GLOSSOLOGIA 7-8 (1988-1989) 117-130

COMPOUNDS IN A TRANSFER BASED MACHINE TRANSLATION SYSTEM*

ANGELA RALLI

In this paper, we deal with the treatment of compounds, whether one-word or multi-word, in a transfer based machine translation system. We examine the issue of keeping the hierarchical structure of compounds below the level of syntactic representations. Such a treatment is not only motivated by the general hypothesis of an autonomous morphological component, but it also allows compounds to be translated in a simple way across European languages. Our claims are supported by examples taken from Greek, English, German and a certain number of the Romance languages such as French, Italian and Spanish.

0. Introduction

Machine Translation constitutes a procedure according to which the translation of a text is obtained by the computer. During this procedure, the computer must be able to produce a text into another language whose meaning is the same as that of the original text. The language of the input text is called «source language» (henceforth SL) and the language of the output text is usually known by the name of «target language» (henceforth TL).

Research and development in machine translation are guided by different strategies. For the last two decades, three strategies prevail in the design of machine translation systems: direct translation, the transfer approach and the interlingua approach.

According to the first approach, every sentence of the text is translated through a series of principal stages (usually about 10), such as dictionary lookup, morphological analysis, processing of idioms, subject-predicate identification, etc. (cf. Nirenburg (eds) 1987 for more details about this approach).

The transfer strategy consists of parsing every SL sentence into an abstract representation. From this representation, the sentence must be translated into a corresponding structure of the target language.

Finally, the interlingua strategy consists of developing a language-free conceptual representation onto which the source text is mapped. From this representation, the target text is obtained with the help of a natural language generator.

In this paper, we will formulate a proposal for the translation of compounds accord-

* I would like to thank Manuel Espagnol and Celia Roniotes for their helpful comments on an earlier draft of this paper.

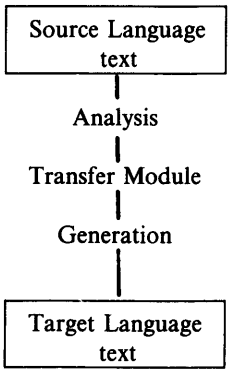
ing to the transfer translation approach: we are most familiar with this particular method from working for the transfer based EUROTRA machine translation project.

1. About the transfer strategy[1]

In a machine translation system, the transfer strategy depends mainly on defining a language independent representation by which the SL sentences are parsed. On the basis of this representation, transfer is achieved at both the lexical and structural levels into TL corresponding structures. The source sentences are monolingually analyzed according to the analysis component of the source language. Their translational counterparts are generated after being composed by the generation component of the target language. The content of the transfer representation varies from purely syntactic to syntactico-semantic information[2].

The transfer module is a bilingual component where language pairs involved have to be as similar as possible in order to keep the transfer procedure simple. Therefore, the burden of the work is shifted from the transfer part to the stages of the SL analysis and the TL generation. Schematically, the process of translation in a transfer based machine translation system can be illustrated as follows:

(1)



Generally, transfer can be kept simple if lexical items from the source language are translated into lexical items of the target language without any substantial change in the information involved (e.g. change in the grammatical category of an item with all the implications that such a change would have on the representation of the target language sentence). Transfer can also be simple if a SL structural representation is mapped onto an identical TL structural representation. However, more often than not, transfer is

1. The transfer approach has been developed and adopted by Machine Translation groups such as GETA in Grenoble (Vauquois 1975, Boitet et al. 1985), SUSY in Saarbrücken (Maas 1984) and the EUROTRA project supported by the European Community.
2. The content of the transfer representation could be augmented by inference mechanisms which apply world knowledge.

complex because, either there is a radical change in the structure of translated sentences or an expansion of a lexical item into a structure.

Compounds, or multi-element morphological structures represent a typical case of complex transfer because, between language pairs, a number of one-word compounds require translations as syntactic constructions. For example, the one-word compound *μηλιά* is translated into the English correspondent structure «*apple tree*». In the same way, the German compound word *Nachkriegsphase* «post-war period» has a correspondent noun phrase in Greek «*μεταπολεμική περίοδος*».

In this paper, we will focus on these cases of complex transfer, which among other things, also address the general theoretical question about the grammatical status and the structure of words. We will propose a specific treatment for the items considered to be compounds, which, in our view, contributes to the reduction of problems caused by complex transfer cases. We will support our claims with examples from a certain number of European languages such as Greek, English, German, French, Italian and Spanish[3].

