Restriction and Correspondence-based Translation

Ronald M. Kaplan
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, California 94304 USA
Kaplan.Parc@Xerox.Com

Jürgen Wedekind
Institute for Computational Linguistics
University of Stuttgart
Azenbergstr. 12
D-7000 Stuttgart 1, FRG
Juergen@ims.uni-stuttgart.de

Abstract: Kaplan *et al.* (1989) present a framework for translation based on the description and correspondence concepts of Lexical-Functional Grammar (Kaplan and Bresnan, 1982). Certain phenomena, in particular the head-switching of adverbs and verbs, seem to be problematic for that approach. In this paper we suggest that these difficulties are more properly considered as the result of defective monolingual analyses. We propose a new description-language operator, restriction, to permit a succinct formal encoding of the informal intuition that semantic units sometimes correspond to subsets of functional information. This operator, in conjunction with an additional recursion provided by a description-by-analysis rule, is the basis of a more adequate account of head-switching that preserves the advantages of correspondence-based translation.

1. Introduction

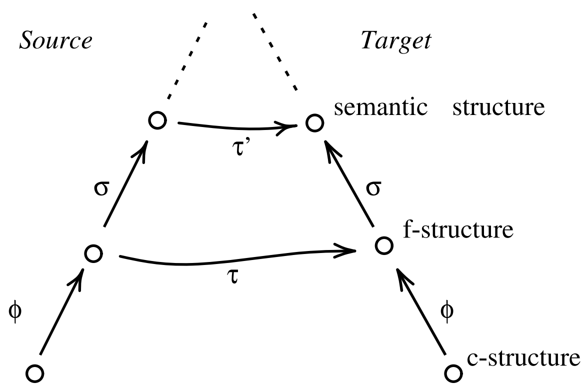
Kaplan *et al.* (1989) present a framework for translation based on the description and correspondence concepts of Lexical-Functional Grammar (Kaplan and Bresnan, 1982). LFG formulates the syntactic dependencies and generalizations of natural languages in terms of the properties of formal structures of different types: ordinary phrase-structure trees represent the surface constituency of sentences while hierarchical finite functions represent their underlying grammatical relations. The structures for a particular sentence are those that satisfy descriptions produced from annotated phrase-structure rules and lexical entries. The description of the more abstract functional structure is determined by the dominance and precedence relations of the superficial constituent structure, given the assumption that there is a piecewise correspondence function that maps the nodes in the c-structure tree into the units of the f-structure. Kaplan (1987) and Halvorsen and Kaplan (1988) extend this structure/description/correspondence architecture to

This paper appeared in the Proceedings of the Sixth European Conference of the Association for Computational Linguistics, Utrecht, 1993, pages 193–202.

provide modular and declarative characterizations of the relationships between syntactic structures and other levels of linguistic representation. Kaplan *et al*. (1989) suggest that this architecture can provide a formal basis for specifying complex source-target translation relationships in a declarative fashion that builds on monolingual grammars and lexicons that are independently motivated and theoretically justified. In particular, the approach permits features from different linguistic levels to affect translation without requiring that reflexes of those disparate features appear together in an otherwise unmotivated transfer or interlingual representation.

Kaplan *et al*. (1989) offer several examples to illustrate the effectiveness of this approach to translation. These examples involve changes in grammatical functions from source to target, differences in control, and differences in embedding (or head-switching). The Kaplan *et al*. solutions depend on monolingual representations of phrasal, functional, and semantic information related by the correspondences ϕ and σ, with translation correspondences τ and τ' mapping source to target structures, as shown in the configuration in (1):

(1)



These solutions utilize the formal device of *codescription* to specify the target structure constraints in terms of simple compositions of the τ and τ' mappings with the monolingual source correspondences. For instance, the fact that the object of the German *beantworten* corresponds to the AOBJ of the French *répondre* is indicated by associating the following transfer equations with the normal monolingual lexical entry for *beantworten*:

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

The last line asserts that the AOBJ in the target f-structure is the translation of the source OBJ. The metavariable $\uparrow$ in LFG is an abbreviation for $\phi(M(*))$ and thus denotes the f-structure that corresponds to the mother of the *beantworten* lexical node in the German c-structure (indicated by *). The expression $\tau\uparrow$ can then be seen as $\tau(\phi(M(*)))= \tau \circ \phi(M(*))$ and thus as incorporating the composition $\tau \circ \phi$. Significantly, the $M(*)$ term is also present, which means that the target constraints are determined in the same recursive analysis of the source c-structure that is used to derive the source f-structure description. The codescription device crucially involves both a composition of correspondences and a single recursive analysis of common ancestor structures. This contrasts with *description-by-analysis*, another technique mentioned by Halvorsen and Kaplan (1988) and Kaplan *et al.* (1989) for deriving descriptions of abstract structures.

2. Difficulties with the correspondence approach

This proposal for correspondence-based translation has been scrutinized by a number of researchers (e.g. Sadler *et al*., 1989, Sadler *et al* ., 1990, Sadler and Thompson, 1991), and several difficulties have been pointed out. These difficulties arise particularly in cases where the independently motivated source and target structures are not very closely aligned, where single units in a source structure map to multiple units in the target (so-called splitting) or where hierarchical relationships are interchanged in mapping from source to target (switching). If such discrepancies are both locally bounded and predictable, then they can in principle be handled by means of codescription statements involving the ordinary monolingual description-language constructs of function-application and equality. But even if possible, such conservative treatments may not permit obvious generalizations about the translation relation to be naturally expressed. Sadler *et al*. (1990) demonstrate this point by examples in which the translation of a lexical head differs according to its dependents in the source sentence (English *commit suicide* translates to French *(se) suicider* whereas *commit a crime* translates to *commettre une crime*). They suggest refining the basic correspondence approach by separating such idiosyncratic source-target interactions into a separate transfer lexicon whose stipulations will override (perhaps via the priority union operator) the generic transfer specifications that might still be associated with the source-language predicates.

Sadler *et al*. (1989) and Sadler and Thompson (1991) focus on another case of structural misalignment in translation, as illustrated in (3):

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

The syntactic head of the English sentence (*fell*) corresponds to the head of the French embedded complement (*tomber*), while the English adjunct *just* corresponds to the head of the French matrix. Other well-known contrasts show syntactic embeddings in English corresponding to sentential adjuncts in Dutch (and German):

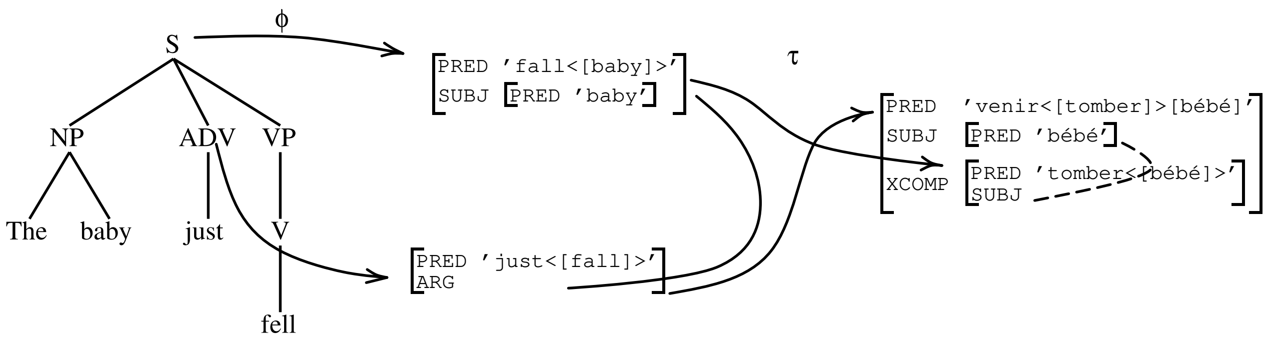
(4) (a) John likes to swim.
    (b) Jan zwemt graag.

