

Exocentricity in Compounding

SERGIO SCALISE ANTONIO FÁBREGAS FRANCESCA FORZA

Univ. of Bologna *Univ. of Tromsø/Bologna* *Univ. of Verona/Bologna*

Abstract: The identification of a compound as endocentric or exocentric depends on the notion of head: if a compound has a head (or two), it is called endocentric; if it has no head, it is called exocentric. Exocentricity, however, has been usually assumed as a unitary notion, exactly because the notion of head has been generally interpreted as a unitary notion. In this paper we will first provide typologically based data on the dimension and limits of exocentricity, and then we will argue that the notion of head can be split into three different subparts: categorial head, semantic head and morphological head. Correspondingly, the notion of exocentricity can be split into categorial exocentricity, semantic exocentricity and morphological exocentricity. Our approach, based on features of the constituents and not on constituents as a whole, will hopefully provide a new analysis of exocentricity in compounding.*

Key words: compounding, head, exocentricity, parameters, VN compounds

1. Identification of the Head

There are at least three points of view according to which the head can be identified. One point of view relies on semantics. Bloomfield (1933: 235), for example, states that an endocentric compound denotes a hyponym of its head: a *door knob* is a kind of *knob* and therefore the compound is endocentric. The same reasoning is adopted by many other scholars, among which Zwicky.¹ Other scholars instead claim that the head must be identified only on formal grounds. Kageyama (2008, 2009), for example, states that “The head should be defined as a category determinant”. Finally, still other scholars, such as Namiki (2001), claim that the head should be identified on the basis of several parameters: semantic, morphological and furthermore it should be taken into consideration also the claim that the head is the morphosyntactic *locus*. This last position deserves some observations. In the first place, it has been shown (Guevara and Scalise 2009) that the head is not always the *locus inflectionis*: there are endocentric compounds that do not

* This research has been carried out also thanks to funds of the Italian Ministry of University and Research (PRIN 2005). Antonio Fabregas’ research has been financed by postdoctoral grant EX2006-0968. We would like to thank two anonymous reviewers for commenting a previous version of this paper. Usual disclaimers apply.

¹ According to Zwicky (1985: 4), “We could take the head/modifier distinction to be at root semantic: in a combination X+Y, Y is the ‘semantic head’ if [...] X+Y describes a kind of the thing described by Y”.

put any inflection on the head (e.g. It. *rosso mattone* ‘red + brick = brick red’ is not inflected in the plural (*due maglioni rosso mattone* / **due maglioni rossi mattone* ‘two brick red sweaters’) and there are exocentric compounds that do have inflection (e.g. *pickpockets*). In the second place, while Namiki claims that the head should be both formal and semantic, he does not attempt to establish any link between these two notions. In the following pages, we will adopt Namiki’s position (without considering the head as *locus inflectionis*) but we will try to set a relationship between semantic head and the so-called formal head.

2. Dimensions of Exocentricity

Exocentricity is not a marginal fact (neither from a quantitative point of view, nor as far as productivity is concerned). One could be tempted to think that exocentricity is a relic of the past and that productive contemporary word formation is always endocentric. While this seems to be true for derivation it is not true for compounding, as it will be shown below. Before facing the theoretical side of the matter, it is interesting to have a general idea of the distribution of exocentric phenomena in compounding across languages. In fact, the empirical observation leads to some considerations, the first one being that all languages exhibit some degree of exocentricity but this degree can vary across languages.

There are in fact languages (such as Turkana: Dimmendaal 1983) which are reported to have almost only exocentric compounds. Furthermore, there are languages where the most productive pattern is endocentric but where exocentric compounds are fully productive, such as V+N compounds in Italian (e.g. *portlettere* ‘carry letters = mailman’).

As for the distribution of exocentricity in compounding in the languages of the world, consider the following Table.²

(1)

Headedness	Mean % ³
Right headed	65.02
No heads	18.9
Left Headed	5.2
Two-heads	4.5

² The data (Guevara and Scalise 2008) are taken from Morbocomp, a database, developed at the University of Bologna, that includes some 3000 compounds from 24 languages (Basque, Bulgarian, Catalan, Chinese, Czech, Dutch, English, Finnish, French, German, Greek, Hungarian, Italian, Japanese, Korean, Latin, Norwegian, Polish, Portuguese, Russian, Serbo-Croatian, Spanish, Swedish, Turkish). Each compound is analyzed in 18 searchable fields (compound, script, internal structure, compound category, first constituent, first constituent category, second constituent, second constituent category, third constituent, third constituent category, categorial head, semantic head, linking element 1, linking element 2, gender, number, gloss and translation, remarks).

³ In this and in the following tables, the total sum of the percentages reported is not 100

We find thus a clear hierarchy of the distribution of heads in compounding:

(2) Right > No Head > Left > Both

Overall, there are clearly more endocentric types than exocentric types and a strong preference can be observed for right-headed types (barely two thirds of all the types).

(3)

Headedness	Mean %	RO %	GE %	EA %	SL %
Right headed	61.1	36.6	86.5	60.5	62.8
No heads	19.1	35.4	8.4	18.9	13.9
Left Headed	7.1	14.7	1.8	6.8	5.3
Two-headed	5.6	5.7	1.5	11.4	3.2

As it can be seen, exocentricity is attested in every considered language group of the database.

As already said, the total number of compounds in the database is 3222. The exocentric compounds are 617, representing a percentage of 19%, namely the fifth part of all compounds. This information alone supports a very simple consideration we have already introduced, i.e.: exocentricity is not an idiosyncratic, possibly ignorable, phenomenon in compounding.

The distribution across the compound classes identified in Bisetto and Scalise (2005) is the following:

(4)

ATTRIBUTIVE	5.1%
COORDINATE	4.6%
SUBORDINATE	8.5%

The attributive and the coordinate classes do not show a dramatically different behaviour as far as headedness is concerned. The subordinate class, on the contrary, has a higher incidence of exocentricity. This could reflect a specific pattern, or even a generalization, supporting the idea that one possible trigger of exocentricity would be a complement-head relation between the constituents, probably—as it will be shown ahead—depending on some typological properties of the licensing of arguments.

Yet, the most pervasive compound type across the corpus is not a subordinate but an attributive one: $[A+N]_N$, such as the English *pale face* or $[N+A]_N$, such as It. *viso pallido* ‘face pale = pale face’ in left headed languages.⁴ The *pale face* type has

because spurious cases, accidental mistakes and problematic entries have deliberately been left-out. Hopefully, the overall coherence of the results is not affected.

⁴ It is also interesting to notice a series of this type of compound with a slightly different flavour, e.g. 奇麗所 *kirei dokoro* ‘beautiful + place = geisha’ in Japanese.

been considered by some authors (e.g. Booji 2007) as endocentric on the grounds that the semantic shift can be explained in terms of metonymy. However, it should be noticed that semantics is not the only problem with this type of compounds since, for example in Romance, they seem also to be exocentric in such formal properties such as gender and number. For example the compound *testa rasata* ‘head + shaven = skin head’ can be masculine or feminine, singular or plural while the constituent *testa* is obligatorily feminine singular.

This type of compound is found in all the 16 fusional languages of the corpus (Bulgarian, Catalan, Dutch, German, Czech, English, French, Greek, Italian, Latin, Norwegian, Polish, Portuguese, Russian, Serbo-Croatian, Spanish and Swedish). In the agglutinative languages of the database this type is also well represented, since it is present in Turkish, Finnish, Hungarian and Japanese. The isolating language of the corpus, Chinese, also contains this type.

3. Limits of Exocentricity

It is not at all clear what are the limits of exocentricity, but we may safely assume that there could be at least three dimensions along which the limits of exocentricity could be investigated: typological, categorial and semantic.

3.1. Typological distribution

Let us now see the distribution of exocentricity among the main groups of our database:

(5)

Group	%
Romance	35.4
Germanic	8.4
Slavic	13.9
East Asia ⁵	18.9

As it is possible to notice from the preceding table, Romance languages tend to show a high rate of exocentricity. The most common exocentric types are $[V+N]_N$, $[P+N]_N$, $[V+V]_N$, and $[A+N]_N$. Instead, Germanic languages are not significantly characterised by exocentric compounding. The most wide-spread headless structure is, as already mentioned, $[A+N]_N$, such as *pale face*.

East Asian languages, as well as Slavic languages, contain an average amount of exocentric types. They share some types with other language groups, such as $[A+N]_N$ or $[V+V]_N$, but some types seem more language-specific, such as the Chinese antonymic compound *dàxiǎo* 大小 ‘large + small = size’ (although we will challenge this view in the last section).

From a typological point of view, we may assume that different types of lan-

⁵ Japanese, Korean and Chinese are grouped under the geographical label ‘East Asia languages’, which cross-cuts the typological and genealogical distinctions between them.

guages may exhibit different kinds of exocentricity, a conclusion that is already implicit in some data seen in (3). Romance languages, in fact, exhibit more exocentricity than Germanic languages (35.8% vs. 8.9%). If we compare two languages, for example Italian and English, we find the following situation:

(6)	Italian	English
	V+N] _N	— ⁶
	P+N] _N	—
	N+A] _N	A+N] _N
	V+V] _N	—
	N+N] _N	N+N] _N

As expected, there are more exocentric patterns in Italian than in English. In both languages the default category for exocentric compounds is N.

These typological differences, as a matter of fact, must find explanation in some deeper reason, a reason concerning the very essence of compounding. As we have seen in the preceding sections, exocentricity is spread among the world's languages.

Exocentric structures can be very different from each other, all the lexical categories are involved, and there seems to be different ways in which a compound can be classified as exocentric.

