

Compounding in French

Wiecher Zwanenburg

The core of this contribution is formed by sections 2 and 3, where I examine the morphological structure and the meaning of French compounds respectively. In the remaining sections, 1 and 4-6, I examine the basic units of compounding and the relation between compounding and derivation, inflection and the lexicon in light of the distinctions made in sections 2 and 3.

As to morphological structure, examined in section 2, I distinguish coordinate, subordinate and exocentric structures. Following to a large extent di Sciullo-Williams (1987), I suggest a possible analysis in which coordinate structures and most subordinate structures are considered to be fixed syntactic phrases listed in the lexicon. This explains their left-hand headedness and their often internal inflection, which they share with phrases. I analyze the remaining subordinate structures, having a preposition or an adverb as their first constituent, as real, right-hand headed compounds.

As to exocentric structures, I suggest, extending a proposal by di Sciullo-Williams (1987), to analyze them as conversions of noun phrases, verb phrases and preposition phrases respectively. This explains the left-hand headedness of the bases of these conversions and, as a consequence of their being conversions, the fact that in general these structures do not present internal inflection.

As to semantic interpretation, examined in section 3, the meaning of phrases listed in the lexicon is, apart from possible semantic specialization on an individual basis, parallel to that of free phrases, and thus does not need discussion. The meaning of real compounds follows from their right-hand headedness and the modifier status of their first constituent.

The meaning of exocentric structures follows from their conversion status, conversion of complex bases expressing the same meaning as suffixes like *-ier* and *-eur* with simplex bases.

A general characterization of French compounding as opposed to Germanic compounding might be that in French real, right-hand headed compounding is very restricted, and that French makes extensive use of left-hand headed lexicalized phrases and conversions with left-hand headed internal phrases.

Lack of space prevents me from discussing a possible alternative analysis in which syntax and morphological right-hand headedness play a less prominent role (see the end of section 2 for a hint at the relation between the two alternatives and the tradition in French morphology). But I am confident that the overview and the classification of the structures proposed here remain valid in such an alternative analysis, and thus can be useful independently of the framework which I have chosen.

o. Introduction.

It is often said that one of the main characteristics of compounds in French, as well as in the other Romance languages, is that they have their governing constituent or head at the left-hand side. This contrasts with right-hand headed compounding in Germanic languages like English. Thus French *timbre-poste* has its head *timbre* on the left-hand side, whereas in its English equivalent *postage stamp* the head *stamp* is on the right-hand side. Another property which opposes Romance compounding to Germanic compounding is that many compounds resemble syntactic phrases. Compare for example French *billet d'entrée*, which has phrasal structure, with its English equivalent *admission ticket*, which does not have such a structure at all.

These two properties of French compounds are the main issues discussed in section 2 on the morphological structure of French compounds. This section, together with section 3 on their semantic structure, constitutes the heart of this contribution.

But I will first discuss the basic units of French compounding in section 1. Compounds are supposed to be made up of, generally uninflected, words. We will see, however, that they may be made up of roots also, particularly in so-called learned compounding, and that they sometimes may contain inflected words.

In sections 4 and 5 I will discuss the relation between compounding and derivation and the inflection of compounds respectively.

Throughout section 1 to 5 I will treat compounding as a form of internal word grammar which is in principle independent from the lexicon. Thus it is appropriate that in section 6 I discuss the relation between compounding and the lexicon, including some observations on the productivity of compounding.

1. Basic units.

In Romance and Germanic languages compounds are in general made up of words. But French, like the other languages of these families, has a large class of compoundlike words for which this does not hold. These are what may be called nonnative or learned compounds. Compare for their English equivalents Selkirk (1982: 99-100). The most common classes of such learned compounds in French are the following, with Greek elements in the left-hand column and with Latin ones in the right-hand column:

(1)	N	NIN:	aéro-drome	agri-culture
	A	NJN:	hiéro-glyphe	rect-angle
	N	VJN:	giéo-graphe	fratri-cide
	N	NJA:	anthropo-morphe	cunéi-forme
	A	NJA:	hétéro-doxe	magn-anime
	N	VJA:	patho-gène	centri-fuge

Many constituents which we may find in such learned compounds are often analyzed as either prefixes or suffixes. This is comprehensible, because, like affixes, these constituents do not occur in general as words, as do the constituents of nonlearned compounds. But nevertheless we must distinguish them from affixes.

In the first place, if in *giéo-graphie*, for example, *giéo-* were a prefix and *-graphie* a suffix, the word in question would consist solely of affixes, which is a contradiction in terms, an affix being characterized by the fact that it combines with a base in order to form a word. Moreover, such constituents, as is shown for example by *philo-logue* and *franco-phile*, can often be either the first or the second constituent of a word, contrary to affixes, which are either prefixes or suffixes, but never both.

I will not discuss learned compounding in detail. Suffice it to say here that the categorial constellations found in learned compounding parallel in general those found in nonlearned compounding. Compare for example the learned noun *agri-culture* made up of two nominal constituents with nonlearned *timbre-poste* with two nominal constituents. These examples are characteristic in that learned compounding is almost unexceptionally right-hand headed, as shown by *agri-culture*, whereas their nonlearned equivalents are usually left-hand headed, as argued in the introduction on the example of *timbre-poste*. I will argue in section 4 that the two kinds of compounding have a different status in (morpho) phonological structure.

Nonlearned compounds are in general made up of uninflected words. But French has many compoundlike expressions with internal inflection marks, such as *basse-cour* with the feminine gender mark *-e* or *beaux-arts* with the plural number marks *-x* and *-s*. I will argue in section 2 that these can be considered to be phrases with the expected distribution of such marks.

