

# Automatic translation of support verb constructions

Morris SALKOFF  
Université Paris-7, L.A.D.L.  
2, Place Jussieu, 75005 PARIS, FRANCE

Verbs like *have, do, take, be*, etc., appear with verb nominalisations in sentences like the following:

- (1)a Max alluded to the crime → b Max made an allusion to the crime
- (2)a Max harmed the child → b Max did harm to the child
- (3)a Max walked along the street → b Max took a walk along the street
- (4)a Jim's story contradicts Max's confession → b Jim's story is in contradiction with Max's confession

M. Gross (1981) calls such verbs 'support verbs', and I shall adopt his terminology. These verbs exhibit many interesting properties which have been studied systematically for several French support verbs: *faire (make), avoir (have), prendre (take), être (be)*, etc. An examination of the results indicates that support verbs must be taken into account in the parser and in the lexicon of a program of automatic translation. I shall discuss the problem in the context of a possible program for French- English translation.

In what follows, *Vsup* designates the support verb, and *Nsup* stands for the 'supported' noun associated with *Vsup*. *N0* and *N2* designate the subject noun and the indirect object noun, respectively. The symbol '+' is used to indicate a choice.

## I. The support verb construction

In order to be able to parse and then translate correctly such sequences as *make an allusion, do harm*, etc., the nominalizations *allusion* and *harm* must be marked in the lexicon as appearing with the support verbs *make* and *do*. There are several reasons for this:

(i) The translation of a simple verb (e.g., *allude*) or of a support verb construction may or may not be an equivalent simple verb or support construction in the target language. For example, (1)a cannot be translated into French with a simple verb:

- (5) Max alluded to the crime = Max (\*alludait + a fait une allusion) au crime

Conversely, there are French support constructions which can more easily be translated into English by a simple verb than by a support construction :

- (6) Max fait une beuverie chaque samedi = Max carouses every Saturday

(ii) A support verb can be zeroed, leaving behind an elliptical construction that cannot be translated as it stands. Consider the following sentences :

- (7) People who are on drugs (are..) = Les gens qui prennent des drogues (sont..)

The English support construction *be on drugs* is related to the sentence *Max drugs himself*. In French, the support verb must be *prendre (take)*. When *who are* is deleted in (7), we obtain:

- (8) People on drugs (are..)

But this sentence can only be translated into French as (7) was, since the French support *qui prennent (who take)* is not deletable. Note that it is not a question here of *on* having a different meaning from its more common locative sense (usually translated as *sur*). Rather, the support verb *be* can be deleted in English, and the elliptical construction (8) has no direct equivalent in French where the support verb is not zeroable.

(iii) More complicated cases of support verb deletion occur. Certain support verb constructions contain a predicate noun that is not the nominalization of a verb; when the support verb is deleted, a problem in translation may arise.

- (9)a L'attention que Max a prêté au problème (est..) = The care that Max gave to the problem (is..)
- (9)b L'attention de Max au problème (est..) = \*Max's care to the problem (is..)
- (10)a L'entorse que Max a fait aux règlements (est..) = The violence that Max did to the rules (is..)
- (10)b L'entorse de Max aux règlements (est..) = \*Max's violence to the rules (is..)

The French support verb constructions *prêter attention* in (9)a and *faire une entorse* in (10)a translate without difficulty into English. However, the support verb can be deleted in French, as in (9)b and (10)b, but not in English. The latter sentences can only be translated by re-introducing the deleted support verb.

To handle all these difficulties, a link must be established in the lexicon between the support verb and all the nouns associated with it (since each support verb can appear typically with many hundreds of nouns). Thus, *beuverie* (in 6) and *entorse* (in 10) must be marked as appearing with *faire*. In English, *drugs* must be marked as *be on*. Then, after parsing, *faire une beuverie* will be a sequence of marked, hence linked, categories and the French-English transfer lexicon furnishes a translation specific to that sequence. Similarly in (8), (9)b and (10)b the mark associated with the nouns *drugs*, *attention* and *entorse* will enable the program to reconstitute the deleted support verb.

Cases like (10)b require more lexical markers than the simpler cases of (1) to (4). The problem is that there are sentences similar to (9)b and (10)b which yet do not require the addition of the deleted support verb in order for a translation to be possible. This is the case in such sentences as the following:

- (11)a L'acquiescement que Max donne au projet (est..) = The consent that Max gives to the project (is..)  
 (11)b L'acquiescement de Max au projet (est..) = Max's consent to the project (is..)