3.2. Limits of categorial combinations

We will now consider Chinese exocentric compound structures:

(7)

Chinese	Script	Example
V+N] _N	天葬	<i>tiānfáng</i> '(to) fill+room=second wife (of a widower)'
V+N] _A	缺德	<i>quēdé</i> 'lack+morals=immoral'
A+N] _N	软卧	<i>ruǎnwò</i> 'soft+(to) lie=(of a train) soft sleeper'
V+V] _N	裁缝	<i>cáifēng</i> 'cut+sew=tailor'
N+N] _N	江湖	<i>jiānghú</i> 'river+lake=vagabond'
A+A] _N	大小	<i>dàxiǎo</i> 'big+small=size'

Chinese seems to exhibit more exocentric compound types than both Italian and English. Furthermore, in Chinese, the same compound structure (V+N) can give rise to compounds with different categories⁷—a very rare fact probably tied to the existence of hybrid categories.

Both Chinese and Italian, but not English, have what we would like to call 'Absolute Categorial Exocentricity (ACE)', when the output is completely differ-

⁶ In English there are some VN compounds such as *pickpocket*, *killjoy*, but this class is not productive.

⁷ This fact has been studied and for [VN]_N and [VN]_V, cf. Ceccagno and Scalise (2005).

ent from the input categories: e.g. Italian $[V+V]_N$ and Chinese $[A+A]_N$. We could express ACE with a schema such as $[X+Y]_Z$.

The question is whether there are limits to ACE. In other words, given three major lexical categories such as A, N, V, which of the following logically possible absolute exocentric combinations are attested in the languages of the world?

- (8)
- $A+A]_N$
 - $A+A]_V$
 - $N+N]_A$
 - $N+N]_V$
 - $V+V]_A$
 - $V+V]_N$
 - $V+A]_N$
 - $N+A]_V$
 - $V+N]_A$
 - $A+V]_N$
 - $A+N]_V$
 - $N+V]_A$

An ACE-oriented research in the database shows that some of these combinations are empirically attested.

Let us see which combinations are found in the database. Consider for instance the structures where none of the two constituents is A, while the whole compound is A. We have four structures of this kind.

(9)

Structure	Language	Example
$[N+N]_A$ ⁸	Latin	<i>auricomus</i> 'gold+hair=golden-haired'
	Serbo-Croatian	<i>ribòlik</i> 'fish+shape=fish-shaped'
	Turkish	<i>paragöz</i> , 'money+eye=greedy for money'
$[V+V]_A$	Turkish	<i>yapış yapış</i> 'stick+stick=sticky'
$[V+N]_A$ ⁹	Chinese	缺德 <i>quēdé</i> 'lack+morals=immoral'
	French	<i>lève-blocs</i> 'lift+block=block lifter'
	Serbo-Croatian	<i>vijòglav</i> 'twist+head=head shaker'
$[N+V]_A$	Korean	<i>neknek-hata</i> 'sufficiency+to be=sufficient'

⁸ These examples, however, can be assimilated to the English *blue-eyed*, even though no suffix is visible in the above-mentioned languages. Discussion on this kind of compounds is in progress (cf. Bisetto and Melloni 2008, among others), but, in any case, these examples do not seem to fit perfectly in ACE.

⁹ Apart from the Chinese example, we consider compounds like *lève-blocs* and *vijòglav* $[V+N]_N$, as having an N output, and only subsequently used as attributes or predicates, a grammatical function prototypically assigned to adjectives, but not exclusive of them. These examples, too, do not perfectly represent ACE.

It has to be noticed that none of the structures in the preceding table occurs with a high rate in the database.

Instead, there are just three types of ACE structures with Nominal output, but the total number of compounds is extremely higher.

(10)

Structure	Language	Example
[A+A] _N	Chinese	大小 <i>dàxiǎo</i> 'large+small=size'
	Japanese	大小 <i>daisho</i> 'big+small=size'
[V+V] _N	Romance languages	It. <i>bagnasciuga</i> 'soak+dry=strand'
		Sp. <i>subibaja</i> 'climb+descend=ups and downs'
	Turkish	<i>yapboz</i> 'construct+destroy=jigsaw puzzle'
	Chinese	裁缝 <i>cáifeng</i> 'cut+sew=tailor'
[A+V] _N	Chinese	广告 <i>guǎnggào</i> 'wide+(to) announce=advertisement'

Notice that the first structure ([A+A]_N) is very common in Chinese and has some examples in Japanese. [V+V] structures with Nominal output are present in basically all Romance languages, of which we provide just two examples for brevity.

The AA and VV structures forming Ns are decidedly, and surprisingly, strongly attested in the database, across all genealogical and typological groupings. They will be discussed in 5.2.

Let us furthermore notice that no Germanic language is found among the ones exhibiting ACE. Chinese, on the contrary, is present with most structures. Therefore, the typological pattern mentioned in 2.1 is coherently reflected in Absolute Categorical Exocentricity.

By looking at the data, then, we notice that some of the ACE structures which were considered 'logically possible' (8) are not attested:

- (11) [A+A]_V
 [N+N]_V
 [N+A]_V
 [A+N]_V
 [V+A]_N

We will leave aside the non-existing structure [V+A]_N, whose absence might be due to the lack of some specific languages in the database, since the reverse structure [A+V]_N is found. So, all the other non-existing ACE combinations have a V as output. Crucially, V cannot be formed if none of the constituents is a V.

Why doesn't ACE allow V-forming processes?¹⁰ The reason for this systematic

¹⁰ As a matter of fact, compounding with Verbal output in general is not well attested (7.8%). Despite the low rate, it is present in Bulgarian, Catalan, German, Dutch, Chinese, Japanese, English, French, Greek, Finnish, Latin, Korean, Norwegian, Portuguese, Serbo-Croatian, Spanish, Swedish and Turkish. i.e. 18 out of 24 languages. Thus, the absence in the database cannot be safely attributed just to the general absence of V output compounds.

lack would have something to do with the nature of grammar. It is as if the [+V] feature could not be present in the compound if it were not already in one of the constituents. The only available counterexample, $[N+N]_A$, has already been demonstrated not to represent a real candidate for ACE. This pattern, at any rate, calls for further investigation.

In this subsection, we have proposed the notion of Absolute Categorical Exocentricity, which is found when none of the constituents impose its category on the compound. We have seen that it is empirically attested and all the major categories are involved, in different combinations.

Moreover, although several typologically and genealogically unrelated languages show ACE, it is still true that some languages, such as Chinese, tend to be more exocentric than others, e.g. the Germanic ones. Then, a clear, possibly universal, limit has emerged, regardless of typological or genealogical differences: No Verbs can be formed when ACE is at work.

3.3. Limits of semantic combinations

If there is such a thing as Absolute Categorical Exocentricity, there could be the semantic counterpart: Absolute Semantic Exocentricity.

There are in fact compounds where the semantics of the input constituents is very different from the semantics of the output. A clear example would be the following Chinese compound:

(12) 东西 *dōng xi* ‘east + west = thing’

Here the ‘distance’ between the constituents and the output is by no means categorial since we have $[N+N]_N$, but the two constituents are semantically distant from the output, with no apparent relationship, neither synonymic nor metonymic. In terms of the Lexical Semantics *à la* Lieber (2004), we would say that in the given example, the bodies of the two constituents are very similar to each other while the body of the output is very different. To appreciate this, consider a compound such as *red skin*. In this compound there is a metonymic relation between the head and the output and therefore it is, so to speak, semantically less exocentric than the Chinese example in (12).

According to Booij (2007) the type *red skin* is not to be considered exocentric because of the metonymic relationship between the head and the output. Unfortunately, the semantic relationships both between the constituents and between the constituents and the output has not yet been adequately studied and, even worse, we are far from having a taxonomy of semantic relationships in compounding.

There are relationships similar to the metonymic type, such as in Cat. *aigua-sal* ‘water + salt = pickle’, or ‘descriptive’ types such as It. *millepiedi* ‘thousand + feet = centipede’, or ‘narrative’ (Soegaard 2005) such as Ch. *jiānghú* 江湖 ‘river + lake = vagabond’. There are also antonymic compounds such as Ch. *hū xī* 呼吸 ‘to expire + to inspire = to breathe’, It. *saliscendi* ‘go up + go down = elevator’, but apparently

no exocentric synonymic compounds.¹¹

Our provisional conclusion is that semantic exocentricity is well attested and well distributed across the languages of the world. All the languages of our database exhibit this type of exocentricity, which is also well distributed across the three different classes: attributive (Cat. *pit roig* ‘breast + red = robin’), coordinate (Cz. *nosorožec* nose + horn = rhinoceros) and subordinate (Sp. *palabrimujer* ‘word + woman = man that speaks like a woman’).

4. Three Types of Exocentricity

As we have seen, exocentricity cannot be considered a simple phenomenon with purely incidental relevance for morphological theory. In this section, we will try to understand the functioning of exocentricity from a new perspective that implies defining it in terms of features and not in terms of constituents and ‘enlarging’, so to speak, the notion of exocentricity also to morphological exocentricity. The motivations for this enrichment of the notion of exocentricity will be clarified in the following sections. Consider the previous approaches to the notion of headedness and, therefore, exocentricity:

- (13) 1. The head is only categorial.
 2. The head is only semantic.
 3. The head must be categorial and semantic.

These three solutions share the presupposition that heads are atomic units which cannot be decomposed in smaller entities. In fact, while previous studies such as those mentioned in section 1 have concentrated on the idea that all properties of a single constituent were transmitted to the whole compound if that constituent is the head, in this paper we will argue that it is possible that only features of a certain kind are transmitted from the head to the compound. In other terms, we argue that headedness does not depend on a single constituent giving all its properties to the whole compound. We will provide evidence to support the idea that we gain a better understanding of the notion of headedness and exocentricity if we consider the role that different kinds of features play in defining the properties of the whole compound. The features that seem to be cross-linguistically relevant to define headedness can be of three types: categorial, semantic and morphological. This leads us to differentiate between three kinds of exocentricity: categorial, semantic and morphological.