There are, however, compoundlike expressions which cannot be thus interpreted, and which may nevertheless sometimes show internal inflection, as in *compte-gouttes* and *entre-deux-guerres*. I will argue in section 2 that these are words derived by conversion or zero affixation from verbal or prepositional phrasal bases, and thus have a particular status as compared to ordinary compounds made up of uninflected words. See also section 5 on compounds and inflection.

2. Morphological structure.

Word formation may be considered as internal word syntax (cf. Selkirk (1982)). And this holds particularly for compounding as the linear combination of words within a word. Derivation is more complex in this respect, because besides linear affixation it operates with many forms of non-linear processes, such as vowel or consonant alternation as in *to thief* from *thief*.

When we look at French compound structure, we can distinguish right from the start, as in syntax, between coordinate structures like *guide-interprète* and subordinate structures like *timbre-poste*. They are called endocentric compounds, because they have their governing constituent or head within (see below for the coordinative ones). Besides these there are so-called exocentric compounds, which seem to have their head outside. An example would be *peau-rouge* 'someone with a red skin, redskin', where 'someone' as the kernel of the meaning seems to correspond to a nominal head somewhere outside of the word.

In (2)-(4) below I illustrate the main French compound classes. They are of the categories N(oun), A(djective) and V(erb) containing either constituents of the same categories or of the categories P(reposition) and Adv(erb). These categories are considered to be the major lexical categories, with some hesitation as to the status of adverbs. I will not discuss compounds of any other category, which are extremely limited in number and in general idiosyncratic.

In (2) I illustrate the classes of coordinate compounds, in (3) the subordinate ones, and in (4) the exocentric ones. In (3) the head of the compounds is italicized, and in (4) the head of the compoundlike part of the words, which I will argue to be words derived by conversion or zero affixation from that compoundlike part, is italicized:

(2)	NN	AA	
	guide-interprète	aigre-doux	
(3)	N	A	V
	<i>timbre-poste</i>	<i>blanc comme neige</i>	(maintenir)
	<i>carte de visite</i>	<i>frais-rusé</i>	<i>rendre compte</i>
	A		
	<i>coffre-fort</i>		
	V		
	<i>basse-cour</i>		
	<i>machine à</i>		
	coudre		
	P		
	<i>sous-chef</i>		
	Adv		
	<i>arrière-pays</i>		
(4)	NP	A	V
	<i>peau-rouge</i>	<i>vert-sapin</i>	<i>sous-lou-(er)</i>
	<i>rouge-gorge</i>	<i>bleu-clair</i>	<i>maltrait-(er)</i>
	VP		
	<i>brise-glace</i>		
	PP		
	<i>en-tête</i>		<i>en-chaîn-(er)</i>

Let us first look at the least complex cases, which are the headed structures in (3). Restricting ourselves to the combinations of N, A and V in the first five rows, we see that they have the structure of syntactic phrases. This includes the fact that they are almost all left-hand headed like phrases, contrary to derivation and also to compounding in the Germanic languages, both of which are overwhelmingly right-hand headed. Moreover they inflect like phrases, forming for example such plurals as *timbres-poste*

and *cartes de visite*. So we may analyze them with di Sciullo-Williams (1987: 83) as syntactic phrases which are fixed to a certain extent, and thus listed in the lexicon on a par with idiomatic phrases.

This holds even for *timbre-poste*. We may consider, with Kleiber (1984) and Riegel (1989), that the prepositionless combination of a noun with a complement noun, as in *le projet Delors*, is in modern French a syntactic structure which, when used for name-giving, becomes more or less fixed and can take a specialized meaning to be listed in the lexicon. The combination of two adjectives is particular in that the first adjective position is restricted to a number of adjectives which can be used in an adverbial position. Among the verbs there are some very rare idiosyncratic right-hand headed examples like *maintenir*, which do not have syntactic structure. But we find masses of verbal expressions like *rendre compte*, which can be analyzed as phrases listed in the lexicon. They too inflect like phrases, which makes linguists often hesitate to consider them as compounds although they resemble them.

On the whole the proposed analysis accounts for the fact that there is much hesitation in the literature on the status of such expressions as *carte de visite*. Is it a compound or a phrase? In light of the proposed analysis I have not hesitated to introduce in (3) examples which are more seldom considered as compounds, but which are nevertheless more or less fixed in character, such as *blanc comme neige* and *bon à jeter*.

The remaining examples in (3) are of two kinds: those in the second last row have a real preposition: *sous-chef*, *surfin* and *sous-lou-(er)*, and those in the last row have an adverb: *arrière-pays*, *bienbeureux* and *maltrait-(er)*. They are right-hand headed. The head status of the right-hand constituent manifests itself in that it determines always the category, and in the case of nouns the gender, of the complex word. We cannot analyze these words as phrases, and thus must consider them as real compounds.

There is much hesitation in the literature concerning the analysis of words like *sous-chef*, *surfin* and *sous-lou-(er)*. They are often treated as prefixed words on a par with such words as *dé-faveur*, *in-capable* and *parler*. We will see in section 3 that these two classes of words resemble each other from a semantic point of view, but morphologically speaking we must divide them into compounds with prepositions occurring independently and prefixed words with prefixes not occurring independently.

