

Compounding in Italian

Sergio Scalise

In this paper, Italian compounds are analyzed within a theoretical framework best known as Lexical Morphology. Given the hypothesis that Italian Compounds are formed with words belonging to the four major lexical categories, a complete list of compounds (both possible and not possible) is discussed. Italian contemporary productive compounds have their head on the left (those with head on the right are, usually, remnants from Latin).

The problem of the relationship between compounding rules and derivation rules is also addressed. The conclusion is that one must distinguish between strict and loose compounds (in the former there are cases of derived compounds, while in the latter the opposite order applies). Inflection of compounds is usually regular in "new" compounds: the head is inflected, but there are many exceptions to be dealt with.

In the last part, the issue of the distinction between compounds and other constructions is addressed (compounds and derivatives, compounds and phrases). The conclusion is that, given a consistent theoretical background, one can always distinguish between compounds and other constructions, especially considering that morphology is basically different from syntax.

1. Introduction.

Compounding is a morphological process which, like derivation, makes it possible to form new words starting from existing ones. Nonetheless, while a derived word is made up of one free and one bound form, a compound word is normally composed of two free forms; that is by two words (or constituents) which, for the sake of convenience, will here be called *Word₁* and *Word₂*. An example of this can be seen below:

(1)	<i>Word₁</i>	<i>Word₂</i>	→	Compound
	[campo] _N 'field'	[santo] _A 'holy'	→	[[campo] _N [santo] _A] _N 'cemetery'
	[alto] _A 'high'	[piano] _N 'plain'	→	[[alto] _A [piano] _N] _N 'plateau'

[lava] _V 'wash'	→	[[lava] _V [piatti] _N] _N 'dishwasher'
[sali] _V 'rise'	→	[[sali] _V [scendi] _V] _N 'dumbwaiter'
[senza] _P 'without'	→	[[senza] _P [tetto] _N] _N 'homeless'
[capo] _N 'head'	→	[[capo] _N [stazione] _N] _N 'station-master'
[agro] _A 'bitter'	→	[[agro] _A [dolce] _A] _A 'bitter-sweet'

From a formal point of view, compounding consists of the concatenation of two words with a given syntactic category. The result is another word with another syntactic category which may or may not be the same as that of the two constituent words. This aspect of compounding may be illustrated through the following rule (cf. Allen 1978):

$$(2) []_X, []_Y \rightarrow [[]_X \# []_Y]_Z$$

The syntactic category of the resulting compound (that is Z) may be "derived" from the syntactic category of the constituent words¹ (i.e. X and Y). In Italian, however, the compound is always a noun unless both constituents are adjectives, in which case the resulting compound will be an adjective. Therefore, the following generalization can be drawn:

$$(3) \begin{array}{l} a. X + Y \rightarrow N \\ b. A + A \rightarrow A \end{array}$$

which can be read as follows: (a) in Italian compounding the combination of any two categories (X and Y) gives rise to a noun; (b) only by combining two adjectives does one obtain an adjective.²

As was mentioned above, compounds are made up of two free forms. In effect compounds do exist which have been made up of two bound forms or by one free form plus a bound form as is the case of those compounds made by joining together so-called "semi words".³

(4)	dattilo + scritto	(semiword + noun)	'typed manuscript'
	bio + grafo	(semiword + semiword)	'biographer'
	astro + nauta	(noun + semiword)	'astronaut'

¹ The syntactic category of the compound, in fact, is the same as the category of the head of the compound. Cf. section 3.

² There is one exception to the generalization just given; the case of compounds made up of an adjective describing a color such as *verde bottiglia* or 'bottle green' where the combination A + N gives rise to an adjective. Nonetheless, as previously mentioned this type of compound is rather restricted and limited to adjectives of color (*giallo canario* 'canary yellow', etc.).

³ Cf. Scalise (1983). Migliorini (1957) called these bound forms "affixoids".

Again in this case the above generalization holds true; that is, the compound is always a noun.

2. Italian compounds.

The table below lists the combinations possible when forming Italian compounds (with indication of the major categories: N, V, A, P and Adv).⁴ In the following columns further specification is given or, more precisely, a) whether the type of compound exists, b) if it is productive, and finally some examples are given.

(5)	category	existing	productive	examples
i.	N + N	yes	yes	crocevia, pesceccane 'crossroad' 'shark'
ii.	A + A	yes	yes	dolceamaro, verdeazzurro 'bitter-sweet', 'blue-green'
iii.	V + V	yes	no	andirivieni 'comings and goings'
iv.	P + P	no	no	*dicon, *sottoper
v.	Avv + Avv	yes	no	malvolentieri 'unwillingly'
vi.	V + N	yes	yes	scolapasta, cantastorie 'colander', 'ballad singer'
vii.	V + A	no	no	*pagacaro, *vedibello
viii.	V + P	no	no	*metticon, *saltasopra
ix.	V + Avv	no	no	*sentimale, *vedimeglio
x.	N + A	yes	no	camposanto, fededegno 'cemetery', 'trustworthy'
xi.	N + V	yes	no	manomettere, crocefiggere 'tamper with', 'crucify'
xii.	N + P	no	no	*scalasotto, *abitosenza
xiii.	N + Avv	no	no	*casamale, *tavolobene
xiv.	A + N	yes	no	altopiano, gentiluomo 'plateau', 'gentleman'
xv.	A + V	no	no	*gentileparla, *caropaga
xvi.	A + P	no	no	*bellocon, *biancosenza
xvii.	A + Avv	no	no	*bellobene, *biancooggi
xviii.	P + N	yes	no	sottopassaggio, oltretomba 'underpass', 'next world'
xix.	P + A	no	no	*congentile, *soprabello

⁴ The categorial status of adverbs is, indeed, controversial.

xx.	P + V	yes	no	? contraddire, ? sottomettere ⁵ 'contradict', 'subjugate'
xxi.	P + Adv	no		*combene, *senzaieri

As one may note, not all the combinations of categories theoretically available are possible.⁶ For example, (5 xii), (5 xvii), (5 viii) (5 vii) and (5 xv) — repeated here below for convenience — are not found:

- (6) *N + P *passaggiosotto
 *A + P *bellosopra
 *V + P *baticontro
 *V + A *girasvelto
 *A + V *bellovedì

there are structural reasons for the non-existence of these compounds. For example, one immediately notes that for the N and A categories both orders (A + N and N + A) are possible while for the categories N and P only P + N is possible and not *N + P. Actually, the combinations *A + P and *V + P do not exist either. Therefore, generalizing, it can be said that in Italian no compound can be found which is composed of any syntactic category X followed by P:

- (7) *X + P

Thus (7) rules out all compounds with P in the second position. The reason would appear to be due to the relationship between syntax and morphology: in syntax the order is 'preposition + noun' and not the other way around (*con pazienza* 'with patience' and not **pazienza con* 'patience with').

