

TRANSLATION BY STRUCTURAL CORRESPONDENCES

Ronald M. Kaplan*
Klaus Netter**
Jürgen Wedekind**
Annie Zaenen*

*Xerox Palo Alto Research Center,
3333 Coyote Hill Road
Palo Alto, CA 94304, USA
Kaplan.pa@xerox.com

** Institut für Maschinelle Sprachverarbeitung,
17 Keplerstraße
D-7000 Stuttgart 1, FRG
Bualis@rus.uni-stuttgart.dbp.de

ABSTRACT

We sketch and illustrate an approach to machine translation that exploits the potential of simultaneous correspondences between separate levels of linguistic representation, as formalized in the LFG notion of codescriptions. The approach is illustrated with examples from English, German and French where the source and the target language sentence show noteworthy differences in linguistic analysis.

INTRODUCTION

In this paper we sketch an approach to machine translation that offers several advantages compared to many of the other strategies currently being pursued. We define the relationship between the linguistic structures of the source and target languages in terms of a set of correspondence functions instead of providing derivational or procedural techniques for converting source into target. This approach permits the mapping between source and target to depend on information from various levels of linguistic abstraction while still preserving the modularity of linguistic components and of source and target grammars and lexicons. Our conceptual framework depends on notions of structure, structural description, and structural correspondence. In the following sections we outline these basic notions and show how they can be used to deal with certain interesting translation problems in a simple and straightforward way. In its emphasis on description-based techniques, our approach shares some fundamental features with the one proposed by Kay (1984), but we use an explicit projection mechanism to separate out

and organize the intra- and inter-language components.

Most existing translation systems are either transfer-based or interlingua-based. Transfer-based systems usually specify a single level of representation or abstraction at which transfer is supposed to take place. A source string is analyzed into a structure at that level of representation, a transfer program then converts this into a target structure at the same level, and the target string is then generated from this structure. Interlingua-based systems on the other hand require that a source string has to be analyzed into a structure that is identical to a structure from which a target string has to be generated.

Without further constraints, each of these approaches could in principle be successful. An interlingual representation could be devised for example, to contain whatever information is needed to make all the appropriate distinctions for all the sentences in all the languages under consideration. Similarly, a transfer structure could be arbitrarily configured to allow for the contrastive analysis of any two particular languages. It seems unlikely that systems based on such an undisciplined arrangement of information will ever succeed in practice. Indeed, most translation researchers have based their systems on representations that have some more general and independent motivation. The levels of traditional linguistic analysis (phonology, morphology, syntax, semantics, discourse, etc.) are attractive because they provide structures with well-defined and coherent properties, but a single one of these

levels does not contain all the information needed for adequate translation. The D-structure level of Government-Binding theory, for example, contains information about the predicate-argument relations of a clause but says nothing about the surface constituent order that is necessary to accurately distinguish between old and new information or topic and comment. As another example, the functional structures of Lexical-Functional Grammar do not contain the ordering information necessary to determine the scope of quantifiers or other operators.

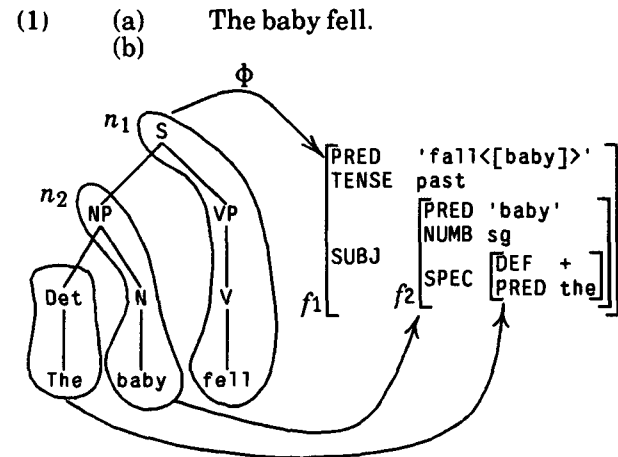
Our proposal, as it is set forth below, allows us to state simultaneous correspondences between several levels of source-target representations, and thus is neither interlingual nor transfer-based. We can achieve modularity of linguistic specifications, by not requiring conceptually different kinds of linguistic information to be combined into a single structure. Yet that diverse information is still accessible to determine the set of target strings that adequately translate a source string. We also achieve modularity of a more basic sort: our correspondence mechanism permits contrastive transfer rules that depend on but do not duplicate the specifications of independently motivated grammars of the source and target languages (Isabelle and Macklovitch, 1986; Netter and Wedekind, 1986).

A GENERAL ARCHITECTURE FOR LINGUISTIC DESCRIPTIONS

Our approach uses the equality- and description-based mechanisms of Lexical-Functional Grammar. As introduced by Kaplan and Bresnan (1982), lexical-functional grammar assigns to every sentence two levels of syntactic representation, a constituent structure (c-structure) and a functional structure (f-structure). These structures are of different formal types—the c-structure is a phrase-structure tree while the f-structure is a hierarchical finite function—and they characterize different aspects of the information carried by the sentence. The c-structure represents the ordered arrangement of words and phrases in the sentence while the f-structure explicitly marks its grammatical functions (subject, object, etc.). For each type of structure there is a special notation or description-language in which the properties of desirable instances of

that type can be specified. Constituent structures are described by standard context-free rule notation (augmented with a variety of abbreviatory devices that do not change its generative power), while f-structures are described by Boolean combinations of function-argument equalities stated over variables that denote the structures of interest. Kaplan and Bresnan assumed a correspondence function mapping between the nodes in the c-structure of a sentence and the units of its f-structure, and used that piecewise function to produce a description of the f-structure (in its equational language) by virtue of the mother-daughter, order, and category relations of the c-structure.