Let us next consider the coordinate structures like *guide-interprète* and *aigre-doux* in (2) (observe that verbal analogs are missing). These cannot all be straightforwardly interpreted as listed syntactic phrases. It is true that the nominal ones can often be interpreted as containing a second noun modifying the first one as in syntax, compare *homme-grenouille* 'man who is (like) a frog, frogman' with an unsuspected phrase like *un avocat ami* 'a lawyer who is a friend, a lawyer friend'. They inflect like such phrases: *hommes-grenouilles*, and resemble them also in taking always the gender

of the first noun, as shown by masculine *homme-grenouille* with masculine *homme* and feminine *grenouille*. But the really coordinative ones, such as *guide-interprète*, would have as their syntactic counterpart a phrase with *et* like *guide et interprète*. On the other hand, even these always take the gender of the first noun and inflect like the other ones. Thus I will consider them to be fixed left-hand headed phrases with a modifying structure, whose interpretation leaves room for a really coordinative meaning.

Adjectival expressions of this kind like *aigre-doux* are much rarer. They are more problematic, since their syntactic counterpart can only be a coordinative adjectival phrase with *et* like *aigre et doux* or two adjectives stacked at different levels and modifying a noun as in *[un [[homme sourd] muet]]*. Given their rarity, the fact that their constituents inflect independently and the fact that they sometimes alternate with listed nominal phrases of the same form, like *sourd-muet / un sourd-muet*, I will interpret them for the time being on a par with such listed nominal phrases. To the extent that this is right, the expressions with coordinative meaning in (2) reduce to subordinative phrases listed in the lexicon like the ones in (3).

The exocentric expressions in (4) are in a way less problematic in that they simply cannot be interpreted as phrases. In *brise-glace*, for example, the noun *glace* must be the complement of the verbal constituent *brise*, but the entire expression is a noun and not a verb. According to di Sciullo-Williams (1987: 78-88) we have in *brise-glace* and *brise-tout* word-internal verbal phrases, which are made into nouns and adjectives by conversion. Apart from their complex verbal base, they resemble agent nouns derived from a verb by conversion like the noun *garde* from the verb *garder* or the English noun *bore* from the verb *to bore*. These conversion structures alternate with derived nouns with an overt suffix like *arrang-er* from *arranger* or English *work-er* from *to work*. As these examples show, the affixless conversion of a verb into a noun may have the same headlike function as a suffix, and this accounts their so-called 'exocentric' character.

The same may be argued for nouns derived by conversion from prepositional phrases like *en-tête*, and also for verbs like *en-chain-er*, where *en-chain-* is a verbal stem derived by conversion from a prepositional phrase, as I have proposed in Zwanenburg (1991). Nouns like *en-tête* take gender independently of the noun which they contain. That is, they take masculine for inanimates, like *un après-guerre*, and masculine or feminine according to biological gender for animates, like *un* or *une sans-coeur* 'a heartless person'. Moreover in the standard case these nouns take no inflection mark for plural: *les hors-bord*, *les sans-coeur*. Di Sciullo-Williams (1987: 83-84) argue that their analysis holds not only for such nouns, but also for nouns like *homme de paille*, *boule de neige* and *bon-à-rien*, *haut-de-forme* containing a nominal or an adjectival phrase. In my view, however, such expressions can be analyzed as endocentric syntactic phrases listed in the lexicon, with a nominalized adjective as head in *bon-à-rien* and *haut-de-forme*. They inflect

like phrases and always take the gender of the left-hand constituent, which is the head.

But di Sciullo-Williams' analysis holds for exocentric nouns and adjectives containing a nominal phrase, like *peau-rouge* and *vert-sapin*. Animate nouns of this kind take masculine or feminine gender depending on the sex of the beings they denote, and inanimate ones take unmarked masculine gender. It must be admitted, though, that in general they inflect like phrases, for example *peaux-rouges*. Adjectives of this kind, like *vert-sapin*, remain uninflected. They have as their syntactic analogs phrases such as (*des jupes d'un vert marron* 'dresses of a fir-like green'). The nominal phrase which they contain is headed by a nominalized adjective. This explains its lack of adjectival inflection, which is paralleled by simplex modifying nouns with adjectival function which remain uninflected, such as *des jupes marron* 'chestnut-colored dresses'. The nominal head is combined with a nominal complement, as in *vert-sapin*, or an adjectival one, as in *bleu clair*. In *peau-rouge* as well as in *vert-sapin* the conversion of a nominal phrase into a noun or an adjective has the same headlike function as an overt suffix in nouns like *basse-courrier* from *basse-cour* and adjectives like *long-courrier* from *long-cours*.

If this is on the right track, we can analyze the words in (4) as conversions with nominal, verbal and prepositional bases, left-hand headed by nouns, verbs and prepositions respectively. A particular class is that of *sous-marin*, which does not present conversion, but an ordinary suffix. I will discuss this further in section 5 when examining the relation between compounding and derivation.

This accounts in principle for the morphological structure of the tentative French compounds in (2)-(4). What we have found, in fact, are nominal, adjectival and verbal phrases listed in the lexicon in (2) and (3), real compounds with nonhead prepositions and adverbs in (3) and conversion structures with word-internal phrases in (4).

The listed phrases as well as the word-internal phrases are left-hand headed, as is French syntax in general. The real compounds are right-hand headed, like derivation. This means that in the proposed analysis French word formation observes the right-hand head rule, like word formation in the Germanic languages (cf. Selkirk (1982) for English).

The proposed analysis corresponds to a large extent to the analysis in Darmester (1875), who distinguishes juxtaposition and compounding. His juxtaposed structures correspond to our listed phrases. Noailly (1989) shows that since Bally (1932) there has been a tendency in French morphology to integrate listed phrases into compounding, and even to consider them as the most typical kind of compounds. In such an analysis French compounding would be dominantly left-hand headed.

