

# Modes of compounding in Bantu, Romance and Chinese

Bianca Basciano, Nancy Kula & Chiara Melloni

This article presents a cross-linguistic survey of endocentric root NN and exocentric VN compounds in a set of typologically and historically unrelated languages, with a special focus on Bemba (Bantu), Italian (Romance) and Mandarin Chinese.

The aim of the paper is twofold. From an empirical viewpoint, we intend to offer a contrastive analysis of comparable phenomena in a word-formation domain which has been neglected, especially in Bemba and to a lesser extent in Romance. Our study uncovers striking similarities between Bantu and Romance, hence reinforcing a connection independently established in other morpho-syntactic domains. Conversely, Mandarin strongly contrasts with Bantu and Romance and has more affinities with Germanic languages, in particular in the formation of NN compounds.

From a theoretical perspective, the overall picture of compounding emerging from our descriptive account strongly challenges a 'protolinguistic' (or syntax-free) view of these phenomena (contra Jackendoff and Progovac). Our cross-linguistic study reveals a number of empirical facts pointing to an underlying "syntax" of compounding both in the domain of NN and VN compounds, and shows that such underlying morpho-syntactic principles can be held responsible for the range of variation attested among these languages.\*

## 1. Introduction

In this paper, we develop a comparative analysis of complex nominal constructions in Bantu languages (with special reference to Bemba), in Romance (with special reference to Italian) and in Mandarin Chinese. Data from other languages, English and German in particular, are taken into account for a better understanding of the

\* We would like to thank all the members of the Research Network "Bantu, Chinese, Romance nouns and noun phrases" (2007-2010) for precious suggestions and discussion on several aspects of this work. We gratefully acknowledge the organizers and audience of the "Bantu, Chinese, Romance Nouns and Noun phrases Network Final Conference" (October 27-29, Leiden), where a previous version of this paper was presented. Further, we are indebted to the editors of this *IJL* issue, Lisa L.-S. Cheng, Denis Delfitto and Lutz Marten, for their detailed comments and suggestions on a previous version of this paper.

This research project was funded by the NWO Council for the Humanities (D36900: language and literature studies of other language groups).

properties of the constructions at issue. On the one hand, we intend to explore the established Romance-Bantu connection in a fairly unexplored domain, i.e. nominal compounding; on the other, we try to understand what the properties of compounding are in a highly isolating language such as Chinese and in what respects they show similarities/differences with respect to the other languages under examination.

Two major classes of compound nouns (NPs) in particular will be examined, as illustrated in the following table:

**Table 1.** NN and VN compounds in Bemba, Italian and Mandarin.

	BEMBA	ITALIAN	MANDARIN
[N+N] <sub>N</sub>	<i>shikùlú-bántù</i> 'elderly (distinguished) man'	<i>casa-famiglia</i> 'care home' lit. 'house-family'	鸡毛 <i>jī-máo</i> 'chicken-feather'
[V+N] <sub>N</sub>	<i>mùsóngá-nsàlà</i> 'appetizer' lit. 'trigger-hunger'	<i>apribottiglie</i> 'bottle-opener' lit. 'open-bottles'	领事 <i>lǐng-shì</i> 'consul' lit. 'lead-business'

The first group is represented by NN root compounds. The exploration of root compounding highlights many commonalities between Italian and Bemba, which strongly contrast with Mandarin. (Most) Italian and Bemba compounds are left-headed, disallow recursion (with the exception of coordinate compounds in Italian), have markers of inflection on the non-head (suffixes in Italian vs. prefixes in Bemba). Further, while the interpretation of Italian compounds is constrained along compositional criteria,<sup>1</sup> the interpretation of Bemba compounds is completely lexicalised/fossilised. It is noteworthy that these structural and interpretive properties of Romance (and Bantu) languages contrast to those of Germanic languages where compounding is right-headed, allows recursion and rules out inflection markers (English), or has only uninterpreted inflection on the non-head (German), while allowing free interpretation of the compound (see Delfitto and Melloni 2009). The present analysis shows that Mandarin resembles Germanic languages in this respect, the only difference consisting in the general lack of inflection markers, which is the hallmark of Chinese morphosyntax.

Furthermore, despite the afore-mentioned similarities between Bemba and Italian, a main difference is in the productivity of root compounding, which is a regular word formation phenomenon in the latter language (even though interpretively constrained) while being completely unproductive in the former. Bantu languages in general have a small set of NN compounds often restricted to kinship terms

and phytonyms (hence, relics of an ancestral lexicon), but, in synchrony, they tend to disallow the bare concatenation of two nouns, preferring other strategies of word formation (affixation or reduplication) over root compounding.<sup>2</sup> Again, Chinese is different in this respect from both languages because the productivity of the phenomenon is similar to that of Germanic languages, where NN compounding is by far the most productive word formation phenomenon.

The second class of compound nouns is represented by [VN] nominal constructions, which are attested in all the language groups and share comparable structural properties. In these compounds, N is the direct argument of V and can be a plural or mass noun (also singular forms are admitted, though generally less common). In Chinese, where number is not marked on nouns, N surfaces in its bare form. V surfaces as a root in Chinese, but it takes a final vowel not directly referable to any transparent derivational process in both Bantu and Romance. Since neither V nor N is the head of the construction, VN compounds are *exocentric* formations in these languages. Further, in contrast with the standard synthetic compounds attested in Germanic, these compounds do not have overt marking of nominalisation, apart from some cases of Bemba compounds taking overtly realized prefixes such as *ka-* and *mu-*. Though homophonous with Class Markers, such prefixes are *nominalising* prefixes. Our analysis is based on synchronic, diachronic and comparative evidence attesting that *ka-* and *mu-* behave differently from the homonymous class 12 and class 1 prefixes, and are instead derivational rather than inflectional affixes.

Chinese too, besides exocentric VN compounds, has a kind of compounds akin to synthetic compounds in Germanic languages, which are characterized by the presence of an overt nominalizing suffix or a nominal root head.

As far as interpretation is concerned, VN compounds convey an agentive meaning and they can refer to humans, animals and objects. Likewise root compounds, the main contrast between Bemba and Italian hinges on the (un)productivity of VN compounds, which are extensively attested and fully productive in Romance, while being restricted to a limited amount of idiomatized constructions in Bemba and many other Bantu languages (a similar kind of synthetic compounding is instead fully productive in other Bantu languages, such as Gĩkũyũ, see Mugane 1997 and Bresnan and Mugane 2006). In Chinese, VN compounding is widely attested but non very productive, while synthetic compounding seems to be a rather productive and little constrained phenomenon.

The paper is organized as follows: while section 2 focuses on NN compounds, with subsections focusing on each language (group), section 3 is entirely dedicated to VN compounds. General conclusions are drawn in section 4.

## 2. NN compounds

Root or primary compounding can be defined as the combination of two roots, stems<sup>3</sup> or fully inflected lexical items<sup>4</sup> which would otherwise be freestanding forms in a given language. This phenomenon is often contrasted with synthetic compounding since root compounds lack either a deverbal head or the argumental relation between the constituents, i.e. the defining properties of synthetic compounds (e.g. *truck-driver* vs. *Sunday driver*; see Selkirk 1982, Lieber 1983, a.o.).

Although several analyses have tried to reduce compounding phenomena to standard morphological or syntactic phenomena, recently, both classes of NN and VN compounds have been analysed as ‘protolinguistic fossils’, that is, as instances of a protolinguistic style of computation, which does not strictly obey morpho-syntactic principles.

Concerning (English) NNs, Jackendoff (2009: 113-114) asserts:

[...] compounding is actually not a *grammatical* phenomenon, but a *protogrammatical* one. Even the right-headedness of (English) compounds, their most grammatical feature, really only relies on a language specific correlation of linear order with semantic headedness, not on X-bar head-argument structure. In this respect it resembles the Agent-First and Focus-Last principles of BV. [...] In short, compounding is only barely syntactic.

In other words, Jackendoff claims that the underlying morpho-syntactic structure of compounds is only apparent and, instead, their structure merely reflects semantic/ (discourse) principles. Thus, this analysis – even though restricted to English data – should hold universally and apply cross-linguistically to the languages displaying comparable phenomena.

However, *contra* Jackendoff, we are going to discuss several empirical facts pointing towards an underlying morphosyntax of nominal compounding in the languages at issue.

### 2.1. NN compounds in Bantu

If Romance and Mandarin compounds have been subject to

several theoretical and empirical analyses, comparatively, NN compounding in Bantu languages has been largely neglected. The phenomenon, in fact, is even more constrained and unproductive than in Romance (notable exceptions are attested in some Bantu languages, e.g. Swahili). Therefore, let us start with a brief description of Bantu NN compound properties.

The general schema of Bantu root compounding conjoins two nominal stems, each preceded by its Class Marker (henceforth, CM), as shown in (1) (for discussion on CMs, see Crisma, Marten and Sybesma, this volume). Schadeberg (2003: 86-87) observes that most Bantu languages have compounds consisting of two nouns, with the second noun sometimes dropping its CM. The augment (or pre-prefix) only appears on the first element, where the syntactic/semantic context requires it.

- (1) [CM + N] + [(CM) + N]

Some standard examples in proto-Bantu are given below (proto-Bantu data in this section are from Schadeberg, *ib.*).

- (2) \**mbúlú-kútú* ‘an animal (?) + ear = ear’ (*kútú* > *kú-tú*)  
 (3) \**nkólo-tíma* ‘heart, breastbone + heart = breastbone’ (*tíma* > \**mu-tíma* ‘heart’)

In general, the nominal stems occurring in these complex forms tend to undergo grammaticalization. These morphemes, in fact, often occur as quasi-prefixes (without CM), in particular in the formation of complex kinship terms, person names and ethnonyms, or as quasi-suffixes, again typically employed in kinship terms (see the proto-Bantu data in 4 and 5).

- (4) \**na-* or \**nya-* ‘mother’ / \**ca-* or \**cí-* ‘father’  
 (5) \**-ntu* ‘person’, \**-lúme* ‘husband, male’, \**-kádí* ‘wife, woman, female’,  
 \**-jána* ‘child’,  
 \**-kúlú* ‘adult, old, big’.

The form \**-kádí* also appears as *kái>ké*. Some of these stems may combine with each other, as shown in the examples below.

- (6) \**mu-kái-ntu* ‘wife/woman + person = woman’  
 \**mu-kái-kúlú* ‘wife/woman + adult/old = old woman’  
 \**mu-kádi-jána* ‘wife/woman + child = girl, daughter in law’

Although most Bantu languages have complex nouns formed by the concatenation of two nominal stems, this pattern of compounding tends to be no longer productive and the attested compounds might have rather idiomatic meaning, as in the Chichewa examples below (Mchombo 2004: 117):

- (7) *bókó-munthu* ‘hippo’ + ‘person’ = ‘human hippo’  
*mnyanjá-sanga* ‘lakedweller’ + ‘savanna grass’ = ‘lake dweller lacking fishing expertise, like someone from the savanna hinterland’  
*mtóngá-chiwále* ‘a Tonga person’ + ‘palm fronds’ = ‘a Tonga person who lives in the palm fronds, i.e. not a real Tonga’

Bemba (a central Bantu language spoken in Zambia) does not represent an exception in this respect, since compounds involving nominal roots are rare and mainly consist of kinship terms and phytonyms with fixed/idiomatized meanings. Some examples are given in (8) and will be discussed in the next sections.

- (8) *nàkùlú-bántù* [1a]  
*nàkùlú* [1a] ‘elderly woman’ + *àbá-ntù* [2] ‘people’ = ‘elderly (distinguished) woman’

*nókó-lùmé* [1a]  
*bà-nókó* [1a] ‘your mother’ + *ùmú-lùmé* [1] ‘husband’ = ‘uncle (on mother’s side)’

*ìmpùndù-búsùshí* [9/10]  
*ìn-pùndù* [9] ‘fruit’ + *ìcí-sùshí* [7] ‘fart’ = ‘fruit’

## 2.2. NN compounds in Romance

NN compounding in Romance is a common strategy of word formation, though it is more constrained than in Germanic (see Delfitto and Melloni 2009). It is acknowledged that Germanic compounds are often expressed in Romance languages by ‘prepositional compounds’, i.e. complex structures conjoining two nouns by means of a preposition-like element (see Delfitto, Fiorin, Kula, this volume), as in (9).

- (9) a. *bread knife* = It. *coltello da pane*, lit. knife + P + bread  
b. *country house* = Sp. *casa de campo*, lit. house + P + country

Prepositional compounds will not be considered in this account since they display mixed lexical/syntactic properties, hence they

are not directly comparable to the NN compounds of the other languages at issue. Keeping such complex words out of the picture, the occurrence of NN compounds in Romance is certainly lower than in Germanic languages. Interestingly, recent studies reveal that the formation of root compounds in Romance is restricted by semantic constraints, explaining the apparent low productivity of the phenomenon in these languages (see, in particular, Delfitto and Melloni 2009).

Structurally, the general schema comprises the combination of two nominal stems, each followed by its final vowel, that is an obligatory inflectional ending expressing gender and number (even though several analyses assign to this vowel the function of marking the declension class rather than expressing phi-features).

(10) [N +fv]<sub>N</sub> + [N+ fv]<sub>N</sub>

Some Romance NN compounds are shown in (11).