2. Definition of compounds

Compounds are *morphological objects*, the structure of which is generated (or analyzed) by a restricted set of word formation rules. They contain *at least* two lexemes; lexeme is taken to be either a word or a stem/root constituent; one of its main characteristics is the lack of subcategorization frame which applies only to affixes[4].

Compounds are headed objects. The location of the syntactic head of a compound depends on the language. For example, Greek one-word compounds are always right-headed. Identification of the head is critical for proper percolation of feature values to the dominating node of the structure of compounds.

Compounds that are included in the scope of the proposed treatment must be compositional in meaning and productively derived. As there are no formal criteria in establishing semantic compositionality, only fully opaque compounds as far as their interpretation is concerned, will be excluded from this treatment (e.g. English *dumb-waiter*, Greek *λεωφορείο* «school-bus»). As a matter of fact, the latter will be considered as single atomic units without an internal structure. Productivity is in turn defined by morphological and grapho-phonological regularity in word formation[5]. That is to say, exocentric non-headed compounds will also be treated as listed atomic units (e.g. *redhead*).

Phrasal syntactic rules never operate within the structure of compounds; this follows

---

3. The position adopted here for the treatment of compounds in a transfer based machine translation system can be traced in a paper written for EUROTRA by Ananiadou and Ralli 1989 (cf. References).
4. A subcategorization frame contains information relevant to the change of grammatical category which is performed through the process of derivational affixation. That is to say, when a specific affix is added to a stem (or to a word) of a specific grammatical category, the created object assumes the category of the affix. For example, the English verb *standardize* is produced by the addition of the verbal suffix *-ize* to the nominal basis *standard*.
5. We refer to grapho-phonological regularity because machine translation deals with language in its written form.

from the Lexical Integrity Hypothesis (Lapointe 1979), which forbids any interaction between syntax and the word formation component.

Recently, the Lexical Integrity Hypothesis has been reformulated by A.M. Di Sciullo and E. Williams (1987) as «Syntactic Atomicity Principle». According to this principle, «morphological objects are 'atomic' at the level of phrasal syntax or phrasal semantics: they have 'features' or properties, ..., but the relation of these features to the internal composition of words cannot be relevant to syntax» (DS&W 87: 49,53). Therefore, words are opaque to syntactic descriptions and operations.

It is worth noting that the Syntactic Atomicity Principle also holds for a class of entities which do not have a traditional morphological form (i.e. one-word units), but behave like words in syntax. These entities can be multi-word structures, but, they display the general properties of words (i.e. X^0s) with respect to phrasal syntax. Di Sciullo and Williams call them 'syntactic atoms' and they provide the following definition:

«Let us call anything that can be inserterd into an X^0 position in syntax a syntactic atom».

The Syntactic Atomicity Principle is more general than the Lexical Integrity Hypothesis because it also refers to entities bearing more than one-word constituents (e.g. fruit market). Note that the internal structure of these items appears to be syntactically derived, but they also share many common properties with a certain number of morphological objects whose form is of one-word:

a) they behave as atomic units with respect to phrasal syntax and to a certain number of syntactic operations. For example, they do not contain any referential material and it is difficult to construct a syntactic phrase without including any referential material.

b) In most of these constructions, the meaning is not fully compositional; that is, it does not follow from the meaning of their constituent parts,

i.e. English : information technology
French : mot clé «key word»
Greek : nomos-plesio
lit. «law-frame»
Italian : uomo scimmia
lit. «man-monkey».

Syntactic atoms are closely related to morphological objects and particularly to compounds. Therefore, both will be treated here as compounds, no matter whether they are one-word or multi-word constructions. It should be stressed that it is critical to determine the status of these constructions with respect to syntax and morphology, i.e. whether they are phrasal or lexical entities. Multi-element compounds forming one-word are conventionally seen as clearly belonging to morphology. There is considerable debate over the status of multi-word compounds. That is why, results from recent research (i.e. the syntactic atomicity notion) enable us to view multi-word compounds as atomic with respect to syntax and treat them as morphological constructions[6]. Thus,

---

6. As far as Modern Greek is concerned, multi-word compounds were first located by A. Αναστασιάδη-Συμεωνίδη (1986). In her study, the author stresses the peculiar behavior of these