In the list given in (6) one will note that P can only be linked with nouns and not, for example, with adjectives; this is due to the fact that in syntax (and thus likewise in compounding) a preposition can be "followed by" a noun but not by an adjective (*con amore* 'with love' vs. **con gentile* 'lit. with kind'). As is the case with prepositions, verbs too are "followed by" a noun but not an adjective (*recitare sonetti* 'recite sonnets' vs. **recitare gentile* 'lit. recite kind') and thus in compounding we find forms of verb plus noun but not verb plus adjective. This leads one to conclude that the rule governing compounding is in agreement with the rules governing syntactic constructions. This is obviously the norm. Each language will, however, deviate from the norm in various ways around a fringe of irregularities — often called "periphery" — found in all linguistic systems.

⁵ The question mark indicates the fact that analysis of this form is anything but simple. Are we dealing with a compound verb (i.e. made up of a preposition plus a verb) or a derived verb (i.e. made up of a prefix plus a verb)? This issue will not be dealt with here, but see section 11.

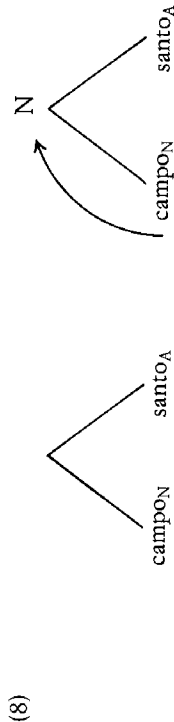
⁶ A comprehensive list of possible compounds should also include the morphologically more complex forms such as: [N + V + ente] (*nullatenente* 'have-not'), [Adv + V + ente] (*malbrivente* 'gangster'), *benestante* 'well-to-do'), [V + P + N] (*salimbanco* 'tumbler'), [N + P + N] (*pomodoro* 'tomato'), [V + Cong + V] (*avvenuti* 'come and go'), [Adv + P part] (*soprattutto* 'above-mentioned', *benvenuto* 'welcome'). These types, however, do not appear productive.

3. Head of compounds.

From what has been said above, the structure of a compound such as *camposanto* is the following:

- (8) [[campo]_N # [santo]_A]_N

This compound has the same syntactic category (noun) as one of its constituents, *campo* which is also a noun. It can be said that *campo* is the head of the compound and that the category N of the compound 'percolates' from the head:



In other words, *camposanto* "IS A" noun because *campo* "IS A" noun. The "IS A" test is applicable both for the syntactic category and for the semantics of the compound. Let us consider a compound such as *cassaforte* 'safe':

- (10) [[cassa]_N [forte]_A]_N

From the syntactic category point of view *cassaforte* "IS A" noun and thus the head is *cassa* 'box'. From the semantic point of view *cassaforte* "IS A" *cassa* (or type of 'box'): thus one can conclude that *cassa* is the head of the compound (both in terms of syntactic category and semantics). Now let us take another compound, *capostazione* 'station-master' and apply the "IS A" test for the syntactic category:

- (11) [[capo]_N [stazione]_N]_N

In this case the category test does not give a clear response: *capostazione* "IS A" noun, although both *capo* and *stazione* 'are' nouns. More in depth analysis, however, indicates that *capo* "IS A" [+masculine], [+animate] noun and that *stazione* "IS A" [-masculine], [-animate] noun. *Capostazione* "IS A" [+masculine], [+animate] noun just like *capo* and unlike *stazione*. *Capo* is thus the "head" of *capostazione*.

Therefore: if it proves impossible to identify the head solely on the basis of syntactic category one can resort to other types of information such as, for example, syntactic-semantic features which are part of the word's lexical representation. It is worth noting that application of the semantic test brings us to the same conclusion: *capostazione* "IS A" *capo* and "IS NOT A" *stazione*.⁷

⁷ Only semantic features can be used in the case of "reversible" compounds such as the English *Christmas cookie* vs. *cookie Christmas* or the Dutch (*de*) *veld-sport* 'outdoor sport' vs. (*het*) *sport-veld* 'sports area'.

Thus we can conclude that it is from the "head" element that all the following linguistic information are passed to the compound itself: a) information regarding category; b) syntactic-semantic features; c) gender. In the clearest cases the "syntactic" and "semantic" heads of the compound coincide.

4. Loose and strict compounds.

In the rule for compounding given in (2) above it has been assumed that between Word₁ and Word₂ there is a "strong boundary", '#'. This boundary separates the phonological and semantic identity of the two constituents and blocks semantic and/or phonological amalgamation processes. As time passes the strong boundaries can weaken in ways which will not be dealt with here; they can, however, be indicated as follows:

(13) # → +

In other words: the 'strong' (word) boundary weakens, becoming a 'morpheme boundary'.

Now let us consider the following compounds:

- (14) a. quinta essenza → quintessenza
'fifth essence' 'quintessence'
sopra abito → soprabito
'over dress' 'overcoat'
lungo Arno → lungarno
'long Arno River' 'street along the Arno'
- b. prete operaio → *pretoperaio
'priest worker'
valigia armadio → *valigiarmadio
'valise wardrobe'
dolce amaro → *dolcamaro
'sweet bitter'

In (14a) Vowel Deletion (VDel) (a form of phonological amalgamation) has taken place while in (14b) this is not the case; VDel was blocked by a 'strong boundary'. In truth, VDel is not only blocked by the type of boundary. Indeed, VDel may take place or it may not take place in any given class of compounds as can be seen below (both 15a and 15b are compounds of the type V + N):

- (15) a. spartiacque → *spartiacque
'watershed'
portaabiti → *portabiti
'clothes-stand'
scolapasta → *scolpasta
'colander'

- b. girarrosto → girarrosto
'rotisserie'
portaerei → portaerei
'aircraft carrier'
portombrelli → portombrelli
'umbrella stand'

VDel applies when no stress clash is produced as is the case in (15b). On the other hand, it does not apply when a clash would be produced. In fact, in *sparti* + *acque* deletion of the vowel would join together two adjacent tonic syllables (*spart + a*).

We can say that the presence of a strong boundary is indication of a loose compound while a weak boundary indicates a strict compound. One might think that a rule for compounding always forms loose compounds and that diachronic factors can weaken the boundary between the two components.