- |          |                       |   |
|----------|-----------------------|---|
| (11) It. | <i>discorso fiume</i> | ‘speech river = long speech’                |
| Fr.      | <i>bateau-mouche</i>  | ‘boat + fly = excursion steamer’            |
| Pt.      | <i>palavra-chave</i>  | ‘word + key = key word’                     |
| Sp.      | <i>falda pantalón</i> | ‘skirt + trousers = pantskirt’ <sup>5</sup> |

### 2.3. NN compounds in Chinese

In Mandarin Chinese, where compounding is the most productive means of word formation (approximately 80% of Chinese words are compounds; Xing 2006: 117), NN compounding is a very productive and creative process: the semantic relations between the constituents cannot be exhaustively listed and novel compounds can be created at will (Li and Thompson 1981; see the discussion below). NN compounds are formed conjoining two bare nominal stems (12); these are generally: roots (13a-b), either free or bound; compound nouns, as in (13c), where the first constituent, 斑马 *bān-mǎ* ‘zebra’, is a compound (‘stripe-horse’). Sometimes, the constituents can contain affixes, as in (13d), where the first constituent, 果子 *guǒzi* ‘fruit’, is formed by the root 果 *guǒ* ‘fruit’ plus the empty (non-productive) nominal suffix 子 *-zi*.

(12) [N + N]

- |         |                   |                                       |
|---------|-------------------|---------------------------------------|
| (13) a. | 鸡毛                | ‘chicken + feather = chicken feather’ |
|         | <i>jī-máo</i>     |                                       |
| b.      | 灯船                | ‘light + ship = lightship’            |
|         | <i>dēng-chuán</i> |                                       |

- |    |                         |                                       |
|----|-------------------------|---------------------------------------|
| c. | 斑马鱼<br><i>bānmǎ-yú</i>  | ‘zebra + fish = zebra-fish’           |
| d. | 果子酒<br><i>guǒzi-jiǔ</i> | ‘fruit + liquor = fruit-based liquor’ |

#### 2.4. NN compounds in Germanic

NN root compounding is extremely productive in Germanic languages, where it has been extensively studied from several viewpoints and framed in different models and research fields. The root compounding pattern consists of the concatenation of two nominal stems, as schematically represented below:

- (14) [N + (LE)] + N]

As we will show in the sections below, a formal peculiarity of Germanic compounds is the presence of the so-called linking element, LE, (or *Fugenelement*), i.e. a declension marker which often coincides with the nominative plural or the genitive singular or plural of the first constituent. This element is practically absent in English, where a (genitival) *s* only appears in fossilized items - see *doomsman* among the data below - and where the stems conjoined in the compound are often bare roots.

- |          |   |                 |               |
|----------|---|-----------------|---------------|
| (15) De. | <i>Hund-e-futter</i>                    | dog+LE+food     | ‘dogfood’     |
| Du.      | <i>boek-en-kast</i>                     | book+LE+case    | ‘bookcase’    |
| En.      | <i>doom-s man</i><br><i>table-cloth</i> | doom+LE+man     | ‘pessimist’   |
| No.      | <i>arbeid-s-dag</i>                     | work+LE+day     | ‘working-day’ |
| Sw.      | <i>jord-a-färd</i>                      | earth+LE+voyage | ‘burial’      |

#### 2.5. Structure and interpretation of NN compounds in a comparative perspective

In what follows, we will compare the main structural and semantic properties of NN compounds in Italian, Bemba and Mandarin, focusing, in particular, on the properties of Bemba compounds and taking into account the acknowledged properties of Germanic compounds, which strongly parallel those of Mandarin NN compounds.

##### 2.5.1. Head

Structurally, Bemba compounds resemble Romance compounds in being left-headed. The following examples show that the CM of the left

member determines the agreement pattern of the whole noun phrase in Bemba, showing that the grammatical properties of the compound (CM in Bemba and gender in Italian) depend on those of the leftmost element. Semantically, the left member is the hyperonym of the compound, ensuring the correspondence of structural and semantic headedness (i.e., a *ímpùndù-búsùshí* ‘fart fruit’ IS A kind of *ímpùndù* ‘fruit’, as well as It. *uomo pesca* ‘man peach’ IS A *uomo* ‘man’).

(16) Be. *ímpùndù-búsùshí* (9/10)

*ín-pùndù* (9) ‘fruit’ + *ící-sùshí* (7) ‘fart’ = ‘fruit’

Agreement:

*ímpùndù-búsùshí íshí-sùmá*

‘fruit’ (9) + ‘fart’(7)‘nice’(9), ‘the nice fart fruit’

(17) It. *uomo pesca*

Agreement:

*il famoso uomo pesca*

the<sub>masc</sub> famous<sub>masc</sub> man<sub>masc</sub> peach<sub>fem</sub> ‘the famous peach-man’

Chinese illustrates the opposite situation: the head, in fact, is on the right, at least if we look for the element that is the hyperonym of the compound. Formally, neither of the morpho-syntactic features hitherto considered (noun class, gender and number) is marked on the compound members. However, in cases where the two noun constituents have different classifiers, the whole compound tends to take the same classifier as the head noun, which proves the matching of formal and semantic head.

(18) a. 一 根 鸡毛

*yī gēn jī-máo*

one CLF chicken (CLF: 只 *zhī*) - feather (CLF: 根 *gēn*)

‘one/a chicken-feather’

b. 一 条 斑马鱼

*yī tiáo bānmǎ-yú*

one CLF zebra (CLF: 匹 *pǐ*) - fish (CLF: 条 *tiáo*) ‘one/a zebra-fish’

c. 一 双/只 草鞋

*yī shuāng/zhī cǎo-xié*

one CLF straw (CLF: 根 *gēn*) - shoe (CLF: 双 *shuāng*/只 *zhī*)

‘a pair of/one straw sandal(s)’

d. 一 盏 车灯

*yī zhǎn chē-dēng*

one CLF vehicle (CLF: 辆 *liàng*) - light (CLF: 盏 *zhǎn*)

‘one/a vehicle headlight’

e. 一 艘 灯船

*yī sōu dēng-chuán*

one CLF light (CLF: 盏 *zhǎn*) - ship (CLF: 艘 *sōu*) ‘one/a lightship’

Germanic is like Chinese in this respect: the agreement pattern of Germanic compounds proves their structural right-headedness. In this case too, there is a matching of syntactic and semantic head (see the German example below).

- (19) De. Hundefutter  
 Agreement: ein gutes Hund-e-futter  
                   one<sub>neut</sub> good<sub>neut</sub> dog<sub>masc</sub>-LE- food<sub>neut</sub>  
                   ‘a good dog-food’

### 2.5.2. Inflection markers

Another structural property shared by the Romance and Bantu NN compounds is the presence of inflection markers. Bantu compounds contain the CM – usually associated with Gender (and Number) (see Corbett 1991, Carstens 2008, Crisma, Marten and Sybesma this volume, a.o.) – of both constituents, even though the prefix of the non-head may be dropped or modified in virtue of phonological processes.

- (20) Be. *shíkùlù-bántù* (1a) ‘elderly (distinguished) man’  
*shíkùlù* (1a) ‘elderly man’ + *aba-ntu* (2) ‘people’

In Romance compounds, besides a root and one or more derivational affixes, we find an inflectional morpheme too, more precisely, a final vowel, which – although subject to many different analyses<sup>6</sup> in the literature – is generally put in relation with gender/number features.

- (21) It. *uom-o ran-a* ‘frogman’  
*tren-o merc-i* ‘train, goods = freight train’  
*regist-i attor-i* ‘directors, actors = actor directors’

Although Bantu CMs have been analysed as Gender (and Number) markers in previous literature (but see Crisma, Marten and Sybesma, this volume, for discussion), an important distinction with respect to Romance final vowels must be drawn, because it is not possible to establish a straightforward mapping between these morphemes and gender features in the Romance languages at issue. This fact is due to the existence of *declension classes* in Romance (and most I.E. languages), which are often analysed as abstract features idiosyncratically associated with nominal lexemes and determining their declensional patterns. These declensional paradigms often obscure the relation between morphosyntactic features such as Gender/Number and their phonological expressions (see Acquaviva 2009 for a recent analysis of Italian Declension Classes).<sup>7</sup>

On the grounds of these data, we argue that the compound mem-

bers of Bantu and Romance compounds are morphosyntactically full nouns. However, a crucial observation on Bemba compounds concerns the absence of the augment on the non-head (see 22a). In fact, the augment can only be preposed to the whole compound, i.e. prefixed to the head member (see 22b).

- (22) a. Be. *shíkùlù-(\*a)-bántù*  
b. Be. *ĩ-mpùndũ-búsùshĩ*

Although a unified analysis of the augment is still pending, the lack of this morpheme is not surprising assuming that it is arguably related to the D position or to the left periphery of the sentence, thus escaping the NP domain of NN compounding (see Visser 2007, de Dreu 2008, Buell 2009). In other words, the lack of augment in Bemba compounds can be compared to the lack of determiners inside both Romance and Germanic compounds.

- (23) It. \**uomo la rana*

- (24) En. \**[the frog] man*

Contrary to what happens in Bemba and Italian, Mandarin compounds completely lack inflection markers. Like the other languages at issue, in Mandarin determiner-like elements cannot appear inside the compound; e.g. the first constituent cannot be modified by a sequence ‘demonstrative+classifier’, as shown by the ungrammatical example below, paralleling the data in (23) and (24).

- (25) Ch. \**[这只鸡]毛* \**[zhè zhī jī] máo* ‘*[this CLF chicken] feather*’

It is worthwhile observing that the lack of inflection markers is the only big difference between Chinese and Germanic languages. While NN compounds in German and Dutch, which are highly inflecting languages, especially in the nominal domain, exhibit so-called Linking Elements,<sup>8</sup> i.e. case and number suffixes on the non-head (see the morpheme *-e-* in *Hund-e-futter*), Chinese and English, given their isolating morphology, display bare concatenation of two nouns.

### 2.5.3. Recursion

A final note on the structure of root compounds takes into account recursion, i.e. a property that, notoriously, characterizes Germanic compounds with their complex subordinative structures, see the German and English examples below.

- (26) Donau dampf schiff fahrt s gesell schaft s kapitän s mütze  
 Danube steam ship journey LE journeyman SUFF LE captain LE cap  
 ‘Cap of the captain of the Danube steam ship company’
- (27) a. Volume Feeding Management Success Formula Award  
 b. winter weather skin troubles  
 c. health management cost containment services

While head and complement recursion in Romance is restricted to NN coordinative and some cases of subordinative NN compounds (see It. examples in 28), and it is in general highly constrained (see 29), in Bemba recursion is virtually impossible (see 30).

- (28) *bar ristorante pizzeria* ‘pizzeria bar restaurant’  
*capo [ufficio acquisti]* ‘purchase office manager’
- (29) \**vacanza [studio pilota]* vacation study pilot  
 \*[*viaggio lampo*] *chiave* travel lightening key
- (30) \**shíkùlú ìmpùndù-búsùshí* ‘elderly man’s fart fruit’

Examples (31) and (32) show Bemba complex nouns formed by a deverbal head (*kásùngà* and *kásàbà*) taking as their complement an attested root compound. These are not recursive root compounds, but instances of synthetic compounding, i.e. a phenomenon attested in Bemba (see *infra*). Although they might be assigned semantically plausible interpretations, these compounds, i.e. recursive structures where root compounds are the non-head of synthetic compounds, are completely out.

- (31) \**kásùngà [shíkùlú-bántù]*  
 (1a) keeper elderly man
- (32) \**kásàbà [ìmpùndù-búsùshí]*  
 (1a) picker fruit

Recursion is another property that sets apart Bemba and Italian from Chinese. Again, Chinese patterns with Germanic languages in allowing free recursion in compound nominal structures.

- (33) 房地产企业管理信息服务 *fángdìchǎn qǐyè guǎnlǐ xìnxī fúwù*  
 ‘real-estate enterprise management information service’

- (34) 人民政府台湾事务办公室主任职务 *rénmín zhèngfǔ Táiwān shìwù*  
*bàngōngshì zhǔrèn zhíwù*  
the-people government Taiwan general-affairs office head post/job  
'the position of head of the local people's government office for  
Taiwan affairs'

#### 2.5.4. Semantics

To conclude this comparative overview, we take into account the semantic properties of NN compounds. It is commonly acknowledged that Germanic root compounds have arbitrary readings because they admit any logically plausible semantic relation between the constituents (including any meaning relation that can be established on the grounds of the context of utterance).

- (35) *Fischfrau* 'fish + woman' (Heringer 1984)
- a. woman that sells fish
  - b. woman that has brought fish
  - c. woman standing close to fish
  - d. woman eating fish
  - e. woman looking like a fish
  - f. spouse of a fish
  - g. woman and fish at the same time (i.e. mermaid)
  - h. woman having Pisces as zodiac (German *Fisch*)
  - i. woman as cold as a fish, etc.

In Romance, instead, the meaning is restricted along strictly compositional criteria, related to the predicative structures encoded in Pustejovsky's Qualia Structure (see Pustejovsky 1995) and, specifically, in the Formal Quale (see Delfitto and Melloni, 2009). The mechanism that allows the interpretation of root compounds in Romance is constant-variable unification in the predicative structure encoded in the head noun Formal Quale (ranging over several formal properties such as colour, dimension, form, etc.). Let us consider *uomo albero*, a novel compound denoting a man affected by a malformation of hands and feet that make them resemble the roots of a tree.

- (36) It. *uomo albero* 'man + tree = man looking like a tree, i.e. sharing some distinctive, formal properties of a tree'

In plain words, the non-head features a salient interpretive property of the head. In Qualia Structure terms, the non-head *albero* provides the predicative constant according to which the variable in the Formal Quale of the head noun can be specified.

The few NN compounds attested in Bantu have instead a highly lexicalized meaning, not always transparently related to the constituent meaning. Their semantic opacity is due to the lexicalization phenomena typically occurring within the ancestral lexicon and is thus not particularly relevant to determine the parametric variation we have highlighted so far.