The formal picture developed by Kaplan and Bresnan, as clarified in Kaplan (1987), is illustrated in the following structures for sentence (1):



The c-structure appears on the left, the f-structure on the right. The c-structure-to-f-structure correspondence, ϕ , is shown by the linking lines. The correspondence ϕ is a many-to-one function taking the S, VP and V nodes all into the same outermost unit of the f-structure, f_1 .

The node-configuration at the top of the tree satisfies the statement $S \rightarrow NP VP$ in the context-free description language for the c-structure. As suggested by Kaplan (1987), this is a simple way of defining a collection of more specific properties of the tree, such as the fact that the S node (labeled n_1) is the mother of the NP node (n_2). These facts could also be written in equational form as $M(n_2) = n_1$,

where M denotes the function that takes a tree-node into its mother. Similarly, the outermost f-structure satisfies the assertions $(f_1 \text{ TENSE}) = \text{past}$, $(f_1 \text{ SUBJ}) = f_2$, and $(f_2 \text{ NUMB}) = \text{sg}$ in the f-structure description language. Given the illustrated correspondence, we also know that $f_1 = \phi(n_1)$ and $f_2 = \phi(n_2)$. Taking all these propositions together, we can infer first that $(\phi(n_1) \text{ SUBJ}) = \phi(n_2)$ and then that $(\phi(M(n_2)) \text{ SUBJ}) = \phi(n_2)$. This equation identifies the subject in the f-structure in terms of the mother-daughter relation in the tree.

In LFG the f-structure assigned to a sentence is the smallest one that satisfies the conjunction of equations in its functional description. The functional description is determined from the trees that the c-structure grammar provides for the string by a simple matching process. A given tree is analyzed with respect to the c-structure rules to identify particular nodes of interest. Equations about the f-structure corresponding to those nodes (via ϕ) are then derived by substituting those nodes into equation-patterns or schemata. Thus, still following Kaplan (1987), if $*$ appears in a schema to stand for the node matching a given rule-category, the functional description will include an equation containing that node (or an expression such as n_2 that designates it) instead of $*$. The equation $(\phi(M(n_2)) \text{ SUBJ}) = \phi(n_2)$ that we inferred above also results from instantiating the schema $(\phi(M(*)) \text{ SUBJ}) = \phi(*)$ annotated to the NP element of the S rule in (2a) when that rule-element is matched against the tree in (1b). Kaplan observes that the \uparrow and \downarrow metavariables in the Kaplan/Bresnan formulation of LFG are simply convenient abbreviations for the complex expressions $\phi(M(*))$ and $\phi(*)$, respectively, thus explicating the traditional, more palatable formulation in (2b).

$$(2) \text{ (a) } S \rightarrow \begin{array}{cc} \text{NP} & \text{VP} \\ \phi(M(*)) \text{ SUBJ} = \phi(*) & \phi(M(*)) = \phi(*) \end{array}$$

$$\text{ (b) } S \rightarrow \begin{array}{cc} \text{NP} & \text{VP} \\ (\uparrow \text{ SUBJ}) = \downarrow & \uparrow = \downarrow \end{array}$$

This basic conception of descriptions and correspondences has been extended in several ways. First, this framework has been generalized to additional kinds of structures that represent other subsystems of linguistic

information (Kaplan, 1987; Halvorsen, 1988). These structures can be related by new correspondences that permit appropriate descriptions of more abstract structures to be produced. Halvorsen and Kaplan (1988), for example, discuss a level of semantic structure that encodes predicate-argument relations and quantifier scope, information that does not enter into the kinds of syntactic generalizations that the f-structure supports. They point out how the semantic structure can be set in correspondence with both c-structure and f-structure units by means of related mappings σ and σ' . Kaplan (1987) raises the possibility of further distinct structures and correspondences to represent anaphoric dependencies, discourse properties of sentences, and other *projections* of the same string.

Second, Kaplan (1988) and Halvorsen and Kaplan (1988) discuss other methods for deriving the descriptions necessary to determine these abstract structures. The arrangement outlined above, in which the description of one kind of structure (the f-structure) is derived by analyzing or matching against another one, is an example of what is called *description-by-analysis*. The semantic interpretation mechanisms proposed by Halvorsen (1983) and Reyle (1988) are other examples of this descriptive technique. In this method the grammar provides general patterns to compare against a given structure and these are then instantiated if the analysis is satisfactory. One consequence of this approach is that the structure in the range of the correspondence, the one whose description is being developed, can only have properties that are derived from information explicitly identified in the domain structure.