If loose compounds are formed by productive rules and strict compounds result from the permanence of said units in the lexicon, one might expect loose compounds to have compositional meaning while strict compounds might exhibit a certain degree of semantic opacity. Thus a compound such as *galantuomo* 'man of honor' or *gentildonna* 'gentlewoman' (presenting a phonological amalgamation indicated by vowel deletion) are not fully compositional; by this we mean that the meaning of *gentildonna* cannot be drawn from the meaning of its components, rather it takes on a new meaning which must be learned separately.

Finally, looking again at the same examples, one notes that the order A + N in *gentiluomo* and *gentildonna* is marked with respect to syntactic order in Italian:

- (16) Gori e' un uomo gentile N + A (unmarked order)
Gori is a 'man kind'
- (17) ?Gori e' un gentile uomo A + N (marked order)
Gori is a 'kind man'

In conclusion, loose compounds differ from strict compounds in the following ways: a) they do not allow phonological amalgamation; b) they have compositional meaning; c) the order of the constituents is not marked. On this basis, we could think that loose compounds are formed by rules and strict compounds are no more generated by rules but stored in the lexicon as such.

5. More on compound heads.

Let us once more consider compounds such as the following:

- (18) a. N + N pescecane head to the left: IS A fish
'fish dog (shark)'

- N + A camposanto head to the left: IS A field
 'field holy (cemetery)'
- b. A + N gentiluomo head to the right: IS A man
 'kind man (gentleman)'

The data given in (18) suggest that in Italian the head can be either to the right or to the left. In truth, however, compounds such as the one in (18b) — with the head to the right — exhibit some of the features of "strict" compounds seen above; notice, in fact, the contrast between *gentiluomo* and *camposanto*: the first is A + N and the second N + A, i.e. respectively having an unmarked and a marked order. We can, therefore, conclude that the productive synchronic rule for the formation of compounds in modern Italian generates compounds with the head to the left.

When studying the position of the head in Italian compounds one must recall that there exist "new" compounds, formed by means of productive rules, together with "older" compounds which are not formed by application of rules. Examples of the latter are:

- (19) a. parricida 'parricide'
 b. sanguisuga 'bloodsucker'

Both in (19a) and (19b) the first constituent is a "Latin" word as can be seen from the form (not *padre* or *sanguie* for 'father' and 'blood'). Moreover, the order is noun + verb, an order which reflects the basic order in Latin which is a SOV language. Such data indicate that compounds with a Latin origin have survived in the Italian language. Thus, when analyzing compounds one must separate this type of compound from those formed by application of productive rules. Hence there are two types of compounds⁸ which are to be kept separate:

(20)	Latin	Italian
syntactic order	SOV	SVO
order in the compounds	OV	VO
position of the head	right	left
inflection	right	left

6. Subordinate and Coordinate compounds.

When discussing only endocentric compounds, on the basis of the relationship existing between the two constituents, it is possible to distinguish between "endocentric subordinate compounds" and "endocentric coordinate compounds".

Those endocentric compounds where it is possible to identify a modified

⁸ In this regard one must also add the recent borrowing from English such as *scuola bus* 'school bus' having the head to the right. In fact, in English the head is normally to the right.

element — the morphological head — and a subordinate modifier are called "subordinate compounds". *Camposanto* is a subordinate compound since its head *campo* is modified by the adjective *santo* which is subordinate to the head. In other words, by "subordinate compounds" we mean those compound words where the relationship between the constituents is that of Head/Complement (cf. *acquavite* 'spirits', *capostazione* 'station-master') or Head/Attribute (cf. *pescecane* 'shark', *casaforte* 'safe').

The "coordinate compounds", on the other hand, are those endocentric compounds whose two elements are coordinated rather than one being subordinate to one another. In other words for these compounds it is impossible to identify which of the elements has been modified and which is the modifier; rather both elements are simultaneously modified and modifiers. Those words created by linking together two words with the same syntactic category are found among the endocentric coordination compounds. Therefore, in this class, one will find words created by joining together two nouns (cf. *caffelatte* 'light coffee', *cassapanca* 'box seat') or two adjectives (cf. *agrodolce* 'bitter-sweet', *grigoverde* 'grey-green'). The linking of two verbs, on the other hand, does not produce coordinate compounds since, in Italian, the V + V combination always forms exocentric compounds (cf. *bagnasciuga* 'shoreline', *andirivieni* 'comings and goings'). Finally, the linking of two prepositions is not normally possible (cf. **disenza*, **sopraverso*).⁹

7. Endocentric and exocentric compounds.

Above it was tacitly assumed that all compounds have a head and that all the linguistic information required for the compound to take on a syntactic function within a sentence derive from this head. In truth, however, not all compounds do have a 'head'. Those compounds having a 'head' are called 'endocentric' compounds while those lacking a 'head' are called 'exocentric'. Let's take a compound such as the following:

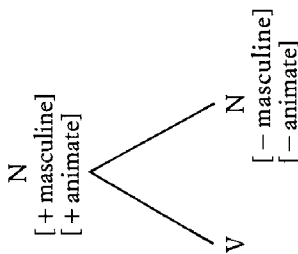
- (22) [[dormi]_v [veglia]_v]_N 'half-asleep'

Should we attempt to apply the "IS A" test to this compound we only get negative responses: in fact, from a semantic point of view, *dormiveglia* "IS NOT A" *dormi* 'sleep' nor is it a *veglia* 'wake' and in terms of category it "IS NOT A" verb: neither *dormi* nor *veglia* are possible candidates for the 'head' of this compound. In such cases the compound is called exocentric. It is worth noting, however, that the conclusion is not always so clear-cut. Let's consider the following compound:

- (23) [[porta]_v [lettere]_N]_N 'postman'

⁹ A form such as *sottosopra* 'upset' is most likely made up of two adverbs.

At first glance the 'syntactic' application of the "IS A" test could lead to the conclusion that the constituent *lettere* 'letters' should be considered the head of the compound, at least from a categorial point of view, since *lettere* is a noun as is *portallettere*. From the semantic point of view, however, this conclusion proves unacceptable: *portallettere* "IS NOT A" *lettere*. If we extend our analysis as mentioned above, besides information regarding category, we obtain the following picture:



As can be seen, there is no identity between the noun *lettere* and the noun *portallettere* since the first is feminine and inanimate while the second is masculine and animate. Other examples of exocentric compounds are:

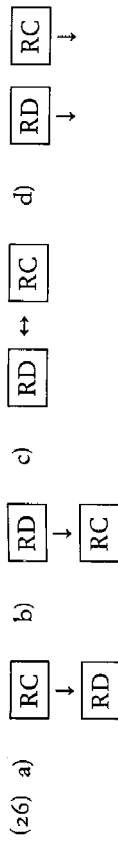
- (25) *pellerossa* 'red-skin'
purosangue 'thoroughbred'
senzatecto 'homeless'

The above case makes it possible to draw some conclusions: a) there are compounds without a head (the so-called exocentric compounds);¹⁰ b) an apparent similarity in category alone is not enough to determine the head of a compound; c) in order to determine the head of a compound the categorial and semantic criteria must be in agreement (in other words: a categorial head must also be a semantic head).