Before considering the empirical necessity of making this distinction, consider the theoretical advantage that it presents. In previous approaches, a compound to be classified as endocentric is required to be endocentric according to all the three criteria. The result of this rough classification is that different compounds that did not have anything in common among them ended up being in the same class

¹¹ Synonymic compounds seem to be in general endocentric. Cf. Ch. *bǎo shòu* 饱受 ‘protect + defend = preserve’. Synonymic compounds have to be distinguished from ‘reduplicative compounds’ such as It. *lecca lecca* ‘lick + lick = lollipop’.

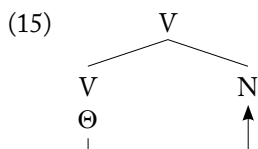
of exocentric compounds. This prevented any analysis from being able to make generalisations about exocentricity. We will indeed show that, on the contrary, our proposal has the consequence that different kinds of exocentricity can be identified and interesting generalisations can be made about them.

4.1. Three levels of exocentricity: An example

Let us first empirically show that three levels need to be distinguished when we talk about exocentricity. Consider, for illustration, a productive type such as VN compounds in Romance, here illustrated in (14):

- (14) a. It. *asciuga-capelli*
 'dry+hairs = hair dryer'
 b. Sp. *corta-úñas*
 'cut+nails = nail clipper'

VN compounds are exocentric in the first sense: what looks as the head of the construction in the internal structure of the compound has a particular lexical category, V, but the compound is a noun, not a verb. According to the standard assumptions about argument structure, from a structure VN where the V takes N as an argument, we expect projection of V, and, instead, we get a compound which is a N—alternatively, taking the rule that compounds in Romance are left-headed, V is also expected.



We will refer to this kind of exocentricity as categorial exocentricity, which we can define as in (16):

- (16) A compound is categorially exocentric if the constituent in the head position does not impose its categorial features on the whole construction.

We can differentiate here between structures where the category of the whole compound coincides with the category of the non-head and structures where the category of the whole compound is not present in any of its constituents (Absolute Categorial Exocentricity, presented in the previous section). VN compounds would be categorially exocentric, but they do not exhibit absolute categorial exocentricity, for the category of the compound, N, is the category of the non-head in the structure of (15).

VN compounds are also exocentric in the second sense. Notice that the compound contains a noun inside the structure. This noun, in Romance languages, is inflected for some morphological properties, such as gender and number. However, these properties contained inside the compound are not projected to the whole construction. Notice, as seen in (17), that the gender and number of the internal N

do not determine the gender and number of the whole compound:

- (17) el [limpia [botas]_{Npl.fem.}]
 the-sg.masc. clean boots

This second type of exocentricity affects purely morphological features,¹² such as gender, number, declension class or conjugation class, so we will refer to it as morphological exocentricity. We can define it in the following way:

- (18) A compound is morphologically exocentric if the morphological features of the compound are not identical to the morphological features of any of its internal constituents.

Notice that this exocentricity is logically and empirically independent from categorial exocentricity. Some Romance compounds are morphologically exocentric, but not categorially:

- (19) It. [testa_N rasata_A]_N
 'head + shaven = skin head'

Even though *testa* is feminine singular—and so is the adjective *rasata* which agrees with it—the whole compound can be masculine plural, as signalled by the determiner in (20):

- (20) i [testa rasata]
 the-pl.masc head shaven 'the skin heads'

Finally, VN compounds are exocentric in the third sense: even though they are typically composed of a verbal stem which denotes an event or an action and a noun that denotes an object which is passively involved in this action, the compound as a whole is interpreted as an agent or instrument. In other words, the semantic type of the compound cannot be derived from the semantic type of any of its constituents. We will refer to this third sense of exocentricity, in which the semantic class denoted by the compound cannot be predicted from the semantic

¹² The independent question here is what we mean by morphological features. As it is well-known, this depends on the general theoretical framework that is adopted. As we want to remain as neutral as possible in our analysis, we will not adopt the so-called Separation Hypothesis (Beard 1995; known as the Feature Disjointness Hypothesis in Distributed Morphology, cfr. Embick 2000) and we will allow some features with semantic or syntactic interpretation to be included in this class, provided that they prove to be relevant for morphological processes in this language. Therefore, in a wide sense, we consider the feature [number] as a morphological feature, because even if it is primarily a semantic feature, it is relevant in Romance for restricting several word formation processes. However, it is worth mentioning that some features seem to be exclusively morphological in the sense that they do not intervene in clear semantic or syntactic processes. Among these morphological features we find the declension class of a noun—the fact that the feminine noun *mano* 'hand' is marked with *-o* in Italian or Spanish—and the conjugation class of a verb or features that mark dependency of a particular form.

class of their constituents, as semantic exocentricity:

- (21) A compound is semantically exocentric if it denotes a class which cannot be derived from the classes denoted by its constituents.

This notion of exocentricity is the one that violates any rule of semantic compositionality; Allen's (1978) IS A rule is, maybe, the clearest exponent in modern generative linguistics of this kind of rule, and therefore of the procedure to define semantic exocentricity.¹³

Once again, notice that semantic exocentricity is logically independent from the other two kinds of exocentricity identified above. Consider the compound in (22), taken from Spanish. Categorially it is endocentric, as its category label is the same as one of its constituents—crucially, the one that would be the head in a normal combination of adjective and noun; it is also morphologically endocentric, because the gender and number of this constituent is also the gender and number of the whole compound:

- (22) la [media_A naranja_{Nsg,fem}]_{Nsg,fem}
 the half orange 'Mr./Mrs Right'

However, it is exocentric in the sense that its meaning is not derived from the meaning of its constituents. The compound in (22) refers to 'Mr./Mrs. Right', but this meaning does not derive from the noun *naranja* 'orange', nor from the adjective *media* 'half'.

Now that we have determined what kinds of exocentricity we have at play and why they are empirically distinct, let us analyse the interaction between them.

4.2. Interaction between categorial and semantic exocentricity

Let us explore whether there is any implication between categorial and semantic exocentricity along the proposal put forth in Guevara and Scalise (2009). There

¹³ Although the semantic endocentricity of a compound is typically characterised by a subset relationship, i.e. the compound denotes a subset of the class denoted by its head, as in Allen's procedure, our definition is made in such a way that so-called co-compounds (Wälchli 2005) are also included here. Co-compounds (cf. *infra*) denote a superset or a hyperonym of the classes denoted by their two constituents; they do not follow Allen's rule, but their denotation is predictable from the denotation of the two constituents. Therefore, in our description, co-compounds have two semantic heads, like coordinative compounds of the type *poet painter*, but, differently from them, denote supersets and not subsets of the constituents. By the same reasoning, coordinative compounds such as *Bosnia-Herzegovina*, that were classified as semantically exocentric in previous work, are classified here as endocentric, as the meaning of the compound, much like in the case of co-compounds, is derived by adding the meaning of the two constituents; this way of calculating the meaning of the compound complies with rules of semantic compositionality, so there is no need to appeal to semantic exocentricity.

are the following logical combinations¹⁴ (abbreviations: C = categorial (head); S = semantic (head); CH = categorial head; SH = semantic head; CRD = coordinate).

(23)

Structure	Example	Logical possibilities	Comment
[[C] []] [[S] []]	It. <i>capostazione</i> 'head+station=stationmaster'	SH and CH head coincide	Traditional endocentricity
[[C] [C]] [[S] []]	Unattested	Two CHs one SH	Necessarily CRD compounds
[[C] []] [[S] [S]]	Unattested	Two SHs, one CH	They would be coordinate compounds formed by two different categories
[[S] ([S])]	Unattested	SH no CH	One or both constituents impose its semantic type, but not its category.
[[C] ([C])]	It. <i>testa rasata</i> 'head+shaven=skin head'	One or two CH, no SH	The category is predictable from the constituents, but not the semantic denotation
[[C] [C]] [[S] [S]]	Sp. <i>poeta pintor</i> 'poet painter' Mordvin <i>penč.t'-vakan.t</i> 'spoon+plate=cutlery'	Two CHs and two SHs	CRD compounds or co-compounds
[[] []] [[] []]	Ch. 大小 <i>dàxiǎo</i> 'big+small=size' It. <i>bagna-asciuga</i> 'soak+dry=strand'	No SH no CH	Absolute exocentricity

Notice that we have not found two logically possible combinations in our database: (i) coordinative compounds with two categorial heads but with only one semantic head and (ii) compounds with a semantic head which is not at the same time the categorial head. Two generalizations are possible depending on the value that we want to give semantic exocentricity in linguistic theory. In the recent literature on lexical semantics a distinction has been generally made: the one between the structural aspects of the meaning of an expression, which are reflected in its grammatical behaviour, and the conceptual aspects of meaning, which depend on world knowledge and are unpredictable. This distinction has been referred to as skeleton vs. body (Lieber 2004) or structural vs. conceptual semantics. Imagine that we wish to make a distinction between these two types of semantic exocentricity. The crucial point here is whether compounds of the type *testa rasata*, or the better known *red skin* in English, can be analysed as semantically endocentric. Proposals such as Booi (2007) try to reduce its semantic interpretation to the normal denotation of the categorial head combined with an operation of semantic metonymy: given that humans have heads and skin, it is easy to denote a human by referring to this part of his or her body. Notice, however, that other semantic features are added to the

¹⁴ In the following table, we exemplify the structure only for left-headed compounds, but the same arguments are valid for right-headed compounds.

denotation of the compound; in the case of *skin head*, one of them is that the person referred to by the compound sympathises with the nazi ideology. This feature cannot be predicted via metonymy, but it is clearly a feature of conceptual semantics. Depending on the orthogonal assumptions about how relevant conceptual semantics is in linguistic analysis, therefore, *skin head* can be reanalysed as semantically endocentric or not.

The topic, in any case, is more complex than we can capture in this paper. Assume, for the sake of the discussion, that semantic exocentricity based on conceptual knowledge—such as that a *red skin* is a type of race or a *skin head* is a nazi—does not affect the grammar of a language in a relevant sense. In that case, *testa rasata* or *skin head* has a semantic head which coincides with the categorial head, and then we can propose the following strong principle which follows from the table:

- (24) If a constituent is CH, then it must also be SH.
(However the reverse is not necessarily true: if a constituent is SH it does not need to be also CH).

