Claudio Iacobini - University of Ferrara (Italy) - e.mail: ibc@dns.unife.it

DISTINGUISHING DERIVATIONAL PREFIXES FROM INITIAL COMBINING FORMS

Abstract: The definition of prefix is based upon formal characteristics which do not allow for a clear distinction between prefixes and initial combining forms (icf's). Even though the differentiation between prefixes and icf's can be represented on a continuous scale, arguments are given in favour of a clear-cut distinction between typical prefixes (derivational elements) and typical cf's (compounding elements). A detailed definition of prefix is provided, and derivational affixation is compared with compounding. An explanation is given of the reasons why some icf's can lose their stem-like characteristics and lean towards prefixes to such an extent that they are fully identified with them. This paper ends with a list of Italian prefixes, and of the icf's which are most similar to them.

0. Introduction

This paper focuses on the definition of derivational prefix and on the criteria distinguishing prefixation from compounding and those which distinguish prefixes from initial combining forms (also called prefixoids, formants, neo-classical elements, confixes, etc.).

The commonly accepted definition of prefix — "a bound form that can be productively attached to the left of a word (in the sense of a free lexical element)" — is unsatisfactory because of its vagueness. Such a definition is not sufficient to identify the prefixes of a language in a consistent way. For example, the collection of elements defined as "prefix" in four authoritative grammars and in five dictionaries of contemporary Italian amounts to 106 items, but there is unanimous agreement only on 10 of them in spite of the fact that each source mentions approximately 80 to 90 items.

The main disagreement concerns the classification of elements which some regard as prefixes and others as initial combining forms (henceforth icf's). The principal reason for this disagreement is that the definition of prefix is essentially a formal one. Since prefixed words and words formed by an icf plus a word are similar, both from a formal (bound form / free form) and a functional point of view (non-head / head), we need more precise criteria to characterize prefixes and distinguish them from icf's, in addition to the positional criterion and the syntactic autonomy of the elements.

For this purpose, I will suggest a set of properties defining (proto)typical derivational prefixes. I will then compare the characteristics of affixal derivation to those of compounding. My definitions of prefix and of compound refer to the Italian language, but they may also be extended to the other Romance languages. The distinction between prefixes and icf's applies more generally and concerns other languages as well.

Even though the differentiation between prefixes and icf's can be represented on a continuous scale, I will argue in favour of a clear-cut distinction between typical prefixes (that is, affixes that operate according to derivational rules) and typical cf's (that is, stem-like elements that operate according to compounding rules).

Unlike words and affixes, cf's do not form a natural class; rather, they are a heterogeneous set of foreign elements interacting with the elements and the formative processes proper to a specific language. The wide usage of technical and scientific terms in everyday language, and processes of grammaticalization can make some icf's lean towards prefixation. The icf's which are at a more advanced stage of the grammaticalization process lose their original categorial identity, they show a tendency to play a subordinate role with respect to the other element in a complex word, and tend to be more general in meaning, expressing cognitively relevant, recurring components of meaning which are organized in paradigmatic relations with the meaning expressed by prefixes.

The suggested set of properties defining (proto)typical derivational prefixes makes it possible to draw a list of prefixes in the Italian language. By working out the criteria distinguishing prefixation from compounding, as well as the characteristics typical of cf's, we will be able to reduce to a minimum the number of elements for which there is still a margin of arbitrariness in distinguishing between prefixes and icf's. We will also be able to justify the decision to assign them to either the one kind of element or the other.

1. Theoretical disagreements

The disagreements in defining prefix and prefixation concern both the kind of formative process to which prefixation belongs (derivation vs. compounding), and the kind of lexical element to which prefixes can be ascribed (affix vs. lexeme). With regard to this, there are two main positions within generative morphology. The one we could call "lexicalist" (represented by Aronoff (1976), Booij (1977), Scalise (1984)) clearly distinguishes lexemes from affixes and states that prefixation belongs to derivation (together with suffixation) and not to compounding. The other, which we could call "syntacticist" (represented by Siegel (1974), Allen (1978), Lieber (1981, 1992)), tends not to distinguish derivation from compounding, and to equate lexemes and affixes. Items like *electromagnet* are considered compounds by the former and derivatives by the latter, like words such as *unhappy* or *happiness*. For instance, Lieber (1981) states that *hydro* and *macro,* are "Level II affixes", like the privative prefix *un-*. Therefore, the distinction between prefixes and cf's is of the least importance in the syntacticist view, whereas it is established in principle — but not in detail — in the lexicalist view.