8. Composition and derivation

The Rules of Composition (RC) and Rules of Derivation (RD), as used here, are two groups of different rules operating in the same subcomponent (i.e. the lexical component). Theoretically, there are four possible relationships between the two groups of rules: a) all RCs come before RDs, b) all RDs come before the RCs, c) RCs and RDs interact freely, d) RCs and RDs do not interact at all.

¹⁰ Exocentric compounds are also called *bahuvrīhi* from sanscrit: that has much [bahu] rice [vrīhi] meaning not an 'abundance of rice' but rather a 'person having an abundance or rice'.



The hypotheses in (26c) and (26d) are not supported by the evidence we have. In fact, (26d) is falsified by the fact that words can be found which are both derived and compounded,¹¹ while (26c) is falsified by the fact that compounds cannot be freely derived nor can derived words be freely composed.¹² This leaves (26a) and (26b). (26a) predicts that only words which are "first" compounded and "then" derived can exist. In this case only two structures would be possible given that the derivation rules cannot "enter" into the compound. Thus only compounds with suffixes (27ai) and prefixes (27aii) would exist. On the other hand, (26b) asserts just the opposite; that is that complex words exist which are "first" derived and "then" compounded. In this case various structures are possible depending on whether Word₁ is suffixed or prefixed (respectively 27bi and 27bii) or whether Word₂ is suffixed or prefixed (respectively 27biii and 27biv) and so on, leading to highly complex structures such as the one in (27bvii) which is most likely only hypothetical.

- (27) a) i. [[[] # []] + Suf]
 ii. [Pre + [[] # []]]
 b) i. [[[] + Suf] # []]
 ii. [[Pre + []] # []]
 iii. [[] # [[] + Suf]]
 iv. [[] # [Pre + []]]
 v. [[[] + Suf] # [[] + Suf]]
 vi. [[Pre + []] # [Pre + []]]

 vii. [[[Pre + [] + Suf] # [[Pre + [] + Suf]]]

There are, in effect, examples of structure (27ai), such as the following:

- (28) *ferrovìa* → *ferroviario*
 'railroad' → 'railroad (adj.)'
crocerossa → *crocerossina*
 'red cross' → 'red cross nurse'
malavita → *malavitoso*
 'low-life' → 'one who lives the bad life'
guardaroba → *guardarobiera*
 'wardrobe' → 'cloakroom attendant'

¹¹ *Crocerossina* 'red cross nurse', *guardarobiera* 'cloakroom attendant' are such words.
¹² A compound word such as *guardaboschi* 'forester' cannot be further derived; **guardaboscaglia* (but cf. *bosco* 'forest' → *boscaglia* 'brushwood'). Derived words such as *pellicina* 'lit. little skin' and *rossastra* 'reddish' cannot be composed: **pellicinarossastra* (but cf. *pellerossa*).

- dopolavoro
'official employees club'
→ dopolavoristico
'relating to the employees club activities'
- verosimile
'likely'
→ inverosimile
'unlikely'
- fotocopiare
'photo-copy'
→ fotocopiatrice
'photocopier'

It can be proved that the above compounds have a structure like the one given in (27a) and not like the one in (27bii). Let's take the word *crocerossina* 'red cross nurse'. It is a word made up of three elements: *croce* 'cross', *rossa* 'red' and *ina*. The suffix *-ina* is not a diminutive suffix but it is, in fact, the suffix (here in the feminine form) used to form nouns indicating trades (cf. *postino* 'postman'). Such suffix is added to nouns (i.e. *posta* 'mail' + *ino*) and not to adjectives. The structure of the word *crocerossina* is, therefore, as follows:

- (29) [[croce] # [rossa] + ina]

This is further proved by the fact that the word **rossina* does not exist, nor do the words **vuario*, **vitoso*, **robiera*, etc. We must not, however, let the above data mislead us. Indeed, one cannot conclude from this that compounds can normally be derived. If we consider compounds with the same structure as those above we will find that they cannot be derived:

- (30) crocevia 'crossroad' → *croceviario
 camposanto 'cemetery' → *camposantina
 altopiano 'plateau' → *altopianura
 lavapiatti 'dishwasher' → *lavapiattiera
 saliscendi 'dumbwaiter' → *risaliscendi
 retroterra 'hinterland' → *retro terrestre
 dattiloscivere 'type' → *dattiloscrittore

The case in which Word₂ is a semi-word is quite different. In this case it appears that the compound can be derived (even if quite often with the suffix *-ico*):

- (31) filantropo → filantropico
 'philanthrope' → 'philanthropic'
 talassografo → talassografico
 'thalassographer' → 'thalassographic'
 biografo → biografico
 'biographer' → 'biographic'
 astronauta → astronautico
 'astronaut' → 'astronautical'

If, on the other hand, we pass from the area of strict compounds to loose compounds we will find another tendency. Just consider data such as the following:

- (32) a. aiuto macchinista
 fondo assistenza
 città dormitorio
 b. trasmissione radio
 formazione base
 c. disegnatore progettista
 marxista leninista
- 'driver's mate'
'aid fund'
'commuter city'
'radio program'
'basic training'
'designer-engineer'
'Marxist Leninist'

These compounds are instances respectively of the structures in (27biii), (27bi) and (27bv). In other words the compounds given in (32a) present suffixes tacked onto Word₂, those in (32b) suffixes added to Word₁ while those in (32c) have suffixes added to both Word₁ and Word₂. Evidence that the structures are those given is found in the fact that the intermediate structures (**aiuto macchina*, **trasmettere radio* and **disegna-progetta*) do not exist.

The above leads to the conclusion that grammar treats strict compounds in similar fashion to the way it treats simple words (that is morphological material can be added to the right).¹³ On the other hand, it deals with loose compounds as complex words which are, in effect, made up of two components. One should note that in loose compounds there exists a structure which is not possible:

- (33) *[Word₁] # [Word₂] + Suf

This is explained by the fact that in loose compounds Word₁ constitutes the head. One can, therefore, deduce that morphological material (derivation and/or inflection) cannot be added 'too far' from the head.