Another description mechanism is possible when three or more structures are related through correspondences. Suppose the c-structure and f-structure are related by ϕ as in (2a) and that the function σ then maps the f-structure units into corresponding units of semantic structure of the sort suggested by Fenstad et al. (1987). The formal arrangement is shown in Figure 1 (next page). This configuration of cascaded correspondences opens up a new descriptive possibility. If σ and ϕ are both structural correspondences, then so is their composition $\sigma \circ \phi$. Thus, even though the units of the semantic structure correspond directly only to the units of the f-structure and have no immediate connection to the nodes of

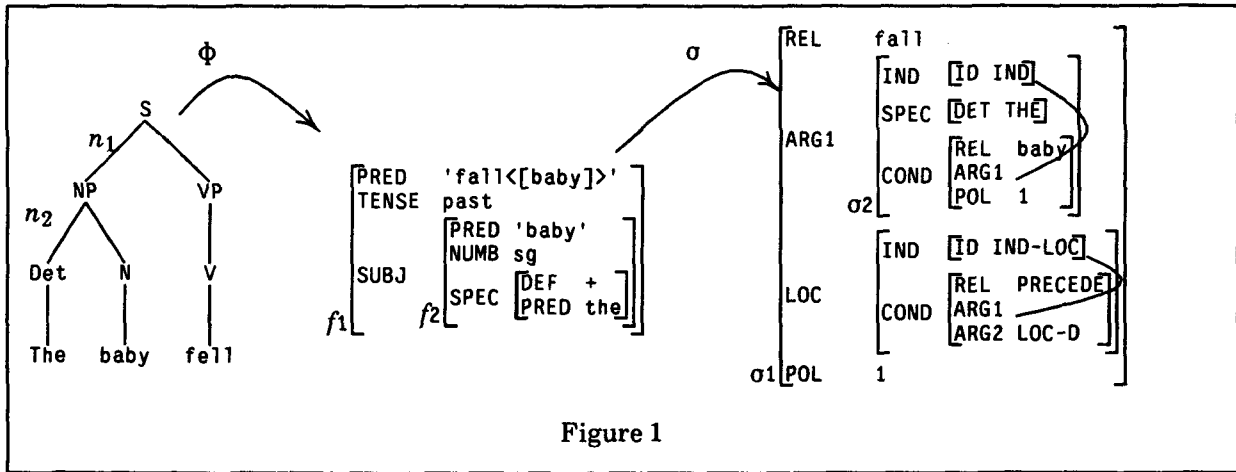


Figure 1

the c-structure, a semantic description can be formulated in terms of c-structure relations. The expression $\sigma(\phi(M^*))$ can appear on a c-structure rule-element to designate the semantic-structure unit corresponding to the f-structure that corresponds to the mother of the node that matches that rule-element. Since projections are monadic functions, we can remove the uninformative parentheses and write $(\sigma\phi M^* \text{ ARG1}) = \sigma(\phi M^* \text{ SUBJ})$, or, using the \uparrow metavariable, $(\sigma \uparrow \text{ ARG1}) = \sigma(\uparrow \text{ SUBJ})$. Schemata such as this can be freely mixed with LFG's standard functional specifications in lexical entries and c-structure rules. For example, the lexical entry for *fall* might be given as follows:

- (3) *fall* V $(\uparrow \text{ PRED}) = \text{'fall'}$
 $(\sigma \uparrow \text{ REL}) = \text{fall}$
 $(\sigma \uparrow \text{ ARG1}) = \sigma(\uparrow \text{ SUBJ})$

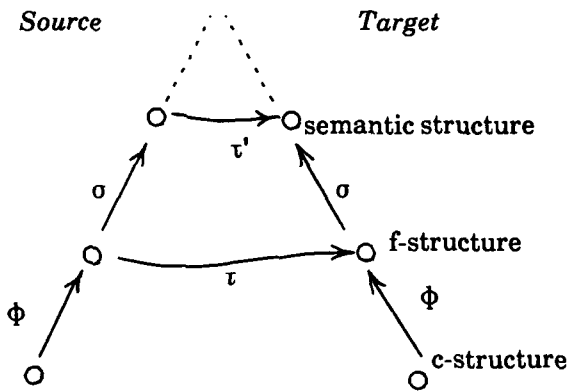
Descriptions formulated by composing separate correspondences have a surprising characteristic: they allow the final range structure (e.g. the semantic structure) to have properties that cannot be inferred from any information present in the intermediate (f-) structure. But those properties can obtain only if the intermediate structure is derived from an initial (c-) structure with certain features. For example, Kaplan and Maxwell (1988a) exploit this capability to describe semantic structures for coordinate constructions which necessarily contain the logical conjunction appropriate to the string even though there is no reasonable place for that conjunction to be marked in the f-structure. In sum, this method of description, which has been called *codescription*, permits information from a variety of different levels to

constrain a particular structure, even though there are no direct correspondences linking them together. It provides for modularity of basic relationships while allowing certain necessary restrictions to have their influence.

The descriptive architecture of LFG as extended by Kaplan and Halvorsen provides for multiple levels of structure to be related by separate correspondences, and these correspondences allow descriptions of the various structures to be constructed, either by analysis or composition, from the properties of other structures. Earlier researchers have applied these mechanisms to the linguistic structures for sentences in a single language. In this paper, we extend this system one step further: we introduce correspondences between structures for sentences in different languages that stand in a translation relation to one another. The description of the target language structures are derived via analysis and codescription from the source language structures, by virtue of additional annotations in c-structure rules and lexical entries. Those descriptions are solved to find satisfying solutions, and these solutions are then the input to the target generation process.