The interpretive relations in Chinese NN compounds do not display the strong restrictions detected in Romance. If lexicalization/idiomatization of compound forms is a frequent phenomenon, Li and Thompson (1978) state that Chinese is very similar to Germanic in that nominal compounds can be created at will. Li and Thompson (1981: 49-53) try to single out the most common semantic relations between the constituents of a nominal compound:<sup>9</sup>

**Table 2.** Semantic relations in Chinese NN compounds.

SEMANTIC RELATION	EXAMPLE
N <sub>1</sub> denotes the place where N <sub>2</sub> is located	床单 <i>chuáng-dān</i> 'bed + sheet = bed sheets'
N <sub>1</sub> denotes the place where N <sub>2</sub> is applied	唇膏 <i>chún-gāo</i> 'lip + ointment = lipstick'
N <sub>2</sub> is used for N <sub>1</sub>	枪弹 <i>qiāng-dàn</i> 'gun + bullet = bullet'
N <sub>2</sub> denotes a unit of N <sub>1</sub>	铁原子 <i>tiě-yuánzǐ</i> 'iron + atom = iron-atom'
N <sub>2</sub> denotes a protective device against N <sub>1</sub>	太阳镜 <i>tàiyáng-jìng</i> 'sun + lens/glass = sun glasses'
N <sub>2</sub> denotes a piece of equipment used in a sport, N <sub>1</sub>	乒乓球 <i>píngpāng-qiú</i> 'ping pong + ball = ping pong-ball'
N <sub>2</sub> is caused by N <sub>1</sub>	油迹 <i>yóu-jì</i> 'oil + mark/trace = oil stains'
N <sub>2</sub> denotes a container for N <sub>1</sub>	书包 <i>shū-bāo</i> 'book + bag = schoolbag/satchel'
N <sub>1</sub> and N <sub>2</sub> are parallel	花木 <i>huā-mù</i> 'flower + tree = vegetation'
N <sub>2</sub> denotes a product of N <sub>1</sub>	蜂蜜 <i>fēng-mì</i> 'bee + honey = honey'
N <sub>2</sub> is made of N <sub>1</sub>	草鞋 <i>cǎo-xié</i> 'straw + shoe = straw-shoe'
N <sub>2</sub> denotes a place where N <sub>1</sub> is sold	药店 <i>yào-diàn</i> 'medicine/drug + shop = drug-store'
N <sub>2</sub> denotes a disease of N <sub>1</sub>	肺病 <i>fèi-bìng</i> 'lung + disease = tuberculosis'
N <sub>1</sub> denotes the time for N <sub>2</sub>	东夜 <i>dōng-yè</i> 'winter + night = winter-night'
N <sub>1</sub> is the source of energy of N <sub>2</sub>	电灯 <i>diàn-dēng</i> 'electricity + lamp = electric lamp'
N <sub>1</sub> is a metaphorical description of N <sub>2</sub>	龙船 <i>lóng-chuán</i> 'dragon + boat = dragon boat'
N <sub>2</sub> is a component of N <sub>1</sub>	鸡毛 <i>jī-máo</i> 'chicken + feather = chicken-feather'
N <sub>2</sub> is a source of N <sub>1</sub>	水源 <i>shuǐ-yuán</i> 'water + source = headwaters/ source of water'
N <sub>2</sub> is an employee or an officer of N <sub>1</sub>	公司经理 <i>gōngsī-jīnglǐ</i> 'company + manager = company-manager'
N <sub>1</sub> denotes a proper name for N <sub>2</sub> , which may be a location, an organization, an institution or a structure	北京大学 <i>Běijīng dàxué</i> 'Beijing + university = Beijing University'
N <sub>2</sub> denotes a person who sells or delivers N <sub>1</sub>	盐商 <i>yán-shāng</i> 'salt + merchant = salt merchant'

However, the list above cannot be exhaustive, since the freedom in the interpretation of Chinese compounds is such that it is possible to build novel compounds on the fly by assigning an interpretive relation between the two Ns that only holds in the context of utterance, in full agreement with what has been pointed out by Downing (1977) for English NN compounds:

- (37) *apple-juice seat*, i.e. the seat in front of which an apple-juice had been placed  
*bike girl*, i.e. a girl who left her bike in the vestibule (Downing 1977)

### 2.5.5. *Interim conclusions*

To conclude, an overview of the structural and semantic properties of root compounds discussed in this section is given in the table below.

**Table 3.** NN properties in Bemba, Italian and Mandarin.

	BEMBA	ITALIAN	MANDARIN
HEAD	Left	Left	Right
RECURSION	No	Constrained (coordinates)	Unconstrained
INTERNAL INFL.	Class Prefixes (Gen/N)	Word Markers	-
INTERPRETATION	Fixed meaning	Constrained	Unconstrained
PRODUCTIVITY	No	Yes (semantic constrains)	Yes

While Italian and Bemba data show many commonalities, the analysis of Mandarin NNs reveals opposite features along the structural, semantic and productivity dimensions. Mandarin, in fact, patterns with Germanic in many respects, apart from the occurrence of inflection markers inside the compound. However, the complete lack of productivity of the phenomenon in Bemba represents a relevant point of divergence with respect to the other languages at issue, where the phenomenon is still productive.

Some remarks can be drawn on the basis of the empirical facts described so far.

Root compounding phenomena can hardly be analyzed as a-syntactic or a-grammatical word formation patterns, as put forward by Jackendoff (see section 2). On the grounds of what is found in these languages, it clearly emerges that headedness is a strictly grammatical phenomenon affecting not only the semantics but also the structural properties of the compound word. The position of the head determines the agreement pattern of the DP even in isolating languages such as Chinese, which displays syntactic evidence in support of this claim (i.e. the choice of the classifier).

Mandarin and Germanic show that right-headedness patterns with freedom in interpretation, a property that seems to go hand in hand with recursion and (robust) productivity of root compounding. Bemba and Italian, on the other hand, show that left-headedness is related to a constrained interpretation and low or null productivity.

Significantly, not only structural headedness, but also the different degree of productivity and the presence of interpretive constraints that depend on the internal structure reveal that compounding cannot be the result of direct interface between semantics and phonology, as put forward by Jackendoff. In particular, no phonological or semantic constraints can explain the lack of compounding in Bemba. Also, the limited productivity of root compounding in Romance has been related to semantic constraints arguably reducible to the feature endowment (declension class and gender) of the lexical items and not to independent LF constraints on the interpretation of these structures (for an analysis along these lines, see Delfitto, Fábregas and Melloni 2008, forthcoming). A non proto-linguistic standpoint becomes more compelling if we consider recursion, which is banned in most NN compounding in Italian, even though recursive compounds would be compositionally interpretable (see 29). Hence, recursion is arguably restricted by syntactic rather than semantic constraints in Romance languages. We thus propose that morphosyntactic principles must be called into play for explaining the attested variation. However, a detailed analysis of NN compounding would exceed the limits of the present contribution, whose purpose is to offer a preliminary and comparative overview of two distinct classes of compounding phenomena in typologically unrelated languages. We leave more detailed language-specific studies for future investigation.

### *3. VN compounds cross-linguistically*

VN compounds represent a challenging phenomenon for contemporary grammar theory due to several reasons. First, they are usually defined as exocentric because they lack a structural and semantic head, thus they defy the endocentricity principle at the basis of Merge (see Chomsky 1995). Further, although in most cases the V constituent seems to select its direct object (N), assigning it a Theme/Patient role, instances can be found where N is not a direct object but an adjunct/complement expressing various relations with V (in particular, locative/temporal relations). Finally, the form of the verb is not clearly defined since it has been alternatively analyzed as

a stem or theme (a root taking a thematic vowel), as an inflected form (an imperative, in particular), as a bare root, etc. (see *infra*), making a uniform morphotactic analysis of its constituent morphemes quite difficult.

The peculiar properties of VN compounding seem to be shared cross-linguistically, and Bantu languages too (at least, those exhibiting this phenomenon) manifest similar properties. But let us start considering some English examples, illustrated in (38).

(38) *pickpocket, turncoat, daredevil, hunchback, wagtail, tattletale, sawbones, cutthroat, Shakespeare*

These data constitute a representative sample of the few VNs attested in contemporary English. Indo-European languages in general attest the existence of these compounds, even though they tend to be no longer productive in most Germanic and Slavic languages. In these languages, in particular, VN compounding has been replaced by synthetic compounding, i.e. a more complex morphosyntactic structure including an additional layer of structure, instantiated by the nominalising affix, and featuring incorporation, as signalled by the reversed N-V order of the compound constituents (see Progovac 2009).

Progovac (2006), who offers a suggestive analysis of VNs, claims that “exocentric compounds used to be productive, but are now, in most languages, only preserved as unproductive fossils, mostly in names, nicknames, and in derogatory expressions which probably derive from nicknames [...]”. Her analysis, building on Jackendoff (2002 and 2009), treats VN compounds as relics of a proto-linguistic style of computation, which is no longer active in modern language. Progovac grounds her account on some peculiar properties of this class of compounds.

First, the (alleged) exocentricity of these compounds seems to suggest a simpler syntax, reflected in the non-hierarchical structure and absence of functional categories in these complexes. From this, other properties follow, such as the ‘vague thematic role assignment’ invoked by Progovac to explain the fact that N can bear other thematic roles than Theme/Patient. The second point raised by Progovac is the imperative form of V. Building on several classes of data (see the Slavic data in 39), she argues that V is an imperative form, i.e. what she claims is a protolinguistic verb form (Rolfe 1996), as also suggested by acquisition data (the imperative mood is among the first productive verbal forms used by young children, see a.o. Bar-Shalom and Snyder 1999).

(39) Se.	<i>ispi-čutura</i>	‘empty-flask = drunkard’
	<i>probi-svet</i>	‘break-world = wanderer’
Po.	<i>goli-broda</i>	‘shave-beard = barber’
	<i>rzezi-mieszek</i>	‘cut-purse = pickpocket’
Ru.	<i>sorvi-golova</i>	‘cut-off head = daredevil’
	<i>verti-hvostka</i>	‘wag-tail = a bird’
Mc.	<i>isturi-čorba</i>	‘stick-out broth = tactless person’
	<i>zajdi-sunce</i>	‘set-sun = sunset’

Another significant property of VNs concerns their vocabulary. While V often denotes concrete actions and N refers to body parts and other ‘basic/core’ lexical items, VNs are often derogatory, coarse and vulgar and/or playful/humorous (see Marchand 1969), in accordance with the hypothesis that protolanguage features a very basic, concrete vocabulary (see Progovac 2006).

On the grounds of these morpho-syntactic and lexical-semantic properties of VNs, Progovac (2009b) concludes that

(t)o the extent that we can find ‘living fossils’ of simpler stages of syntax in present-day languages (see e.g. Bickerton 1990, 1998, Jackendoff 1999, 2002, Progovac 2008, 2009), VN compounds are the best candidate. Little about these compounds makes sense except in the light of evolution.

Her analysis is strongly supported by the lack of productivity of this phenomenon in the languages at issue, so that she argues “VN compounds are now on their way to extinction.”

A notable exception, however, is represented by Romance languages, where VN compounding is a productive phenomenon, synchronically employed for the formation of agent and instrument nouns. Romance VNs will be examined in section 3.2.

The phenomenon is also attested in several Bantu languages and Chinese. In what follows, we will examine a heterogeneous class of VN compounds in Bemba (see 3.1), as well as in Italian and other Romance languages and in Mandarin Chinese, trying to identify the properties which VNs share in these languages (see 3.2 and 3.3).

### 3.1. VN compounds in Bantu / Bemba

VN compounding is an attested phenomenon in Bantu and usually exhibits the following structure:

(40) (CM/prefix) + [ $\sqrt{V}$  + fv] + [(CM) +  $\sqrt{\text{Noun}}$ ]

VNs are formed conjoining a verbal stem, i.e. a verb root followed by a final vowel (formally matching the vowel attested in fully inflected verbs), with a nominal stem, i.e. a noun element which might retain (Chichewa, Bemba) or drop (Swahili) its CM. Since the output of this operation is a noun, the compound is usually preceded by a CM expressing the ‘gender’ of the whole compound.

These compounds are attested in several Bantu languages. In Chichewa, for instance, Mchombo (2004) observes that the commonest form of compounding takes a verb and its unmodified object noun or locative noun to create a noun by adding a prefix (/CM):

- (41) *ph-a dzúwa* ‘kill the sun’ → *chi-phadzúwa* ‘beautiful woman’  
*sw-a bumbu* ‘smash vulva’ → *chi-swábumbu* ‘vulva-breaker  
(large penis)’  
*tol-a nkhâni* ‘pick up news’ → *m-tolankhâni* ‘reporter’  
*pal-a matábwa* ‘scrape timber’ → *m-palamatabwa* ‘carpenter’  
*low-a m’málo* ‘enter in place’ → *m-lowammalo* ‘substitute, pronoun’  
*gon-á m’báwa* ‘sleep in bar’ → *chi-gonambáwa* ‘a drunk, an alcoholic’  
(Mchombo 2004: 117)

In Swahili too, a common pattern of compounding is the formation of agentive or instrumental VN nouns preceded by a noun class prefix:

- (42) *m-pita*      *njia*      ‘passerby’ (lit. ‘street-passer’)  
CM1-pass      street  
*ki-choma*      *mguu*      ‘herb with barbered seed’ (lit. ‘foot-piercer’)  
CM7-pierce      foot/leg  
(Contini Morava 2007: 1131)

Other instances of VN compounds can be found in Ganda and Gikũyũ, among other Bantu languages (see Schadeberg 2003). In this paper we only focus on Bemba, whose compounding phenomena manifest similarities with Romance VNs and Germanic synthetic compounding.