In conclusion, the composition/derivation relationship is to be analyzed separately depending on whether the word in question is a strict or a loose compound.¹⁴

- (34) loose compounds strict compounds



Finally, one will note that the derivation would appear to be almost exclusively represented by the addition of a suffix. In other words, a prefixed compound is a very rare construction.¹⁵

¹³ It is worth noting that, when possible, any inflection is also added only to the right.
¹⁴ One problem is found with V + N compounds where there can be cases of derivation of the entire compound as well as cases of Word₂ derivation. This can be seen, respectively, in the examples *paracadutista* 'parachutist' and *portatovagliolo* 'napkin-holder'.
¹⁵ Among the highly unusual examples of prefix/composition interactions one will recall *inverosimile* 'unlikely' and *andirivieni* 'commotion'.

9. Compounding and inflection.

Inflection of compounds is a highly complex morphological area, at least in Italian, and it is impossible to state regularities without finding a multitude of exceptions. The following theoretical possibilities exist (where "Infl" stands for 'inflection'):

- (35) a. W_1 + W_2 + Infl
 b. W_1 + Infl + W_2
 c. W_1 + Infl + W_2 + Infl
 d. W_1 + W_2
 e. W_1 + (W_2 + Infl)
 f. (W_1 + Infl) + W_2

The following cases would, therefore, be possible: (35a): inflection at the end of the compound; (35b): inflection after the first word in the compound; (35c): inflection after both constituents; (35d): compounds where inflection is not possible; (35e): compounds where the inflection is the inflection of the second word but not of the compound as a whole; (35f): inflection of W_1 where the inflection is the inflection on the first word but not of the overall compound.

All of these possibilities can be found in Italian with the exception of (35f), as can be seen below:

- (36) a. mezzogiorni b. navi traghetto
 ferrovie mobili bar
 camposanti capi stazione
 c. cassepanche d. andirivieni
 mezzelune tritacarne
 terrefemè voltafaccia
 e. portallettere f. *filivia
 guardasigilli *maniscritto
 rompiscatole *palchiscenico

Plurals of the type in (35c), that is the 'double' plural where both of the constituents have been pluralized, appear to have a twofold nature: either it is a sort of 'agreement' (where there is agreement in number between the noun and the adjective) or it is a "double-head" plural. This is the case of the word *cassapanca* which "IS A" *cassa* 'case' and "IS A" *panca* 'bench'; thus the inflection is encountered on both constituents (cf. the paragraph on coordinate compounds).

As has been previously mentioned, however, it is difficult to "predict" the plural of the compound in a uniform manner. Still the above considerations regarding the "head" can, in some cases, be of some help. Let us consider the following three compounds: *capostazione* 'station-master', *capogiro* 'dizziness' and *capomastro* 'master builder'. All three form their plurals in different ways; that is: *capistazione* (type 35b), *capogiri* (type 35a) and *capimastri* (35c).

Now, type (b) is a compound the head of which is to the left and with inflection only on the head. Closer examination reveals that type (a) is different because the head is not *capo* but rather *giro* and thus the inflection remains to the right. Finally *capimastri* can be explained if one assumes it is a coordinate compound (as was seen above with *cassapanca*) essentially composed of two heads, each followed by its own inflection.

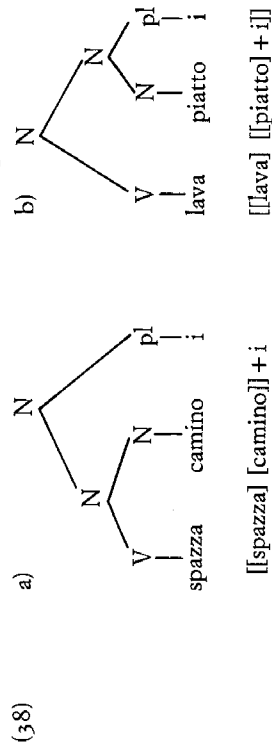
One important step is to identify the productivity of a given type of compound; only for productive compounds can a true and proper "rule" be proposed. Leaving a certain margin of approximation, one can say that, today, productive compounds are those of type (35b); that is those compounds with the head to the left and followed by inflection only on the head. In the course of time (and most likely in relation to extra-linguistic factors such as frequency of use) compounds tend to lose transparency. In this case it becomes more difficult to identify the head and the compound, perceived as lacking an inner structure, takes on its own inflection following the general rules governing inflection in Italian; that is, inflection to the right.¹⁶

9.1. Word₂ inflection and inflection of the entire compound.

Consider two compounds such as those below:

- (37) a. *spazzacamini* 'chimney(pl) - sweep'
 b. *lavapiatti* 'dish(pl) washer'

Although they appear to be built with the same constituents (Verb - Noun - Plural marker) these two compounds are structurally different. It is worth noting that while the former corresponds to a singular form (*lo spazzacamino* 'the chimney-sweep') this is not the case for the second (**la lavapiatto*). In (37a) the plural morpheme is added to the compound as a whole and thus one obtains the structure shown in (38a). On the other hand, in (37b) the plural morpheme is only added to W_2 and not to the compound as a whole (which is singular and feminine) thus giving the structure in (38b):



¹⁶ An emblematic case is that of the word *pomodoro* 'tomato' which should be inflected as *pomidoro*, but, instead, in the spoken language, it is inflected as *pomodori*, and, sometimes *pomidori*.

The reason why compounds have a pluralized Word₂ may, in some cases, be pragmatic: in examples such as *stuzzicadenti* 'toothpick', *asciugamani* 'towel', etc it would appear that the plural is dictated by the fact that human beings have more than one tooth and more than one hand.¹⁷ Plural forms such as found in the word *lavapiatti* 'dishwasher' (and there are many such examples)¹⁸ are problematic in an "ordered" morphological model.

10. Structure within compounds.

In the previous paragraph it was mentioned that in compounds such as *crocerossina* and *guardarobiera* the words *rossina* and *robiera* do not exist. Let us consider once more the representation of one of these compounds:

(39) [[croce]_N [rossa]_A]_N + ina

The suffix *-ino/-ina* used to form nouns indicating jobs and trades is added to nouns (and at times verbs) (cf. 40b) but not to adjectives (cf. 40a):

- (40) a. *gentile + ino/a
 *allegro + ino/a
 *veloce + ino/a
 b. bagno + ino/a (bath + suffix, 'lifeguard')
 stagno + ino/a (tin + suffix, 'tinsmith')
 contado + ino/a (countryside + suffix, 'farmer')

This means that the suffix *-ina* 'sees' the node N created by the compounding process and thus does not attach itself to node A of Word₂ (*rossa*). We can, therefore, conclude that morphological processes create structure (represented here either by a morphological tree or by parentheses) and that said structure is clearly present, even if it is not 'visible'.