In this case, it suffices to indicate that *acquiescement* can be supported by *donner*, so that (11)a can be parsed correctly; then the transfer lexicon contains the translation of *donner acquiescement à* = *give consent to*. This entry further indicates that *acquiescement*, without *donner*, translates as *consent*. This part of the entry is sufficient for the translation of *acquiescement à* in (11)b.

In (9)b and (10)b, the nouns *attention* and *entorse* are marked for the *Vsup* (*prêter*, *faire*) with which they can appear, just as *acquiescement* is marked for *donner*. However, if the sequence *de N0 à N2* appears after the nominalizations, it has to be transformed into *que N0 Vsup à N2* before it can be translated correctly. Therefore, the lexical entries of *attention* and *entorse* must contain more than just the translation of *faire attention* and *faire entorse*: they must additionally be marked as requiring this transformation of a prepositional phrase into a relative for the purpose of translation.

## II. Parsing support verb constructions

The principal feature of the support verb construction is that the verbal slot in the sentence is occupied by the combination of a verb, *Vsup*, plus a noun, *Nsup*. Thus, if I schematize a sentence as follows:

$$(12) \text{ Sent} \longrightarrow \text{S V O}$$

then the verb slot *V*, when parsing (1), can contain either *allude* or *Vsup Nsup* = *make an allusion*. This means that the verb *V* must have at least two options in its definition: a simple verb, or a support verb construction. This yields the following kinds of analyses:

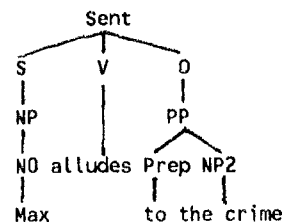


FIG 1

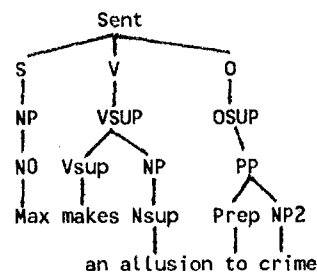


FIG 2

The symbol *VSUP* in Fig. 2 designates the whole support construction. In these two sentences, the object *O* and the support verb object *OSUP* contain the same prepositional phrase *to the crime*. In other cases they may differ, e.g., as in (2), where *O* is an NP (*child*) and *OSUP* is a PP (*to the child*).

In order to obtain just the two analyses of Figs. 1 and 2, many details must be attended to:

(1) The noun *Nsup* associated with *Vsup* in Fig. 2 (*allusion*) must be analyzed as part of a complete noun phrase *NP*, and not just as a bare noun *N*. In effect, left and right modifiers may appear with this noun:

Max makes a brief allusion in Latin to the crime

In this sentence, *allusion* takes an adjective left modifier, and a prepositional phrase as a right modifier.

(2) Since the *NP* appearing after *Vsup* has a slot for the prepositional phrase as a right modifier, the parser will always try to place the sequence

*OSUP* = *to the crime* in that slot, as a modifier of *allusion*. This can be prevented by attaching the following test to the set of right modifiers in the *NP*: a prepositional phrase is not allowed as a right modifier of the noun *Nsup* if the preposition is precisely the one associated with *Vsup*.

(3) Conversely, the parser will try to analyze the sentence of Fig. 2 with the schema of Fig. 1. That is, it tries *Nsup* = *allusion* as a simple direct object of *makes*. This should be disallowed, even if the sentence is such that *OSUP* is empty, e.g., *Max made a strange allusion*. Whenever a *Vsup* appears with an *Nsup* it should be analyzed as a verb support construction, and not as an ordinary verb followed by a direct object.

This is accomplished by adding a test which examines the markers attached to the verb and object in their lexical entries. If the test can verify that the verb carries the mark *Vsup*, as *make* does, and the noun object carries the mark of an *Nsup*, as *allusion* does, then the option 'ordinary verb' of *V* is disallowed. This forces the parser to choose the option *VSUP* for the verb. All these tests and constraints on the parsing can easily be incorporated into the string analyzer of French that I have developed for the IBM 9370 (1973, 1979), and which is currently being rewritten in ADA.