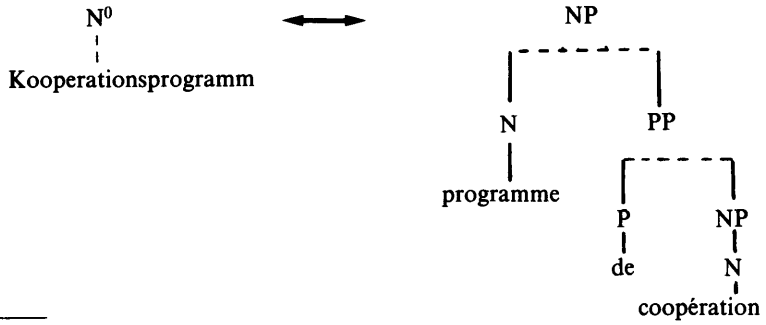
for syntax, all compounds have a dominating major lexical category and syntax has no access to the compound's substructure. Therefore, they will be considered as creations of an autonomous morphological component[7].

3. Machine translation and compounds

3.1 Morphology in machine translation

In machine translation, morphological analysis is usually seen as an uninteresting task for the level of transfer because words can be listed in a dictionary. They can be used as units without an internal structure and there is no need to identify their constituents. As we have already seen (§ 1.), as far as compounds are concerned, this strategy may cause a lot of difficulties in transfer: noun phrases in a language are structurally represented whereas in another language, their translation counterparts may correspond to one-word compounds usually treated as a lexical item. For example, when the German one-word compound *Kooperationsprogramm* is translated into the French nominal structure *programme de coopération* transfer becomes complex because a leaf in the structural representation of the source language has to be expanded into a structure in the target language or a SL structure has to be condensed into a TL leaf:

(2)



items with respect to syntax and she calls them «lexical phrases». According to her analysis, Greek multi-word compounds (lexical phrases for Αναστασιάδη-Συμεωνίδη) are distributed in three different structural patterns:

a. Adjective + Noun, e.g. διαστημικό λεωφορείο «space-bus»

b. Noun + Noun in genitive, e.g. ομάδα εργασίας «working team»

c. Noun + Noun in nominative: e.g. παράγοντας άνθρωπος «human factor».

At the time the book was written, no specific morphological treatment was available for multi-word compounds. Therefore, Αναστασιάδη-Συμεωνίδη provides a specific syntactic analysis for them.

In this paper, all Greek compounds productively derived will be considered as belonging to an open set of structures related to morphological constructions. They will be distinguished into one-word (i.e. morphological objects) and multi-word compounds (i.e. syntactic atoms). The two categories are different in some ways, but they are both words, in a well defined word formation sense of the word. For a detailed account of one-word and multi-word compounds in Greek, cf. Ralli '90 in Proceedings of the 11th Meeting of the Dept. of Linguistics, Univ. of Thessaloniki (forthcoming).

7. This treatment however does not cover the class of idiomatic phrases; the latter belong to a close set of lexicalized syntactic constructions and their meaning is not compositional (e.g. παίρνω μέτρα «take measures»).

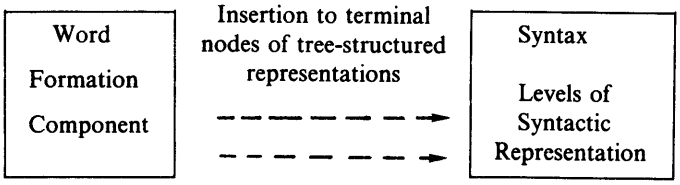
Problems of translating leaves into structures and vice-versa could be avoided if we follow a different strategy, according to which, all compositional compounds are treated as structured representations. As a consequence, it will be easier to map representations than being obliged to transfer lexical items into structures.

Every principled treatment of SL morphological objects is linguistically motivated, since morphology is now considered to be a grammatical domain with its own principles and rules. Furthermore, many corresponding TL structures could be considered as syntactic atoms and as such they could receive a morphological treatment. As we have already seen, syntactic atoms refer to the notion of word conceptually distinct from the traditional view of the word as one typographical unit because a well defined number of multi-word units have an apparent syntactic form, but their internal structure is produced and affected by rules and principles which are different from the phrasal syntactic ones. Following these remarks, every transfer procedure which involves translation of morphological structures, should keep this process separate from the subprocess of translating sentential structures. This position is in accordance with the hypothesis that morphology constitutes an autonomous grammatical component.