The above is not only true in cases of the derivation of compounds but it also holds true for inflection of compounds. Let us consider a compound such as *portaceneri* 'ashtray'. The regular plural should be the same as the singular since *ceneri* 'ash' is a mass noun (when pluralized *ceneri* takes on an idiosyncratic meaning of 'mortal remains'). In the spoken language, however, the plural *portaceneri* is commonly found; two *portaceneri* just means more than one *portaceneri* and not an 'urn to hold ones mortal remains'. This also means that the plural morpheme has been added to the more external node N (which is a [+count] N) and not to the node N of the word *ceneri* (which is a [-count] N):

(41) [[porta]_V [ceneri]_N]_N + i
 [-count] | [+count]

¹⁷ Cf. *baciamano* 'obsequiousness'. The singular forms *stuzzicadente*, *asciugamano* are derived from the plural and not viceversa.

¹⁸ Cf. *guardasigilli* 'chancellor', *portalettere* 'mullman', *scacciacani* 'blank pistol', *portaombrelli* 'umbrella-stand', *rompicatole* 'nuisance', *appendiabiti* 'clothes-rack', etc.

Therefore, the rules of compounding are structure creating rules (labelled parentheses) and such structure is accessible to other morphological rules.

11. V + N compounds.

One type of highly productive compound is the [V + N] compound which is always a noun. Below follows a brief discussion of some questions such as: the noun/verb relationship, the order of the two constituents, the form of the verb and the nature of the outcoming words.

In regard to the first point one will note that the noun appears to be the internal argument of the verb:

- (42) a. portallettere
 'letter carrier'
 spazzacamino
 'chimney sweep'
 b. *portasinghiozzi
 'lit. carry hiccups'
 *spazzaamarezza
 'lit. sweep bitterness'

In the word *portallettere* the noun is the internal argument of the verb, as it would appear in a syntactic construction (G. *porta le lettere a casa* 'G. carries the letters home'). The form *porta singhiozzi* is not grammatical since *singhiozzi* 'hiccups' cannot be the object of the verb *portare* 'carry'.

Still further it must be noted that only the direct object of the verb can appear in a compound.

- (43) *portacasa (with house as locative element)
 'lit. carry-house'
 *spazzascopa (with broom as instrument)
 'lit. brush-broom'
 *lavacura (with cure as modal)
 'lit. wash-cure'

Moreover, the external argument of the verb does not appear to play any role in compounds:

- (44) *portaamico 'lit. carry friend'
 *spazzauomo 'lit. sweep man'
 *lavaragazzo 'lit. wash boy'

The above forms are not possible with the noun as the external argument of the verb.¹⁹ The above data indicate that in compounding the verb can — so to speak — saturate fewer 'positions' than in syntax. In other words, the V-N relationship in syntax differs from what is found in morphology.

In regard to the second point, it is worth noting that the word order is V + N and not vice versa; that is not N + V. In fact, V + N is the basic word order in Italian contemporary syntax, i.e. SVO. This means — as

¹⁹ Exceptions can be found in non productive forms such as *batticore* 'heartbeat', *battiscopa* 'skirting-board', *marciapiedi* 'sidewalk'.

we have already seen above — that the order of the constituents in compounding is in agreement with the order of the constituents in syntax. Nonetheless, once again, while morphology exhibits one and only one possibility, syntax can exhibit different surface orders through application of movement rules (cf. Ieri ho visto Gianni 'I saw G. yesterday' and *Gianni ho visto ieri* with topicalization). In morphology it would appear that movement rules are impossible and thus the V + N word order can never be inverted (cf. **lettereporta*, **caminospazza*, etc.).

Finally, in regard to the third point, it must be underlined that there are at least three different proposals for the verb form: the verb is a) imperative; b) third person singular of the present indicative; c) the stem (i.e. root plus thematic vowel). It should be pointed out here that the third solution is the most coherent from a semantic point of view (in compounds such as *rompicatole* 'nuisance' there does not appear to be any temporal-imperative type of determination). Moreover, it is more coherent with the data found in regard to derivation:

(45)	<i>conjugation</i>	I	II	III
	thematic vowel	a	e	i
	imperative	a	i	i
	III pers pres	a	e	e
	derivation	a	i	i
	composition	a	i	i

As can be seen, from a formal point of view, the least 'costly', the easiest solution would support the hypothesis of an 'imperative' form although, from a semantic viewpoint, both in derivation and in compounding there is no trace of the "imperative" whatsoever. The "thematic vowel" solution requires the operation of an Adjustment Rule (and only one for all morphology, both in terms of derivation and compounding) which would change the thematic vowel *e* into *i* for all verbs of the IIInd conjugation.

I. I. I. *Endocentric reading of exocentric compounds?*

According to some, all morphological structures are endowed with a head; the term 'exocentric compound' would be a contradiction. In this light proposals have been put forth for 'endocentric' analysis of compounds which are generally considered 'exocentric'.

Let's take the compound *portalelettere* once more. It has been mentioned that *lettere* cannot be the head. And what can we say of *porta*? Obviously the answer is to be found in analysis of this constituent. In this type of compound the first element is, by mutual consent, always considered a verb. If *porta* is a verb, the compound must be considered 'exocentric'. Another possible type of analysis is to assume that *porta* is a noun with a sort of "zero suffix" or "unrealized suffix"; as if it were *porta(tore di) lettere*

'carry(ier of) letters'. There is not much data in support of this type of analysis; one must resort to observing whether noun forms such as "you are a) *rompi*", "you are a) *mangia*" exist. In this case a compound such as *rompicatole* would be analyzed as a N + N form with the head to the left.

This analysis, if confined (cf. Bisetto (ms.)), would have the advantage of eliminating from morphology unwanted constructions such as "headless morphological complex unities" and, furthermore, it would make possible to consider the compounds under discussion as "normal" Italian compounds, that is compounds with the head to the right.