In that case, *testa rasata* would be re-classified in our table, with the effect that everything that counts as a categorial head also counts as a semantic head.

If, on the other hand, conceptual semantics is a crucial part of grammar and we need to consider *testa rasata* as semantically exocentric, then the principle has to be weakened as a tendency, not as a generalisation, in the following terms:

- (25) If a constituent is the CH then, preferably, it will also be the SH. A constituent can be the CH but not the SH only if the language has some specific typological properties such as the existence of hybrid categories.

Independent evidence of the second version of the principle has been offered in the literature. Consider, for example, Kageyama's (2008) analysis of adjectival nouns in Japanese, where it is argued that two different constituents can co-define the category label of the output even if only one of them is the semantic head. Notice, also, that following our definition of semantic exocentricity, the compounds of the type *testa rasata* cannot be considered endocentric, as the class denoted by the compound is unpredictable from the denotation of its parts.

4.3. Interaction between morphological exocentricity and semantic/categorial exocentricity

In the previous section we have analyzed the relationship between the categorial and the semantic head; in this section, we will consider the interaction between the morphological head and the other two kinds of heads. The ultimate goal of this part of the paper is to try to understand whether the morphological features of a compound are associated to the category, the semantics, both or none of them.

In order to answer this question, let's add the information about the morphological head to the previous table.

(26)

Left Headed	Morphological head	Example	Logical possibilities	Comment
[[C] []] [[S] []]	Coincide	It. <i>capostazione</i> 'head+station=stationmaster'	The same element is the CH and the SH	Normally called endocentricity
	Not coincide	Unattested		
[[C] [C]] [[S] []]	Coincide	Unattested	Two CHs one SH	
	Not coincide	Unattested		
[[C] []] [[S] [S]]	Coincide	Unattested	Two SHs one CH	
	Not coincide	Unattested		
[[S] []]	Coincide	Unattested	SH no CH	
	Not coincide	Unattested		
[[C] []]	Coincide	Unattested	CH no SH	
	Not coincide	It. <i>testa rasata</i> 'head+shaven=skin head'		
[[C] [C]] [[S] [S]]	Coincide	Sp. <i>poeta pintor</i> 'poet painter'	Two CHs and two SHs	Necessarily coordinative compounds
	Not coincide	Mordvin <i>penč.t'-vakan.t</i> 'spoon+plate=cutlery'		Co-compounds typically change their morphological gender to neuter, independently of the gender of the base.
[[] []] [[] []]	Not computable in Chinese (Chinese has an extremely reduced morphology); in Italian, necessarily not coincident.	Ch, 大小 <i>dàxiǎo</i> 'big+small=size' It. <i>bagnaasciuga</i> 'soak+dry=strand'	No SH, no CH	Absolute exocentricity

It is now possible to formulate some hypotheses about how the morphological head interacts with the other types of head.

Once we factor out the cases that are unattested because they are non-existent combinations of categorial and semantic heads, two generalisations emerge. The first one is that whenever a compound is semantically exocentric, it is also morphologically exocentric—the inverse not being true, as witnessed by co-compounds. The second one is that if a compound has only one categorial head which is also the semantic head, it must also be morphologically endocentric.

If the second generalisation is not surprising, giving the frequent assumption on Western linguistics that morphological features are markers of different categories, the first generalisation may lead to some surprising implications: a compound may be endocentric categorially and then exocentric morphologically, but cannot

be exocentric semantically and endocentric morphologically. These two generalisations, combined, seem to relate, in some sense, morphological features more closely with interpretative notions than with the definition of the grammatical category. We believe that this conclusion is challenging for theories of grammar that treat morphological exponents as dummy markers at phonological interface of categorial distinctions, such as, for example, Distributed Morphology. According to our database, it seems to be the case that semantic exocentricity may be a trigger for morphological exocentricity, while categorial endocentricity does not trigger morphological endocentricity.

4.4. Different types of exocentricity: Causes and consequences

The first result that we have obtained so far is that differentiating three levels of exocentricity, apart from being empirically adequate, pays off in an analytical sense, because it makes it possible to establish generalisations that can help to determine the status of morphological features or the connection between category and semantics in the morphology of natural languages.

There is a second way in which the distinction can pay off analytically, as we will argue in the next section. One of the crucial questions about exocentricity is what causes it from a synchronic perspective. If different kinds of exocentricity are not distinguished, there are very few things to say about this, but if we differentiate different levels of exocentricity, we will see that some generalisations arise.

More in particular, we will show in the next section that exocentricity which is driven by categorial or morphological features is relative to the typological properties of the language; in other words, it is not universal. We will illustrate this with an analysis of exocentric VN compounds in Romance, which, as we will see, are exocentric due to a particular property of NP licensing in this family of languages.

On the other hand, we will argue that exocentricity which is driven by semantic features is not relative to the parametric or typological properties of the language. It is universal, and, as such, we will see that identical semantic problems are solved in an identical way by languages typologically unrelated such as Chinese, English, Italian and Spanish. Let's proceed, therefore, with the analysis.

5. Categorially Driven Exocentricity and Semantically Driven Exocentricity

In this section we will show what are the advantages of considering separately semantic, categorial and morphological exocentricity. We will show that exocentricity caused by morphological or categorial requisites of the lexical items involved in the compound varies from one language to another, as it depends, crucially, on parameterised properties of the lexical items involved. We will illustrate this with the case of VN compounds in Romance (5.1). In contrast, we will show that exocentricity that derives from semantic conditions on the lexical items involved is universal and leads to comparable results in typologically unrelated languages. Semantic properties, therefore, do not seem to be parameterisable in different languages. We will illustrate this second point with a number of different exocentric compounds in Chinese, Italian and Spanish (5.2).

The general conclusion drawn from this last part of the paper is compatible with Kageyama's (2008) observation that the notion of categorial head depends on independent typological properties of the language; we add to this finding the generalisation that the semantic requisites of a compound are not parameterised in any language.

5.1. Exocentricity in Romance VN compounds and the licensing of categorial features

As we have seen, VN compounds are potentially very problematic cases, and as such they have extensively been studied in Romance (Zuffi 1981, Varela 1990, Bisetto 1999, 2008, Bok-Bennema and Kampers-Mahne 2005, Booij 2005, Ricca 2008). They illustrate exocentricity in three levels, as these authors noticed and as we have shown in the previous section.

We will add to the description presented in section 4 some notes about the semantic denotation of VN compounds. VN compounds typically denote agents and instruments—but not exclusively—and there is no formal marking of this relationship (27). They can also denote event nouns (28) and places (29). Furthermore, the argument relationship is not always patient; some other relationships are possible, including some restricted cases of agents (30).

- (27) a. Sp. *limpiabotas* → *limpiador de botas*
 'clean + boots = bootcleaner'
 b. It. *portacenera* → *portatore di cenere*
 'bring + ash = ashtray'
- (28) a. It. *batticuore*
 'beat + heart = heartbeat'
 b. It. *passaparola*
 'pass + word = order passed by word of mouth'
- (29) Sp. *rompe-olas*
 'break + waves = breakwater'
- (30) Sp. *canta-lobos*
 'sing + wolves = name of a village'

Some previous analyses concentrated on the possibility that the verbal form has been nominalised by a zero affix (Bisetto 1999) or by the theme vowel (Varela 1990), or proposed that the whole compound is nominalised by a zero affix.¹⁵ We believe that appeal to zero affixes is problematic, and we will instead try to propose an analysis where the apparent exocentricity of the compound derives from independent principles.

Here is the intuition which underlies the analysis:

- (31) In Romance, nouns which are arguments require a determiner.

¹⁵ Actually, Marchand (1969) proposed a zero suffix 'external' to the compound and Zuffi (1981) and subsequently Bisetto (1999) proposed a sort of an agentive suffix internal to the compound, attached to the Verb.

these compounds in Romance follow from here. In an intuitive sense, the N inside the compound, being an argument, requires a determiner, but for some independent reason the determiner is not available inside the word. The only way in which the noun can get the determiner it needs is having access to the one that will dominate the whole compound outside its internal morphological structure. But the only way of performing this is having access to the structure upside, which requires the categorial feature N to percolate in such a way that it is contained in the head of the word. This imposes its categorial label to the whole compound and explains why the whole word becomes a noun.

Our proposal is that the following principle determines which features are percolated to the higher node of the compound structure.

(35) Percolation of features inside the compound

Only active features of the constituents of a compound percolate to the head of the word and are accessible for syntax.

What counts as an active feature in our analysis is expressed in the principle of (36).

(36) Condition on the activation of features.

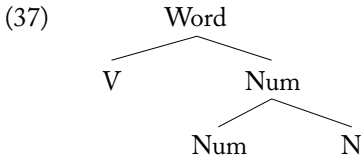
Only features which have not been formally licensed are still active in the representation.

This condition on the activation of features follows the spirit, if not the letter, of Chomsky's (2004) Activation Condition in the Minimalist Program. While Chomsky restricts activation to features which are still uninterpretable and do not have a value, we extend it to any kind of feature, semantically interpretable or not, which has not yet established a formal relationship with another unit. If the activation of features in syntax is indeed restricted to uninterpretable features, then our proposal indirectly defines a formal difference between the activation of features in syntax and in the lexicon.