I will not discuss here at length the fact that compounds can be embedded in compounds. Rohrer (1967: 148) mentions such examples as *lave-pare-*

brise or *écran timbre-poste* (*de la télévision*). Compounds with coordinative meaning can have more than two members at the same level, such as *rouge-blanc-bleu*. The relation between compounding on the one hand and derivation and inflection on the other will be discussed more at length in sections 4 and 5.

3. Semantic or conceptual structure.

Discussing the meaning of French compounds, I will suppose, with Jackendoff (1983, 1990), that there is a level of conceptual structure, independent from, but linked to morphological structure. The principal relations in conceptual structure are that of function-argument and restrictive modification. Certain conceptual arguments are marked for obligatory realization in syntactic or morphological structure, and are thus syntactic arguments. The discussion of this section will be to a large extent a tentative one, as much work has to be done in order to develop a full-fledged conceptual analysis.

In this section I will not consider phrases listed in the lexicon, because they can be analyzed as phrases to the extent to which they are not semantically specialized. I will thus restrict myself to words, that is either real compounds in the last two rows of (3) or words with word-internal syntactic structure in (4), repeated here for convenience sake:

(3)	N	A	V	
P	<i>sous-chef</i>	<i>surfin</i>	<i>sous-lou-(er)</i>	
Adv	<i>arrière-pays</i>	<i>malpropre</i>	<i>maltrait-(er)</i>	
(4)	N	A	V	
NP	<i>peau-rouge</i>	<i>vert-sapin</i>		
VP	<i>rouge-gorge</i>	<i>bleu-clair</i>		
PP	<i>brise-glace</i>	<i>brise-tout</i>		
	<i>en-tête</i>	<i>sous-mar-(in)</i>	<i>en-chaîn-(er)</i>	

3.1. Real compounds.

Let us first consider the real compounds in (3). In order to analyze them it is useful to contrast them with the superficially resembling words in the last row of (4). The latter will be discussed for their own sake in section 3.4 below.

In accordance with its status of nonhead, the initial prepositional or adverbial constituent in the expressions in (3) can be interpreted as a modifier of the head, serving to denote a particular variety of what the head denotes. Thus *arrière-pays*, *malpropre* and *maltraiter* with an initial adverb denote particular varieties of *pays*, *bereux* and *traiter* respectively. In principle this holds also for the words beginning with a preposition, but some discussion is needed here.

According to Guilbert (1971) the prepositions we can find in nominal compounds like *sous-chef* in (3), as well as in the superficially resembling expressions like *en-tête* in (4), are *après*, *arrière*, *avant*, *contre*, *entre*, *bors*, *ouïre*, *sans*, *sous* and *sur*. All of these are straightforwardly prepositions, except for *arrière*. This was a preposition in older French, and occurs nowadays as a noun and in some fixed expressions like the adverbial expression *en arrière*. In the kinds of words I am interested in here it has in general the meaning of the locative preposition *derrière* (which does not occur itself in such words), as *avant* has the meaning of temporal *avant* as well as that of locative *devant* (which does not occur in such words either).

In the nouns under consideration most of the prepositions can express together with a noun different semantic relations. Four of such relations can be distinguished, which I will analyze for one of the prepositions presenting them all, namely *avant*.

Here is a brief characterization of the four different semantic relations expressed by *avant* and some examples of each (N represents the noun of the structure, X represents an unexpressed nominal concept):

- (5) a. N before X:
avant-poste "outpost", avant-train "forequarters (of horse)"
- b. forepart of N:
avant-bras "fore-arm", avant-port "outer harbor"
- c. N before N:
avant-cour "forecourt", avant-goût "foretaste", avant-projet 'preliminary plan'
- d. X before N:
avant-guerre "pre-war period", avant-main "forehand (of horse)", avant-scène "stage-box"

In (5) one can distinguish, to start with, two opposite meanings, namely 'N before X' as in *avant-poste* 'poste before something' and 'X before N' as in *avant-guerre* 'something before guerre'. In between is the case of 'N before N' as in *avant-bras* 'forepart of bras'. The fourth meaning, 'forepart before N', applied to parts of N only: an *avant-bras* is 'a part of the arm in front of the rest of the arm'.

Types (5 a-c), as opposed to type (5 d), show the same morphological behavior, which is that of *sous-chef* in (3). They take the same gender as the noun which they contain, like *un avant-poste*, *un avant-bras* and *une avant-cour*, and inflect regularly for plural, like *des avant-postes*. We thus may analyze them all as right-hand headed compounds.

The morphological behavior of nouns of type (5 d), on the other hand, is that of *en-tête* in (4). They take masculine gender for inanimates, and masculine or feminine according to biological gender for animates. In the standard case they remain uninflected for plural. We can correspondingly analyze them, like *en-tête*, as nominal conversions of a prepositional phrase.

There are nouns of type (5 d) which are exceptional with relation to gender and number. Thus for certain nouns gender, independent in principle, is determined after the gender of the N complement of the preposition, like *une avant-main*. Moreover nouns of type (5 d) inflect sometimes for plural like those of (5 a-c). These exceptions may be due to a more general indecision concerning the gender of certain—mostly inanimate—nouns in French, and to some semantic fuzziness of the distinction between the four meanings of (5) caused by the two intermediary meanings.

In accordance with the fact that the noun is the (right-hand) head in the case of (5 a-c) and the preposition a modifier, the meaning of (5 a), 'N before X', can be considered to be the canonical meaning of these kinds of nouns, and those of (5 b-c) can be interpreted as extensions of it. In the case of (5 d), on the other hand, where the conversion of a prepositional phrase to a noun has a headline function, we can derive naturally the meaning 'someone who / something which is PP'.

According to Guilbert's examples all the prepositions mentioned serve to form words of the class of (5 d). But contrary to his statement, *après*, *hors*, *outre* and *sans* seem to serve only to form these, the others can form also words of one or more of the classes (5 a-c).