3.2 Autonomous morphological component

In theoretical linguistics, Borer (1988) has proposed to regard morphology as an independent module, operating in parallel with the syntactic component. In such a system, word formation takes place outside syntax but at different levels of syntactic representation resulting in slightly different syntactic properties. Consequently, some morphological constructions may be available prior to a number of syntactic operations (e.g. all derivational structures triggering a grammatical category change) and some of them may follow (e.g. multi-word constructions, involving word internal inflection (cf. Borer 1988 for details), providing their output does not violate any independent well-formedness conditions which are applicable at the syntactic levels. Schematically, these views could be represented as follows:

(3)



If these views are applied to machine translation, then, every treatment of one-word and multi-word compounds would consist of establishing an independent structural representation of morphologically created objects which are compositional in meaning and generated by productive word formation mechanisms.

Representations of sentences will then be structured in two levels:

a) the upper level, which will be dedicated to phrasal syntax and phrasal semantics, and

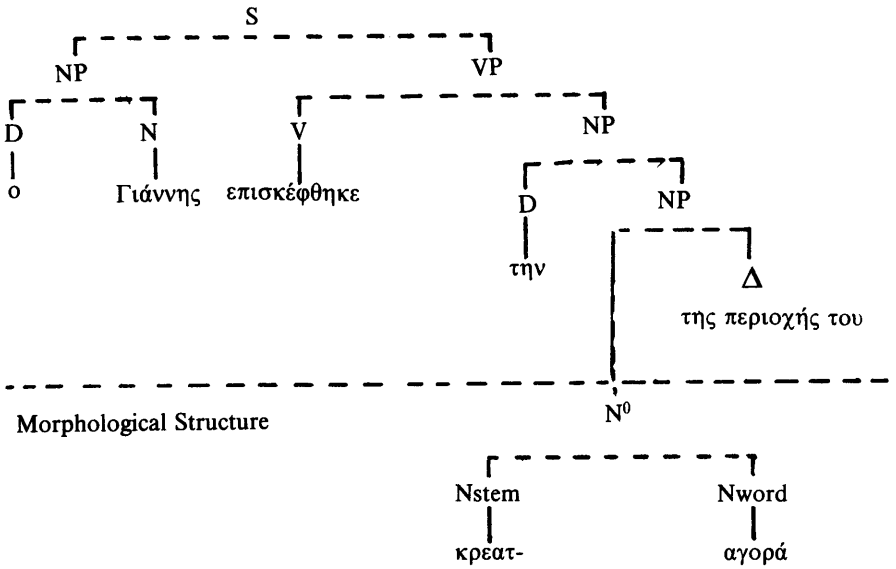
b) the lower level, which will be dedicated to morphological constructions.

For a phrasal syntactic representation, a multi-word compound would have a dominating major lexical category and phrasal syntactic operations will not be free to apply to the compound's substructure.

A morphological structured representation is opaque to phrasal syntax if the rules of phrasal syntax can only inspect the topmost feature bundle of this morphologically built object. In order to computationally implement this notion of opacity, one can adopt a special device to characterize the topmost nodes which expand into structured representations output by the sublevel of morphological constructions. This device would act as a special blocking mechanism preventing syntactic rules from applying in word internal constructions; it could be nothing else than the well known zero barlevel feature (X°) which characterizes all items appearing as terminal nodes in syntactic tree-representations.

According to the remarks outlined above, a Greek sentence like *Ο Γιάννης επισκέφθηκε την κρεαταγορά της περιοχής του* «John visited the meatmarket of his area» would be monolingually analyzed as follows:

(4) Syntactic Structure



4. Why the proposed representation is adequate for translating compounds

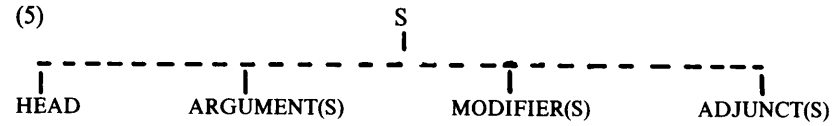
4.1 Representations in transfer modules

Before we offer any arguments about the adequacy of separating morphological and syntactic representations for the translation of compounds, we should place a number of general remarks concerning the structure of representations in transfer modules.