A few words are in order to clarify in which sense our proposal differs from Di Sciullo and Williams' (1987: 26) notion of Relativised Head, which also proposes that there may be different heads for different features relevant for the word. First, notice that unlike these authors, who define the head for a feature G as the rightmost element containing G, we are not defining a head in positional terms, that is, we do not impose that the head must be the rightmost element carrying a property inside a configuration. In fact, as it will become apparent in the following pages, our proposal defines the head of a word as the highest node inside the word's internal structure, and its label depends on the features that are transmitted from the different constituents to that higher node. A second, and perhaps more important, point of difference is that we are proposing an independent motivation to determine which features from which constituents will still be active for syntax and, therefore, percolate as part of the label of the highest node of the word. Our goal is to motivate, rather than to stipulate, percolation of features in morphological structures. Via this procedure we aim to reduce exocentricity to a situation in

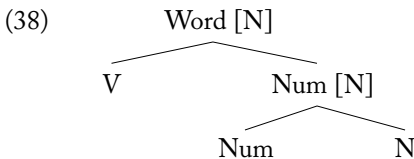
which there are active features in more than one constituent of the word, and, therefore, it cannot be said that only one constituent provides all the features that percolate to the highest node of the word.

Let us see how the principle of percolation of features combined with the requisite that argumental nouns require a determiner explain the categorial exocentricity of VN compounds in Romance. Following the data presented in (34), the nominal constituent that combines with the verb inside the compound is represented as headed by a Number morpheme:



Notice that in (37) we do not assign a label to the higher node of the compound, which we represent neutrally as *Word*. This is the one that will be visible for syntax, and, therefore, the one that will define the behaviour of the word in syntax. The reason for not assigning a label to this node is that this node will only contain the features that percolate from each of the constituents of the compound, and, therefore, its label will depend on which features are still active in each of its constituents. Let's analyse each kind of feature in turn.

We start with categorial features. We have two categorial features here: *V* and *N*. The feature *V* requires an internal argument,¹⁶ and this requisite is satisfied inside the compound, because this argument is satisfied by the nominal constituent. This means that the *V* categorial node will be satisfied, as it has already established a thematic relationship with a constituent labelled *N*. Following our proposal about the activation of features, then, this feature is inactive and, therefore, will not percolate to the upper node of the compound. In contrast, the constituent that contains the *N* feature is an argument and, as such, requires a determiner in Romance languages, but this determiner cannot be projected inside the compound for independent reasons. This implies that the feature *N* is not licensed, because inside the compound it has not established a formal relationship with a determiner. Therefore it is still active, and, as such, it percolates first to the Number head and then to the highest node of the compound (38):



¹⁶ We assume in this article Hale and Keyser's (2002) proposal that unergative verbs are actually derived in lexical-syntax with a nominal internal argument which is used to define the lexical specification of the verbal head.

As the *V* categorial feature is licensed inside the compound, but the *N* categorial feature is not, only *N* percolates to the whole word. The result is that the highest node, accessible for syntax, is defined as an *N* and therefore the compound will be treated as a noun in syntax.

This solves categorial exocentricity without the need to propose a specifically designated zero affix that changes the category of the verb and without treating the theme vowel of the verb as a categorially-ambiguous head.

Notice that this way of explaining the categorial exocentricity of *VN* compounds is dependent on the proposal that *N* needs a determiner and has to look for it outside the compound. The prediction is that, if there are languages where either argumental nouns can be licensed without a determiner or nouns do not keep an argumental relationship with verbs, a combination *VN* would be well-formed when the *V* projects its categorial label, as no licensing of the *N* feature of the noun would be necessary. This indeed seems to be the case. Notice that in polysynthetic languages (Baker 1988, 1996) there is the phenomenon called ‘noun incorporation’, whereby a complex word is formed by combining the verb with the noun (39).

- (39) wa’-ke-nakt -a-hninu-’
 FACT-1st.sg- bed-epenthetic vowel -buy-PUNC
 ‘I bought a bed.’ (Oneida, Iroquian)

Notice that in these languages, the label of the complex word is *V*, as shown, among other things, by the fact that verbal inflection appears in the word (factive, punctual and subject agreement morphemes). The reason for this is to be found in Baker’s (1996) own analysis of polysynthetic languages: in these languages, the arguments of the verb are the so-called agreement markers, which are actually pronouns, and the nouns are predicates of these pronouns. As the noun itself is not an argument but a predicate, the *N* feature does not need to be dominated by a *D* and licensing does not have to take place, so the label of the whole construction is *V*. This shows that this kind of exocentricity is dependent on the morphological and categorial parameters of the language.

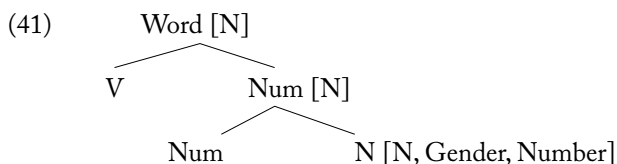
Going back to the tree structure in (38), let us now consider morphological exocentricity. Nouns in Romance are associated with two different morphological features: gender and number. Assume that both are represented as features contained inside the category *N*, which captures the intuition that all nouns in Romance have a gender and a number to become well-formed morphological objects. The question is whether these features are still active, and therefore percolate to the highest node for syntactic processing, or have been satisfied inside the compound. It becomes apparent from the description of compounds in Romance that the internal *N* is already combined with a head for number. For example, in Spanish, even mass nouns, which normally do not inflect for plural number, take a plural marking inside the compound (40).

- (40) a. agua ‘water’
 b. para-aguas ‘stop + waters = umbrella’

This empirical property provides independent evidence that a head for Number is present inside the morphological structure of the compound.

As for gender, remember that in Romance the gender of a noun is not predictable by semantic or syntactic means (Harris 1991, Alexiadou 2004). Gender is a lexical property that nouns carry with them in the lexicon, and, as such, it is not imposed by syntax. This suggests that gender is satisfied lexically in Romance, that is, that gender is licensed simply by being inside a head which contains a noun categorial feature.

Therefore, as gender is satisfied lexically by nouns and number is satisfied by the number constituent internal to the compound, these features will not percolate to the higher node of the compound, as represented in (41).



In (41) we see that among the three features contained in N, only the categorial noun is active, as it has not been licensed in any way inside the compound. Therefore, only N, and not Number or Gender, will percolate to the highest node of the word, with the result that in syntax the compound will be read as a noun, but its gender and its number will not be imposed by the internal noun.

Let us now consider the semantic exocentricity associated to these compounds. There are a number of questions that have to be addressed at this point. 1) Why are these compounds typically agent-instrument denoting? 2) Why are there cases of VN compounds that denote action nouns? 3) Why are there (very restricted) cases in which the argumental relationship inside the compound is an agentive one?

To answer these questions, we have to consider the theta roles that a particular verb is compatible with. On this point, we stand against extremely exo-skeletal proposals such as Borer (2005) where verbs as lexical items do not impose any particular theta grid on their representations. We will assume that each verb lexically requires some particular theta roles to be satisfied by its arguments. We will, however, be neutral with respect to whether these theta roles are imposed on the syntactic configuration by the verbal head—projectionist perspective—or the theta roles are imposed by the lexical meaning of a late-inserted item, as this is, again, an orthogonal problem for our discussion.

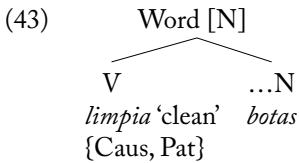
Consider why instruments and agents are typically denoted by these compounds. Instruments and agents belong to the general semantic class of causers of an event, with a difference in the animacy of the referent that acts as causer of the event. In fact, most VN compounds are ambiguous between an instrumental and an agentive reading (42).

- (42) a. *salva-vidas*
 ‘save + life = life vest or baywatch’

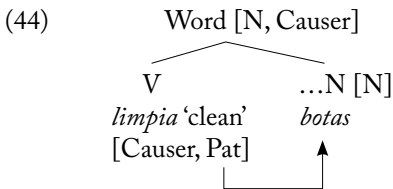
- b. *guarda-sellos*
 ‘keep + stamps = object to keep the stamps ordered or philatelic’

Some of these VN compounds are specialised in one of the readings, but here they are due either to the fact that the action denoted by the verb cannot be performed by a human or to some other world-knowledge notions. If we create the compound *broncea-mejillas* ‘tan + cheeks = cheek tanner’, it will be interpreted as an instrument because people cannot cause tanning; *par(a)-aguas* ‘stop + water’ will be interpreted as an instrument because people normally are not used to performing this action. Conversely, *porta-voz* ‘carry + voice = spokesperson’ will not be an instrument because objects do not talk. In any case, the internal structure of the compound predicts that the denotation will be a causer, leaving instrument or agent to world-knowledge.

Consider under this light the structure in (43).



From the two semantic features that the verb has, the Patient feature has already been satisfied when the N is taken as an argument of the verb; however, inside the compound no second noun which can be interpreted as an agent is introduced. Therefore, the agent feature is still active and percolates to the highest node. {Causer}, therefore, will be read by syntax and this explains that syntax treats this kind of compounds as agents (44).



Therefore, as we see, to the highest node of the compound features of both constituents percolate: categorial features of the N and semantic features of the V. This is, in a sense, reminiscent of the co-definition of the grammatical category of several compounds in Japanese identified in Kageyama (2008).

Let's consider now the cases where the Romance VN compound denotes an action. A generalisation that can be done in these cases is that systematically the verb which is used in the compound is either inchoative (45) or allows for an inchoative variant (46)

- (45) a. It. *batticuore*
 ‘beat + heart = heartbeat’
 b. It. *passaparola*
 ‘pass + word = order passed by word of mouth’

- (46) It. *alza-bandiera*
 ‘rise + flag = flag-rising’

The proposal is that these compounds are constructed from verbs that do not contain a Causer role and, therefore, when the time arrives that the compound’s referential role has to be coindexed with an open position, the only available position is the eventive theta (Davidson 1967) role of the verb itself (47).