Kaplan *et al*. (1989) discussed such differences in embedding and offered two alternative analyses that rely only on codescriptive specifications. On one account, head-switching is accomplished by mapping the source S node to an f-structure that contains information about the central clausal relations but excludes adjunct information. The ADV node maps to an f-structure that has the adverb as its main predicate with the central clausal f-structure appearing in argument position. The 'just' f-structure, though not accessible from the S node, maps through τ to the outermost target f-structure. This complex interchange is specified in the lexical entry and rule in (5) and is diagrammed in Figure 1 (ignoring such details as person, number, case, and tense).

(5) (a) *just* ADV $(\uparrow \text{PRED})=\text{'just}\langle(\uparrow \text{ARG})\rangle\text{'}$
$(\tau\uparrow \text{PRED FN})=\text{venir}$
$(\tau\uparrow \text{XCOMP})=\tau(\uparrow \text{ARG})$

(b) $\text{S} \rightarrow \underset{(\uparrow \text{SUBJ})=\downarrow}{\text{NP}} \quad \underset{\uparrow=(\downarrow \text{ARG})}{(\text{ADV})} \quad \text{VP}$

Figure 1. Head-switching with inaccessible adjunct



In the second proposal, a completely integrated source f-structure maps to a single integrated target structure. As specified by the rule (6), the adverb is assigned an adjunct grammatical function in the source and its translation includes the translation of the enclosing source f-structure as its XCOMP, as shown in Figure 2.

(6) $\mathrm{S} \rightarrow \underset{(\uparrow\ \mathrm{SUBJ}) = \downarrow}{\mathrm{NP}} \quad \underset{\substack{(\uparrow\ \mathrm{ADJ}) = \downarrow \\ \tau\uparrow = (\tau\downarrow\ \mathrm{XCOMP})}}{\mathrm{ADV}} \quad \mathrm{VP}$

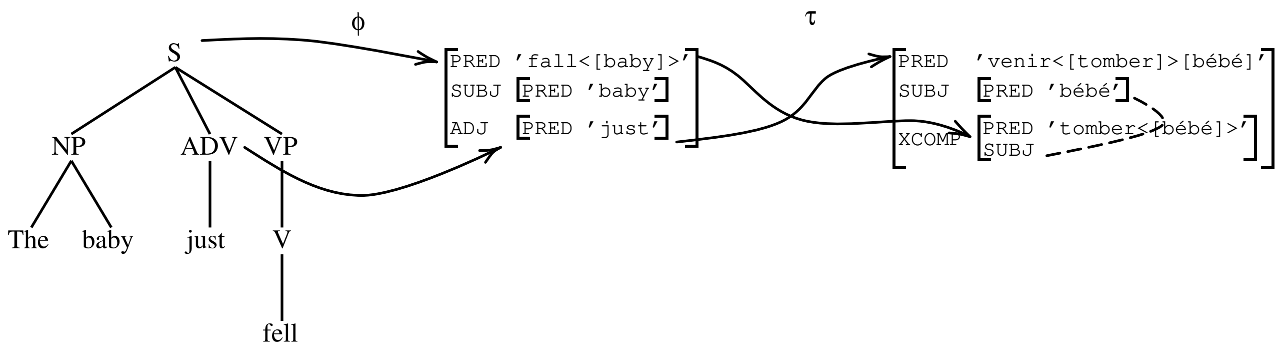
Sadler *et al*. (1989) and Sadler and Thompson (1991) point out a significant inadequacy of both these arrangements. Even though the proper target embeddings are derived under both correspondence configurations, in neither case does the translation of the f-structure of the source S node include the translation of the adverb. This shows up as a problem when such examples are embedded as complements in larger sentences:

(7) (a) I think that the baby just fell.
(b) Je pense que le bébé vient de tomber.