Bemba has several instances of VN compounding, but, as in the case of NN compounds, such phenomenon is no longer productive. As shown in Kula (forthcoming), VN compounds can take the following prefixes:

- (43) *mu-* / *ka-* (1a)      (‘one who’)  
*aka-* (12) / *ici-* (7)      (‘thing that’)  
*ulu-* (11)  
*n-* (9)

The following data constitute a representative sample of Bemba VNs and are grouped on the basis of the CM they take:

- (44) *mu-* [1a]  
a. *mùsùngá-bántù* [1a] ‘hospitable person’  
*ùkú-sùngà* ‘keep’ *aba-ntu* [2] ‘people’  
b. *mùsóngá-nsàlà* ‘appetizer’  
*ùkú-sóngà* ‘whet’ *ín-sàlà* [9] ‘hunger’  
c. *mùpètá-ndùpé* ‘kind of shrub for making baskets’  
*ùkú-pètà* ‘to fold’ *ùlú-pé* [11] ‘kind of basket’  
d. *mùkwíílílábwàmbà* ‘person who works only when necessary’  
*ùkú-kwíílà* ‘to earn a living’ *ùbú-àmbà* [14] ‘nakedness’
- (45) *aka-* [12]  
a. *akààkántélé-màfwéésà* [12] ‘insect that lives under rocks’  
*ùkú-sélà* ‘to move’ *àmá-fwéésà* [6] ‘rocks’  
b. *àkàshìndá-béèni* [12] ‘small stinging insect’  
*ùkú-shìndà* ‘to sting’ *àbá-ínì* (*abeeni*) [2] ‘visitors’  
c. *àkápàlá-pùté* [12] ‘pimple’  
*ùkú-pàlà* ‘to resemble’ *ìcì-pùté* [7] ‘boil’
- (46) *ici-* [7]  
a. *ìcìlèngwá-lésà* [7] ‘abnormal person’  
*ùkú-léngà* ‘to draw’ *lésà* [1a] ‘God’  
b. *ìcìsénsé-ngàndà* [7] ‘cricket (insect)’  
*ùkú-sénsà* ‘to make a noise’ *ín-ngàndà* [9] ‘house’
- (47) *ka-* [1a]  
a. *kámíná-mísà* [1a] ‘drunkard’  
*ùkú-mínà* ‘to swallow’ *ímí-sà* [4] ‘gulps’  
b. *kàténshá-màbúúlà* ‘authoritative person’  
*ùkú-téntà* ‘to shake’ *àmá-búúlà* [6] ‘leaves’
- (48) *n-* [9]  
a. *ntíntá-mùkòshí* [9] ‘tattoo marks along spine/neck’  
*ùkú-tíntà* ‘to pull’ *ùmú-kòshí* [3] ‘neck’  
b. *ntùnká-máfí* [1a] ‘dung beetle’  
*ùkú-tùnkà* ‘to push’ *àmá-fí* [6] ‘faeces’

(49) *lu-* [11]

*lúbùmbá-nóngó* [11] ‘mud-building hornet’  
*úkù-bùmbà* ‘to mould’ *ín-óngó* [9] ‘clay pot’

These compounds, in fact, do not represent a uniform class, since they display different structural properties, which will be discussed in the remainder of this paper. In the next sections, we present the analysis of Bemba VNs with a description of their general properties.

### *3.1.1. Properties of Bemba VN compounds*

#### *3.1.1.1. Interpretation*

From the interpretive point of view, Bemba VN compounds follow the pattern described in section 3.1. In fact, the Bemba examples in the preceding section show that the interpretation of these compounds is agentive, even though their meaning is always highly lexicalised. The data show that VNs can refer to humans (44a), animals (45a) and objects (44b). To the best of our knowledge, there are no VNs with eventive or locative meanings, which are instead attested in Romance and other Bantu languages (see *infra*).

#### *3.1.1.2. The verb constituent*

The V constituent in VNs is typically bimorphemic, specifically, composed of a verb root and a final vowel, most often (high-toned) *-á-*. The role of this element is not uniform across Bantu languages; however, it can be argued that *-a-* generally marks the indicative in most of the languages of this family (*-a-* is defined by Nurse 2008 as ‘neutral vowel’ and contrasted with subjunctive *-é-*). The same vowel is high-toned when it signals second person imperative. Therefore, since *-á-* is the final vowel found in VNs, the imperative analysis suggested by Progovac for I.E. languages is a plausible analysis in Bemba too. However, more fine-grained morpho-phonological investigation is needed in order to ascertain whether the high-tone on the final vowel is instead a prosodic feature resulting from a readjustment rule in the VN template.

Further, between the root and the final vowel a derivational morpheme can be found, which modifies verb arity, as in the case of applicative constructions. Some cases of arity-changing morphemes are found in VNs (see *infra*).

#### *3.1.1.3. The Noun constituent*

From the structural point of view, N is almost always the direct object (or complement) of V and bears a Theme/Patient role. Among the counterexamples, we find the case of *mùkwíílílábwàmbà* ‘per-

son who works only when necessary'. The N constituent (*ùbù-àmbà* 'nakedness') does not play the role of the direct object of *ùkú-kwíílà* 'to earn a living', which is an intransitive verb, but, in the context of a metaphorical interpretation, indicates the extreme condition under which someone is forced to work (see 44d). Actually, this example provides suggestive evidence for an analysis of these complexes as nominalized VPs. The N constituent is in fact licensed by the applicative *-íl-*, which is attached to the verb *ùkú-kwíílà* 'to earn a living' (*mùkwíílá*), thus paralleling the corresponding VP in the syntax. A more complex case is *ìcísénsé-ngândà* 'insect': the N element *ín-ngândà* [9] 'house', a locative, is not licensed by any verbal morphology. However, it is interesting to note that V is, again, an intransitive verb: to purport the VP analysis it could be argued that even (certain) adjuncts can be sisters of V, if V does not take a direct object (see the *First Sister Principle* by Roeper and Siegel 1978 and, in general, Larson 1988 on the possibility of an indirect complement base-generated as the first sister of V, when the direct object is in *spec* VP).

While V does not take inflectional or derivational suffixes, except for the final vowel and, in few cases, the applicative, N almost always appears with the CM, which expresses gender and number features (with few exceptions, like in 45c). In most cases, N is mass or plural (like in Romance VNs, see *infra*). For instance, in *àkántélé-màfwéésà* [12] 'insect that lives under rocks', *àmá-fwéésà* 'rocks' is class 6, which is a plural class. In this case too, there are exceptions, see *ntíntá-mùkòshí* [9] 'tattoo marks along spine/neck' where *ùmú-kòshí* [3] 'neck' is neither plural nor mass.

Finally, it should be noticed that N cannot appear in VNs with the augment (see Romance VN compounds below, where N never appears with D). This fact provides perhaps the strongest evidence against the VP analysis of these compounds, which has been suggested above for the case of the VN with the applicative.

#### 3.1.1.4. *Headedness*

A non-trivial question concerning Bemba VNs regards their headedness. As put forward in the preceding sections, these compounds are exocentric, since neither V nor N determines the morpho-syntactic properties and the semantic type of the compound. An endocentric analysis is viable, however, if one considers the role of the affix occurring in these constructions. A phonologically overt morpheme, i.e. the CM, is the prefix of the VNs in Bemba and, as suggested above, it signals the 'gender' of the compound. The occurrence of a (head) affix

might suggest an account of VN compounding along the lines of synthetic compounding in English (see, among many others, Fabb 1984 and Lieber 1983). Therefore, an endocentric analysis must take into consideration the role played by the prefix in VNs. In particular, is the prefix occurring in these VNs an ‘inflectional’ affix, or does it function as a nominalising morpheme, i.e. as a derivational affix, akin to the nominalising affixes in Germanic synthetic compounds? An answer to this question is attempted in the next section.<sup>10</sup>

### 3.1.2. Two classes of VN compounds in Bemba

An analysis of the Bemba data in (44) through (49) reveals that VN compounds with *ka-* and *mu-* (both in class 1a) – repeated below as (50) and (51) – have distinct properties with respect to the others:

(50) *ka-* [1a]

- a. *kámíná-mísà* [1a] ‘drunkard’  
*ùkú-mínà* ‘to swallow’ *ìmí-sà* [4] ‘gulps’
- b. *kàtéshá-màbúúlà* ‘authoritative person’  
*ùkú-téntà* ‘to shake’ *àmá-búúlà* [6] ‘leaves’

(51) *mu-* [1a]

- a. *mùsùngá-bántù* [1a] ‘hospitable person’  
*ùkú-sùngà* ‘keep’ *aba-ntu* [2] ‘people’
- b. *mùsóngá-nsàlà* ‘appetizer’  
*ùkú-sóngà* ‘ignite/trigger’ *ín-sàlà* [9] ‘hunger’
- c. *mùpètá-ndùpé* ‘kind of shrub for making baskets’  
*ùkú-pètà* ‘to fold’ *ùlú-pé* [11] ‘kind of basket’
- d. *mùkwíílílá-bwàmbà* ‘person who works only when necessary’  
*ùkú-kwíílà* ‘to earn a living’ *ùbú-àmbà* [14] ‘nakedness’

First, even though these constructions are strongly lexicalized, it seems that *ka-* and *mu-* VN compounds are more compositional from an interpretive point of view than the others.

More significant, however, are the formal properties of these compounds. *Ka-* and *mu-* compounds are peculiar since they belong to a ‘special’ class, 1a, whose members are kinship terms, names of gods, personifications, proper names, etc. Moreover, this class contains agentive nouns, formed through derivational processes. This class is exceptional in many respects: nouns in class 1a do not take the augment, do not have an (overtly realized) CM, but act as class 1 as far as plural and concord are concerned (the plural is class 2, *ba-*).

Crucially, *mu-* and *ka-* are not CMs but agentive nominalising affixes (as *-er* in the Eng. *truck driver*).<sup>11</sup> This analysis, pro-

posed elsewhere in Bantu word formation studies (for instance, van Sambeek 1955: 99), is supported not only by the lack of augment but also by the test of plural formation. While pluralization typically requires for the substitution of a singular CM with the respective plural CM, in this case the plural (class 2) is *added* to the singular class prefix in an agglutinative fashion (see the examples below: in 52b, in particular, the CM *ka-* [12] (singular) is substituted by the plural CM *tu-*).

- (52) a. *mù-sùngá-bántù* (1a) ‘hospitable person’ → ***ba-mùsùngá-bántù*** (pl)  
 b. *àká-ntélé-màfwéésà* (12) ‘insect that lives under rocks’ → ***utu-ntélé-màfwéésà***(pl)

This analysis finds support in derivational phenomena which are still synchronically productive in Bemba. Prefix *ka-* is the affix employed for the formation of agentive deverbal nominalizations (see 53); all these nominals are assigned to class 1a.

- (53) *ká-fúndishà* ‘teacher’      *úkù-fúndà* ‘to teach’  
*ká-lúngà* ‘hunter’      *úkù-lúngà* ‘to hunt’  
*ká-pángà* ‘maker’      *úkù-pángà* ‘to make’

Synchronically, *mu-* is not an agentive nominalizer in Bemba. However, both diachronic and cross-linguistic (among other Bantu languages) evidence speaks in favour of the fact that *mu-* is a nominalizer rather than a nominal class prefix.

Relics of this kind of nominalization are still found in the contemporary Bemba lexicon:

- (54) *úkù-lima* ‘cultivate’ → *mu-lim-i* ‘cultivator’.

Nouns such as these suggest that in previous stages of the language there was a nominalization process by means of a derivational suffix, *-i*; the resulting words were assigned to class 1 (*mu-*), which still was a CM at this stage:

- (55) *mu-* + verb stem + final vowel *-i* → [*mu-...-i*]<sub>N</sub>.

Note that in some Bantu languages the [*mu-...-i*]<sub>N</sub> type of nominalization is still productive, as for example in Gĩkũyũ (see Mugane 1997, Bresnan and Mugane 2006). Examples below are from Bresnan and Mugane (2006: 4):

- (56) a. *mũ-in-i*  
 1-SING-NOM  
 ‘singer’  
 b. *mũ-thũnj-i*  
 1-slaughter-NOM  
 ‘slaughterer’

A feasible hypothesis is that, with the loss of the *-i* suffix as a nominalising morpheme, the *mu-* prefix has been reanalysed as a nominalising affix and employed for the formation of VN compounds. Further evidence for the nominalising analysis of *mu-* comes from other Bantu languages. Synthetic compounds, formed by concatenation of a [*mu-...-i*] deverbal noun and a noun are very productive in Gikũyũ (see Mugane 1997). See the examples in (57), from Mugane (1997: 42):

- (57) a. *mũ-end-i*            *andũ*  
 1Agt-love-NOM    2people  
 ‘one who likes    people’  
 b. *mũ-thũnj-i*            *mbũri*  
 1Agt-slaughter-NOM    10goats  
 ‘one who slaughters goats’  
 c. *mũ-tah-i*            *maaĩ*  
 1Agt-draw-NOM    6water  
 ‘one who draws water’

Mchombo (2004: 117) observes that in Chichewa there are rare cases of nominal compounds formed by a deverbal noun taking a nominal complement, hence similar to synthetic compounds (normally Chichewa forms exocentric VN compounds rather than synthetic NN compounds; see 41):

- (58) *msungi-chũma* ‘keeper-wealth = treasurer’ (*msungi* < *sung-a* ‘keep’)

Schadeberg (2003: 88) claims that similar forms of synthetic compounding are productive in several Bantu languages. A proto-Bantu example is given below:

- (59) \**mu-lind-a* + *mi-nue* ‘ring finger’  
 watch+er    (pl.) finger

According to Schadeberg, nouns such as \**mu-lind-a* are agent nouns taking complements (like *minue* ‘finger’ in 59), though not nec-

essarily direct objects in the traditional syntactic sense. Similar structures are found in Swahili. In (60), the agent noun *mwenda* – derived from an intransitive V – takes an indirect complement, similar to the Bemba example in (44d) discussed above.