- (47)
-
- ```

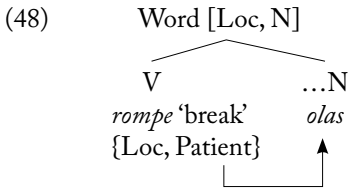
graph TD
 Root["Word [E, N]"] --- V["V"]
 Root --- N["...N [N]"]
 V --- V_ex["passa, 'pass'"]
 N --- N_ex["parola"]
 V_ex --- Role["{E, Patient}"]
 Role --> N_ex

```

This percolation of the eventive role of the verb seems to be a last-resort strategy to the extent that it is unavailable when there is a causer theta role still unassigned. The restriction seems to be that some semantic element has to percolate to the highest node of the compound, and when there is no normal theta role still active, then morphology percolates, by default, the eventive role associated to the verb. However, as far as the eventive role is concerned, it seems that nothing is wrong with it not being percolated to the upper node. Notice that in compounds which denote agents or instruments, the eventive role of the verb will still be active, as in these cases, and there is nothing ungrammatical about it not being coindexed with the referential role of the compound. It seems, therefore, that the restriction imposed by grammar is that the whole compound must be referentially coindexed with a meaning constituent of the verb, but coindexing with the eventive role is only possible in the absence of a prototypical argument position.

Additional evidence that coindexation with the eventive role of the verb is a last-resort option comes from the empirical fact that it competes with coindexation with a locative argument. Consider the unequivocally place-denoting VN compounds of identical meanings such as Sp. *rompe-olas* ‘breakwater’ or It. *frangiflutti* ‘break + waves = breakwater’, *sparti-acque* ‘divide + waters = watershed’. The verb *romper*, ‘break’, allows for an inchoative and a causative variant, but the VN compound is clearly constructed over the inchoative variant, to the extent that it means ‘place where waves break’, not ‘place where someone or something breaks the waves’. Therefore, it seems that the VN compound is associated to the same structure that makes grammatical *Aquí rompen las olas* in Spanish (literally ‘Here break the waves’), with a locative external argument. It seems, therefore, that ‘break’ is one of the verbs that allow for a theta grid that contains a location and a patient; (48) represents the VN counterpart of the sentence just mentioned, constructed with a verb ‘break’ that has a theta grid with a locative argument.





The remaining problem that we have to address are the few cases where the internal argument structure of the compound involves the notion of *Agent/Causer*. Even though these cases are unusual and have a very restricted semantics, they do exist, as shown in (49).

- (49) a. Cat. *Canta-llops* (name of a village in Girona, Spain)  
 ‘sing + wolves’  
 b. Sp. *Canta-lobos* (name of a village in Huesca, Spain)  
 ‘sing + wolves’  
 c. Sp. *Salva-diós* (name of a village in Ávila, Spain)  
 ‘save + god’  
 d. Fr. *Chante-loup* (name of several villages in France)  
 ‘sing + wolf’

The denotation of these VN compounds appears extremely systematic: they are all toponyms, that is, they all give proper names of places. This property can be related to another property of agentive verbs, noticed by Torrego (1989). When a locative subject is present in the structure, the agent-denoting DP behaves as an internal argument. The crucial data are presented in (50).

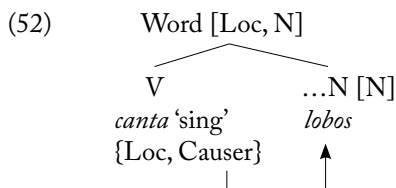
- (50) a. \**Juegan niños.*  
 play-3pl. children.  
 b. *Juegan los niños.*  
 play-3pl. the children.  
 c. *Aquí juegan niños.*  
 here play-3pl. children.

The crucial contrast refers to (50a) and (50c). In the absence of a locative, the post-verbal subject, as all agentive subjects in Spanish, cannot be a bare noun. When the locative subject is introduced, the agent can be expressed by a bare NP. This is a property of internal arguments (51), and as such a standard test to determine the unaccusativity of an intransitive verb.

- (51) a. *Juan come manzanas.*  
 Juan eats apples  
 b. *Nacen niños.*  
 are-born children.

Torrego’s analysis is that the locative subject in (50c) is introduced in the position of the external argument, which forces the agent to be introduced as an internal argument, if anything.

This analysis could be transferred to the VN compounds. If the compound is constructed with the theta grid that the verbs have in (50c), where the external argument takes a locative theta role and the internal argument takes an agent theta role, then the representation in (52) explains why VN compounds with an internal argument need to denote names of places: for the internal argument to get an agentive theta role, the external argument must be a location; the location theta role is not assigned to anything and, therefore, it percolates to the whole word.



In summary, we have shown that with the same internal structure, combined with a principle of activation of lexical features which states that a feature only percolates when it has not been satisfied inside the word-internal structure, we have motivated an analysis of the exocentricity of Romance VN compounds. We would like to emphasise that, as we have seen through the comparison with noun incorporation in polysynthetic languages, the fact that Romance VN compounds act as exocentric is a result of a language-particular property of argumental nouns, and, therefore, we do not expect VN compounds to be universally exocentric. A different case, where we expect universal exocentricity, is presented in the next section.

## 5.2. Exocentricity in compounds formed by antonymic predicates and universal retraction to noun

As we have seen, some particular types of exocentricity can be reanalysed as operations on features driven by language-particular requisites on the licensing of categorial or morphological properties of words. In contrast, in this section we will analyse a type of exocentricity which is driven by the semantics of the construction and not by the formal licensing conditions on categorial or morphological features. We will see that exocentricity caused by the semantics of the construction gives very similar results in typologically unrelated languages. This contrast suggests that semantics of natural languages is not subject to parametric variation, in such a way that equally problematic situations for the semantics are solved in similar ways, while the morphology and syntax is subject to parametrisation. Notice that in the latest developments in Minimalism this claim has also been generally supported (Chomsky 2005).

Chinese, Italian and Spanish can construct nominal compounds by combining two predicates—adjectives or verbs—with opposite meanings.

- (53) a. 大小 *dàxiǎo* 'big + small = size'  
 b. 长短 *chángduǎn* 'long + short = length'

- c. 深浅 *shēnqiǎn* ‘deep + shallow = depth’  
 d. 宽窄 *kuānzǎi* ‘wide + narrow = width’  
 e. 快慢 *kuàimàn* ‘fast + slow = speed’
- (54) a. It. saliscendi  
 ‘climb + descend = latch’  
 b. It. bagna-asciuga  
 ‘soak + dry = strand’  
 c. Sp. subibaja  
 ‘climb + descend = lift’

A crucial property of this kind of compounds is that there is a semantic primitive which underlies them in Spanish, Italian and Chinese: the notion of path. The path, taken as an ordered series of values, underlies both the notion of trajectory and the notion of scale. In the first case, the path is an ordered series of locative points and in the second case, the path is constructed over a series of degree values.

To gather evidence for this generalisation and to refine its nature, consider the Chinese cases. Notice that Chinese can designate the name of an abstract property just by taking two adjectives with opposite meanings, but this is not possible with all kinds of properties, but only with those which can be defined inside an open scale (Kennedy 1999, Kennedy and McNally 2005) and have two opposite extremes inside that scale. Therefore, the abstract property of colour, which is not constructed as an ordered scale (for example, yellow is not a degree of red) and does not make oppositions of values (red is not the opposite of green), cannot be referred to by an AA compound.

- (55) 颜色 *yán-sè* ‘colour + expression’

However, having an underlying scale is a necessary but not a sufficient condition. Scales which are closed in one of the extremes are not fit for this kind of compounds. Take, for example, the scales that underlie the adjectives ‘clean’ and ‘dirty’. The scale of clean is closed in the lowest point, as can be shown by the fact that a minimal quantity of dirt is enough to claim that something is dirty. As for the scale associated to clean, it is also closed, but in the highest point, as only an entity which is maximally clean can be said to be clean—a minimal quantity of dirt, again, is enough to say that the thing is dirty. Consequently, there is no AA compound which refers to the property of being dirty or the property of being clean.

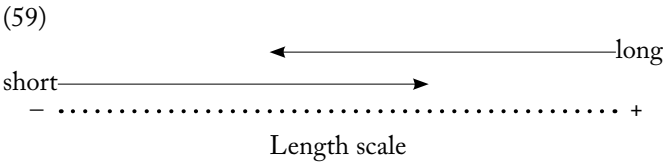
- (56) a. 明 *míng* ‘cleanness’  
 b. 污秽 *wūhuì* ‘dirt + dirt = dirtiness’.

A final condition on the nature of the scale seems to be that for a property to be referred to by an AA compound it requires to be associated with two lexical items which designate opposite values inside the scale. For example, an adjective associated with an open scale but without an opposite value, such as ‘bored’, does

not express its abstract property by an AA compound (57). Notice that the scale of being amused is independent from the scale of being bored, as shown by the fact that the negation of bored does not imply any value of amused, while the negation of short implies a certain value of long (58).

- (57) 无聊性 *wúliáoxìng* ‘boring + suff = property of being boring’  
 (58) a. John is not bored  $\nrightarrow$  John is amused.  
 b. This stick is not short  $\rightarrow$  This stick is long.

Therefore, AA compounds in Chinese can only denote abstract properties when Chinese has two lexical items which lexicalise opposite directions in the values of the same scale. Consider, for example, the scale associated with length. Inside this scale, the lexical item ‘long’ refers to the interval that goes from the upper side of the scale to a point where the arbitrary standard of comparison is defined; the lexical item ‘short’, in contrast, goes from the lower side of the scale to the point where this very same standard of comparison is defined. There is an arbitrary overlap between the two lexical items, which, out of context and therefore without a standard of comparison arbitrarily defined, denote opposite directions inside the same scale, as represented in (59).

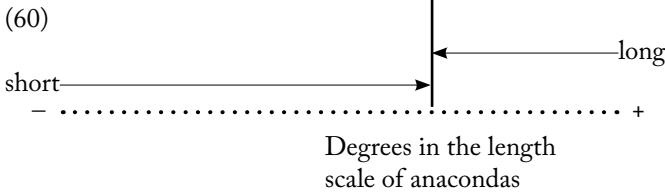


Now that we have provided evidence that the semantic primitive of path underlies always this class of Chinese exocentric compounds, let us consider how the semantic properties of this path explain the exocentricity in the construction.