Monolingual representations, ready to be automatically translated into another language, should combine representational adequacy with computability. Representational adequacy does not necessarily mean to engage in particularly complex linguistic solutions because this would increase complexity in computability. In transfer, major differences between source and target language representations, triggered by language

specific application of linguistic principles, could result in several conflicting analyses that have to be mapped. That is why the transfer based approach to machine translation implies that source and target language structures should be normalized for translation purposes. Normalization means to minimize surface structural differences (such as differences in syntactic non-terminal categories and differences in word order) by providing representations whose constituents are placed in a canonical order. For example, the following could be a canonical order of constituents in a structured tree-representation, which is established on criteria related to notions of predicate-argument structure and functional roles of the different constituents[8]:

(5)

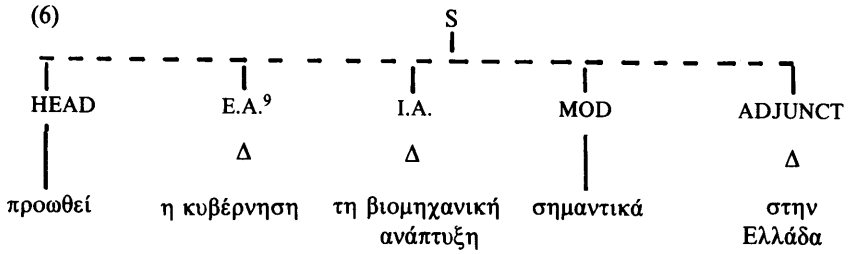


The structure of this representation can be seen as a product of the subcategorization properties of predicates, plus properties deriving from other operations on predicate argument structures. In such a way, representations between translation pairs reach a suitable level of generality and transfer becomes a simple procedure independently of phenomena related to the treatment of word order and complex syntactic structures.

To illustrate the above remarks, let us give an example from Greek which is a language with a relatively free-word order:

The sentence *Στην Ελλάδα, η κυβέρνηση προωθεί σημαντικά τη βιομηχανική ανάπτυξη* «In Greece, the government promotes industrial development considerably» can be represented as follows:

(6)



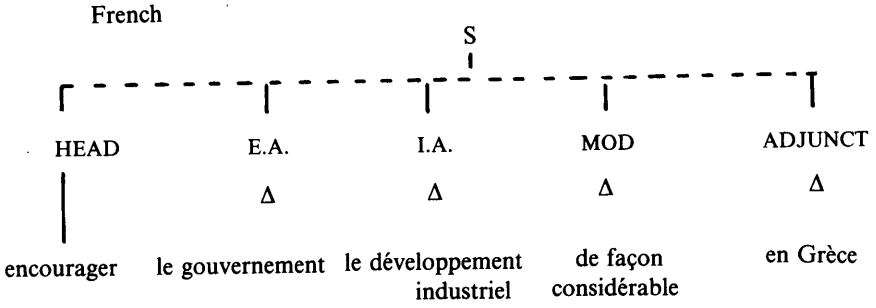
This tree-structure corresponds to isomorphic representations in several European languages, such as French and English for example, and complex transfer is avoided:

---

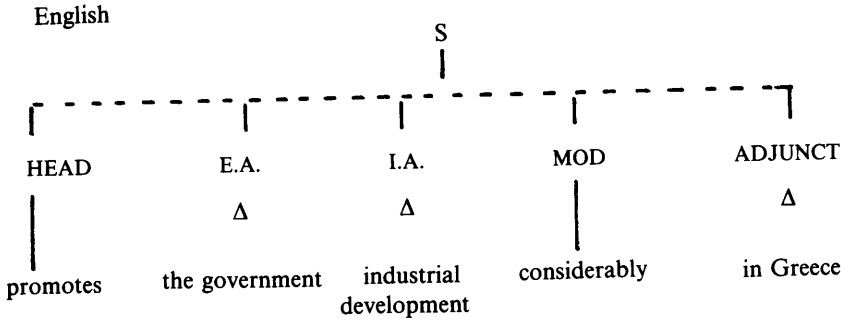
8. Such an order is also adopted by the Eurotra framework (cf. Steiner E. 1989). However, constituents are labelled in a different way.
9. E.A. and I.A. are abbreviations for External Argument and Internal Argument respectively. MOD stands for Modifier.