Thus the meaning of the preposition or adverb initial nouns in (3), which is one of restrictive modification, results naturally from their morphological status of right-hand headed compounds. And the same holds for preposition or adverb initial adjectives and verbs like *surfin* and *malpropre*, *sous-lou-er* and *maltrait-er*.

As to verbs beginning with a modifying preposition like *sous-lou-er*, we have seen in section 2 that they must, like the nouns in (3), be distinguished from conversions with a word-internal prepositional phrase in (4), like *en-chaîn-er*, which resemble them superficially. The former may present the following prepositions:

- (6) a-batt-(re)
 contre-mand-(er)
 en-lev-(er)
 entre-v-(oir)
 par-achev-(er)
 pour-chass-(er)
 sous-estim-(er)
 sur-abond-(er)

The preposition has a modifying meaning. Thus *sous-lou-er*, like *maltrait-er*, denotes a variety of what the base verb denotes. In general the realization of the conceptual arguments of these verbs in syntax remains the same. But in some, relatively rare, cases the preposition may bring with it a modification of the syntactic argument structure of the verb. Here are those which I have found, and which may have to be supplemented with some more, but not many:

- (7) en-dorm-(ir) trans. < dorm-(ir) intrans.
 en-cour-(ir) trans. < cour-(ir) intrans.
 en-dur-(er) trans. < dur-(er) intrans.
 en-ferm-(er) X dans Y < ferm-(er) Y
 en-fou-(ir) X dans Y < fou-(ir) Y
 par-cour-(ir) X < cour-(ir) à travers X
 par-sem-(er) X de Y < sem-(er) Y dans/sur X
 sur-pass-(er) X < pass-(er) par-dessus X
 sur-vol-(er) X < vol-(er) par-dessus X

We have seen in section 2 that preposition and adverb initial right-hand headed compounds like *sous-chef*, *surfin* and *sous-lou-er* alternate with prefixed words like *dé-faveur*, *in-capable* and *re-parler*. These two classes of words are morphologically different, but they receive the same kind of semantic interpretation in terms of restrictive modification.

3.2. Words with internal nominal phrases.

Let us now have a look at the words with internal nominal phrases in the first two rows of (4), repeated here once more for convenience sake:

- (4) N A
 NP *peau-rouge* *vert-sapin*
 rouge-gorge

We have seen in section 2 that nouns like *peau-rouge* and *rouge-gorge* and adjectives like *vert sapin* can be analyzed as nominal and adjectival conversions of a nominal phrase. Nominal *peau-rouge* and *rouge-gorge* differ among each other only in the order of the constituents, NA and AN, both allowed by syntax. We have seen in section 2 that in the standard case the gender of the nouns is independent of that of the noun which they contain. This confirms their status as conversion structures, where conversion of a nominal phrase to a noun or an adjective has a headline function.

This function is comparable to that of the suffixal head in derived words like *boutiqu-ier* or *voil-ier*. There is a set of nominal suffixes, of which *-ier* is one of the most wide-spread, which can form nouns denoting persons or things by combining with a nominal base denoting a characteristic object with which these persons or things entertain some relation. And the same holds for the conversion structure in the words under consideration.

The semantic relation expressed by suffixes like *-ier* or the corresponding conversion is a very general one, and one which may be found also in so-called compounds like *timbre-poste* or their real compound analogs in the Germanic languages, like *postage stamp*, as well as in the possessive genitive of the Germanic languages, like *John's book*. Williams 1981 (88-89), considering examples like *John's* in *John's book*, characterizes it as "any relation whatsoever to the head noun; this is a great exaggeration, but it is a first approximation that is difficult to improve upon."

Along these lines we can analyze the nominal conversion structures *peau-rouge* and *rouge-gorge* as 'somebody in relation with *peau rouge*' and 'something in relation with *rouge gorge*' respectively. We have seen in section 2 that adjectives like *vert-sapin* contain the conversion of a nominal phrase to an adjective, and they establish the same general relation between the word-internal nominal phrase and the noun which they can modify, as in *un rideau vert-sapin*. Moreover they contain the same relation between the head noun of the nominal phrase and the complement noun: '*vert* in relation with *sapin*'.

3.3. Words with internal verbal phrases.

We must now consider the words, mostly nouns, with internal verbal phrases in (4), repeated here for convenience sake:

- (4) VP *brise-glace* *brise-tout*
 N A

We have seen in section 2 that we can analyze them, with di Sciullo-Williams (1987: 78-88), as conversions of verbal phrases into nouns and adjectives, respectively. The nouns are thus comparable to agent nouns with an overt suffix like *arrang-eur* from *arrang-(er)*, and the adjectives to adjectival participles like *charm-ant* from *charm-(er)*.

This analysis allows us to interpret nouns like *brise-glace* semantically along the same lines as agent nouns like *arrangeur*. Now that we are dealing with verbal bases, it is important to state that conceptual arguments of verbs may be syntactic arguments, that is arguments obligatorily realized in syntactic (or morphological) structure. Thus the conceptual arguments of *to put* in *John puts the book on the table*, for example, are obligatorily realized as syntactic arguments, called external argument (subject), direct internal argument (direct object) and indirect internal argument (indirect object) respectively. Williams (1981) argues that nouns, although lacking in general obligatory complements, have an external argument corresponding to their subject in predicative use, as *John* in *John is a fool*.