2. Research model

For my research I have adopted the "word based morphology" model, where word means "lexeme" as proposed by Aronoff (1994). However, unlike Aronoff (1976), I believe the lexicon includes units of different kinds, which word formation rules combine to form new words (cf. Selkirk 1982, Corbin 1987). The lexicon, besides lexemes, includes stems and affixes with a proper lexical entry (cf. Lieber 1992). Consequently, affixes do not identify with word formation rules, but more than one affix can be associated to a single word formation rule. The separation between affixes and

word formation rules allows us to maintain the distinction between derivational and compounding processes and at the same time to handle intermediate phenomena which otherwise would be difficult to classify. The classification of prefixes I am suggesting and the distinction between prefixes and icf's is not based on a number of examples selected from those which are clearest, but is the result of an analysis covering all the bound-elements preposed to words in the word formation system of contemporary Italian (cf. Iacobini, 1992).

3. The distinction between prefixation and compounding

The distinction between derivation and compounding, between affixes and lexical elements centred on the syntactic autonomy criterion shows its limits when we consider words formed by two cf's (e.g. *anthropology, geography*) which are rightfully regarded as compounds even if they violate the definition of compounding based on the syntactic autonomy of components.

Scalise (1984: 75) gives several arguments — which I have summarized in the following five points — for these elements to be considered as learned stems, combined in compounds, by postulating properties of affixes that these elements do not share:

a. An affix cannot be a prefix in some words and a suffix in others.

b. Affixes cannot be "factored out".

c. The position of affixes is fixed, that of cf's is not.

d. A free element (= a word) cannot consist only of affixes.

e. Cf's have a suppletive semantic relation with words.

Let us now look carefully at Scalise's arguments one by one.

a) The first point is valid, but it has a limitation. It only applies to the limited number of cf's that can occupy both the initial and the final position with no difference in meaning (e.g. It. *bio* in *bioritmo* "biorhythm" and *anfibio* "amphibious"; *geo* in *geologia* "geology" and *ipogeo* "hypogean"). However, it does not help us to distinguish prefixes from icf's combined with words.

b) Prefixes too can be factored out. For example, Engl. *pro-* and *enclitics*, Germ. *be- oder entladen* ("charge or discharge"), Dutch *ge- en verboden* ("orders and bans").

c) Scalise gives examples of copulative compounds, a kind of compounding in which the order of elements is not significant. He makes the example *Italo-Anglo-Soviet* compared to *Anglo-Italo-Soviet*. However this property is the consequence of this specific type of compounding and not of the use of cf's. In compounds having the structure specifier/head, the order is meaningful: see It. *logopatia* "logopathy", *logopedia* "logopedics" as compared to *patologia* "pathology",*pedologia* "pedology"; or the order is determined by the fact that the cf's can only occupy one position, either initial or final: see the English *pseudomorph* vs. the impossible word *morphopseudo.

d) If it is true that outside the foreign-learned vocabulary it is not possible that a word is formed by two affixes only, it is also true that in Italian it is impossible to find words formed by two or more stems (Italian compounds are generally formed by the combination of two free words). Moreover, being foreign to common lexicon is not a feature pertaining only to cf's: in many languages there is a subset of the vocabulary that behaves unusually in certain morphological processes, and consists of loan-words having a "learned" connotation. Besides cf's, it contains clear affixes (e.g. the English suffix

134

-ity, which is usually joined to stems of Latin origin, as in *curiosity, profanity*), and words (like the English *curriculum:* pl. *curricula* besides *curriculums*).

e) Scalise's examples (*bio-* meaning "life" and *anthropo-* meaning "man") point to semantic relationships between cf's and words; but there are also semantic relationships between affixes and words, e.g. *pre-* and *before, post-* and *after, re-* and *again*; and between cf's and words belonging to closed classes (for example, locative prepositions and adverbs): *eso-, ecto-* "out", *endo-* "in", *cata-* "down", *opistho-* "behind".

To sum up, Scalise's criteria, apart from the first (which is useful, but of limited scope), are not very helpful in clearly distinguishing between affixes and cf's. However these criteria give a very useful, yet often overlooked insight: cf's are a heterogeneous set with different characteristics and behaviours.

4. Further complications

Three more phenomena make the distinction between compounding and prefixal derivation more intricate.

First and most important: some icf's can be employed as determiners even before words (and not only in combination with other cf's), that is, in conformity with the typical structure and function of prefixation (bound form / free form) (specifier / head) (e.g. It. *narcotrafficante, euroburocrate, pornoeditore*). Second: prefixes, which are productively combined only with free words, may appear to be combined with bound stems (e.g. Engl. *reduce, produce, remit, commit*). Third: some prefixes are used as free words (e.g. It. *super* meaning "gasoline, petrol").