To maintain modularity, the codescriptive lexical entry for *think* must provide a direct mapping to the French *penser* that is not sensitive to the internal structure of the complement, along the lines of (8):

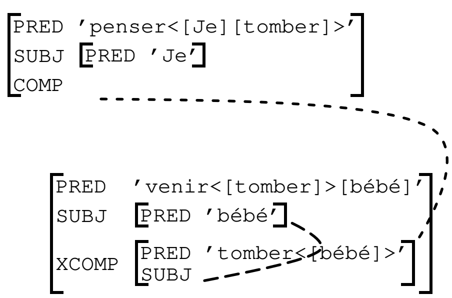
(8) *think* V
$(\uparrow \mathrm{PRED})=\text{'think}<(\uparrow \mathrm{SUBJ})\ (\uparrow \mathrm{COMP})>\text{'}$
$(\tau\uparrow \mathrm{PRED\ FN})=\text{penser}$
$(\tau\uparrow \mathrm{COMP})=\tau(\uparrow \mathrm{COMP})$

Figure 2. Head-switching with integrated f-structure



But then the translation f-structure constraints will characterize the pair of structures (9); these share the common *tomber* substructure but are otherwise unrelated.

(9)



The τ constraints thus leave unspecified the relative scopes of the *penser* and *venir* predicates. As Sadler *et al*. (1989) observe, the problem is even worse when several adverbs appear together: there is no obvious way to modify either of the rules (5b) or (6) to account for the scope interactions among the adverbs, let alone their relations to higher predicates.

Sadler *et al*. (1989) and Sadler and Thompson (1991) consider several ways in which the translation constraints might be modified to circumvent these difficulties yet remain faithful to the spirit of the correspondence-based approach. While each of their proposals is carefully worked out, they conclude (and we agree) that none of them is completely satisfactory or particularly compelling. The reason, we believe, is that the head-switching with adverbial modifiers that shows up as a problem in correspondence-based translation is actually a symptom of a more fundamental error in the syntactic and semantic analysis of the source language. In sentences with adverbial modifiers, the syntactic head (which controls subcategorization and enters into agreement relations) is not the same as the semantic head (the predicate with widest scope). Moreover, normal linguistic arguments would assign a flat f-structure to sentences with several adverbs while meaning relations would be represented in a hierarchical semantic structure. Thus on this view, if translation codescriptions map from the proper hierarchical semantic structures via τ' instead of from flat f-structures via τ, adverbial head-switching disappears as a special problem for correspondence-based translation.

Sadler and Thompson (1991, footnote 1) mention the arrangement we are proposing, and observe that the translation codescription problems may then merely be displaced to equivalent difficulties in characterizing the monolingual σ instead of τ: the problems may be moved around and renamed, but not solved. This may be so, but any conceptual clarification in such a murky domain must be regarded as an advance, if only because it helps to spotlight the issues that are relevant to a solution and to connect them to other related phenomena. Indeed, we now suggest that adverbial head-switching is a special case of the general problem of mapping flat syntactic structures to hierarchical semantic ones. So-called "light verbs", complex-predicates, and clause-union phenomena in many languages are similarly difficult to handle in LFG using only codescription, attribute-value function- application, and equality constraints (or using the analogous formal devices of other theories, such as attribute-value unification over signs or categories). In the next section of this paper, we extend LFG's f-structure description language by introducing a new formal operator, called *restriction*. We illustrate its properties by applying it to a simple light-predicate sentence in Urdu. In Section 4 we combine restriction with description-by-analysis to characterize the appropriate hierarchical semantic structures for English sentential adverbs. At the end, we return to the head-switching problem of correspondence-based translation, providing a simple solution in terms of the τ' correspondence.