### III. Additional problems

Various difficulties have to be taken care of when a support verb is added to the source sentence for the purposes of a smoother translation.

(i) We have seen in (10)b and (11)b that some modifiers of an *Nsup* of the form *de N0 à N2* require the addition of the deleted support verb if a correct translation is to be obtained. Adding the support verb implies that a tense must be supplied for it. This tense can be derived, but not unambiguously, from the tenses of the verbs in which the support construction is embedded. The problem is that more than one tense may be acceptable, even in the source language. Hence, only an approximate solution can be given.

When the support verb construction is the subject, a tense can be added to the support verb as follows.

a) If the tense of the principal verb is present, future or conditional, then the support verb can take the present tense:

(13) *L'entorse de Max aux règlements (est + sera + serait) nuisible* = The violence Max *does* to the rules (is+ +will be + would be) harmful

b) If the tense of the principal verb is some form of the past, then the support verb takes a past tense:

(14) *L'entorse de Max aux règlements (était + a été + aurait été) nuisible* = The violence Max *did* to the rules (was + has been + would have been) harmful

When the support verb construction is in the direct or indirect object, then either the past or present tense can be added. The French construction is ambiguous:

(15) (*Je remarque + J'avais remarqué*) *l'entorse de Max aux règlements* = (I notice + I had noticed) the violence Max (does + did) to the rules

(ii) Some support constructions, like (11)b, do not require the addition of the deleted support verb. The difference between those that do, as in (13) and (14) above, and those that do not must appear as a mark in the lexical entry of the supported noun.

Thus, the *Nsup entorse (violence)* has a first mark indicating that its *Vsup* is *faire (do)*. A second mark indicates that the following transformation must be effected in order to obtain a correct translation:

(16) *Nsup (entorse) de N0hum à N* → *Nsup (entorse) que N0hum fait à N*

In addition to checking for the presence of this second mark, the program must also verify that the transformation of (16) is applicable, i.e., that the noun appearing after *de* is a 'human' noun that can be an acceptable subject of *faire une entorse*.

(iii) In many support constructions, the noun phrase consists of a frozen expression, i.e., an invariable sequence of words:

(17)a *Max fait grand cas de cela* = Max sets a high store on that

(17)b *Max fait fi de vos conseils* = Max scorns your advice

There are many hundreds, or thousands, of expressions of this kind in which no substitutions can be made, nor any words added.

To parse this type of support construction, the string *VSUP*, schematized in Fig. 2, must contain a second option which is not a noun phrase *NP*, but which consists only of a sequence of fixed words. Then, another test must be added to *VSUP* to ensure that *grand cas* or *fi* is not analyzed as a variable *NP*, i.e., as an occurrence of the first option. A mark can be added to the lexical entry of *cas* and *fi* to indicate invariability. In addition, one distinguished entry must contain the exact sequence of words constituting the

supported expression. For example, the entry for *cas* contains the sequence *grand*, standing for the supported expression *grand cas*.

The passive transformation is possible for some of these support constructions, but its applicability cannot be predicted:

(18)a *Grand cas a été fait de cela* = High store was set on that

(18)b \**Fi a été fait de vos conseils*

Hence, this possibility must also be marked in the lexical entry, so that the absence of an article in the subject *grand cas* can be taken into account by the test for the well-formedness of the subject.

#### IV. Conclusions

Systematic studies of French support verbs have shown that there are many thousands of support constructions of the type discussed here, for each such verb (cf. J. Giry-Schneider, 1987; G. Gross,

1989). The translations of such expressions is irregular, but this is to be expected, and such irregularity will be confined to the lexicon. The regularities of support verb constructions can be handled in the manner discussed here, and this treatment simplifies considerably the parsing module of a program for automatic translation.

#### References

- J. Giry-Schneider, "Les prédicats nominaux en français", Droz, Genève, 1987  
G. Gross, "Les constructions converses du français", Droz, Genève, 1989  
M. Gross, "Les bases empiriques de la notion de prédicat sémantique", in *Langages*, No. 63, Sept. 1981, pp. 7-51, Larousse, Paris  
M. Salkoff, "Une grammaire en chaîne du français", Dunod, Paris, 1973 (out of print)  
M. Salkoff, "Analyse syntaxique du français", J. Benjamins, Amsterdam, 1979