(7)

French



English



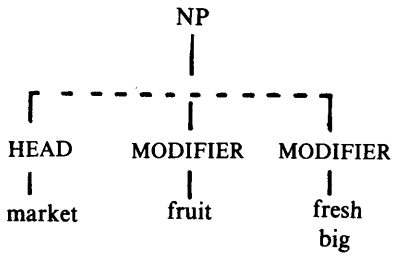
4.2 Arguments in favor of separating the morphological from the syntactic representation

Let us look now at a number of arguments in favor of the option for separating morphology from syntax. We believe that the approach described above for the analysis of compounds is adequate for the following reason (§ 3):
There is a clear distinction between rules applying to phrasal syntax and rules relevant to morphology. Mixing all rules together will cause two serious problems:
a) complication to the grammar, and
b) an unwanted overgeneration in transfer.

An argument for not mixing together syntactic and morphological rules comes from considerations concerning the scope of an adjective functioning as a modifier to one of the constituents of a compound. Let's take for instance the English examples of *fruit market, fresh fruit market* and *big fruit market.*

Given the fact that in a transfer based system representations should be isomorphic (according to the position described above (§ 4.1)), the structural representations for the English examples *fresh fruit market* and *big fruit market* will be the same if they receive a phrasal treatment:

(8)

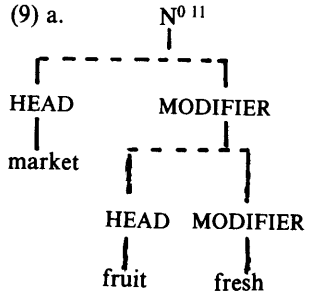


In such a system where no distinction is made between the levels of syntax and morphology, there is no way of structurally representing the fact that in the case of *fresh fruit market* the adjectival constituent *fresh* bears upon the modifier *fruit* whereas in the third case (*big fruit market*), the adjective *big* refers to the whole complex structure *fruit market*[10].

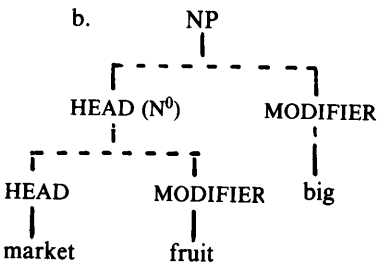
As the English representations do not contain any information for gender, this lack of information could cause wrong translations in languages where adjectives agree with their head nouns. In Spanish for example, *fresh fruit market* could be automatically translated as *fresco mercado de fruta* as well as *mercado de fruta fresca*. This problem could be avoided though if we keep morphological structures separate from phrasal syntactic ones.

Along the lines of the proposal formulated in paragraphs 2 and 3, the representations of *fresh fruit market* and *big fruit market* would involve different levels of representations, the syntactic and the morphological one, allowing scope of adjectives to be expressed. Different levels of representation would result in the following structures:

(9) a.



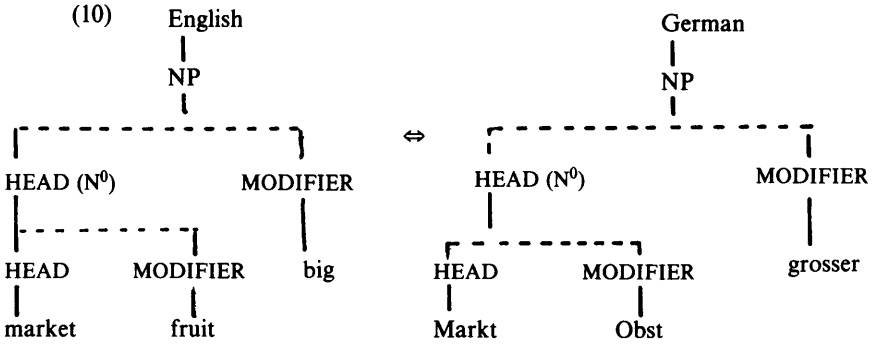
b.



10. Scope matters could be expressed with the help of a very rich system of semantic features referring to the domain of lexical semantics. This system, however, is still far from being established.

11. Syntactic atoms like the one in (9a.) do not have a «pure» morphological form but they display the general properties of X^0s. According to Di Sciullo & Williams (1987) syntactic atoms are created by a word creating rule assigning a X^0 category to a phrasal category (p. 84, 87).