The third point concerns idiosyncratic behaviours of an extremely small number of prefixes. Rather than prefixes, they can actually be seen as truncated forms. The word-like meaning these self-standing prefixes express is that of the former complex word which underwent a truncation process. The second point stems from diachronic causes. These are words with a Latin origin, not formed according to productive word formation rules. The first point deserves our consideration since it is the kind of formation process involving cf's which most resembles prefixation.

5. The need for a definition of the prefix

We have seen that, in order to distinguish between prefixes and icf's, and to single out a set of homogeneous elements which can be defined as prefixes, the property of syntactic autonomy is not sufficient. Neither is it helpful to work out another criterion that might serve the same purpose on its own, for example, a criterion of etymological or semantic nature. (Reasons of space do not allow us to discuss these matters here).

The reason for this failure is that attempts were made to identify icf's by comparing them with prefixes without having an adequate definition of prefix. Whenever it has not been possible to identify a cf by showing that it may occur as the first and last element of a word with no difference in meaning, an attempt was made to prove that it is a bound element other than a prefix. But this reasoning demands independent criteria for "real" prefixes, as well as a comparison between prefixation and compounding.

In order to establish whether we are dealing with two different processes and kinds of lexical elements, or with one process and one kind of element, it is therefore necessary to draw up the set of properties defining a prototypical model of prefix and of

prefixed word. It is then necessary to compare the behaviour of bound initial elements according to these properties (taking into account tendencies in usage and quantitative and distributional data), so as to see which of those elements conform to the proposed model and thus can be defined "real prefixes". On the other hand, this will enable us to see which elements having behaviours and properties typical of compounding, gradually depart from it, and lean towards the definition of prefix.

In paragraphs 6 to 9 I will suggest a definition of prefixation and of compounding, and I will discuss how combining forms relate to the two formative processes. I have adopted this way of reasoning rather than an intensional definition of cf's because cf's are not a natural class in languages. They are bound stem-like elements combined in compounds belonging to technical and scientific registers. Many of them do not become part of common language, and stay apart from the word formation system. Some cf's become part of the language through grammaticalization and lexicalization processes, and their behaviour range from that of autonomous words to that of derivational affixes (See table in paragraph 8).

6. The defining criteria of prefixes

The following list suggests a set of properties identifying the prefixes productively used in the formation of words of common usage in the Italian language. (It can be extended to the other Romance languages with slight modifications).

1) They are affixes (= bound elements) without a syntactic category (= N, A, V) that attach to the left of a lexical base in order to modify its conceptual meaning.
2) They form new words by attaching to words; they do not productively form words if combined with affixes or other bound elements.
3) They can play a role in the formation of parasynthetic verbs, that is, verbs like the Italian *arrivare* "to arrive", *ingabbiare* "to cage", for which the corresponding non-prefixed verbs — *rivare, *gabbiare — do not exist.
4) They occur only in the initial position of a word (a prefix can be preposed to an other prefix within certain restrictions).
5) They act as determiners (they are not in coordinative relation with the word which they are attached to, nor with other prefixes).
6) They select bases mainly according to semantic criteria. They can thus violate the Unitary Base Hypothesis, even in the modified version proposed by Scalise (1984).
7) They express functional-relational meanings (as opposed to both lexical and grammatical ones); pragmatic or encyclopedic knowledge is not necessarily required to interpret the compositional meaning of prefixed words.
8) Their meaning is generally known to all speakers.
9) They form endocentric words of which they are not the head: the category, the gender and other inherent features of prefixed words (e.g. inflection, animacy) remain the same as those of the base word.
10) They can affect all the meanings of the base word or only part of them; in this second case, the characteristics inherited by the prefixed word are a subset of those proper to the base and match the selected meanings.
11) Some prefixes can be used recursively and have (like all prefixes) their semantic and syntactic scope over the whole complex word.

12) They do not move the primary stress of the base.
13) They do not modify the word onset of the base.
14) They can undergo slight phonotactic modifications.
15) They have restrictions in length.

7. Compounding

Let us now look briefly at the definition of compounding. (Based on Italian primary compounding).

- Compounding combines two stem-like elements.
- The elements of compounds have an independent lexical meaning.
- The semantic relation between the elements is not explicitly signalled; consequently, the meaning of compounds allows for several possible interpretations.
- A head-element can be a non-head in another compound.
- The typical relation between constituents is subordinative, but coordinative or exocentric constructions are also possible.

8. Different kinds of combining forms

By comparing the properties of compounding with those of affixal derivation, it is possible to work out a set of criteria which help us to identify cf's and their behaviour with respect to word formation processes.