3. The restriction operator and structural misalignments

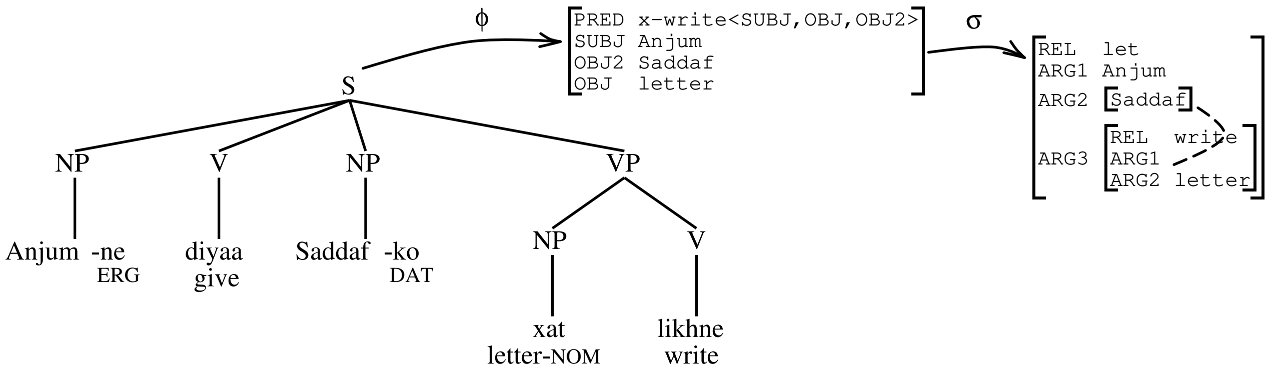
A simple kind of syntax/semantics misalignment is exemplified by constructions involving light-verbs or complex predicates. These are constructions formed by two or more verbs involving two or more semantic relations, but with the notable peculiarity that the complex behaves as a single, monoclausal syntactic unit according to standard tests of subcategorization and agreement. Butt *et al*. (1990) have argued persuasively that Urdu complex predicates are syntactically monoclausal and further, that the complex predicate cannot be formed in the lexicon. The details of the argument do not matter for present purposes, and we will simply accept their analysis of sentence (10a) whose English translation is given in (10b). The c-, f-, and semantic-structures, according to their analysis, are shown in Figure 3.

(10) (a) Anjum-ne diyaa Saddaf-ko xat likhne.

(b) Anjum let Saddaf write a letter.

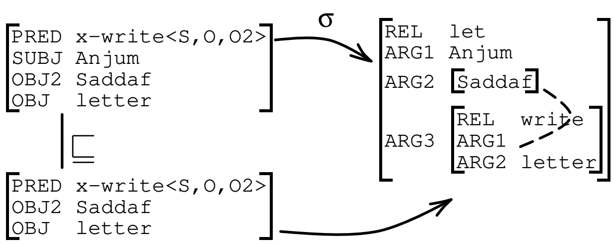
The crucial feature of this analysis is that there is a single set of governed grammatical functions in the f-structure, and these are derived systematically from the normal subcategorization of the main predicate *likhne* 'write'. The governable functions of *diyaa* when it stands as an independent predicate (usually glossed as the ditransitive predicate 'give') are not represented in the f-structure. The second obvious feature is that the flat f-structure maps to the hierarchical semantic structure, where the outer predicate has the permissive reading conveyed by *diyaa* in its light-verb sense and the inner proposition contains the main predicate and its arguments. This analysis cannot be formulated using standard codescription, function-application, and equality because there is no separate level in the f-structure that the inner semantic proposition can correspond to and thus no way to describe its properties.