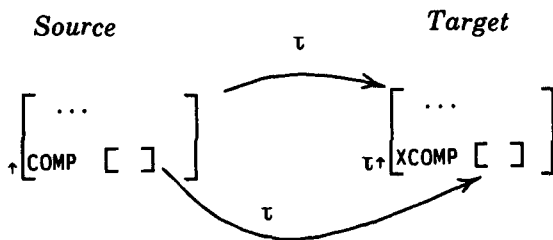
In the two language arrangement sketched below, we introduce the τ correspondence to map between the f-structure units of the source language and the f-structure units of the target language. The σ correspondence maps from the f-structure of each language to its own corresponding semantic structure, and a second transfer correspondence τ' relates those structures.

(4)



This arrangement allows us to describe the target f-structure by composing ϕ and τ to form expressions such as $\tau(\phi M^* \text{COMP}) = (\tau \phi M^* \text{XCOMP})$ or simply $\tau(\uparrow \text{COMP}) = (\tau \uparrow \text{XCOMP})$. This maps a COMP in the source f-structure into an XCOMP in the target f-structure. The relations asserted by this equation are depicted in the following source-target diagram:

(5)



As another example, the equation $\tau(\sigma \uparrow \text{ARG1}) = (\sigma \uparrow \text{ARG1})$ identifies the first arguments in the source and target semantic structures. The equation $\tau\sigma(\uparrow \text{SUBJ}) = \sigma(\tau \uparrow \text{TOPIC})$ imposes the constraint that the semantics of the source SUBJ will translate via τ' into the semantics of the target TOPIC but gives no further information about what those semantic structures actually contain.

Our general correspondence architecture thus applies naturally to the problem of translation. But there are constraints on correspondences specific to translation that this general architecture does not address. For instance, the description of the target-language structures derived from the source-language is incomplete. The target structures may and usually will have

grammatical and semantic features that are not determined by the source. It makes little sense, for example, to include information about grammatical gender in the transfer process if this feature is exhaustively determined by the grammar of the target language. We can formalize the relation between the information contained in the transfer component and an adequate translation of the source sentence into a target sentence as follows: for a target sentence to be an adequate translation of a given source sentence, it must be the case that a minimal structure assigned to that sentence by the target grammar is subsumed by a minimal solution to the transfer description. One desirable consequence of this formalization is that it permits two distinct target strings for a source string whose meaning in the absence of other information is vague but not ambiguous.

Thus this conceptual and notational framework provides a powerful and flexible system for imposing constraints on the form of a target sentence by relating them to information that appears at different levels of source-language abstraction. This apparatus allows us to avoid many of the problems encountered by more derivational, transformational or procedural models of transfer. We will illustrate our proposal with examples that have posed challenges for some other approaches.

EXAMPLES

Changes in grammatical function. Some quite trivial changes in structure occur when the source and the target predicate differ in the grammatical functions that they subcategorize for. We will illustrate this with an example in which a German transitive verb is translated with an intransitive verb taking an oblique complement in French:

- (6) (a) Der Student beantwortet die Frage.
 (b) L'étudiant répond à la question.

We treat the oblique preposition as a PRED that itself takes an object. Ignoring information about tense, the lexical entry for *beantworten* in the German lexicon looks as follows:

- (7) *beantworten* V
 (\uparrow PRED) = 'beantworten < (\uparrow SUBJ) (\uparrow OBJ) >'

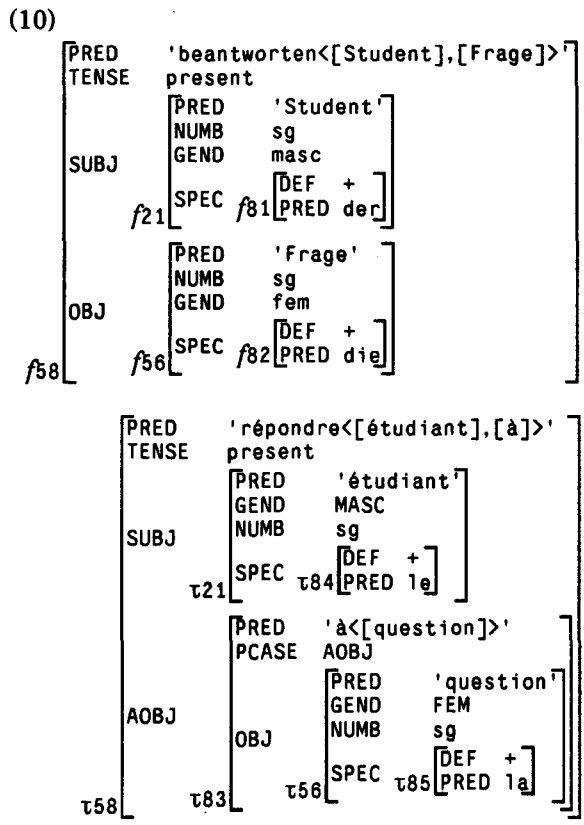
while the transfer lexicon for *beantworten* contains the following mapping specifications:

- (8) $(\tau \uparrow \text{PRED FN}) = \text{répondre}$
 $(\tau \uparrow \text{SUBJ}) = \tau(\uparrow \text{SUBJ})$
 $(\tau \uparrow \text{AOBJ OBJ}) = \tau(\uparrow \text{OBJ})$