- (60) Sw. *mwenda kimya mla nyama*  
walker silence eater meat  
'the one (lion) who walks silently will eat the meat'

In some cases, the noun seems to be the subject of the preceding verb and the compound assumes a locative (or object) interpretation:

- (61) Ganda *e-bugwa-njuba* 'west' > *-gu-* 'fall' + *i-júba* 'sun'  
Umbundu *u-kujá - mbámbi* '(sp. of) tree' > *-ku-n* 'gnaw' + *o-mbambi* 'gazelle'  
(Schadeberg 2003: 88)

Locative (or eventive) VN compounds do not occur in Bemba, but Romance has similar cases, although no longer productive (see *infra*).

To sum up, among the Bemba data we can distinguish at least two different classes: actual VNs, taking a CM establishing their 'gender' feature, also on the basis of inherent semantic properties (VN compounds with *aka-*, *ici-*, *ulu-*, *n-*), and cases of synthetic compounding, where a nominalising prefix (*ka-* and *mu-*) can be identified, which might take scope just over V rather than over the VN complex, straightforwardly paralleling the case of Germanic synthetic compounds of the *truck-driver* type. Evidence of this separate class comes from intra-, cross-linguistic and diachronic data, speaking in favor of the overt nominalising function of the *ka-* and *mu-* affixes at stake.

### 3.2. VN compounds in Romance

While Progovac (2006: 4) argues that VN compounds used to be productive in many Indo-European languages but are now mostly preserved as unproductive forms, in Romance languages VN compounding is the most productive compounding phenomenon (see also the abovementioned case of Chichewa, among Bantu languages); see the examples below:

- (62) It. *lavapiatti* 'wash-dishes = dishwasher'  
*posacenere* 'put (down)-ash = ash tray'  
*accendisigari* 'light-cigar = cigar lighter'

- (63) Fr. *grille-pain* 'grill-bread = toaster'  
*ouvre-boîte* 'open-tin = tin opener'  
*tournevis* 'turn-screw = screwdriver'
- (64) Sp. *rascacielos* 'scrape-sky = skyscraper'  
*lavaplatos* 'wash-dishes = dishwasher'  
*limpiabrisas* 'wipe-breeze = windshield wiper'

Along with opaque, fossilized forms, e.g. It. *pungitopo* 'sting-mouse = butcher's broom', Romance languages productively form VN nouns, more precisely agentive/instrumental nouns, as we will see in the next section (see Bisetto 1999, Bok-Bennema and Kampers-Manhe 2006, Desmets and Villoing 2009, a.o.).

### 3.2.1 Interpretation of VN compounds

In Romance languages, VN compounds are productively used to form agentive/instrumental nouns, as in the examples below:

- (65) a. It. *lavapiatti* 'wash-dishes = dishwasher'  
*asciugacapelli* 'dry-hair (pl)= hair drier'  
*spazzacamino* 'sweep-chimney= chimney sweep'  
*portalettere* 'bring-mails= mailman'
- b. Fr. *grille-pain* 'grill-bread = toaster'  
*sèche-linge* 'dry-wash = dryer'
- c. Sp. *lanza cohetes* 'throw rockets = rocket launcher'  
*buscapersonas* 'look for-persons = pager'

This seems to match the behaviour of VNs in Bantu (in particular in Bemba), which, as we have seen, are nouns with an agentive/instrumental interpretation, even though their meaning is lexicalised/frozen; they can refer to humans, animals and objects.

While – to the best of our knowledge – no eventive or locative interpretation of VN compounds is attested in Bemba, in Romance languages such interpretations are available (locative (66) and eventive (67)). VNs with these interpretations do not seem to be productive.

- (66) a. It. *corrimano* 'run-hand = hand rail'  
*battiscopa* 'beat-sweep = skirting board'
- b. Fr. *coupe-gorge* 'cut-throat = dangerous back alley'

- (67) a. It. *alzabandiera* 'raise-flag = flag raising'  
*battimano* 'clap-hand = hand clap/applause'
- b. Fr. *leche-vitrine* 'lick-window = window shopping'  
*baise-main* 'kiss-hand = kiss on the hand'

### 3.2.2. Argumenthood

Generally speaking, in Romance languages the noun forming a VN compound is the direct object of the verb. However, other patterns are found as well: the noun can be the subject (but, thematically, not a volitional Agent) (see 68) or an adjunct, such as a locative (either spatial or temporal, see 69), of the verb constituent.

- (68) a. It. *batticuore* 'beat-heart = heartthrob'  
*bollilatte* 'boil-milk = (milk) kettle'  
(Bisetto 1999)
- b. Fr. *trotte-bebé* 'toddle along-baby = baby-walker'  
*gobe mouton* 'swallow-sheep = kind of poisonous plant'  
(Desmets and Villoing 2009)
- (69) a. Fr. *traîne-buisson* 'hang around on-bush = animal'  
*veille-matin* 'wake up-morning = alarm clock'  
(Desmets and Villoing 2009)
- b. Sp. *saltamontes* 'hop-woods/mountains = grasshopper'  
(Rainer and Varela 1992)

This situation is similar to what we find for Bantu VNs, where the noun constituent, as we have seen, usually represents the direct object of the verb, but sometimes can be a locative or another kind of complement licensed by an applicative.

### 3.2.3. The noun and verb constituents

The noun constituent of Italian VN compounds is usually a plural or a mass noun, in the same way as in Bemba. However, sometimes the noun constituent can be singular, as in It. *spazzacamino* 'sweep-chimney= chimney sweep' or *battitappeto* 'beat-carpet = carpet sweeper/cleaner'. In Spanish too the noun constituent is usually plural, e.g. *limpiabotas* (clean-boots 'shoeblock') and, sometimes, a mass noun, e.g. *chupasangre* (suck-blood 'blood sucker'). Rainer and Varela (1992: 130) highlight that there seems to be a growing tendency in Spanish to extend this plural morpheme even to cases where it is not semantically plausible, e.g. *quitapelos* (take away-hair 'barber').<sup>12</sup> According to

Rainer and Varela, if this tendency continues, it could be the case that final *-s* may end up as an empty morpheme constitutive of the type as such, very much as linking morphemes.

The situation seems to be slightly different in French where, according to Desmets and Villoing (2009), the noun constituent of VN compounds can, in most cases, be analyzed as a stem, but sometimes it appears to be a word form marked for plural, as in *essuie-mains* (dry-hands ‘hand towel’), *presse-fruits* (press-fruits ‘squeezer’). However, Desmets and Villoing argue that this is not a syntactic marking but an inherent inflection (see Booij 1986), which is required by semantics and not by syntax; the choice of singular or plural does not really change the semantics of the whole VN compound.

The noun constituent of Romance VN compounds never appears with a determiner (70), whereas the corresponding sentences require a determiner (71):<sup>13</sup>

(70) a. Fr. Cet objet est un grille-pain. vs. b. \*Cet objet est un grille-le-pain.  
 this object is a grill-bread this object is a grill-the bread  
 ‘This object is a toaster’

(Adapted from Desmets and Villoing 2009)

c. It. Quest’oggetto è un asciugacapelli vs. d. \*Quest’oggetto è un asciug  
 -i-capelli  
 this object is a dry-hair this object is a dry-the-hair  
 ‘This object is a hair-drier’

(71) a. Fr. Cet objet grille le pain.  
 this object grill the bread  
 ‘This object grills the bread/bread’

(Adapted from Desmets and Villoing 2009)

b. It. Quest’oggetto asciuga i capelli  
 this object dry the hair  
 ‘This object dries the hair’

The lack of a determiner in Romance VN compounds is in line with the lack of the augment (often related to D position, as we have seen; see Visser 2007, de Dreu 2008, Buell 2009) on the noun constituent of VN compounds in Bemba.

As in Bantu VN compounds, the verb constituent of Romance VNs is formed by a verbal root plus a final vowel. In the literature on the topic there is a great debate on the form of the verb constituent in Romance VN compounds (for an overview, see Bisetto 1999; see also Rainer and Varela 1992): it is regarded to be either a verbal theme<sup>14</sup> (Pagliaro 1930) or the second person singular form of the imperative mood (Prati 1931, Progovac 2006; see also the discussion in the previ-

ous section), or the third person singular form of the present indicative (Tollemache 1945). Desmets and Villoing (2009), in their analysis of French VN compounds, argue that, since the verbal constituent is not marked for inflection, it is a stem.

#### 3.2.4. Headedness and recursion

As for headedness, VN compounds are generally considered as exocentric formations, since they are nouns but nevertheless the noun constituent is not the head of the compound (as in the case of the first class of Bemba VN compounds). First of all, the noun constituent cannot be considered as the head of the compound semantically. It is clear that, for example, a *spazzacamino* (sweep chimney ‘chimney sweep’) is not a type of *camino* ‘chimney’. Moreover, the noun constituent and the whole VN compound have different features; for example, a compound like *lavapiatti* (wash-dishes ‘dishwasher’) is singular, whereas the noun constituent is plural. To give another example, a compound like *portacipria* (carry-(face) powder ‘compact’) is masculine, whereas the noun constituent *cipria* ‘(face) powder’ is feminine (see Bisetto 1999).

However, some endocentric analyses have been proposed for these compounds. According to some of these, the first constituent is actually an agentive nominal (see Rainer and Varela 1991, Bisetto 1999): Coseriu (1977) proposes that it is the result of the deletion of the agentive suffix; Zuffi (1981) argues for the existence of a (phonetically not realized) agentive suffix, i.e. *-tore* ‘-er’ (*portalettere* bring-mails ‘mailman’ → *portatore lettere* ‘bringer mails’). According to Bisetto (1999), Italian VN compounds are actually NN compounds: the first constituent contains an agentive suffix *-tore* ‘-er’, which is deleted and, thus, does not appear overtly. Varela (1990), in contrast, puts forth a hypothesis according to which there is a reanalysis of the (originally inflectional) final *-a/-e* morphemes of the first constituent of Spanish VN compounds as derivational agentive suffixes. According to Scalise (1983), there is a zero morpheme attaching to the VN compound; the resulting word would be the result of compounding and derivation (through the null suffix), i.e. [[V+N]+0]; thus, the whole word would be right-headed. We will not deal any further with this issue here, for which further investigation is needed.

As we have shown in the previous section, recursion is impossible in Bemba VN compounds, as well as in Germanic languages (72c). Similarly, Romance VN compounds are generally not recursive, even though a few examples showing a modest degree of recursion can be found in some Romance languages (see Progovac and Locke 2009: 342, fn. 10):

- (72) a. Sp. *limpia-para-brisas* ‘wipe-stop-wind = windshield wiper’  
b. It. *porta-stuzzica-denti* ‘carry-pick-teeth = object where toothpicks are placed’  
c. En. \**scare-pick-pocket* ‘one who scares pickpockets’

### 3.3. VN compounds in Chinese

In Mandarin Chinese there are a number of VN compounds similar to those found in Romance languages, as shown by the examples below:

- (73) 镇纸 *zhèn-zhǐ* ‘press.down-paper = paperweight’  
护胸 *hù-xiōng* ‘protect-chest = chest protector’  
司机 *sī-jī* ‘operate-machine = chauffeur, driver’

According to Steffen Chung (1994), this kind of compounds used to be very productive in Chinese, especially those related to ‘profession’ terms; examples of these compounds date back to the Zhou dynasty (11th century BC-221 BC). Later, these compounds have decreased sharply in productivity, even though a number of them are still in use in the contemporary language, where there seems to be evidence of limited productivity of this compounding strategy (see Steffen Chung 1994). The new words section of *The contemporary Chinese Dictionary* (2002) lists a few of such compounds, e.g. 监事 *jiān-shì* ‘supervise-matter = member of a supervisory board’, 攀岩 *pān-yán* ‘climb-rock = rock climbing’. Even though this kind of compound is not as productive as in Romance languages, besides frozen forms (e.g. those indicating animals and plants) as those in (74), the examples found in the contemporary language can be divided in clear-cut categories, basically corresponding to those found for VN compounds in Romance languages, as will be shown in the next section.

- (74) 忍冬 *rěn-dōng* ‘endure-winter = honeysuckle’  
(a plant that can withstand low temperatures)  
防风 *fáng-fēng* ‘guard.against-wind = the *fangfeng* plant’  
(*Saposhnikovia divaricata* – a plant that acts as a windbreak)  
守宫 *shǒu-gōng* ‘guard-palace = gecko’  
守瓜 *shǒu-guā* ‘guard-melon = beetle’ (*Aulacophora femoralis*)  
(Steffen Chung 1994)

#### 3.3.1. Interpretation

Just like in Romance languages, VN compounds in Mandarin Chinese mainly have an agentive/instrumental interpretation (examples from Steffen Chung 1994 and Qiu 2000):

- (75) 护耳 *hù-ěr* 'protect-ear = earflaps, earmuffs'  
绑腿 *bǎng-tuǐ* 'tie-leg = leggings'  
镇纸 *zhèn-zhǐ* 'press.down-paper = paperweight'  
调羹 *tiáo-gēng* 'stir-thick.soup = spoon'  
围嘴 *wéi-zuǐ* 'surround-mouth = bib'  
领事 *lǐng-shì* 'lead-business = consul'  
管家 *guǎn-jia* 'manage-household = housekeeper'  
司机 *sī-jī* 'operate-machine = chauffeur, driver'  
编剧 *biān-jù* 'write-play = playwright'

Moreover, just as Romance, Mandarin Chinese has also instances of VN compounds with a locative or eventive interpretation:

- (76) 扶手 *fú-shǒu* 'support-hand = handrail'  
靠手 *kào-shǒu* 'lean-hand = armrest'  
攀岩 *pān-yán* 'climb-rock = rock climbing'

### 3.3.2. Argumenthood

As for argumenthood, VN compounds in Mandarin Chinese behave similarly to those in Bantu and Romance languages. As emerges from the examples above, the noun constituents of VN compounds are mainly the direct object of the verb. However, examples of locative nouns can be found as well:

- (77) 跑街 *pǎo-jie* 'run-street = errand boy'  
跑堂 *pǎo-táng* 'run-hall = waiter'  
领港 *lǐng-gǎng* 'lead.in-port = river pilot'  
(Steffen Chung 1994)

Furthermore, as in the Romance examples seen above, there are also cases of noun constituents acting as subjects (experiencer/source) of mono-argumental/intransitive verbs:

- (78) 地震 *dì-zhèn* 'earth-shake = earthquake'  
海啸 *hǎi-xiào* 'sea-scream = tsunami'  
头疼 *tóu-téng* 'head-ache = headache'

Interestingly, in such cases the order of the constituents is reversed, i.e. NV, consistently with the S(ubject)-V(erb) syntactic order.<sup>15</sup>

The varying precedence N-V / V-N relations in these compounds (strictly mirroring the syntactic order) might suggest that they are instances of reduced clauses or that they are nominalized VPs, projecting the internal or external argument. In any case, the hypothesis

that these VN (or NV) compounds are protolinguistic fossils is strongly challenged by the following facts: 1. the different position of N, consistent with its syntactic/thematic role; 2. the parallelism between the linear order of word and phrase (see the discussion in 3.4).