The criteria of comparison I have used between compounding and affixal derivation are as follows:

1) Position: the affix has a fixed position, the constituent of compound does not.
2) Combinatorial capacity: the affix does not combine with other affixes; the constituent of compound combines with derivational and/or flexional affixes.
3) Relation between constituents: derivation only allows for a subordinative relationship, compounding also allows for a coordinative relationship.
4) Head position:
   a) derivatives are always right-headed, productive compounds are left-headed;
   b) both derivatives and compounds are typically endocentric.
5) Semantics: affixes have a relational meaning which requires a complementing semantic feature, and they clearly indicate the semantic relation between the elements of the derived word. The constituents of compounds have a denotative lexical meaning, which is semantically autonomous; the relation between the constituents is not clearly expressed; as a consequence, the meaning of compounds is much less precisely determined compared to that of derivatives.
6) Formal criterion: the affix is a bound element; the constituents of compounds are free elements.

On the basis of the above criteria, it is possible to draw the following table (worked out on the basis of Masseroli 1994:283) which depicts the behaviour of cf's as compared to the two poles consisting of prototypical derivational affixes and prototypical constituents of compounds.

137

Criteria	1	2	3	4a	4b	5	6
Prototyp. Affix	a	a	a	a	e	a	a
1) endo-, oligo-	a	a	a	a	e/x	a	a
2) crio-, igro-	a	a	a	a	e	c	a
3) agro-, bio-	a	a	c	a	e	c	a
4) lipo-, melano-	a	c	a	a	e	c	a
5) ippo-, mio-	a	c	c	a	e/x	c	a
6) -teca, -cida	a	a	a	a	e	a/c	a
7) -crazia, -machia	a	a	a	a	e	c/a	a
8) -antropo-, -geo-	c	c/a	c	a	e/x	c	a
9) -blasto-, -teo-	c	c	a	a	e/x	c	a
10) -termo-, -crono-	c	c	c	a	e/x	c	a
11) -socio-, -anglo-	c/a	c	c	a/c	e/x	c	a
12) -auto-, -foto-	c	c	c	a/c	e/x	c	c
Prototyp. Comp.	c	c	c	c	e	c	c

In this table, "a" indicates a characteristic typical of derivational affixes, "c" indicates a characteristic typical of constituents of compounds, except for column 4b, where "e" indicates endocentric formation and "x" exocentric formation. Where "c/a", "a/c", "e/x" appear, the first characteristic prevails over the second.

9. A list of prefixes

For reasons of space, my comments on the table will be limited. The table shows clearly the multiple behaviours possible in the whole set of cf's. Many cf's, though they are bound-forms and occupy a fixed position, have the typical behaviour of constituents of compounds, particularly with regard to semantics and the kind of relationships they have with the other members of a complex word. The conclusion we may draw is that the properties of some cf's cannot be applied to them all, and that cf's as a whole cannot be equated to affixes.

Restricting my comments to icf's, and returning to the issue of their distinction from prefixes, the data in the table show us that the icf's most similar to prefixes are those in group 1).

In addition to the properties included in the table, there are two more that help distinguish icf's from prefixes:
- Whether it is possible for them to combine with words and/or with cf's.
- Which register they belong to: everyday language as opposed to technical and scientific languages.

The great majority of words formed with cf's result from a combination of two cf's and belong to the technical and scientific register. In most cases, the meaning of cf's is only known to the people specialized in a particular field. When a similar word is used in everyday language (e.g. It. *telefono* "telephone", *frigorifero* "refrigerator", *idrogeno* "hydrogen"), the speakers are not usually able to distinguish the elements that make it up, and to give them their proper meaning. However, when a cf is attached to the left of several different free words (e.g. It. *ecoalimento, ecomercato, ecoprodotto, ecoterrorismo*; *protobarocco, protolingua, protoscienza, protoindustriale*) it is more

probable that speakers will be able to segment this kind of complex word. This is because they will recognize at least one familiar unit that is the word, and that is moreover the head of the formation. The compositional meaning of the "Icf + Word" type is easier to understand than the type consisting of two cf's. Since the base word is easier to understand, it will also be easier to interpret the meaning of the cf. So, the more an icf is used before words in formation of common usage, the more likely it is to become part of the competence of speakers. The structural analogy between this kind of formation and prefixation helps to associate (and in some cases identify) the icf's involved and prefixes.

