutch (*-theek*) and German (*-thek*).

¹⁴ The process of grammaticalization of semi-nouns does not involve the same semi-words in Italian and English.

- d. *-logy* ('study'): it is used in words like *Islamology* 'the study of Islam', *Japanology* 'the study of Japan' where the words *Islam* and *Japan* are free forms. *-Logy* is nowadays a suffix.

To conclude, Italian and English semi-nouns are involved in neo-classical root compounding, semi-affixes are involved in derivation with prefixes and affixes. These two groups of semi-words share their origin from Latin and predominantly Ancient Greek free nouns. In spite of the same origin, it seems they do not share the same historical development: several semi-words underwent a process of grammaticalization that made them used as affixes in the modern language.

Despite this difference, they share a high frequency in morphologically productive processes of the modern language. As discussed before, the high frequency and the distinction of semi-words in two groups is not specific of European languages but it characterizes also a typologically and genetically non-related language, Mandarin Chinese.

I will compare Italian and English on one side and Mandarin on the other in the paragraph that follows.

5. Comparison Between Mandarin and European Semi-Words

A comparison between semi-words in Chinese and in European languages was proposed by Packard (1998) who suggests that bound morphemes in Chinese and bound roots in English (which he calls "Latin stems") have similarities. He considers for instance the English [-log-] (that has the meaning 'word, text'), both in word-initial (*logogram*, *logorrhoea*) and in -final position (*monolog(ue)*, *dialog(ue)*)¹⁵ and claims that it seems to behave like a Chinese bound root. In a later study, he also notices (2000:77) that these bound roots are visible not only in English but also in Italian; in both languages, they are morphemes with a lexical rather than a grammatical identity that cannot occur in a syntactic class slot until they are supplemented with additional morphological material that causes them to be 'completed' as words. For this reason, they are lexical rather than grammatical: in English and in Italian a bound root is clearly distinct from an affix. For instance, the English bound root [-*pathy*] is distinct from the suffix [-*ify*] and similarly the Italian bound root [-*grafia*] '-graph' is distinct from the agentive suffix [-*tore*] '-er' (e.g., *suonatore* 'player').

On the other hand, Packard (1998) also underlines the necessity of tracing a difference between European (Italian and English) and Chinese bound roots. First, unlike Italian, Mandarin bound roots may, and in fact do, form words by combining with other bound roots. Second, unlike bound roots in both Italian and English, Mandarin bound roots generally are less positionally restricted, i.e., they may in general occur as either

¹⁵ Cf. Packard (1998: 16, 17).

the first or second constituent of a word whereas in English and Italian a given bound root is generally restricted to occurring as either a left- or right- hand word constituent, but not both. Following Packard's ideas, we could not claim that Mandarin bound roots shows the same behaviour of English and Italian semi-words.

Nevertheless, on the basis of the data discussed above, I propose a different idea: Mandarin bound roots and European semi-words are similar also with respect to their possibility of combining with other bound roots and with respect to their position in the compound. Let's see some cases in (17): the Italian data in (17a.) and (17b.) falsify Packard's first claim and the examples in (17c.), (17d.) and (17e.) falsify his second claim¹⁶. Semi-words are labelled as sW.

(17) a. [[geo]_{sW} [logia]_{sW}]_N 'geology'

b. [[teo]_{sW} [crazia]_{sW}]_N 'theocracy'

Considering [-logo-] '[-log-]'

c. Grafologo/logopedista 'graphologist'/'logopedist'

Considering [-grafo-]

d. Calligrafia grafologia 'calligraphy'/'graphology'

Considering [-morpho-] '[-morph-]'

e. Morfologia/antropomorfo 'morphology'/'anthropomorphous'

The examples a. and b. in (17) present two cases of combination of two semi-words in Italian (and English); (17c. - e.) show that semi-words in Italian (and English) can occur both as a left- or right- hand word constituent. Thus, Mandarin bound roots are similar to English and Italian semi-words.

All considered and in the light of the latest observations, I propose that European "semi-words" resemble Mandarin bound roots on the basis of five similarities:

1. both derive from an old language: Chinese bound roots derive from Old Chinese monosyllabic roots, European bound roots derive from Greek and Latin roots.

¹⁶ See also their correspondent translated words in English in the glosses of the examples in (17).

2. both are still productively used: however, nowadays they are not independent (*i.e.*, free) but they are bound forms that productively create new independent (*i.e.*, free) words;
3. Sproat and Shih's proposal for Chinese bound roots can be adopted also for European neo-classical roots: according to their analysis, bound roots denote KINDS and root compounds denote the canonical instance of a kind. One can represent the relationship between a semi-word and a root compound containing this semi-word in terms of an IS-A hierarchy and the compound is a daughter node; the thick edge to the node the compound represents the fact that this node denotes the canonical instance of the kind. Other words in which the same semi-word is the head would also be representable as bearing an IS-A relation to the semi-word node;
4. two groups of semi-words have to be distinguished on the basis of morphological and semantic criteria: a group that is closer to the derivational usage ("semi-affixes"), a group that is closer to a compositional process ("semi-nouns"). With the first usage, both Mandarin and European semi-words are affixes derived from the grammaticalization of original nouns, are positionally restricted and select only certain semantic and morphological features. With the second usage, semi-words behave more like content words with a lexical meaning (*i.e.*, like nouns in nominal compounds).
5. both the original old meaning of European semi-words and that of Mandarin "semi-nouns" can change according to new contexts and in order to form a more technical vocabulary.