### 3.3.3. The verb and noun constituents

In Chinese there is no gender/number marking in nouns; since these nouns are monomorphemic (and, incidentally, monosyllabic) and coincide with roots, we can argue that the noun constituents of VNs are roots.

As in Romance languages, demonstratives or numeral-classifier sequences can never appear inside Chinese VN compounds (79b), which, in contrast, can be found in the corresponding verb phrase (79c):

- (79) a. 这个人是一个管家  
*zhè ge rén shì yī ge guǎn-jīa*  
 this CLF person be one CLF manage-house  
 ‘This person is a housekeeper’
- b. \*这个人是一个管这/一个家  
*zhè ge rén shì yī ge guǎn-zhè/yī-ge-jīa*  
 this CLF person be one CLF manage-this/one- CLF -house  
 ‘This person is a manage-this/a-house’
- c. 这个人管这/一个家  
*zhè ge rén guǎn zhè/yī ge jīa*  
 this CLF person manage this/one CLF house  
 ‘This person manages this/a house’

Verbs are never inflected in Chinese, since the language lacks inflectional morphology; the verb constituent of a VN compound is a monosyllabic/monomorphemic verbal root. Thus, both the noun and the verb constituents of VN compounds are roots.

Moreover, the constituents of such compounds can be bound roots, i.e. lexical morphemes that are not syntactically free and must combine with another morpheme in order to function as syntactic words. For example, in 护耳 *hù’ěr* (protect-ear ‘earflaps, earmuffs’), 耳 *ěr* ‘ear’ is a bound root (the corresponding free form is 耳朵 *ěrduo*). In the VN compound 司机 *sījī* (operate-machine ‘chauffeur, driver’), both constituents are bound roots. These examples would exclude the phrasal status of these formations.

### 3.3.4. Headedness and recursivity

In Chinese, nominal VN compounds are exocentric formations, just like Romance VN compounds and the first class of Bemba

VN compounds, since they are nouns but, nevertheless, the noun constituent is not the head of the compound. First of all, from the semantic point of view, the noun constituent cannot be considered as the head of the compound. It is clear that, for example, a 镇纸 *zhèn-zhǐ* ('press-down + paper = paperweight') is not a kind of 纸 *zhǐ* 'paper'. Furthermore, from a formal point of view, the choice of classifiers can be considered as further evidence of the exocentricity of these compounds. As a matter of fact, generally speaking, the whole VN compound and the noun constituent require different classifiers:

- |   |  |
|---|--|
| (80) a. 一个/位/名领事<br><i>yī ge/wèi/míng lǐng-shì</i><br>one/a CLF lead-business<br>'a consul' | cf. a <sup>i</sup> . 一件事<br><i>yī jiàn shì</i><br>one CLF business<br>'a business'     |
| b. 一个镇纸<br><i>yī ge zhèn-zhǐ</i><br>one/a CLF press-paper<br>'a paper-weight'               | cf. b <sup>i</sup> . 一张纸<br><i>yī zhāng zhǐ</i><br>one CLF paper<br>'a sheet of paper' |

To the best of our knowledge, VN compounds are not recursive in Mandarin Chinese, which is perfectly consistent with the behaviour of this kind of compounds in Bemba and in Romance languages (although in the latter there actually is a very modest degree of recursion), as shown above.

### 3.3.5. Mandarin VN synthetic compounds

Besides VN compounds, Mandarin Chinese also has a word formation process similar to English synthetic compounds. There are two different patterns for these compounds. The first one consists of a monosyllabic verb and a monosyllabic noun plus the suffix 者 *-zhě*, roughly corresponding to the English suffix *-er*, which acts as the head of the complex word (VN-者 *zhě*); see He (2004, 2006, 2009), Cheng (2005):<sup>16</sup>

- |                               |   |   |
|-------------------------------|---|---|
| (81) 造谣者<br>爱美者<br>求职者<br>骑马者 | <i>zào-yáo-zhě</i><br><i>ài-měi-zhě</i><br><i>qiú-zhí-zhě</i><br><i>qí-mǎ-zhě</i> | 'make-rumour-SUFF = rumour monger'<br>'love-beauty-SUFF = beauty lover'<br>'seek-job-SUFF = job applicant'<br>'ride-horse-SUFF = horse rider' |
|-------------------------------|---|---|

Differently from English, this kind of compounds retains the verb-object order (cf. Eng. *truck driver*). It could be thought that this is due to the fact that VN combinations (with monosyllabic V and N)

tend to be stored as verbs in Chinese, and thus the suffix attaches to the whole complex verb (see Cheng 2005). As a matter of fact, in the examples in (81), the VN bases are verb-object compounds and are usually listed in dictionaries;<sup>17</sup> however, this pattern seems to be fully productive, since synthetic compounds can be formed also with non-lexicalized VN bases. For example, besides 骑马者 *qí-mǎ-zhě* (ride-horse SUFF ‘horse rider’), where 骑马 *qí-mǎ* ‘ride-horse = horse ride’ is a lexicalized VN verb, with a Google search we can find a compound like 骑龙者 *qí-lóng-zhě* (ride-dragon-SUFF ‘dragon rider’, the Chinese name of a Japanese cartoon and the name of a videogame). Other examples of this kind of compounds with VN bases that are not verbal VN compounds are the following:

- (82) 猎鲸者 *liè-jīng-zhě* ‘hunt-whale- SUFF = whale hunter’  
 猎龙者 *liè-lóng-zhě* ‘hunt-dragon- SUFF = dragon hunter’  
 追梦者 *zhuī-mèng-zhě* ‘pursue-dream-SUFF = dream pursuer’

Furthermore, unlike English *-er*, the suffix 者 *zhě* cannot form instrumental nouns, but only agentive nouns. To form instrumental nouns, a root carrying the instrumental meaning is needed instead of 者 *zhě*, as e.g. the bound roots 机 *jī* ‘machine’ and 器 *qì* ‘utensil/ware’:

- (83) 灭火器 *miè-huǒ-qì* ‘extinguish-fire-utensil = fire extinguisher’  
 洗衣机 *xǐ-yī-jī* ‘wash-clothes-machine = washing machine’  
 洗碗机 *xǐ-wǎn-jī* ‘wash-bowl-machine = dish washer’  
 切纸机 *qiē-zhǐ-jī* ‘cut-paper-machine = paper cutter’

There is yet another pattern to form synthetic compounds: a disyllabic (or polysyllabic) noun and a disyllabic verb plus the affix 者 *zhě*, which acts as the head of the complex word (NV-者 *zhě*); see He (2004, 2006, 2009), Cheng (2005):<sup>18</sup>

- (84) 谣言制造者 *yáoyán-zhìzào-zhě* ‘rumour-make- SUFF = rumour-monger’  
 病毒传播者 *bìngdú-chuánbō-zhě* ‘virus-spread- SUFF = virus spreader’  
 展览主办者 *zhǎnlǎn-zhǔbàn-zhě* ‘exhibition-(to)sponsor- SUFF =  
 exhibition sponsor’  
 浪漫文学爱好者 *làngmàn-wénxué-àihào-zhě* ‘romantic-literature-love-SUFF =  
 romantic literature lover’  
 影视爱好者 *yǐngshì-àihào-zhě* ‘movies.and.television-love-SUFF =  
 movies and television lover’  
 文艺爱好者 *wényì-àihào-zhě* ‘literature.and.art-love-SUFF = literature  
 and art lover’  
 北半球观测者 *běibànqiú-guāncè-zhě* ‘Northern.Hemisphere-observe-SUFF =  
 Northern Hemisphere observer’

故事杜撰者 *gùshi-dùzhuàn-zhě* 'story-fabricate/make.up- SUFF = story maker'

(examples from He 2004, Cheng 2005 and the Center for Chinese Linguistics PKU corpus)

This pattern is structurally identical to the one found in English synthetic compounds, where the order is NV-SUFF.

Therefore, two different kinds of VN compounds can be found in Mandarin Chinese: VN exocentric compounds, which have limited productivity, and synthetic compounds, which seem to be largely productive. The presence of both types of compound in the language and their distribution is reminiscent of the situation found in the history of exocentric V-N compounds in English. In English there was probably a number of exocentric VN compounds before the Norman conquest, especially in proper names and epithets (see Gast 2008). Later, the productivity of these forms increased, apparently under French influence, and continued to increase in the 14th and 15th centuries, reaching its peak in the 16th century. Then, from the 17th century on, their productivity started to decline, leading to the situation of the modern language, where this kind of words are attested but hardly any new words are created (Gast 2008). Gast (2008) points out that decline of exocentric VN compounds was accompanied, and possibly partly caused, by a strong increase of synthetic compounds (N-V-*er*); for many centuries the two word-formation patterns existed side by side in the language, sometimes providing alternative terms for one meaning. However, eventually, at the time of the Industrial Revolution, synthetic compounds became the predominant pattern and took over great parts of the denotational domain previously covered by exocentric VN compounds (Gast 2008).

Further studies on the productivity and distribution of VN exocentric compounds and of synthetic compounds in Mandarin Chinese, also from a diachronic perspective, could reveal if a similar evolution as the one described here for English has taken place in this language too.

### 3.4. Concluding remarks on VN compounds

Summing up, VN compounds, as well as NN compounds, can be found in all the languages considered; also, these compounds share many properties, as summarized in table 4.

**Table 4.** Properties of VN compounds in Romance and Bantu.

	ROMANCE	BANTU	CHINESE
ARGUMENTHOOD	N is the direct object of V (but can also be the subject or a locative)	N is the direct object of V (also locative or complement licensed by the applicative)	N is the direct object of V (also locative); N can be the subject: in this case the order is NV
N CONSTITUENT	Plural or mass noun / stem or (sometimes) plural (French) It never appears with a determiner	Plural or mass noun It does not have the augment	Root
V CONSTITUENT	Verb root + final vowel	Verb root + final vowel	Root
HEADEDNESS	Exocentric	Exocentric (but see <i>mu-</i> and <i>ka-</i> VN compounds)	Exocentric (but see synthetic compounds)
DETERMINER ON N	No	No	No
RECURSION	Very limited	No	No
PRODUCTIVITY	Agentive/instrumental interpretation	No (Bemba, but see Chichewa a.o. Bantu languages)	Limited (but see synthetic compounds)

The languages examined show many similarities, only differing in the presence or lack of declension markers (final vowel on V / number morphology on N) – arguably related to the morphology of the language at issue – and in the productivity of the phenomenon.

The verb constituent of this kind of compound is formed by a verbal root and a final vowel in Bantu and Romance languages, while it is a root/stem in Chinese; the noun constituent is usually a plural or mass noun in Bantu and Romance (even though in French it is usually a stem), while in Chinese it is a root. Moreover, the noun constituent in Romance languages never appears with a determiner; similarly, the noun constituent of Bantu VN compounds never takes the augment (which could be seen as equivalent to the determiner in Romance). As far as argumenthood is concerned, the noun constituent is more often than not the direct object of the verb constituent, even though it can sometimes represent other kinds of complements, as e.g. locative. As for the interpretation of the whole compound, in Bantu languages VN compounds denote agents and instruments; Romance productively forms agentive and instrumental VN compounds, even though (non-productive) eventive and locative compounds are attested as well. In Chinese, all the semantic types found in Romance VN compounds are attested, but the agentive/instrumental interpretation

is predominant. Furthermore, VN compounds do not allow modification of the noun constituent and are exocentric formations in all the languages considered. While Romance languages allow a very limited degree of recursion in VN compounds, in Bantu and Chinese recursion is apparently never allowed.

The biggest difference in the formation of VN compounds between the languages at issue, as we have seen, concerns the productivity of these formations. In Romance languages, VN compounding is a very productive means of word formation, commonly used to form agentive and instrumental nouns. In contrast, in some Bantu languages, this phenomenon seems to be limited to a number of fossilized forms, the meaning of which is very opaque; this is the case, for instance, of Bemba. However, in some other Bantu languages, such as Chichewa (see Mchombo 2004), this process seems to be the most common form of compounding and creates nouns putting together a verb and its unmodified object or locative noun. In Chinese there are many examples of exocentric VN compounds, even though apparently their productivity is limited.