Keeping these last two points in mind, the margin of uncertainty between prefixes and icf's narrows. And now the most probables candidates to prefix status are (within those depicted in group number 1) the icf's which mainly combine with words, which are known and used by the majority of speakers, and express a functional and relational meaning. Assigning these elements to the category of prefixes or to that of cf's requires some arbitrary judgement. (However, this only concerns a small number of items). Among these are a few which share all the defining properties of prefixes, even though sometimes they are used in combination with cf's to coin technical and scientific terms. Therefore, in a synchronic perspective, the following items can be considered as prefixes:

auto-, macro-, mega-, micro-, multi-, neo-, paleo-, para-, pluri-, poli-, semi-.

On the other hand, the other icf's belonging to group 1), do not fully match the defining properties of prefixes, even though they have many characteristics in common with prefixes. These icf's differ from prefixes in that they usually combine with other cf's (as well as words), they are not as general and diffuse in meaning as prefixes are, and they have not been mastered by the majority of speakers. They are:

archeo-, ecto-, emi-, endo-, equi-, eso-, etero-, filo-, meso-, oligo-, proto-, pseudo-, tauto-, vetero-.

The full list of Italian prefixes, i.e. of those elements which comply with the 15 characteristics shown in paragraph 6, is therefore as follows:

a(n)-, ad-, ante-, anti-[1] (e.g. *antieroe*), anti-[2] (e.g. *antivigilia*), arci-, auto- (e.g. *autogestione*), avan-, circum-, cis-, co-, con-, contro-, de-, dis-, ex-, extra-, in-[1] (e.g. *ingiallire*), in-[2] (e.g. *ingiusto*), infra-, inter-, intra-, iper-, ipo-, macro-, maxi-, mega-, meta-, micro-, mini-, multi-, neo-, oltre-, paleo-, para-, pluri-, poli-, post-, pre-, pro-, ra-, re-, retro-, ri-, rin-, s-, semi-, sopra-/vra-, sotto-, stra-, sub-, super-, sur-, trans-, ultra-, vice-.

Besides these, we should mention those prefixes which are no longer used to form productively words of common usage, but which are still present and easily detectable in many Italian words. (Some of them, however, are employed to coin technical and scientific terms). A full list follows:

ab-, ambi-, anfi-, archi-, bis-, circon-, citra-, contra-, di-, dia-, e-, es-, estra-, estro-, fra-, giusta-, intro-, mis-, ob-, per-, se-, so-, sor-, su-, tra-, tras-

10. Conclusion

In this talk I have argued that prefixation and composition are two separate processes involving elements of different kinds.

Cf's form a heterogeneous set which can interact both with the lexical elements and the word formation processes of a language. Their presence in compounds and their bound status are favourable conditions for their grammaticalization. The icf's expressing recurring components of meaning in paradigmatic relation with the meaning expressed by prefixes, and that are employed several times as determiners preposed to a word, can become so similar to prefixes as to fully identify with them.

An accurate definition of the notion of prefix, the comparison between such definition and the one identifying the constituent of compound, and a detailed analysis of the behaviour of initial bound-elements have made it possible
- to define the set of Italian prefixes
- to identify, motivate and circumscribe the inevitable overlapping between processes and formative elements that are actually different.

REFERENCES

Allen, M.R. (1978). *Morphological investigations*, Doctoral Dissertation, University of Connecticut.

Aronoff, M. (1976). *Word formation in generative grammar*, Cambridge (Mass.), MIT Press.

Aronoff, M. (1994). *Morphology by itself: stems and inflectional classes,* Cambridge (Mass.), MIT Press.

Booij, G.E. (1977). *Dutch morphology. A study of word formation in generative grammar*, Lisse, Peter de Ridder Press.

Corbin, D. (1987). *Morphologie dérivationelle et structuration du lexique,* Tübingen, Niemeyer.

Iacobini, C. (1992). *La prefissazione nell'italiano contemporaneo*, Unpublished Doctoral Dissertation, University of Rome "La Sapienza".

Lieber, R. (1981). *On the organization of the lexicon*, Doctoral Dissertation, MIT, Cambridge (Mass.), reproduced by the Indiana University Linguistic Club.

Lieber, R. (1992). *Deconstructing morphology*, Chicago, The University of Chicago Press.

Masseroli, A. (1994). *Le semiparole nella lingua italiana contemporanea,* Unpublished Doctoral Dissertation, University of Pavia.

Scalise, S. (1984). *Generative morphology*, Dordrecht, Foris.

Selkirk, E.O. (1982). *The syntax of words*, Linguistic Inquiry, Monograph Seven, Cambridge (Mass.), MIT Press.

Siegel, D. (1974). *Topics in English morphology*, Doctoral Dissertation, MIT, Cambridge (Mass.).