We use the special attribute FN to designate the function-name in semantic forms such as 'beantworten<(\uparrow SUBJ)(\uparrow OBJ)>'. In this transfer equation it identifies *répondre* as the corresponding French predicate. This specification controls lexical selection in the target, for example, selecting the following French lexical entry to be used in the translation:

- (9) *répondre* V
 $(\uparrow \text{PRED}) = \text{'répondre}<(\uparrow \text{SUBJ})(\uparrow \text{AOBJ})>$

With these entries and the appropriate but trivial entries for *der Student* and *die Frage* we get the following f-structure in the source language and associated f-structure in the target language for the sentence in (10):



The second structure is the f-structure the grammar of French assigns to the sentence in (6b). This f-structure is the input for the generation process. Other examples of this kind are pairs like *like* and *plaire* and *help* and *helfen*.

In the previous example the effects of the change in grammatical function between the source and the target language are purely local. In other cases there is a non-local dependency between the subcategorizing verb and a dislocated phrase. This is illustrated by the relative clause in (11):

- (11) (a) ...der Brief, den der Student zu beantworten scheint.
 (b) ...la lettre, à laquelle l'étudiant semble répondre.
 ...the letter, that the student seemed to answer.

The within-clause functions of the relativized phrases in the source and target language are determined by predicates which may be arbitrarily deeply embedded, but the relativized phrase in the target language must correspond to the one in the source language.

Let us assume that relative clauses can be analyzed by the following slightly simplified phrase structure rules, making use of functional uncertainty (see Kaplan and Maxwell 1988b for a technical discussion of functional uncertainty) to capture the non-local dependency of the relativized phrase (equations on the head NP are ignored):

(12) NP → NP S'
 $(\uparrow \text{RELADJ}) = \downarrow$

S' → XP S
 $(\uparrow \text{REL-TOPIC}) = \downarrow$ $\uparrow = \downarrow$
 $(\uparrow \text{XCOMP* GF}) = \downarrow$

We can achieve the desired correspondence between the source and the target by augmenting the first rule with the following transfer equations:

(13) NP → NP S'
 $(\uparrow \text{RELADJ}) = \downarrow$
 $\tau(\uparrow \text{RELADJ}) = (\tau \uparrow \text{RELADJ})$
 $\tau(\downarrow \text{REL-TOPIC}) = (\tau \downarrow \text{REL-TOPIC})$

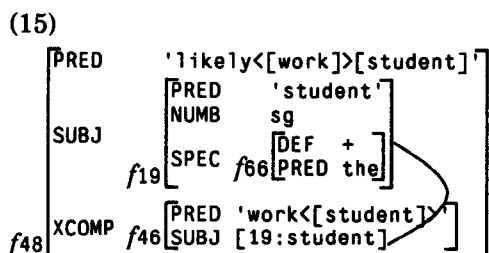
The effect of this rule is that the τ value of the relativized phrase (REL-TOPIC) in the source language is identified with the relativized phrase in the target language. However, the source REL-TOPIC is also identified with a within-clause function, say OBJ, by the uncertainty equation in (12). Lexical transfer rules such as the one given in (8) independently establish the correspondence

between source and target within-clause functions. Thus, the target within-clause function will be identified with the target relativized phrase. This necessary relation is accomplished by lexically and structurally based transfer rules that do not make reference to each other.

Differences in control. A slightly more complex but similar case arises when the infinitival complement of a raising verb is translated into a finite clause, as in the following:

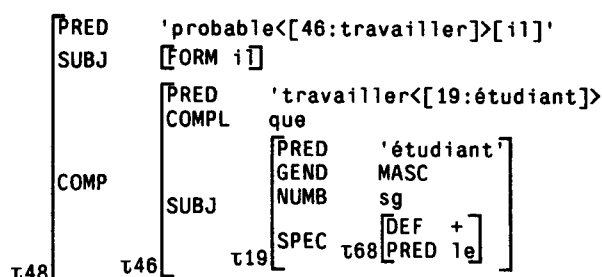
- (14) (a) The student is likely to work.
 (b) Il est probable que l'étudiant travaillera.

In this case the necessary information is distributed in the following way over the source, target, and transfer lexicons as shown in Figure 2. Here the transfer projection builds up an underspecified target structure, to which the information given in the entry of *probable* is added in the process of generation. Ignoring the contribution of *is*, the f-structure for the English sentence identifies the non-thematic SUBJ of *likely* with the thematic SUBJ of *work* as follows:



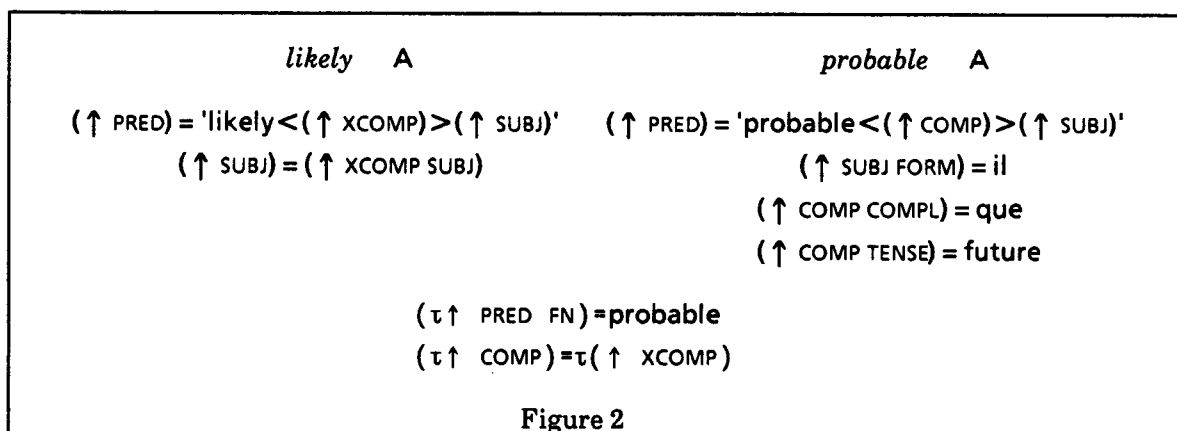
The corresponding French structure in (16) contains an expletive SUBJ, *il*, for *probable* and an overtly expressed SUBJ for *travailler*. The latter is introduced by the transfer entry for *work*:

(16)



Again this f-structure satisfies the transfer description and is also assigned by the French grammar to the target sentence.

The use of multiple projections. There is one detail about the example in (14) that needs further discussion. Simplifying matters somewhat, there is a requirement that the temporal reference point of the complement has to follow the temporal reference point of the clause containing *likely*, if the embedded verb is a process verb. Basically the same temporal relations have to hold in French with *probable*. The way this is realized will depend on what the tense of *probable* is, which in turn is determined by the discourse up to that point. A sentence similar to the one given in (13a) but appearing in a narrative in the past would translate as the following:



(17) Il était probable que l'étudiant travaillerait.

In the general case the choice of a French tense does not depend on the tense of the English sentence alone but is also determined by information that is not part of the f-structure itself. We postulate another projection, the temporal structure, reached from the f-structure through the correspondence χ (from $\chi\rho\nu\nu\iota\kappa\omicron\varsigma$, temporal). It is not possible to discuss here the specific characteristics of such a structure. The only thing that we want to express is the constraint that the event in the embedded clause follows the event in the main clause. We assume that the temporal structure contains the following information for *likely-to-V*, as suggested by Fenstad et al. (1987):

- (18) *likely* V
 $(\chi \uparrow \text{ COND REL}) = \text{precede}$
 $(\chi \uparrow \text{ COND ARG1}) = (\chi \uparrow \text{ IND})$
 $(\chi \uparrow \text{ COND ARG2 ID}) = \text{IND-LOC2}$

This is meant to indicate that the temporal reference point of the event denoted by the embedded verb extends after the temporal reference point of the main event. The time of the main event is in part determined by the tense of the verb *be*, which we ignore here. The only point we want to make is that aspects of these different projections can be specified in different parts of the grammar. We assume that French and English have the same temporal structure but that in the context of

likely it is realized in a different way. This can be expressed by the following equation:

$$(19) \quad \chi \uparrow = \chi\tau \uparrow$$

Here the identity between χ and $\chi\tau$ provides an interlingua-like approach to this particular subpart of the relation between the two languages. This is diagrammed in Figure 3. Allowing these different projections to simultaneously determine the surface structure seems at first blush to complicate the computational problem of generation, but a moment of reflection will show that that is not necessarily so. Although we have split up the different equations among several projections for conceptual clarity, computationally we can consider them to define one big attribute value structure with χ and τ as special attributes, so the generation problem in this framework reduces to the problem of generating from attribute-value structures which are formally of the same type as f-structures (see Halvorsen and Kaplan (1988), Wedekind (1988), and Momma and Dörre (1987) for discussion).

Differences in embedding. The potential of the system can also be illustrated with a case in which we find one more level of embedding in one language than we find in the other. This is generally the case if a modifier-head relation in the source language is reversed in the target structure. One such example is the relation between the sentences in (20):

- (20) (a) The baby just fell.
 (b) Le bébé vient de tomber.