In Germanic languages, instances of VN compounds are found, e.g. Eng. *pickpocket*, *scare-crow*, but represent a restricted non-productive set, often associated with negative connotations. As a point of fact, Germanic languages tend to form so-called ‘synthetic compounds’, such as *truck driver* or *dishwasher*, made up by a root noun and a deverbal noun (an alternative analysis suggests that the underlying structure is  $[[N+V]_{V+SUFF}]_N$  rather than  $[N+[V-SUFF]]_N$ ). Interestingly, some Bantu languages too seem to have processes akin to synthetic compounding, and in some of these languages this phenomenon is very productive, as e.g. in Gĩkũyũ, which apparently does not have VN compounds of the Romance type. This kind of compounds is found in Bemba too, where, as we have mentioned, the class of VN compounds containing the nominalizing prefixes *ka-* and *mu-* are more compositional in meaning with respect to the other VN compounds, whose meaning is lexicalised/frozen. Therefore, Bantu languages present instances of both VN compounds and of compounds akin to synthetic compounds. While some of these languages (e.g. Chichewa) seem to productively form VN compounds, just as Romance languages, other languages (e.g. Gĩkũyũ) productively form synthetic compounds, like Germanic languages. A language like Bemba presents instances of both kinds of compounds; nevertheless, neither of these two processes is synchronically productive. In Chinese too, both exocentric VN compounds and synthetic compounds can be found, but, as we have seen, the productivity of exocentric VN

compounds seems to be limited, showing a situation in many respects similar to that of English.

It would be interesting to find out whether the syntactic word order of a language is connected with the order of the constituents of exocentric VN compounds. The languages considered in this paper are all VO languages; interestingly, we have observed that those instances of compounds where N seems to act as the subject of V in Chinese, e.g. 海啸 *hǎi-xiào* (sea-scream ‘tsunami’), have the NV order, consistently with the syntactic SVO order. But the question seems to be even more interesting if we observe the behaviour of this kind of compounds in OV languages. For example, Japanese forms NV compounds of the kind of e.g. 爪切り *tsume-kiri* (nail-cut (cutting)<sup>19</sup> ‘nail cutter’), whose order of constituents (NV) is consistent with the order of the constituents in its syntax. Moreover, Gast (2009) shows that NV compounds were common in Latin, whose dominating word order is OV, e.g. *agri-col-a* (field-LNK-cultivate-NOM ‘farmer’). He also points out that NV compounds are found in Early Germanic languages, e.g. Old English *mere-far-a* (sea-travel-INFL ‘sailor’). This too seems to be consistent with the syntax of the language, since the dominant syntactic order of Old English is OV (see van Kemenade 1994). Examples of NV compounds can still be found in Middle English (where the syntactic order has already changed in VO, though the order OV is attested in its first stages), e.g. *eu-bruche* (matrimony-break ‘adulterer’), and relics of this pattern can be found in Modern English (often with a negative meaning or connotation), e.g. *chimney sweep*.

Given these facts, it is not clear what a protolinguistic analysis à la Progovac can really tell us about VN compounds. As a matter of fact, if exocentric VN compounds seem to follow the syntactic word order of a language, they can be hardly seen as relics of a protolinguistic style of computation. Moreover, a proto-linguistic analysis may be attractive for those languages where the VN pattern is no longer productive and which show the marked properties put forward by Progovac, e.g. English and Slavic languages. However, this analysis cannot account for the productivity of Romance VN compounds, where, in addition, this pattern is not confined to the formation of words with a negative meaning/connotation or of ancestral lexical items, but forms mainly instrumental/agentive nouns (see also Chinese). It is not clear, for example, why synthetic compounding has not taken over VN “proto-linguistic” compounding in Romance, unless an endocentric, hence non-protolinguistic, analysis of the VN phenomena is feasible, as proposed by Zuffi (1981), Scalise (1983) and Bisetto (1999), among others (see section 3.2.4).

#### 4. Concluding remarks

The study of nominal compounding phenomena in the sample of typologically and genealogically unrelated languages considered here has revealed a number of interesting facts on the alleged protolinguistic nature of these phenomena, a view that has been recently defended by influential authors such as Jackendoff and Progovac.

First of all, the structural makeup of compounds seems to obey strict hierarchical principles and cannot be easily reduced to non-syntactic principles (cf. Jackendoff's standpoint on the alleged headedness of English NN compounds as instantiation of pragmatic principles). Evidence in this direction comes from the fact that NN compounds display robust, yet language-specific headedness patterns in the languages at issue; further, this view also holds for VN compounds which, despite lacking a formal/semantic head, encode a hierarchical relation between the verb and the noun. Specifically, the verb is the head of the VN complex since it is a thematic assignor to the noun constituent and it never "modifies" it (no *exocentric* VN compounds are found where V is a modifier of N).

Moreover, both NN and VN compounds display a number of morphosyntactic features that are "internal" to the complex structure, such as the presence of inflectional markers on the non-head in NN compounds and on the noun in VN compounds. These markers are arguably the exponents of functional structure "inside" the NP domain of these compounds. The cross-linguistic variation attested in this respect is sensitive to the morphological typology of the languages at issue.

Finally, as to VN compounds, the suggested mirror effect of word and phrase syntax – to be supported by broader investigation – could be another interesting aspect of these complex forms, since it might be the strongest indication of the underlying syntax of compounding.

All in all, several aspects of NN and VN compounds point towards an overall picture of compounding that is hardly reconcilable with a protolinguistic view of the phenomenon.

#### *Addresses of the Authors*

Bianca Basciano: Università Cattolica del Sacro Cuore (Brescia), Facoltà di Scienze Linguistiche e Letterature Straniere, via Trieste 17, 25121 Brescia, Italy <bianca.basciano@unicatt.it>

Nancy C Kula: Department of Language and Linguistics, University of Essex,  
Wivenhoe Park, Colchester CO4 3SQ <nckula@essex.ac.uk>

Chiara Melloni: Università di Verona, Dipartimento di Filologia, Letteratura  
e Linguistica, viale dell'Università 2, 37129 Verona, Italy <chiara.mel-  
loni@univr.it>

### *Co-authorship*

All authors, whose names are listed alphabetically, participated in the research. For academic purposes, Bianca Basciano takes responsibility for sections 2.3, 3.2, 3.3, Nancy Kula for 2.1, 3.1 and Chiara Melloni for 1, 2, 2.2, 2.4, 2.5, 3, 4. All authors have approved the final version of this article.

### *Notes*

<sup>1</sup> Compositionality is arguably reducible to operations on the Qualia Structure of the nouns in the compound (see Delfitto and Melloni 2009 and forthcoming).

<sup>2</sup> An exception is represented by Swahili, which has two productive types of NN compounding (see Schadeberg 2003).

<sup>3</sup> We define as stem “the uninflected parts of independent words that do not themselves constitute independent words”, see Lieber and Stekauer (2009: 5).

<sup>4</sup> A cover term for these different morphological elements is ‘lexeme’, used indeed by Bauer (2003: 40) for defining compounds.

<sup>5</sup> Bisetto and Scalise (2005) identify three main classes of nominal compounds on the basis of the grammatical relationship between the constituents:

SUBORDINATE	<i>angolo cottura</i>	corner + cooking ‘cooking corner’
ATTRIBUTIVE	<i>uomo rana</i>	man + frog ‘frogman’
COORDINATE	<i>cantante attore</i>	singer + actor ‘actor singer’

Of the tree classes, only the attributive class is considered here. A class of alleged subordinate root compounds (e.g. It. *centrotavola* ‘centrepiece’, *fondovalle* ‘valley bottom’, etc.) have been convincingly analyzed as (lexicalised) aprepositional genitives (see Delfitto and Paradisi 2009). We refer the reader to Delfitto and Melloni (2009 and forthcoming) for further details.

<sup>6</sup> Word markers in Harris (1991), declension class markers in Alexiadou (2004), gender markers in Carstens (2001, 2008), epenthetic elements / n° exponent in Ferrari Bridgers (2008), among others.

<sup>7</sup> It could be noticed that, while Romance and Bantu have morphemes more or less transparently related to Gender/Number features, Germanic also exhibits case markers on the non-head. The distinction with respect to Germanic, moreover, does not only concern the absence of case marking, but also the interpretability of inflection markers. While number on the non head is not necessarily interpreted in Germanic (see examples below), singular or plural marking is necessarily interpreted in Romance. A similar analysis does not easily apply to the Bantu data, since the meaning of the compounds is often highly lexicalised and cannot be retrieved in the constituents meaning.

De. <i>Sonn-en-schein</i>	‘sunshine’ (from OHG gen.sg. <i>-in</i> )
<i>Kind-er-wagen</i>	‘pram for one child’

*Schwein-e-braten* 'roast consisting of one pig or part of a pig'

*Freund-es-kreis* 'circle of more than one friend' (gen. sg.)

<sup>8</sup> Linking elements are also found in other non-Germanic languages. However, LEs typically occur in right-headed compounds, as in the examples below.

Gr. *Diavol-o-θiliko* devil+LE+woman 'minx'

La. *verb-i-velitatio* word-LE-dispute 'verbal dispute'

Po. *gwiazd-o-zbiór* star-LE-collection 'constellation'

In fact, LEs have been analyzed as default (case or gender/number) markers occurring in incorporation structures (cf. Progovac 2005). A syntactically motivated analysis of these morphemes is also proposed by Delfitto, Fábregas and Melloni (forthcoming), arguing they are exponents of a morphosyntactic feature (Declension Class) attracting the non-head and causing its displacement in German compounds.

<sup>9</sup> Li and Thompson stress the fact that the twenty-one types of nominal compounds they have listed do not constitute an exhaustive categorization, and in fact one can think about more kinds of nominal compounds.

<sup>10</sup> For a derivational, rather than inflectional, analysis of class markers in Bantu see Crisma, Marten and Sybesma (this volume).

<sup>11</sup> The prefixes *mu-* and *ka-* are thus homonymous w.r.t CM *mu-* of class 1 and *ka-* of class 12 respectively.

<sup>12</sup> *Pelo* is a mass noun in Spanish, like *hair* in English.

<sup>13</sup> Desmets and Villoing (2009) observe that lexicalization of verb phrases, including those involving a V and an N, as in VN compounds, normally preserves the functional words of the original syntactic phrase, including prepositions, pronouns and determiners: e.g. Fr. *trompe-la-mort* 'deceives-the-death = daredevil', *boit-sans-soif* 'drinks-without-thirst = drunkard', *rendez-vous* 'go-you = appointment'.

<sup>14</sup> A verbal theme is formed by a verbal stem plus a thematic vowel. In Italian, for example, there are three thematic vowels, *-a*, *-e*, *-i*, corresponding to the three verbal conjugations (infinitive endings: *-are*, *-ere*, *-ire*).

<sup>15</sup> Note that these forms do not go against the universal Verb-Object Constraint (VOC), according to which a nominal that expresses the theme/patient of an event combines with the event-denoting verb before a nominal that expresses the agent/cause does (see Baker 2001, 2010). First of all, the kind of verbs appearing in these compounds are intransitive. Moreover, the nouns in these NVs are not prototypical agents/initiators, but experiencers/sources of the situation expressed by the verb (see 78).

<sup>16</sup> Cases with a nominal root head are attested too, e.g. 生 *shēng* 'student', 长 *zhǎng* 'chief', 师 *shī* 'master' (see He 2004).

<sup>17</sup> Chinese makes wide use of the so-called verb-object (separable) verbal compounds (see e.g. Chao 1968, Li and Thompson 1981, Huang 1984), formed by a monosyllabic verb and a monosyllabic noun constituent. These compounds are often separable, even when, for example, their meaning is lexicalized; thus, VN verb compounds are ambiguous between being words and phrases.

<sup>18</sup> For evidence in favour of the lexical status of these formations, see He (2004). He (2004) highlights that the reversed order, i.e. VN-者 *zhě* (with a disyllabic verb and noun) is possible as well, but the resulting item is a phrase, rather than a compound. For example, He points out that only NV-者 *zhě* forms allow the plural marker *-men*, e.g. 月球探险者们 *yuèqiú-tànxiǎn-zhě-men* (moon-explore-SUFF-PL 'moon explorers'; Center for Chinese Linguistics PKU corpus); for other evidence on the phrasal status of VN-者 *zhě* (with a disyllabic verb and noun) formations, see He (2004). Note that any VN-者 *zhě* with a disyllabic verb and noun can be reversed in the NV-者 *zhě* order, while the reverse is not always true.

<sup>19</sup> It is argued that the continuative form of the verb (as *kir-i*), one of the six conjugated forms of Japanese verbs, is used to nominalize the verb, e.g. 'draw'→'drawing', 'cut'→'cutting', etc. Kageyama (2009) argues that, when compounded with a concrete noun on the left, such form acquires a specific meaning of agent. He states that it is not feasible to derive exocentric compounds directly from the VP structures because their corresponding tensed compound verbs, e.g. \*爪切る *tsume-kiru* (切る *kiru* being the plain, non-past form of the verb), do not exist. However, it should be noted that in Romance languages too the nature of the final vowel of the verb constituent is not clear and one hypothesis is that it is a verbal theme, i.e. an inflected form, thus being a stem. Moreover, note that in Japanese the continuative form of the verb has a number of uses: for instance, the first constituent in compound verbs is in the continuative form; the same form is used also in VN modifier-head compound nouns and in VP coordination, where the first verb is in the continuative form. In any case, it seems beyond doubt that the *-i* ending in the example above cannot be considered as an agentive suffix corresponding to English *-er* in synthetic compounds.