Now, according to Sproat (1985: 166-182), in agent nouns the external argument of the nominal suffix is identified with the external argument of the verbal base. One might represent this as in (8), where x and z are external arguments and <y> is a direct internal argument, and where identification is marked by identical indexes:

- (8) N N
 V
 arr-ang-
 x₁ <y> z₁
 -eur z₁

This accounts naturally for the semantic interpretation of agent nouns, which is 'someone who / something which Vs'. And nouns like *brise-glace*, where conversion has the same headlike function as the suffix in *arrang-eur*, can be interpreted in the same way. Thus *brise-glace* means 'something which breaks the ice, ice-breaker'. And the same holds for the corresponding adjectives.

I will not discuss here at length the status of the internal syntactic arguments of the verbal base, such as y in (8). Levin & Rappaport (1988) show that agent nouns come in two sorts, either with or without realization of the internal arguments of the verbal base, with a corresponding meaning difference. Thus *arrang-eur* may occur alone, or with its internal argument, as in (9):

- (9) l'arrang-*eur* de la *chanson*

Thus we may find *side* by *side* simple agent nouns without realization of the internal argument (*bris-eur*) or with syntactic realization of the internal argument (*bris-eur de glaces*), and complex agent nouns with word-internal realization of the internal argument (*brise-glace*).

3.4. Words with internal prepositional phrases.

This brings us finally to the words with word-internal prepositional phrases in (4), repeated here:

- (4) N A V
 PP *en-tête* *sous-mar-(in)* *en-chain-(er)*

In section 2 we have analyzed the nouns and the verbs as conversions of prepositional phrases to nouns and verbs, the adjectives as derived by means of a suffix from a prepositional phrase.

I have already discussed the semantic interpretation of the nouns in section 3.1., contrasting them with preposition initial right-hand headed compounds. We have seen there that from the headlike function of the conversion of a prepositional phrase to a noun we can derive naturally the meaning 'someone who / something which is PP'.

The adjectives can be interpreted along the same lines, except that in their case there is a suffixal head.

The verbs with word-internal prepositional phrases are different from nouns and adjectives in that they alternate with prefixed verbs where the prefix has a prepositional meaning. Thus the prefix *dé-* in the derived verb *dé-barqu-(er)* 'to disembark' is interpreted along the same lines as the proposition *en* in *em-barqu-(er)* 'to embark'.

According to Guilbert's data, the prepositions which we find in such verbs are *à* and *en*, and the prefixes which we find are *dé-* and *é-*. Because

of their parallel behavior, I will treat here together preposition and prefix initial verbs of this kind.

Here are the relevant transitive and ergative verb classes which we can find according to Guilbert, illustrated on verbs with *en* and *dé*:

- (10) transitive:
- 'to put to': *em-prisonn-(er)* "to imprison", *em-barqu-(er)* "to embark", *en-terr-(er)* "to put in the earth"
 - 'to provide with': *en-racin-(er)* "to root (tree)", *en-gazonn-(er)* "to turf over (piece of ground)", *en-crass-(er)* "to dirty (clothes)"
 - 'to make into': *em-brigad-(er)* "to brigade (troops)", *en-gerb-(er)* "to sheaf (corn)", *en-chaîn-(er)* "to make into a chain"
- ergative:
- 'to go to': *em-barqu-(er)* "to embark"
- (11) transitive:
- 'take out of': *dé-barqu-(er)* "to unship", *dé-boît-(er)* "to remove from its case (watch)", *dé-bourb-(er)* "to haul out of the mire"
 - 'to strip of': *dé-brid-(er)* "to unbridle", *dé-bois-(er)* "to deforest", *dé-bourb-(er)* "to clear (of mud)"
 - 'to make not': *dé-boucl-(er)* "to uncurl"
 - 'to go out of': *dé-barqu-(er)* "to disembark", *dé-boît-(er)* "to go out of a (military) column"

I lack space to discuss these verbs in detail. So I will just sketch their possible meanings in terms of thematic roles understood as relational notions defined structurally over conceptual structure (cf. Jackendoff (1990: 47)).

The standard meanings in (10)-(11) are transitive (10 a) and (11 a), and ergative (10 d) and (11 d). In transitive *em-prisonn-(er)*, for example, the verbal conversion structure introduces the complex predicate CAUSE TO GO TO, the internal syntactic argument is the THEME of TO GO and the incorporated argument is the GOAL, in conformity with its being governed by a locative preposition. In ergative *em-barqu-(er)* we have the simplex predicate GO TO, but apart from that it receives the same interpretation. The semantic difference between (10) and (11) is that in the latter case the predicate is CAUSE TO GO FROM and the incorporated argument is the SOURCE, in conformity with the meaning of the prefix.

The meanings of transitive (10 b) and (11 b) present a swap in the realization of the two arguments: the internal syntactic argument is the GOAL or the SOURCE and the incorporated argument is the THEME. This reminds one of the relation between the two kinds of argument realization of the *load* and *clear* verbs, studied in Rappaport & Levin (1988). These verbs present also two ways of realizing THEME and GOAL or SOURCE, as in (12):

- (12) (a) *charger des marchandises sur un wagon* 'to load goods on a wagon'
 (b) *charger un wagon de marchandises* 'to load a wagon with goods'

The verbs under consideration tend to present each only one meaning with the corresponding argument realization, but some have the two

meanings with the corresponding swap in the realization of the arguments, such as:

- (13) (a) *en-sabl-(er)* 'to put into sand → to strand, run aground (a ship)'
 (b) *en-sabl-(er)* 'to provide with sand → to cover with sand (a river)'

Compare also *em-paill-(er)*, which can mean 'to pack (goods) in straw' or 'to bottom (chairs) with straw'.