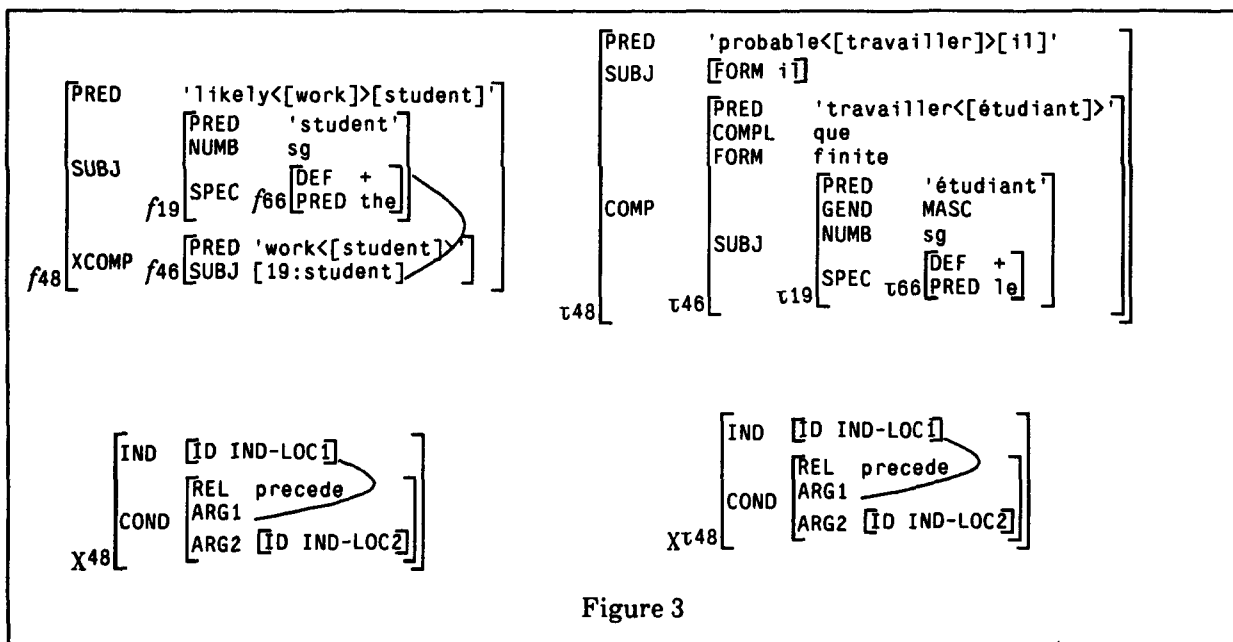


Figure 3

One way to encode this relation is given in the following lexical entry for *just* (remember that all the information about the structure of *venir* in French will come from the lexicon and grammar of French itself):

(21) *just* ADV (\uparrow PRED)='just<(\uparrow ARG)>'
 ($\tau \uparrow$ PRED FN)=venir
 ($\tau \uparrow$ XCOMP)= $\tau(\uparrow$ ARG)

This assigns to *just* a semantic form that takes an ARG function as its argument and maps it into the French *venir*. This lexical entry is combined with phrase-structure rule (22). This rule introduces sentence adverbs and makes the f-structure corresponding to the S node fill the ARG function in the f-structure corresponding to the ADV node.

(22) S \rightarrow NP (ADV) VP
 (\uparrow SUBJ) = \downarrow \uparrow = (\downarrow ARG)

Note that the f-structure of the ADV is not assigned a function within the S-node's f-structure, which is shown in (23). This is in keeping with the fact that the adverb has no functional interactions with the material in the main clause.

(23)
$$\left[\begin{array}{l} \text{PRED} \\ \text{TENSE} \\ \text{SUBJ} \end{array} \right] \left[\begin{array}{l} \text{'fall<[baby]>'} \\ \text{past} \\ \left[\begin{array}{l} \text{PRED} \\ \text{NUMB} \\ \text{SPEC} \end{array} \right] \left[\begin{array}{l} \text{'baby'} \\ \text{sg} \\ \left[\begin{array}{l} \text{DEF} \\ \text{PRED the} \end{array} \right] \end{array} \right] \end{array} \right]$$

The relation between the adverb and the clause is instead represented only in the f-structure associated with the ADV node:

(24)
$$\left[\begin{array}{l} \text{PRED} \\ \text{ARG} \end{array} \right] \left[\begin{array}{l} \text{'just<[fall]>'} \\ \left[\begin{array}{l} \text{PRED} \\ \text{TENSE} \\ \text{SUBJ} \end{array} \right] \left[\begin{array}{l} \text{'fall<[baby]>'} \\ \text{past} \\ \left[\begin{array}{l} \text{PRED} \\ \text{NUMB} \\ \text{SPEC} \end{array} \right] \left[\begin{array}{l} \text{'baby'} \\ \text{sg} \\ \left[\begin{array}{l} \text{DEF} \\ \text{PRED the} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$$

In the original formulation of LFG, the f-structure of the highest node was singled out and assigned a special status. In our current theory we do not distinguish that structure from all the others in the range of ϕ : the grammatical analysis of a sentence includes the complete enumeration of ϕ -associations. The S-node's f-structure typically does contain the f-structures of all other nodes as subsidiary elements, but not in this adverbial case. The

target structures corresponding to the various f-structures are also not required to be integrated. These target f-structures can then be set in correspondence with any nodes of the target c-structure, subject to the constraints imposed by the target grammar. In this case, the fact that *venir* takes an XCOMP which corresponds to the ARG of *just* means that the target f-structure mapped from the ADV's f-structure will be associated with the highest node of the target c-structure. This is shown in (25).