### *Bibliographical References*

- ACQUAVIVA Paolo 2009. The structure of the Italian declension system. In MONTERMINI Fabio, Gill BOYÉ & Jesse TSENG (eds.). *Selected Proceedings of the Sixth Décebrettes*. Somerville, Mass.: Cascadilla Proceedings Projects. 50-62.
- ALEXIADOU Artemis 2004. Inflection Class, Gender and DP Internal Structure. In MÜLLER Gereon, Lutz GUNKEL & Gisela ZIFONUN (eds.). *Exploration in Nominal Inflection*. Berlin: Mouton de Gruyter. 21-50.
- BAKER Mark 2001. *The atoms of language*. New York: Basic Books.
- BAKER Mark 2010. Formal generative typology. In HEINE Bernd & Heiko NARROG (eds.). *Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press. 285-312.
- BAR-SHALOM Eva & William SNYDER 1999. On the relationship between Root Infinitives and Imperatives in Early Child Russian. In GREENHILL Annabel et al. (eds.). *Proceedings of the 23<sup>rd</sup> Boston University Conference on Language Development*. 56-67.
- BICKERTON Derek 1990. *Language and Species*. Chicago: University of Chicago Press.
- BICKERTON Derek 1998. Catastrophic evolution: The case for a single step from protolanguage to full human language. In HURFORD James R., Michael STUDDERT-KENNEDY & Chris KNIGHT (eds.). *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge: Cambridge University Press. 341-358.
- BISETTO Antonietta 1999. Note sui composti VN dell'italiano. In BENINCA Paola, Alberto MIONI & Laura VANELLI (eds.). *Fonologia e morfologia dell'italiano e dei dialetti d'Italia*, Atti del XXXI Congresso internazionale di studi della Società di Linguistica Italiana. Roma: Bulzoni. 503-538.
- BISETTO Antonietta & Sergio SCALISE 2005. Classification of Compounds. *Lingue e Linguaggio* 2. 319-332.

- BOK-BENNEMA Reineke & KAMPERS-MANHE Brigitte 2006. Taking a closer look at Romance VN compounds. In NISHIDA Chiyo & Jean-Pierre Y. MONTREUIL (eds.). *New Perspectives on Romance Linguistics: Morphology, Syntax, Semantics and Pragmatics*. Amsterdam: John Benjamins. 13-26.
- BOOIJ Geert 1986. Form and meaning in morphology: The case of Dutch “agent-nouns”. *Linguistics* 24. 503-518.
- BRESNAN Joan & John M. MUGANE 2006. Agentive nominalizations in Gikūyū and the theory of mixed categories. In BUTT Miriam, Mary DALRYMPLE & Tracy H. KING (eds.). *Intelligent Linguistic Architectures: Variations on themes by Ronald M. Kaplan*. Stanford: CSLI. 201-234.
- BUELL Leston 2009. The Distribution of the Nguni Augment: A Review. Talk given at “Bantu Augment Workshop” (June 17, 2009). Leiden: The Netherlands.
- CARSTENS Vicky 2001. Multiple agreement and case deletion: Against  $\phi$ -(in) completeness. *Syntax* 3. 147-163.
- CARSTENS Vicky 2008. DP in Bantu and Romance. In DE CAT Cécile C. & Katherine DEMUTH (eds.). *The Bantu-Romance Connection: A comparative investigation of verbal agreement, DPs and information structure*. Amsterdam and Philadelphia: John Benjamins. 131-166.
- CHAO Yuanren 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- CHENG Gong 2005. 汉语“者”字合成复合词及其对普遍语法的启示 [Chinese synthetic compounds with “zhě” and their influence on general grammar]. 现代外语 *Xiandai waiyu* 8 (3). 232-238.
- CHOMSKY Noam 1995. *The Minimalist Program*. Cambridge, Mass.: The MIT Press.
- CONTINI MORAVA Ellen 2007. Swahili morphology. In KAYE Alan S. (ed.). *Morphologies of Asia and Africa*. Winona Lake, Indiana: Eisenbrauns. 1129-1158.
- CORBETT Grevill 1991. *Gender*. Cambridge: Cambridge University Press.
- COSERIU Eugenio 1977. *Inhaltliche Wortbildungslehre*. In MARCHAND Hans, Herbert Ernst BREKLE & Dieter KASTOVSKY (eds.). *Perspektiven der Wortbildungsforschungen*. Bonn: Bouvier. 48-61.
- CRISMA Paola, Lutz MARTEN & Rint SYBESMA. The point of Bantu, Chinese and Romance nominal classification. This volume.
- DE DREU M. 2008. The internal structure of the Zulu DP. Master’s thesis. Leiden University.
- DELFIFFITO Denis & Chiara MELLONI 2009. Compounds don’t come easy. *Lingue e Linguaggio* VIII (1). 75-104.
- DELFIFFITO Denis & Chiara MELLONI forthcoming. *Compounding as a symmetry breaking strategy*. In BERTINETTO Pier Marco et al. (eds), *Linguaggio e cervello / Semantica, Atti del XLII Convegno della Società di Linguistica Italiana* (Pisa, Scuola Normale Superiore, 25-27 settembre 2008). Volume I. Roma: Bulzoni.
- DELFIFFITO Denis & Paola PARADISI 2009. For a diachronic theory of genitive assignment in Romance. In CRISMA Paola & Giuseppe LONGOBARDI (eds.). *Historical syntax and linguistic theory*. Oxford: Oxford University Press. 292-310.

- DELFITTO Denis, Antonio FÁBREGAS & Chiara MELLONI 2008/forthcoming. Compounding at the Interfaces. In *Proceedings of 39<sup>th</sup> NELS* (November 2008, Cornell University, Ithaca). Amherst, Mass.: GLSA.
- DELFITTO Denis, Gaetano FIORIN & Nancy KULA 2011. Romance vs. Bantu complex nouns at the morpho-syntax interface. This volume.
- DESMETS Marianne & Florence VILLOING (2009). French VN lexemes: morphological compounding in HPSG. In MÜLLER Stefan (ed.). *Proceedings of the HPSG09 Conference*. Stanford: CSLI Publications. 89-109.
- DOWNING Pamela 1977. On the Creation and Use of English Compound Nouns. *Language* 53 (4): 810-842.
- FABB Nigel 1984. *Syntactic affixation*. Cambridge, Mass.: MIT press.
- FERRARI BRIDGERS Franca 2008. A unified syntactic analysis of Italian and Luganda nouns. In DE CAT Cécile & Katherine DEMUTH (eds.). *The Bantu-Romance Connection: A comparative investigation of verbal agreement, DPs and information structure*. Amsterdam and Philadelphia: John Benjamins. 239-258.
- GAST Volker 2008. V-N compounds in English and German. *Zeitschrift für Anglistik und Amerikanistik* 56 (3). 269-282.
- GAST Volker 2009. Verb-noun compounds in European languages. Hand-out of the talk given at the University of Erfurt, November 4, 2009: <http://www.personal.uni-jena.de/~mu65qev/talks.html>
- HARRIS James W. 1991. The Exponence of Gender in Spanish. *Linguistic Inquiry* 22 (1). 27-62.
- HE Yuanjian 2004. The Words-and-Rules Theory: Evidence from Chinese Morphology. *Taiwan Journal of Linguistics* 2.2. 1-26.
- HE Yuanjian 2006. Lexicon as a Generating System: Restarting the Case of Complex Word Formation in Chinese. *Journal of Chinese Language and Computing* 16 (2). 99-119.
- HE Yuanjian 2009. 论合成复合词的逻辑形式 [On the logical form of synthetic compounds]. *语言科学 Yuyan Kexue* 8 (5). 503-516.
- HERINGER Hans-Jürgen 1984. Wortbildung: Sinn aus dem Chaos. *Deutsche Sprache* 12. 1-13.
- HUANG James C.T. 1984. Phrase Structure, Lexical Integrity, and Chinese Compounds. *Journal of Chinese teacher's association* 19 (2). 53-78.
- JACKENDOFF Ray 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences* 3 (7). 272-279.
- JACKENDOFF Ray 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press. 105-128.
- JACKENDOFF Ray 2009. Compounding in the Parallel Architecture and Conceptual Semantics. In LIEBER Rochelle & Pavol ŠTEKAUER (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 105-128.
- KAGEYAMA Taro 2009. Isolate: Japanese. In ROCHELLE Lieber & Pavel ŠTEKAUER (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 512-526.
- KULA Nancy forthcoming. Nominal compounding and associative phrases in Bemba. In BRENZINGER Matthias & Anne-Maria FEHN (eds.). *Proceedings of the 6th World Congress of African Linguistics*. Cologne: Köppe.

- LARSON Richard K. (1988). On the double object construction. *Linguistic Inquiry* 19 (3). 335-391.
- LI Charles & Sandra THOMPSON 1978. An Exploration of Mandarin Chinese. In LEHMANN Winfred P. (ed.). *Syntactic Typology: Studies in the Phenomenology of Language*. Austin: University of Texas Press. 233-266.
- LI Charles & Sandra THOMPSON 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- LIEBER Rochelle 1983. Argument Linking and Compounds in English. *Linguistic Inquiry* 14. 251-285.
- LIEBER Rochelle & Pavel STEKAUER 2009. Introduction: Status and Definition of Compounding. In LIEBER Rochelle & Pavel STEKAUER (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 3-18.
- MARCHAND Hans 1969. *The Categories and Types of Present-day English Word formation* (second edition). München: Beck'sche Verlagsbuchhandlung.
- MCHOMBO Sam 2004. *The Syntax of Chichewa*. Cambridge: Cambridge University Press.
- MUGANE John M. 1997. *A Paradigmatic Grammar of Gĩkũyũ*. Stanford, California: CSLI publications.
- NURSE Derek 2008. *Tense and Aspect in Bantu*. Oxford: Oxford University Press.
- PAGLIARO Antonino 1930. *Sommario di Linguistica Ario-europea*. Roma: Scuola di Filologia Classica dell'Università di Roma (serie II 2.1).
- PRATI Angelico 1931. Composti imperativi quali casati e soprannomi. *Revue de Linguistique Romane* 7. 250-264.
- PROGOVAC Ljiljana 2005. Synthetic Agent Compounds in Serbian: An Incorporation Analysis. In *Tasseva-Kurktchieva Mila, Steven Franks, and Frank Gladney (eds.) Formal Approaches to Slavic Linguistics 13: The Columbia Meeting 2004*. Ann Arbor: Michigan Slavic Publications. 253-264.
- PROGOVAC Ljiljana 2006. Fossilized Imperative in Compounds and Other Expressions: Possible Implications for Historical and Evolutionary Studies. In *Online Proceedings of the First Meeting of Slavic Linguistics Society (SLS)*. Bloomington, Indiana.
- PROGOVAC Ljiljana 2008. What use is half a clause? In SMITH Andrew, Kenny SMITH & Ramon FERRERI CANCHO (eds.), *Evolution of Language: Proceedings of the 7th International Conference, Barcelona, Spain (March 12-15)*. New Jersey, NJ: World Scientific. 259-266.
- PROGOVAC Ljiljana 2009. Layering of grammar: Vestiges of evolutionary development of syntax in present-day languages. In SAMPSON Geoffrey, David GIL & Peter TRUDGILL (eds.). *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press. 203-212.
- PROGOVAC Ljiljana 2009b. Compounds and Commands in the Evolution of Language. Handout of a talk given at *Ways to Protolanguage* (September 21-23, 2009). Toruń, Poland.
- PROGOVAC Ljiljana & John L. LOCKE 2009. The Urge to Merge: Ritual Insult and the Evolution of Syntax. *Biolinguistics* 3 (2). 337-354.

- PUSTEJOVSKY James 1995. *The Generative Lexicon*. Cambridge, Mass.: MIT Press.
- QIU Liqin 2000. VN复合词初探 [A first exploration of VN compounds]. 徐州师范大学学报 *Xuzhou shifan daxue xuebao* 26 (1). 70-72.
- RAINER Franz & Soledad VARELA 1992. Compounding in Spanish. *Rivista di linguistica* 4 (1). 117-142.
- ROEPER Thomas & Muffy SIEGEL 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9. 199-260.
- ROLFE Leonard 1996. Theoretical stages in the prehistory of grammar. In LOCK Andrew & Charles R. PETERS (eds.). *Handbook of Human Symbolic Evolution*. Oxford: Clarendon Press. 776-792.
- SCALISE Sergio 1983. *Morfologia Lessicale*. Padova: Clesp.
- SCHADEBERG Thilo C. 2003. Derivation. In NURSE Derek & Gérard PHILIPPSON (eds.). *The Bantu Languages*. London and New York: Routledge. 71-90.
- SELKIRK Elisabeth O. 1982. *The Syntax of Words*. Cambridge, MA: MIT Press.
- STEFFEN CHUNG Karen 1994. Verb + Noun Function-Describing Compounds. *Bulletin of the College of Liberal Arts* 41. National Taiwan University. 181-221. Revised version: [homepage.ntu.edu.tw/~karchung/pubs/vncomp\\_rev.pdf](http://homepage.ntu.edu.tw/~karchung/pubs/vncomp_rev.pdf)
- TOLLEMACHE Federigo 1945. *Le parole composte nella lingua italiana*. Roma: Roes.
- VAN KEMENADE Ans 1994. Old and Middle English. In KÖNIG Ekkehard & Johan VAN DER AUWERA (eds.). *The Germanic Languages*. London: Routledge. 110-141.
- VAN SAMBEEK Jan. 1955. *A Bemba Grammar*. London: Longmans. Green.
- VARELA Soledad 1990. Composición nominal y estructura temática. *Revista española de linguística* 1. 56-81.
- VISSER Marianna 2007. Definiteness and specificity in the isiXhosa determiner phrase. Talk handout.
- XING Janet Zhiqun 2006. *Teaching and Learning Chinese As a Foreign Language: A Pedagogical Grammar*. Hong Kong: Hong Kong University Press.
- ZUFFI Stefano 1981. The nominal composition in Italian: Topics in generative morphology. *Journal of Italian Linguistics* 2. 1-54.
- 现代汉语词典 *The Contemporary Chinese Dictionary* (Chinese-English edition) 2002. Beijing: Foreign Language Teaching and Research Press.

**Corpus:**

*Center For Chinese Linguistics PKU (Peking University) corpus of Modern Chinese:* [http://ccl.pku.edu.cn:8080/ccl\\_corpus/index.jsp?dir=xiandai](http://ccl.pku.edu.cn:8080/ccl_corpus/index.jsp?dir=xiandai)