The meanings of transitive (10 c) and (11 c), finally, can be characterized as CAUSE TO BECOME. The internal syntactic argument is the THEME of BECOME, and the incorporated argument is the GOAL or SOURCE of BECOME. The preposition functions here as a copulative preposition in the terms of Walinska (1986: 272-275). In syntax such a preposition turns up in such sentences as:

- (14) (a) We introduced him *as* John's brother.
 (b) He presented the facts *as* irrelevant.
 (c) The children made what we gave them *into* a toy village.

I thus distinguish three transitive and one ergative meaning for the verbs under consideration.

Certain verbs allow more than one interpretation, as illustrated in (13) for 'to put to' and 'to provide with', and in (15) for 'to put to' and 'to make into':

- (15) (a) *en-chaîn-(er)* 'to put into chains → to enchain'
 (b) *en-chaîn-(er)* 'to make into a chain → to string (beads)'

This is a very global overview of verbs with word-internal prepositional phrases. I leave aside all questions concerning supplementary syntactic arguments which they may have, as well as some problematic marginal cases, which must await further study.

4. Compounds and derivation.

Bases of compounds can be derived words, as for example in *sous-entrepreneur* or *surintensité*. But on the other hand, Nyrop (1908: 38-40) observes that in general compounds cannot be bases of derived words in French. This might be the result of a general ordering of (morpho)phonological levels, as suggested in Zwanenbourg (1987) along the lines of Kiparsky (1982) (learned compounding represents the deepest or inner level, nonlearned compounding represents the highest or outer level):

- (16) learned compounding
 learned derivation
 nonlearned derivation
 nonlearned compounding

As to the relation between derivation and compounding suggested by Nyrop and illustrated in (16), Nyrop mentions some ad hoc exceptions as well as two kinds of exceptions which seem to be more principled and which are those in (17). These concern derived words containing right-hand headed nominal phrases and left-hand headed prepositional phrases respectively:

- (17) (a) extrême-oriental, grand-ducal, moyenâgeux
 (b) sous-marin, surnaturel

For the examples in (17 a) I have argued in Zwanenburg (1988) (406-409) that *extrême-oriental*, for example, might have the morphological structure [*extrême-orient-ial*] and the (morpho)phonological structure *extrême-[orient-ai]*. Such mappings between morphological structure and morphophonological structure have been proposed by Sproat (1985: 66-88) and, restricted to morphological structures where the embedded complex constituent is right-hand headed, by Hoeksema (1984: 48). This would mean that the ordering illustrated in (16) is valid on the (morpho)phonological level.

If such mappings are restricted to right-hand headed embedded constituents, the examples in (17 b) would be exceptional. But then the examples which can be given for (17 b) constitute a very restricted class of adjectives with *sous* and *sur*, which are structurally parallel to a large class of adjectives with learned compounding embedded regularly in derivation according to (16), such as:

- (18) inter-communal, sub-alpin

Moreover nonlearned *sous* and *sur* alternate unproblematically with learned *sub-* and *super-* in many non embedded compounds, as in *surfin* and *superfin*. And this may explain the exceptional cases of (17 b).

As to its phonological behavior, compounding seems to be scarcely distinguishable from syntax. Thus, in contrast to derivation, there is little or no phonological readjustment. There seems to be only a slight tendency to maintain more often schwa in compoundlike structures of all kinds discussed, so that *morte-saison* and *porte-bonheur* will be pronounced more often with schwa than the phrases *porte scellée* and *cela porte bonheur*. Compoundlike structures may individually show up unexpected liaison, as *pied-à-terre* and *pot-au-feu* with *t*, but this happens also in unsuspected phrases with idiomatic character such as *joindre le fait à la menace* with *t*.

5. Compounds and inflection.

As we have seen in section 2, most of the expressions in (3) can be considered to be syntactic phrases which are fixed in the lexicon, namely:

- (3) N N timbre-poste
 A carte de visite
 A coffre-fort
 V basse-cour
 V machine à coudre
 A blanc comme neige
 frais-rasé
 bon à jeter
 V (maintenir)
 rendre compte

As we have seen, most of the expressions in (3) are left-hand headed, as in syntax and exactly in the cases where syntax has them left-hand headed. And the same parallelism holds for their inflection: they inflect like phrases, so that we have *timbres-poste*, *cartes de visite*, etc. Thus no further comment is necessary here.

The same holds probably for the coordinate structures in (2):

- (2) NN guide-interprète
 AA aigre-doux

We have seen that these also present inflection as if they were phrases. But let us now consider the expressions in (3) and (4) which we have analyzed as real compounds and conversions with word-internal syntactic phrases respectively:

- (3) N N sous-chef
 Adv arrière-pays
 A surfin
 bienheureux
 V sous-lou-(er)
 maltrait-(er)
- (4) NP peau-rouge
 VP brise-gorge
 PP en-tête
 A vert-sapin
 brise-tout
 sous-mar-(in)
 en-chaîn-(er)

The expressions in (3), containing a nominal, adjectival or verbal head preceded by a preposition or an adverb as a modifier, have regular inflection marks on the head: *sous-chefs*, *surfin-e* / *surfin-s*, *sous-lou-ors* etc.

Among the expressions in (4), *sous-mar-in* is particular in that it presents a regular derivational suffix, which functions as the head and takes the inflection marks: *sous-mar-in-e* / *sous-mar-in-s*. The other nominal and adjectival conversions tend to remain without inflection marks, with the exception of the classes of *peau-rouge* and *rouge-gorge*, which inflect as if they were phrases listed in the lexicon. Verbs of the class of *en-chaîn-er* take inflection.

Thus inflection, when present, is always outside of compounds, except for the unexpected behavior of *peau-rouge* / *rouge-gorge* and the class of *brise-glace*, which takes sometimes internal plural inflection: *compte-gouttes*. The latter fact must be linked, without any doubt, to the syntactic nature of the word-internal phrase.