1.2. *Compounds and derivatives.*

Identifying compounds and distinguishing them from derived words is generally a simple task in light of the above definition (which can be applied in most cases). According to this definition compounds are formed by two free forms and derivatives are made up of one free form plus a bound form:

(46)	(i) <i>free form</i>	+	<i>free form</i>	capo	+	stazione	<i>compounding</i>
	(ii) <i>free form</i>	+	<i>bound form</i>	fama	+	oso	<i>suffixation</i>
	(iii) <i>bound form</i>	+	<i>free form</i>	ex	+	fidanzato	<i>prefixation</i>

This picture, however, is somewhat idealized since various problems can arise as will be seen briefly below.

At times, difficulties are encountered in distinguishing between prefixing and compounding. Just consider the following:

(47)	a.	sottocommissione	'subcommittee'
		sottobiblioteca	'assistant librarian'
		sottoprefetto	'assistant prefect'
		sottosegretario	'under secretary'
	b.	sottoscala	'space beneath the staircase'
		sottogola	'chinstrap'
		sottotetto	'attic'
		sottobicchiere	'coaster'

At first glance the data given in (47) all appear to have the same structure. Yet, if one applies the procedures indicated above, it becomes evident that this is not the case. In fact, the "IS A" test gives different results: in (47a) one can say that *sottocommissione* "IS A" *committee* while in (47b) one cannot say that *sottoscala* "IS A" *staircase* (rather it is a particular space that lies beneath the staircase). Still further: in (47a) the linguistic information of the entire complex word coincides with that of the constituent word to the right (that is *sotto-commissione* is a singular feminine word as is *missione*). On the other hand, in (47b) *sottoscala* is a masculine noun while the constituent word to the right is feminine (*un sottoscala/una scala*). This means the words in (47a) have the head to the right while those

in (47b) are exocentric. Given that Italian compounds have the head to the left we can conclude that the forms in (47a) are words with a prefix while those in (47b) are 'exocentric' compounds.²⁰

Another problematic point regards the so-called prefixoids. It has already been mentioned that if such units are considered as semiwords (as proposed above) it would be possible to have the following combinations:

- (48) a. semiw + semiw nosocomio 'hospital'
 b. semiw + wordz bio + chimico 'biochemical'
 c. word1 + semiw astro + nauta 'astronaut'
 d. semiw + suf fono + ico 'phonic'
 e. Pref + semiw a + fono 'voiceless'

Therefore, semiwords can be used to form both compounds (48.a.b.c.) and derivatives (40.d.e.).

Another problem involving the distinction between compounds and derivatives involves the X-mente 'X-ly' formation. Consider the following examples:

- (49) amaramente 'bitterly'
 storicamente 'historically'
 voracemente 'greedily'

Bosque (1987) and Zagona (ms.) have maintained that these formations are in effect compounds while Scalise et al. (1990) maintain that they are suffixed words. It is likely that the problem involves different diachronic stages. For example, in Spanish *-mente* can be factored out (cf. 50a) while this is not possible in Italian (cf. 50b):

- (50) a. rapida y eficazmente 'quick(ly) and efficaciously'
 b. *rapida e efficacemente

That we are dealing with a problem which is in effect diachronic can be confirmed by the fact that in ancient Italian factoring was possible (*anticha o novamente* [Guittone d'Arezzo], *villana e aspramente* [Novellino]).²¹ The factoring of ancient Italian can be interpreted as an intermediate phase between Latin, where *mente* was an X category (cf. *firma mente*) and contemporary Italian where *mente* is an X-1 category.

13. Morphology and syntax.

Morphology and syntax are two separate grammatical components although they must interact and, to a certain extent, "cooperate" in determining common objectives. The rules of word formation can only create

²⁰ Also in this case, of course, it would be possible to propose an endocentric reading of this type of compounds, but it never was to my knowledge.

²¹ Ancient Italian could be pushed to constructions such as *onestà e pacifica e discretamente* [Statuti volgari]. For these and the other forms, quoted in ancient Italian, see Migliorini (1957).

complex words in agreement with the principles involved in syntax. For example, the basic word order in Italian syntax is SVO; thus one would expect Italian morphology to "build" words which agree and do not contrast with this order. In this section some aspects of the interaction syntax/morphology will be discussed.

Several different criteria have been proposed in order to distinguish between compounds and phrases. Not all have the same 'weight', however. Moreover, one cannot assert that the criteria available today are definitive. A list of these criteria is given here below²² followed by a brief discussion of each:

- (51) 1. Compounds are atoms from a syntactic point of view.
 a. They do not allow insertions
 b. They do not allow movement
 c. They do not allow anaphoric relationships
 2. They involve only maximal projections
 3. They exhibit a limited degree of recursivity
 4. They require specific readjustment rules.
 5. They are subject to semantic drift.
 6. They show lexical limitations.

1. One of the most highly 'articulated' criteria is that of the word as atom. In other words, a compound is (as is a simple word) an atom within which syntax cannot 'enter'. Some indicate the same thing by saying that compounds are 'opaque' to syntax. For instance, while lexical material can be inserted into a phrase (cf. 52a) no material can be inserted into a compound (cf. 52b):

- (52) a. G. porta bagagli → G. porta grandi bagagli
 'G. carries bags' 'G. carries big bags'
 b. portaombrelli → *porta grandi ombrelli
 'umbrella holder'

In like fashion it would appear that the usual syntactic movement rules (cf. 53a) cannot be applied to compounds (cf. 53b):

- (53) a. Maria taglia carte → Cosa taglia Maria t?
 'Maria cuts paper' 'What does Maria cut t?'
 b. M. ha un tagliacarte → *Cosa ha Maria un taglia t?
 'M. has a paper cutter'

A constituent of a compound cannot be an antecedent of a lexical anaphora (cf. 54b):

- (54) a. queste carte; si mettono [le une sulle altre];
 'these papers; are set [one on another];
 b. *questi tagliacarte; si mettono [le une sulle altre];
 'these papercutters; are set [one on another];

²² Cf. Di Sciullo (1991).

2. Compounds only involve major lexical categories. In Italian compounding rules do not involve non major lexical categories. The sole exception to this regards proper names (cf. *Bevilacqua* 'lit. drink the water').

3. In general Italian compounds are not recursive.²³ Once formed a compound cannot generally be subjected to other rules of compounding. Compounds such as *cassapanca*, *capostazione* cannot serve as constituents of other compounds. In English, Dutch and German (i.e. in Germanic languages in general) the exact opposite holds true: compounding is normally recursive.²⁴

It has been observed²⁵ that the rules of syntax usually exhibit free recursivity both to the right and to the left as well as center embedding. Morphological recursivity, in those languages which allow it, is in general restricted to one side.