An additional argument in favor of the approach for separating morphology from syntax is provided by the German translation of the English examples as *Frischobstmarkt* and *grosser Obstmarkt*. The first complex structure constitutes a one-word compound whose adjectival constituent part *frisch* is not inflected. Therefore, is should be considered as a stem, appearing at the left-hand side of the compound. The second case constitutes a noun phrase which contains an adjective (*grosser*) modifying a one-word compound (*Obstmarkt*). That means that in this case, the adjective *grosser*, which is fully inflected, is taken into account by rules of phrasal syntax while, in the first case, *frisch* is taken into account by morphology. It is clear that the internal structure of both examples matches the internal structure of the English corresponding cases. Therefore, the transfer procedure from English to German and vice-versa becomes simple. Naturally, in order to obtain results in this simple way, the German compounds must be morphologically analyzed and not listed as lexical entries:



4.3 Against the lexicalization approach

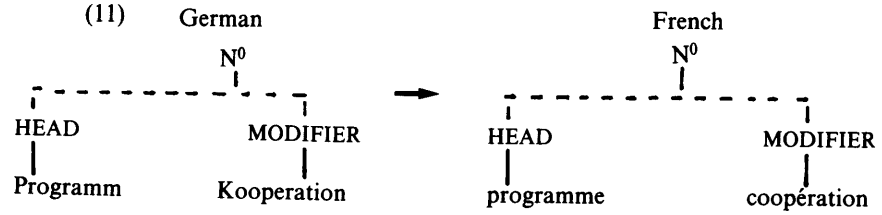
Translation of compounds from one language into another can provide us with good arguments against the lexicalization of one-word compounds. Let's take the German example *Kooperationsprogramm*. In some of the European languages, it has the following translations:

Greek : [[συνεταιριστικό]A [πρόγραμμα]N]
Italian : [[programma]N di [cooperazione]N]
English : [[cooperation]N [program]N]
French : [[programme]N de [coopération]N]

If we follow the lexicalization approach and treat the German example as one unit without an internal structure, then, the problems caused by complex transfer are not avoided because the translation that has to be performed involves the mapping of a lexical unit to NP structures in the other languages.

However, if the German example is compositionally treated, then, the translation procedure is simplified. Moreover, the procedure becomes less complicated if we keep the syntactic structure separate from the morphological one.

(11)



Note that in the German case, the -s- is a binding element. As such, it does not provide substantial information in the transfer module. Therefore, for transfer purposes, it will probably be treated as a feature on the mother node (N^0 [bind.elem. = s]).

In the same way, the French preposition *de* can be treated as an inherent monolingual marker for these kinds of compounds. Following the position we adopt here about distinct treatments between syntactic phrases and syntactic atoms, *de* should be featurized in all syntactic atoms (e.g. 12a.), but should be considered a phrasal constituent in noun phrases such as in (12b.)[12].

(12) a. genre de vie
question de vie
style de vie

b. durée de la vie
période de la vie
expérience de la vie

The Greek case is more difficult because there is a change in the grammatical category of the modifier constituent:

N ⇒ Adj (Kooperation ⇒ συνεταιριστικό).

In order to treat this particular case, we need a special transfer rule performing the change. However, it is easier to write that rule, than being obliged to deal with cases where a one-word unit has to be translated into a structure.

4.4 How will the grammar recognize compounds if there are no orthographical criteria

According to what we said in paragraphs 2. and 3. a compound is defined as a one-word or a multi-word unit. Bearing in mind that machine translation deals with texts, one-word compounds constitute more or less a «clear» case as far as recognition by the rules of the grammar is concerned. The second case, however, (i.e. multi-word compounds), is more complex considering the fact that most of the times they can also receive treatment as noun phrases.

As we already pointed out, multi-word compounds are distinguished from syntactic phrases by being opaque to the application of some specific phrasal syntactic rules. For example, the adjective-noun multi-word compound *προεδρικό διάταγμα* «presidential decree» is not subject to an internal modification of the head, caused by the insertion of another adjective like *σημαντικό* «significant», between the two words *προεδρικό* and *διάταγμα*[13]:

---

12. For a similar account of syntactic atoms in Portuguese, cf. Villalva (1989).
13. Adjectives like *προεδρικό* which appear in these kinds of constructions can also be distin-

*προεδρικό σημαντικό διάταγμα

The presence of an adjective like *σημαντικό* is admitted only if *σημαντικό* is placed in front of the adjective *προεδρικό* (*σημαντικό προεδρικό διάταγμα*) because in that position, it modifies the whole construction *προεδρικό διάταγμα*.

Therefore, it is possible to establish grammatical criteria according to which multi-word compounds can be distinguished from syntactic structures.