As noticed in the semantic literature (Kamp 1975, Zwarts 1992, Kennedy 1999, among many others), an adjective denotes a value of a property. This value is defined inside a scale. The minimal value that qualifies as a value of a particular property depends on the position inside the scale that is picked up in the particular context as the standard of comparison which defines what counts as a relevant value of that property in that particular utterance. If the standard of comparison inside the length scale is anacondas, which has a particular value of length, then a cobra counts as a short object, but if the standard of comparison is worms, which has a different value of length inside the scale, then a cobra is not a short object. Notice, therefore, that, as Kamp (1975) observes, it is not possible to assign truth value to a predicate constituted by an adjective in the absence of a standard of comparison.

However, once that the standard of comparison is fixed as a point in the scale, it is impossible that the same object exhibits a value of a property which at the same time counts as short and long. Once that the standard of comparison is defined, ‘long’ denotes the interval which goes from that standard of comparison value—including it—to the top of the scale, while ‘short’ denotes the values below

that standard of comparison (60).



This semantic property of adjectives explains that no object can satisfy at the same time the truth value of long and short; the same goes for all pairs of adjectives that have the scalar properties of long/short, among them big/small, deep/superficial and narrow/wide.

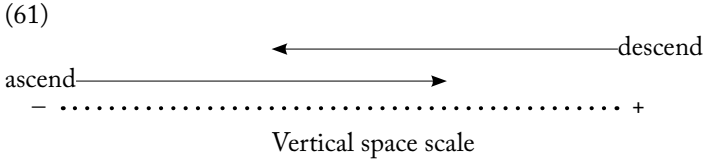
The semantic denotation of a compound composed by two opposite adjectives is, therefore, incompatible with that of an adjective, but notice that it can be naturally used to name the whole scale. As shown in (59), each of the adjectives expresses opposite directions inside the same scale, with an arbitrary overlap between them in the absence of a standard of comparison that distributes the values in two intervals. Let us concentrate in the space where the two adjectives can overlap. There is a part of the scale that can be identified as the interval where the two lexical item overlap. In other words, the overlap of the two lexical items names a part of the scale, and as the scale is open in both extremes, this overlap can be arbitrarily maximal in the absence of a standard of comparison and, therefore, cover the whole scale. Therefore, this kind of compounds naturally can give name to the whole scale. However, if the two lexical items, instead of naming a single value inside the scale, which is impossible given (60), name the whole set of values inside the scale, they must be mapped as nouns and not as adjectives.

Therefore, we have derived categorial exocentricity in Chinese AA compounds and also predicted the category of the whole from a single semantic notion which is associated to the values inside the scales. The semantic operation that we have just presented also explains why only lexical items associated with scales of the type that we have described can take part in this kind of exocentricity. If the lexical item is associated with a scale where the standard of comparison is semantically fixed in the upper end, 'clean', or in the lower end, 'dirty', there will be no arbitrary overlap between the two adjectives and this overlap cannot be used to refer to name the whole scale. Notice also that if there is only a lexical item inside the scale and it does not compete with an opposite lexical item, it is not possible to use that lexical item combined with another one to denote an overlap inside the scale.

Let us go now to Romance languages and see how the same kind of semantic incompatibility inside paths explains that two lexical items with opposite values can only combine as nouns that name a path or a notion very closely related to it. Consider, again, the relevant example *bagnasciuga* in Italian and *subibaja* in Spanish.

The nominal compound *subibaja* in Spanish is composed of two verbs, *subir*, which means 'to ascend', and *bajar*, which means 'to descend'. If we concentrate in

the meaning of these verbs in isolation, notice that they denote opposite transitions inside the same path, in this case a locative path defined vertically. ‘Ascend’ denotes a transition from a lower to a higher value, and ‘descend’ denotes a transitions from a higher to a lower value, always in a series of ordered spatial points (61).

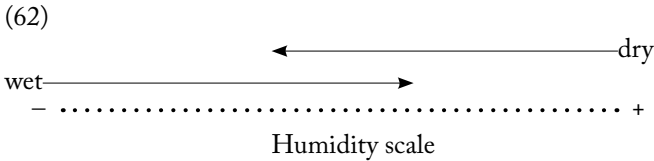


Notice that if we express each one of these lexical items as a verb, the transition that they denote from one point in space to the other is mapped into a temporal sequence. That is, as a verb ‘ascend’ means that in the course of a temporal interval it is true that there has been a transition from a point X to a point Y, such as that the point X is ordered higher in a vertical path than the point Y. ‘Descend’, as a verb, means the opposite, that is, that in the course of the same temporal interval there has been a transition from the point Y to the point X, such as that Y is ordered lower than X in the vertical axis. Notice that, as happened in the case of AA compounds in Chinese, it is impossible that the truth condition of both transitions are met under these circumstances. It is not possible that a transition from X to Y and a transition from Y to X are both met in the same point in time, which clearly shows that if the two lexical items are combined it cannot be inside a category which is bound by a time expression such as the verb. Instead, the two expressions combined have a coherent meaning if, in a way similar to the Chinese examples, they are taken to refer to the interval of the path where the two lexical items can overlap. Consequently, *subibaja* names this interval and therefore has a semantics which means the event of transitioning from one point to the other inside this path.

A similar consideration can be made about other VV compounds in Italian and Spanish that combine two movement verbs which denote opposite transitions inside the same spatial path and are required, therefore, to be interpreted as a noun which names an irregular series of spatial transitions inside that very same path: *metesaca* ‘put-in/take-out’, colloquial Spanish for ‘copulation’, or the Italian *andirivieni*, also meaning ‘swinging’, with a first constituent from *andare* ‘to go’, and a second constituent which comes from the verb *rivenire* ‘to come back’, as second constituent; occasionally, the verbs are shown in a conjugated form, but the semantic operation that takes place is the same and they denote the event noun that corresponds to an irregular movement: Spanish *vai-vén*, composed of ‘go’ and ‘come’, ‘swinging’. A closely related compound in English, where the meaning of opposition is obtained by combining two incompatible temporal forms, is *seesaw*, which also names the event of swinging.

Although the most productive VV combination in Italian and Spanish of this kind seems to be the one which produces event nouns, other interpretations compatible with the notion of path are also available. Consider, for example, the Italian

*bagna-asciuga* ‘the part of the shore where the waves come and go’, which clearly denotes a location. As it is probably clear by now, *bagna* ‘to wet’ and *asciuga* ‘to dry’ denote opposite transitions inside a scale of humidity. *Bagnare* refers to a transition from a lower to a higher degree of humidity and *asciugare* refers to the opposite transition, from a higher to a lower degree of the same scale. Both transitions overlap in an arbitrary interval of the scale, as represented in (62).



As in the other cases, the combination of the two lexical items does not have sense unless it stands for the name of the interval where the two items overlap. However, in this case, this interval is mapped to a location, that is, the interval covered by the two lexical items refers to a spatial area where cyclically an event of drying is followed by an event of wetting.

This type of denotation, where a temporal transition inside a scale is mapped into a spatial path, is reminiscent of some cases discussed in detail by Gawron (2005). Consider (63).

(63) The crack widens from the ceiling to the ground.

The crucial property of (63) is that it does not denote a dynamic event, but describes a state. Gawron's explanation is that the lexical meaning of the verb requires to denote a path, more in particular the path defined inside the scale of wideness which underlies to its interpretation. This path is normally mapped into a temporal domain, and then the verb denotes a dynamic event such as that every step in time is a progression in that scale of wideness. However, it can also be mapped into the spatial domain, as in (63), and then the verb does not denote a temporal transition, but describes the portion of space which is occupied by its subject.

The case of *bagnasciuga* seems to be an instance of the same operation whereby the path that normally denotes a transition in the temporal domain is taken to describe an area covered by these transitions. In connection with cases like *subibaja*, *bagnasciuga* also extrapolates the transition outside from time, but the difference is that, while *subibaja* does not map the transition to the spatial domain but just names the action of moving inside that path, *bagnasciuga* straightforwardly maps the path into the spatial domain.

Finally, another class of VV compounds is created from lexical items with related, but complementary meanings. Although this class of compounds does not strictly belong to the group which we have identified, because there is no notion of path involved in them, yet it is worth noticing that even in these cases the output category is a noun. Typically, they denote machines which perform both complementary actions, which are incompatible with each other in the same time interval,

in a sequential fashion: *lava-asciuga* ‘wash-dry’, Italian for ‘washing machine’, *taglia-cuci* ‘cut-saw’, Italian for ‘sawing machine’, or *sali-scendi* ‘ascend-descend’, Italian for ‘elevator’. To the extent that these nouns take two actions that cannot be performed simultaneously and dissociate them from the temporal domain by projecting the compound as a noun, they are closely related to the case studied here.

Our analysis makes a clear prediction with respect to VV compounds of this kind in Spanish. Notice that parasynthetic verbs in Spanish cannot be taken to denote stative verbs, as opposed to other verbs which do not have a parasynthetic morphological structure (64). This seems to suggest that the presence of the prefix, in a sense, forces the path structure associated to the verb to be mapped into the temporal domain.

- (64) a. \*La grieta se a-larg-a del techo al suelo.  
           the crack SE pref-long-V from.the ceiling to.the ground  
           ‘The crack widens from the ceiling to the ground’  
       b. La grieta va del techo al suelo.  
           ‘The crack goes from the ceiling to the ground.’

The prediction that we make is that there will not be VV compounds of the relevant kind in Spanish where one or both verbs are parasynthetic, for these verbs will have to be mapped into the temporal domain and, therefore, their combination will be uninterpretable for semantics. This prediction, to the best of our knowledge, is confirmed by the data. Compounds as those in (65) are unattested in any reading in Spanish.

- (65) a. \*en-sancha-estrecha  
           from *en-sanch-ar*, parasynthetic ‘widen’.  
       b. \*alarga-acorta  
           from *a-larg-ar*, parasynthetic ‘lengthen’, and *a-cort-ar*, parasynthetic ‘shorten’.  
       c. \*calienta-enfría  
           from *en-fri-ar*, parasynthetic ‘cool down’.