6. Compounds and the lexicon.

In what precedes I have supposed a model of the grammar in which morphology and the lexicon are independent components (cf. di Sciullo & Williams (1987)). The lexicon contains, besides all the morphemes which constitute the input to the morphology, all the unpredictable morphological as well as syntactic structures, but not the regular ones.

This accounts nicely for the fact that much of what is sometimes considered to be compounds, namely the cases illustrated in (2), must be analyzed as phrases, which can nevertheless serve as names in the terms of Kleiber (1984) and Riegel (1989), and as such be listed in the lexicon.

Among the listed phrases which I have illustrated in (2) and (3), only certain classes of the nominal ones seem to be relatively productive, particularly the coordinative ones (*guide-interprète*) and the subordinative ones with conceptual function-argument structure (*timbre-poste*, *carte de visite* and *machine à coudre*).

Among the other structures only the class of *brise-glace* seems to be relatively productive to form instrument nouns (cf. Rohrer (1967: 146-148)).

As we have seen, the fact that there is such a close relation between compounding and syntax in French may account for the fact that there is much uncertainty in general concerning the delimitation of compounding in the literature. It may also account for the fact there is much hesitation as to the written form of compoundlike expressions, as a single word, with a hyphen or as two words: compare *surplace* and *sur-place*, *chef-d'oeuvre* and *chef d'oeuvre*.

7. Conclusion.

A general characterization of French compounding, as opposed to Germanic compounding, might be that many tentative classes of compound nouns, adjectives and verbs, corresponding to real compounds in the Germanic languages, are lexicalized phrases. Besides, there is a series of classes of mostly zero derived nouns, adjectives and verbs which contain word-internal nominal, verbal and prepositional phrases.

Real compounding seems to be restricted to nouns, adjectives and verbs with a modifying preposition or adverb. The most striking outcome of this overview might well be that, contrary to what is normally said in the literature, French compounding, restricted to these classes of real compounds, is right-hand headed.

Address of the Author:

Romaanse talen en culturen
Rijksuniversiteit te Utrecht
Kromme Nieuwe Gracht, 29
3512 HD Utrecht, The Netherlands

Bibliographical references

- Bally, Ch., 1965 (first edition 1932). *Linguistique générale et linguistique française*, Bern: Francke.
- Darmesteter, A., 1875. *Traité de la formation des noms composés dans la langue française comparée aux autres langues romanes et au latin*, Paris: Franck.
- Guilbert, L., 1971. "De la formation des unités lexicales", in L. Guilbert et al. (eds.), *Grand Larousse de la langue française*, Paris: Larousse: IX-LXXXI.
- Hoeksema, J., 1984. *Categorical morphology*, PhD diss. Groningen.
- Jackendoff, R., 1983. *Semantics and cognition*, Cambridge Ma: MIT Press.
- Jackendoff, R., 1990. *Semantic structures*, Cambridge Ma: MIT Press.
- Kiparsky, P., 1982. "Lexical morphology and phonology", in I. S. Yang (ed.), *Linguistics in the morning clam*, Seoul: Hanshin: 3-91.
- Kleiber, G., 1984. "Sur la sémantique et pragmatique des SN le projet *Delors et la camarade Catherine*", *L'information grammaticale* no 27: 3-9.
- Levin, Beth, & Malka Rappaport, 1988. "Nonevent -er nominals: a probe into argument structure", *Linguistics* 26: 1067-1083.
- Noailly, Michèle, 1989. "Le nom composé: us et abus d'un concept grammatical", *Cahiers de grammaire* 14: 109-126.
- Nyrop, Kr., 1908. *Grammaire historique de la langue française* 3. Gyldenkalke Boghandel-Nordisk Forlag.
- Rappaport, Malka, & Beth Levin, 1988. "What to do with theta-roles", in Wendy Wilkins (ed.), *Syntax and semantics 21. Thematic relations*, San Diego etc.: Academic Press: 7-36.
- Riegel, M., 1989. "Vrais et faux noms composés: les séquences binominales en français moderne", in Dominique Weil and Huguette Fugier (eds.), *Actes du Troisième Colloque Régional de Linguistique Strasbourg 1988*, Strasbourg: USHS-Université Louis Pasteur Strasbourg: 371-191.
- Rohrer, Chr., 1967. *Die Wortzusammensetzung im modernen Französisch*, Tübingen: Betz.
- Sciullo, Anna Maria di, and E. Williams, 1987. *On the definition of word*, Cambridge Ma: MIT Press.

- Selkirk, Elisabeth O., 1982. *The syntax of words*, Cambridge Ma: MIT Press.
- Sproat, R. W., 1985. *On deriving the lexicon*, PhD dissertation MIT, Cambridge Ma.
- Walinska de Hackbeil, Hanna, 1986. *The roots of phrase structure*, Seattle Wa: PhD thesis University of Washington.
- Williams, E., 1981. "Argument structure and morphology", *Linguistic review* 1: 81-114.
- Zwanenburgh, W., 1987. "Le statut de la formation des mots savants en français et en anglais", *Meta* 32: 223-229.
- Zwanenburgh, W., 1988. "Morphological structure and level ordering", in M. Everaert at al. (eds.), *Morphology and modularity*, Dordrecht: Foris: 394-410.
- Zwanenburgh, W., 1991. "Morphological heads, French 'compounding' and Germanic 'prefixation'", in Christiana Lauefer and T. A. Morgan (eds.), *Theoretical analyses in Romance linguistics*, Amsterdam: Benjamins: 167-179.