5. Conclusions

In this paper, we tried to show that the proposed representation for compounds which are compositional in structure and meaning is adequate for machine translation purposes. Specifically, we examined the issue of keeping the hierarchical structure of compounds below the level of phrasal syntax; this strategy is not only theoretically motivated (cf. «syntactic atomicity» thesis (Di Sciullo-Williams)), but it also allows compounds, whether one-word or multi-word, to be translated in a simple way across european languages. In order to support our claims, we used examples from different languages, namely, Greek, English, German and a number of the Romance languages such as French, Italian and Spanish.

Angela Ralli
University of Athens
Dept. of French Language and Literature

BIBLIOGRAPHY

Ananiadou E. – Ralli A. 1989: «Compounds and related constructions». *Eurotra Reference Manual* v. 6.0.

Αναστασιάδη-Συμεωνίδη A. 1986: *Η Νεολογία στην Κοινή Νεοελληνική* (Thessaloniki, Greece).

Arnold D.J. – L. Jaspaert – R. L. Johnson – S. Krawer – M. Rosner – L. des Tombe – G. B. Varile – S. Warwick 1985: «A MU1 view of the 'CA', T framework in Eurotra». In *Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation of Natural languages* (Hamilton, N.Y.: Colgate University) 1-14.

Borer H. 1988: «Morphological parallelism between compounds and constructs». In G. Booij – J. van Marl (eds.), *Yearbook of Morphology* (Dordrecht: Foris).

Boitet Ch. – P. Guillaume – M. Quézel-Ambrunaz 1985: «A case study in software evolution: from ARIANE-78 to ARIANE-85». In *Proceedings of the Conference on*

---

guished from other adjectives, by means of specific features; they show a special behavior as far as their syntactic and semantic use is concerned. According to Αναστασιάδη-Συμεωνίδη who adopts Postal's proposal (1969) they are called *pseudo-adjectives* (p. 148).

Theoretical and Methodological Issues in Machine Translation of Natural Languages (Hamilton, N.Y.: Colgate University).
Di Sciullo A. M. – Williams E. 1987: *On the definition of the word* (MIT Press).
Kay M. 1982: «Machine Translation». *American Journal of Computational Linguistics* 8.
Kittredge R. – P. Isabelle – L. Bourbeau 1976: «Design and implementation of an English-French transfer grammar». In *Proceedings of COLING.*
Lieber R. 1980: *On the organisation of the lexicon* (Ph.D. diss. MIT).
Maas H.-D. 1984: «The MT system SUSY». Paper presented at the ISSCO Tutorial on Machine Translation 1984.
Nirenburg S. (eds.) 1987: *Machine translation* (Cambridge University Press).
Postal P. 1969: «Anaphoric islands». *Papers from the 5th Regional Meeting of the Chicago Linguistic Society* (Chicago).
Ράλλη A. (forthcoming): «Λεξική Φράση: Αντικείμενο μορφολογικού ενδιαφέροντος». In *Proceedings of the 11th meeting of the Department of Linguistics* (Thessaloniki, Greece).
Ralli A. 1988a: *Eléments de la morphologie du grec moderne: la structure du verbe* (Montréal: Ph. D. diss. Université de Montréal).
Ράλλη A. 1988b: «Τα ρηματικά σύνθετα της Νέας Ελληνικής». In *Proceedings of the 9th Meeting of the Department of Linguistics* (Thessaloniki, Greece).
Ralli A. – Galiotou E. 1987: «A morphological processor for Modern Greek». In *Proceedings of the ACL Meeting.* European Chapter (Copenhagen).
Reyle U. – Rohrer C. (eds.) 1988: *Natural language parsing and linguistic theories.* Studies in Linguistics and Philosophy (Kluwer Academic Publishers Group).
Scalise S. 1984: *Generative morphology* (Dordrecht: Foris).
Schmidt P. 1989: «Transfer guidelines». In *Eurotra Reference Manual*, v. 6.0.
Selkirk E. 1982: *The syntax of words* (MIT Press).
Steiner E. 1989: «Predicate-argument structure for transfer». *Eurotra-D Working Papers* 11.
Tucker A. B. 1987: «Current strategies in machine translation research and development». In S. Nirenburg (eds.), *Machine Translation* (Cambridge University Press).
Vauquois B. 1975: *La traduction automatique à Grenoble* (Paris: Dunod).
Villalva A. 1989: «Compounding in Portuguese». *Morphology Workshop, EUROTRA-DK* (Copenhagen 4-5 April 1989).
Williams E. 1981: «On the notions "lexically related" and "head of the word"». *Linguistic Inquiry* 12.