Italian, however, behaves differently from Spanish in this sense, for parasynthetic verbs may be taken in non-dynamic readings (e.g., *a-(l)lung-are* ‘to widen’) and, as expected if these two properties are related, parasynthetic verbs can actually give rise to nominal VV compounds in this language.

Therefore, we have shown that, in contrast with the kind of exocentricity found in VN compounds, which is parametrised to the extent that it implies a particular relationship between a noun and a verb which is subject to the nature of syntactic and morphological features, this kind of exocentricity works in the same way and leads to identical results in typologically unrelated languages, among them Chinese, English and Italian. This seems to suggest, as we noticed at the beginning of the section, that semantically-driven exocentricity is universal, while morphological and categorial exocentricity is subject to the morphosyntactic properties of the language.



## 6. Conclusions

In this paper, we have argued for three points that, in our opinion, challenge the traditional interpretation of the phenomenon of exocentricity in an interesting way. The first one is that, against some previous theories that connected productivity with endocentricity, most languages have productive word formation rules that produce an exocentric result. Exocentricity, is, therefore, not marginal, although there are different degrees in which it can be manifested. The second point is that we gain a better understanding of the phenomenon of exocentricity if we distinguish between categorial, morphological and semantic exocentricity, which, as we have shown, are empirically independent from each other. Being exocentric is, in our theory, having features from more than one constituent internal to the compound percolate to the highest node of the word, the one read by syntax. The separation in three levels has led us to identify a generalisation, which is our third point in this article: categorially or morphologically-driven exocentricity is language-particular, while semantically-driven exocentricity is independent of the typological properties of each particular language. We have illustrated this difference with two case studies: VN compounds in Romance and exocentric nominal compounds related to paths in a number of languages.

Some empirical issues stand out as topics for forthcoming research. As we have seen, the most widely documented exocentric compound type is the attributive combination of a noun and adjective (the *pale face* type). We have not provided a case study for this class of compounds, which are relevant, among other things, to understand how attribution is performed inside words. A second empirical issue is to determine whether, as our preliminary research suggests, there are no exocentric coordinative compounds—with the only possible exception of the *taglia-cuci* type in Italian—and, if confirmed, why this should be the case. Finally, we have found out that Absolute Categorial Exocentricity cannot produce compounds of verbal class; again, the question would be what makes verbs special in this sense. But all these questions exceed the limits of this article and will be addressed in forthcoming research.

## References

- Allen, Margaret (1978) Morphological investigations. Unpublished doctoral dissertation, University of Connecticut, Storrs.
- Alexiadou, Artemis (2004) Inflection class, gender, and DP-internal structure. In: Gereon Müller, Lutz Gunkel, and Gisela Zifonun (eds.) *Explorations in nominal inflection*, 21–50. Berlin: Mouton de Gruyter.
- Baker, Mark (1988) *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Baker, Mark (1996) *The polysynthesis parameter*. New York: Oxford University Press.
- Bisetto, Antonietta (1999) Note sui composti VN dell'italiano. In: Paola Benincà, Alberto Mioni, Laura Vanelli (eds.) *Fonologia e morfologia dell'italiano e dei dialetti d'Italia*, 503–538. Roma: Bulzoni.
- Bisetto, Antonietta and Sergio Scalise (2005) Classification of compounds. *Lingue e Linguaggio* 2: 319–332.

- Bisetto, Antonietta and Chiara Melloni (2008) Parasynthetic compounding. *Lingue e Linguaggio* 2.
- Bloomfield, Leonard (1933) *Language*. New York: Henry Holt and Co.
- Bok-Bennema, Reineke and Brigitte Kampers-Mahne (2005) Taking a closer look at Romance VN compounds. In: Chiyo Nishida and Jean-Pierre Montreuil (eds.) *New perspectives on Romance Linguistics, vol. 1: Morphology, syntax, semantics and pragmatics*, 13–27. Amsterdam: John Benjamins.
- Booij, Geert (2007) *The grammar of words*. 2nd ed. Oxford: Oxford University Press.
- Borer, Hagit (2005) *Structuring sense*. Oxford: Oxford University Press.
- Ceccagno, Antonella and Sergio Scalise (2006) Classification, structure and headedness of Chinese compounds. *Lingue e Linguaggio* 2: 233–260.
- Chomsky, Noam (1994) Bare phrase structure. *MIT working papers in linguistics*. [Reprinted in Gert Webelhuth (ed.) *Government and binding theory and the minimalist program*. Blackwell, 1995.]
- Chomsky, Noam (1995) *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam (2004) *The generative enterprise revisited: Discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui and Miboko Zushi*. Berlin: Mouton de Gruyter.
- Chomsky, Noam (2005) Three factors in language design. *Linguistics Inquiry* 36: 1–22.
- Cinque, Guglielmo (2005) The dual source of adjectives and phrasal movement in the Romance DP. Ms., University of Venice.
- Davidson, Donald (1967) The logical form of action sentences. In: Nicholas Resher (ed.) *The Logic of Decision and Action*, 81–95. Pittsburgh: University of Pittsburgh Press.
- Dimmendaal, Gerrit (1983) *The Turkana language*. Dordrecht: Foris.
- Gawron, Jean Mark (2005) Generalized paths. In: Effi Georgala and Jonathan Howell (eds.) *Proceedings of semantics and linguistic theory XV*, 43–55. Ithaca, NY: Cornell Linguistics Publications
- Guevara, Emiliano and Sergio Scalise (2009) Searching for universals in compounding. In: Sergio Scalise, Elisabetta Magni, Antonietta Bisetto (eds.) *Universals of language today*, 101–128. Amsterdam: Springer.
- Hale, Kenneth and Samuel J. Keyser (2002) *Prolegomena to a theory of argument structure*. Cambridge, MA: MIT Press.
- Harris, James (1991) The exponence of gender in Spanish. *Linguistic Inquiry*, 22: 27–62.
- Jackendoff, Ray (1997) *The Architecture of the language faculty*. Cambridge, MA: MIT Press.
- Kageyama, Taro (2008) Semantic effects of left-hand elements on right-hand head structure. *The 136th LSJ meeting handbook*, 68–73. Linguistic Society of Japan.
- Kageyama, Taro (2009) *Isolate*: Japanese. In: Rochelle Lieber and Pavol Štekauer (eds.), 512–526.
- Kamp, Hans (1975) Two theories of adjectives. In: Edward Keenan (ed.) *Formal semantics of natural language*, 123–155. Cambridge: Cambridge University Press.
- Kennedy, Christopher (1999) Gradable adjectives denote measure functions, not partial functions. *Studies in the Linguistic Sciences* 29(1): 65–80.
- Kennedy, Christopher (2001) Polar opposition and the ontology of ‘degrees’. *Linguistics and Philosophy* 24: 33–70.
- Kennedy, Christopher and Louise McNally (2005) Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81: 345–381.
- Lieber, Rochelle (2004) *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
- Lieber, Rochelle and Pavol Štekauer (eds.) (2009) *The Oxford handbook of compounding*.

- Oxford: Oxford University Press.
- Marchand, Hans (1969) *The categories and types of Present-day English word formation*. München: Beck'sche Verlagsbuchhandlung.
- Namiki, Takayasu (2001) Further evidence in support of the righthand head rule in Japanese. In: Jeroen van de Weijer and Tetsuo Nishihara (eds.) *Issues in Japanese phonology and morphology*, 276–297. Berlin: Mouton de Gruyter.
- Ricca, Davide (2008) VN Compounds in Italian: Data from corpora and theoretical issues. Presented at the CompoNet Congress of Compounding, University of Bologna, June 6–9, 2008.
- Scalise, Sergio and Antonietta Bisetto (2009) The classification of compounds. In: Rochelle Lieber and Pavol Štekauer (eds.), 34–53.
- Soegaard, Anders (2005) Compounding theories and linguistic diversity. In: Zygmunt Frajzyngier, Adam Hodges and David S. Rood (eds.) *Linguistic diversity and language theories*, 319–337. Amsterdam: John Benjamins.
- Williams, Edwin (2007) Dumping lexicalism. In: Gillian Ramchand and Charles Reiss (eds.) *The Oxford handbook of linguistic interfaces*, 353–383. Oxford: Oxford University Press.
- Zuffi, Stefano (1981) The nominal composition in Italian. *Journal of Italian Linguistics* 6(2): 1–54.
- Zwarts, Joost (1992) *X'-syntax-X'-semantics, on the interpretation of functional and lexical heads*. Unpublished doctoral dissertation, Utrecht University.
- Zwicky, Arnold (1985) Heads. *Linguistics* 21: 1–29.

*Authors' contact information:*  
 Department of Foreign Languages  
 University of Bologna  
 via Cartolerie 5, 40124 Bologna, Italy  
 Scalise: scalise@lingue.unibo.it  
 Fábregas: webfabregas@gmail.com  
 Forza: francescaforza@gmail.com

[Received 31 October 2008;  
 Accepted 16 December 2008]

【要 旨】

複合語形成における外心性

SERGIO SCALISE    ANTONIO FÁBREGAS    FRANCESCA FORZA

(セルジオ・スカリーゼ) (アントニオ・ファブレガス) (フランチェスカ・フォルツァ)

ボローニャ大学

複合語が外心構造か内心構造かは主要部の概念にかかっている。すなわち、主要部を持つ複合語は内心構造、主要部を持たない複合語は外心構造とされる。主要部の概念は通常、一元的なものとなされるので、したがって外心性も従来は一元的な概念とされてきた。本稿では、まず、外心構造の基準と制限について類型論的なデータを提示し、それに基づいて、主要部の概念が、範疇としての主要部、意味的な主要部、形態論的な主要部という3つの異なる要素に分かれることを論じる。これにより、外心性という概念も、範疇の外心性、意味の外心性、形態的特徴の外心性に分割される。本稿では、複合語全体ではなく複合語の構成要素の性質に基づいて、外心複合語に関する新しい分析を提示する。