(25)
$$\left[\begin{array}{l} \text{PRED} \\ \text{SUBJ} \\ \text{XCOMP} \end{array} \right] \left[\begin{array}{l} \text{'venir<[tomber]>[bébé]'} \\ \left[\begin{array}{l} \text{PRED} \\ \text{GEN} \\ \text{NUMB} \\ \text{SPEC} \end{array} \right] \left[\begin{array}{l} \text{'bébé'} \\ \text{MASC} \\ \text{sg} \\ \left[\begin{array}{l} \text{DEF} \\ \text{PRED the} \end{array} \right] \end{array} \right] \\ \left[\begin{array}{l} \text{PRED} \\ \text{COMPL} \\ \text{TENSE} \\ \text{SUBJ} \end{array} \right] \left[\begin{array}{l} \text{'tomber<[bébé]>'} \\ \text{de} \\ \text{inf} \\ \left[\begin{array}{l} \text{SUBJ} \\ \text{[14:bébé]} \end{array} \right] \end{array} \right] \end{array} \right]$$

The above analysis does not require a single integrated source structure to map onto a single integrated target structure. An alternative analysis can handle differences of embedding with completely integrated structures. If we assign an explicit function to the adverbial in the source sentence, we can reverse the embedding in the target by replacing (22) with (26):

(26) S \rightarrow NP (ADV) VP
 (\uparrow SADJ) = \downarrow
 $\tau \uparrow$ = ($\tau \downarrow$ XCOMP)

In this case the embedded f-structure of the source adverb will be mapped onto the f-structure that corresponds to the root node of the target c-structure, whereas the f-structure of the source S is mapped onto the embedded XCOMP in the target. The advantages and disadvantages of these different approaches will be investigated further in Netter and Wedekind (forthcoming).

CONCLUSION

We have sketched and illustrated an approach to machine translation that exploits the potential of simultaneous correspondences between different levels of linguistic representation. This is made possible by the equality and description based mechanisms of LFG. This approach relies mainly on codescription, and thus it is different from other LFG-based approaches that use a

description-by-analysis mechanism to relate the f-structure of a source language to the f-structure of a target language (see for example Kudo and Nomura, 1986). Our proposal allows for partial specifications and multi-level transfer. In that sense it also differs from strategies pursued for example in the Eurotra project (Arnold and des Tombe, 1987), where transfer is based on one level of representation obtained by transforming the surface structure in successive steps.

We see it as one of the main advantages of our approach that it allows us to express correspondences between separate pieces of linguistically motivated representations and in this way allows the translator to exploit the linguistic descriptions of source and target language in a more direct way than is usually proposed.

ACKNOWLEDGEMENTS

Thanks to P.-K. Halvorsen, U. Heid, H. Kamp, M. Kay and C. Rohrer for discussion and comments.

REFERENCES

- Arnold, Douglas and Louis des Tombe (1987) Basic theory and methodology in Eurotra. In S. Nirenburg (ed.), *Machine translation: Theoretical and methodological issues*. Cambridge: Cambridge University Press, 114-135.
- Fenstad, Jens Erik, Per-Kristian Halvorsen, Tore Langholm, and Johan van Benthem (1987) *Situations, Language and Logic*. Dordrecht: D. Reidel.
- Halvorsen, Per-Kristian (1983) Semantics for lexical-functional grammars. *Linguistic Inquiry*, 14 (3), 567-613.
- Halvorsen, Per-Kristian (1988) Situation semantics and semantic interpretation in constraint-based grammars. *Proceedings of the International Conference on Fifth Generation Computer Systems*, Tokyo, Japan, 471-478.
- Halvorsen, Per-Kristian and Ronald Kaplan (1988) Projections and semantic description. *Proceedings of the International Conference on Fifth Generation Computer Systems*, Tokyo, Japan, 1116-1122.
- Isabelle, Pierre and Elliott Macklovitch (1986) Transfer and MT modularity. *Proceedings of Coling 1986*, Bonn, 115-117.
- Kaplan, Ronald (1987) Three seductions of Computational Psycholinguistics. In Peter Whitelock et al. (eds.), *Linguistic Theory and Computer Applications*. Academic Press, London, 149-188.
- Kaplan, Ronald (1988) Correspondences and their inverses. Paper presented at the Syntax and Semantics Workshop, April, Titisee, FRG.
- Kaplan, Ronald and Joan Bresnan (1982) Lexical Functional Grammar: a formal system for Grammatical representation. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, Mass, 173-281.
- Kaplan, Ronald and John Maxwell (1988a). Constituent coordination in Lexical-Functional Grammar. *Proceedings of COLING 88*, Budapest, 303-305.
- Kaplan, Ronald and John Maxwell (1988b). An algorithm for Functional Uncertainty. *Proceedings of COLING 88*, Budapest, 297-302.
- Kay, Martin (1984) Functional Unification Grammar: A formalism for Machine Translation. *Proceedings of Coling 1984*, Stanford University, 75-78.
- Kudo, Ikuo and Hirosato Nomura (1986) Lexical-Functional Transfer: A Transfer Framework in a Machine Translation System based on LFG. *Proceedings of Coling 1986*, Bonn, 112-114.
- Momma, Stefan and Jochen Dörre (1987) Generation from f-structures. In Ewan Klein and Johan van Benthem (eds.) *Categories, Polymorphism and Unification*, Edinburgh, Amsterdam, 147-168.
- Netter, Klaus and Jürgen Wedekind (1986) An LFG-based Approach to Machine Translation. *Proceedings of IAI-MT 86*, Saarbrücken, 197-209.
- Netter, Klaus and Jürgen Wedekind (in prep.) Transfer by projection. IMS, Stuttgart.
- Reyle, Uwe (1988) Compositional semantics for LFG. In Uwe Reyle and Christian Rohrer (eds.), *Natural language parsing and linguistic theories*. Dordrecht: D. Reidel, 448-479.
- Wedekind, Jürgen (1988) Generation as Structure Driven Derivation. *Proceedings of COLING 88*, Budapest, 732-737.