4. Compounds require a group of specific readjustment rules. There are, in fact, phonological processes which can only be applied to compounds; i.e. some cases of haplology²⁶ such as the following:

- (57) eroico comico → eroicomico 'heroicomic'
fotogramma metria → fotogrammetria 'photogrammetry'

On the contrary, haplology cannot be applied in the domain of derivation nor when prefixing or suffixing are involved (cf. 58c., d.):

- (58) a. antitetanico → *antetanico 'antitetanic'
b. semiminaccioso → *seminaccioso 'half threatening'
c. enormemente → *enormente 'enormously'
d. armamento → *armento 'armament'

We are, therefore, dealing with a rule which (in Italian) is limited to compounds and not to other classes of morphologically complex words.

Another case of readjustment rules which can only be applied in the domain of compounds is the deletion of a suffix under identity:²⁷

- (59) social + ista comun + ista → social-comunista
cartol + eria libr + eria → cartolibreria

²³ Dressler (1988) asserts, instead, that compounds can be recursive even in Italian citing cases such as the following: *sala personale viaggiante*, *sala dirigente capo*, *campo tiro a volo*, *nave pesca d'altomare*. It is possible — although the problem requires further study — that these are not compounds but rather abbreviated phrases (cf. *sala grande personale viaggiante*, *sala grande personale viaggiante sindacalizzato*, *sala grande personale non viaggiante sindacalizzato*, etc. It may be that the so-called phrasal compounds recently studied by Lieber (1992) also fall into this category (e.g. *a pipe and slipper husband*, *an ate too much headache*, *God is dead theology*). Compounds of this sort are also found in Dutch (*lach of ik schiet humor* 'laugh or I shoot humor') and in German (*die Muskel-für-Muskel-Muskel-Methode*). The status of these compounds is, however, dubious (cf. *un marito casa-ufficio*, *un marito tutto casa-ufficio*, *un marito tutto pantofole solo chiesa niente ufficio*). I feel that these constructions are to be defined as 'abbreviated syntax' although further study is required.

²⁴ Cf. Selkirk (1982:15).

²⁵ Cf. Anderson (1988:149).

²⁶ Cf. Scalise (1983).

²⁷ For this rule (and its limits) cf. Scalise (1983).

In general, we know that the Readjustment Rules cannot be formulated once for all complex words. Even the most general rules such as Vowel Deletion must be formulated in specific ways when applied to different complex words; for example it may be optional for compounds (cf. 60a) while it is mandatory for suffixed words (60b):

- (60) a. portaaerei / portaaerei 'aircraft carrier'
giraarosto / giraarosto
b. *bosco + oso / boscoso 'woody'
*mangia + ata / mangiata 'full meal'

5. Compounds are subject to semantic drift. Once formed, a compound "exists" in the lexicon. A lengthy stay therein may obscure the original transparency of the compound.

6. Composition may exhibit lexical 'gaps'. Given, for example, a lexical domain, the same type of compounding rule cannot be regularly applied to all subparts:

- (61) a. mezzogiorno 'midday'
mezzanotte 'midnight'
mezzodi 'midday'
b. *mezzasera
*mezzopomeriggio
*mezzamattina

Compounds are lexical constructions and as such are subject to restrictions which syntactic constructions are not subject to. Just compare a 'syntactic' construction (62a) with a morphological construction (62b):

- (62) a. contatore di giri
contatore di scatti
contatore della luce
contatore del gas
b. contagiri
contascatti
*contaluce
*contagas

When dealing with compounds, the object noun must satisfy the limitations imposed by the verb (after *contare* only [+count] objects can be used).

14. Conclusion.

Compounding in Italian can be thought as a lexical process which is independent from syntax but which, at the same time, must 'cooperate' with it.

Italian compounding rules form in general nouns. Productive compounding rules form endocentric compounds with their head on the

left. Other compounds do exist, however, with their head on the right but these can be thought as lexicalized items.

The relationship between the two constituents of a compound is governed by the head which is, in contemporary Italian, the left constituent. Compounds with their head to the right are either survivals from Latin or loan words from other languages (mainly English).

The relationship between compounding rules and derivation rules are quite straightforward: productive compounding rules can take as their base derived words. Otherwise there are limited examples of derived compounds, but in this case the compound is a lexicalized unit (and therefore treated as a simple lexical unit).

There are some cases in which it is difficult to distinguish between a compound and a derived word: some borderline cases are discussed (both from prefixation and suffixation) showing that one can always propose a principled distinction between these two morphological domains.

Inflection of compounds is very complex area. The main regularity that can be extracted is the following: productive contemporary compounds have their head inflected (i.e. the first constituent). There are, then, other cases such as: double inflection (that is inflection of both constituents) or inflection of the second constituent (the whole compound being singular). It remains, however, a large area of irregularity which is maybe due to the survival of diachronic facts in the present synchronic state.

Finally, some aspects of the relationship between compounds and phrases are discussed and it is shown that these two domains of the grammar must be kept separate: phrases and compounds cannot be generated by the same set of rules.

Address of the Author:

Istituto di Lingue e Letterature Straniere
Facoltà di Magistero
Università degli Studi di Ferrara
Via Savonarola, 27
I-44100 Ferrara, Italia

E-mail: G7MFEVE1@ICINECA

Bibliography

- Allen, M. R. (1975), *Morphological Investigations*, Ph.D. dissertation, University of Connecticut, Storrs.
- Anderson, S. (1988), "Morphological Theory", in F. J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, Cambridge, Cambridge University Press.
- Bisetto, A. (ms.), "Compounding and principles of grammar", Venice.
- Bosque, I. (1987), "Construcciones morfológicas sobre la coordinación", *LEA*, 9: 83-100.
- Di Sciullo, A. (1991), "Deverbal compounds and the external argument", ms.
- Di Sciullo, A. & E. Williams (1987), *On the definition of word*, Cambridge (Mass.), MIT Press.
- Dressler W. (1988), "Preference vs. Strict Universals in Morphology", in M. Hammond and M. Noonan (eds), *Theoretical Morphology*, New York, Academic Press: 143-154.
- Lieber, R. (1992), *Deconstructing morphology*, Chicago, University of Chicago Press.
- Migliorini B. (1957), *Saggi linguistici*, Firenze, Le Monnier.
- Scalise, S. (1983), *Morfologia lessicale*, Padova, Clesp.
- Scalise S. et al (1990), "Il suffisso mente", *Studi di linguistica teorica ed applicata*, 19: 61-88.
- Selkirk E. (1982), *The syntax of words*, Cambridge (Mass.), MIT Press.
- Zagona, K. (ms.), "Mente adverbs, Compound interpretation and the Projection Principle".