To conclude, Mandarin Chinese productively uses bound roots as both nouns and affixes for the creation of new words in the modern lexicon. Moreover, Chinese shares the morphological properties of these elements and the trigger of the morphological process they are involved in with several European languages.

6. Concluding Remarks

This paper investigates a word-formation process in typologically and genetically non related languages, Mandarin Chinese and European languages.

This word-formation process involves forms that are made up at least by a bound root that have been analysed as root compounds by Sproat and Shih (1996) or bound root words by Packard (2000) with respect to Chinese and as neo-classical compounds by Ralli (2005) or as "semi-parole" by Scalise (1994) with respect to English and Italian.

In this work I proposed a different analysis and I argued for treating all of them as similar even if they are used in typologically non-related languages. The reasons for this similarity are several: they all have an old origin, Italian and English "semi-words" derive from Latin and Greek, Mandarin bound roots derive from old Chinese monosyllabic free

forms. They all were free in the past and used as bound morphemes in the current language in order to create a new and often scientific vocabulary. In addition, in most cases, these semi-words undergo a change of meaning according to the contexts. Moreover, in both cases, semi-words do not constitute a close and unique class of words but have to be divided into two sub-classes on the basis of their position, the selectional property [+bound] and their change of meaning in a compound according to the context. More precisely, I claimed for distinguishing within the class of semi-words a class of semi-nouns, that are lexical elements that even as bound morphemes behave like compound constituents and a class of semi-affixes that are the result of a grammaticalization process which involved several semi-nouns; both in Mandarin and in Italian and English semi-words original semi-nouns shifted from lexical to grammatical elements, the former still used in compounding processes, the latter used in derivational. The reasons of this distinction and the usages of the two types are the same for Mandarin Chinese and for European languages probably because these semi-words have the same structure, the same distribution, the same trigger and both derive from previous stages of the language. Semi-affixes are less morphologically productive than semi-nouns.

Finally, as Sproat and Shih (1996) point out, the common Mandarin word related a root would typically denote the canonical instance of that kind. This analysis was elaborated for Chinese, but it seems to be good also for Italian and English semi-words.

The analysis of root compounding presented in this work aims to confirm that a productive morphological process exists in Mandarin Chinese and to describe in detail the morphologically different elements that take part to this phenomenon. In addition it proposes several similarities between Chinese and typologically non related languages like English and Italian.

REFERENCES

- ADAMS, VALERIE. 2001. *Complex Words in English*. Essex: Pearson Education Ltd.
- BAESKOW, HEIKE. 2004. Lexical Properties of Selected Non-native Morphemes of English. *Tübinger Beiträge zur Linguistik* 482.
- BAUER, LAURIE. 1983. *English Word Formation*. Cambridge: Cambridge University Press.
- CARLSON, GREGORY N. 1977. A unified analysis of the English bare plural, *Linguistics and Philosophy*, 1:3, 413-458.
- CHOMSKY, NOAM. 1993. A minimalist program for linguistic theory, In: K. Hale and S. J. Keyser (eds.) *The View from Building 20*. Cambridge, Mass: The MIT Press, pp. 1-52.
- CHOMSKY, NOAM. 1995. *The Minimalist Program*. Cambridge, Mass: The MIT Press.

- HARRIS, ALICE. C. 1990. Alignment typology and diachronic change. In: W.P. Lehmann *Language typology 1987: Systematic balance in language*, 67-90. Amsterdam: Benjamins.
- PACKARD, JEROME. 1998. *New Approaches to Chinese Word Formation*. Berlin: de Gruyter.
- PACKARD, JEROME. 2000. *The Morphology of Chinese*. New York/London, Cambridge University Press.
- RALLI, ANGELA. 2005. *Morphologia*. Athens: Patakes.
- SCALISE, SERGIO. 1984. *Generative Morphology*, Foris, Dordrecht.
- SCALISE, SERGIO. 1994. *Morfologia, le strutture del linguaggio*, Il Mulino, 1-331.
- SPROAT, RICHARD AND SHIH CHILIN. 1996. A Corpus-Based analysis of Mandarin nominal Root Compounds, *Journal of East Asian Linguistics* 5.1, 49-71.
- WARREN, BEATRICE. 1990. The importance of combining forms. In: Wolfgang Dressler et al. (eds) *Contemporary Morphology*. Berlin: Mouton de Gruyter.
- WASZINK, VIVIEN. 2006. Neo classical compounding. Paper presented at the International Conference on Historical Lexicology and Lexicography. University of Leicester.