Figure 3. Structural correspondences for complex predicate



While there may be a technical problem in finding a structure that the inner proposition can correspond to, there is a very clear intuition about what parts of the f-structure carry the information that constrains that piece of the semantic structure, namely, the sub-f-structure obtained by eliminating the SUBJ attribute and value. The following diagram depicts this intuition:

(11)



In this arrangement the semantic correspondence σ relates each level of the semantic structure to a unit in f-structure space, and that unit is the source of constraints on the properties of the corresponding element of the semantic structure. The ARG3 hierarchy in the semantic structure is not the image of an attribute embedding in the f-structure as is usually the case; rather, the semantic hierarchy here corresponds to a subsumption relation in the f-structure lattice. This organization of informational dependencies can be expressed by means of the restriction operator.

Restriction is a new operator in the f-structure description language notated by \ and with the following (partial) definition:

(12) If f is an f-structure and a is an attribute:

$$f \setminus a = f\big|_{\mathrm{Dom}(f)-\{a\}} = \{\langle s, v\rangle \in f \mid s \neq a\}$$

The restriction of a given f-structure f by a particular attribute a is the f-structure that results from deleting a and its value from f. If f is the f-structure in (13a), then the f-structure in (13b) is $f \setminus$SUBJ:

(13) (a) $f = \begin{bmatrix} \text{PRED} & \text{kick} \\ \text{SUBJ} & \text{John} \\ \text{OBJ} & \text{ball} \end{bmatrix}$

(b) $f\backslash\text{SUBJ} = \begin{bmatrix} \text{PRED} & \text{kick} \\ \text{OBJ} & \text{ball} \end{bmatrix}$

Restriction is a designator analogous to ordinary function-application in that it provides a way of referring to elements of f-structure space by virtue of their relations to other f-structures. If f and g are two f-structure designators and a denotes an attribute, then

(14) $(f\ a)=(g\ a)$

asserts that f and g both have an attribute a with exactly the same value; they may or may not have other attributes and values in common.

$f\backslash a=g\backslash a$

asserts that f and g have all attributes and values in common other than a; they may or may not have values for a and those values may or may not be identical.

Thus, restriction and function-application can be used to impose complementary constraints on f-structure values. We note that restriction is associative and commutative in its second (attribute) argument, so that $[f\backslash a]\backslash b = [f\backslash b]\backslash a = f\backslash\{a\ b\}$, and that for any f-structure f and attribute a it is always the case that $f\backslash a$ subsumes f $(f\backslash a \sqsubseteq f)$.

Returning now to the Urdu example, we see that if the top-level f-structure (the one corresponding to the S node) is denoted by f, the subsidiary f-structure to which the inner proposition in (11) corresponds is the restriction of f by SUBJ and can be referred to by the expression $f\backslash\text{SUBJ}$. The restriction operator can be used in codescription statements so that exactly the configuration in (11) is assigned to the Urdu sentence. A lexical redundancy rule can be introduced to systematically modify the lexical entries for normal verbs like *likhne* to make them suitable for combination with a light-verb:

(15) $(\uparrow \text{SUBJ}) \rightarrow (\uparrow \text{OBJ2})$

$\sigma\uparrow \rightarrow \sigma[\uparrow\backslash\text{SUBJ}]$

This rule replaces all references to the grammatical function SUBJ with OBJ2, thus avoiding conflict with the SUBJ introduced by the light verb, and it replaces all occurrences of the term $\sigma\uparrow$ with the term $\sigma[\uparrow\backslash\text{SUBJ}]$. This indicates that the main predicate provides constraints on the semantic structure corresponding to the subject-free f-structure. As a result of this rule, the usual equations for *likhne* in (16a) would give rise to the alternatives in (16b):

(16) (a) $(\sigma\uparrow \text{ARG1})=\sigma(\uparrow \text{SUBJ})$

$(\sigma\uparrow \text{ARG2})=\sigma(\uparrow \text{OBJ})$

(b) $(\sigma[\uparrow\backslash\text{SUBJ}]\ \text{ARG1})=\sigma(\uparrow \text{OBJ2})$

$(\sigma[\uparrow\backslash\text{SUBJ}]\ \text{ARG2})=\sigma(\uparrow \text{OBJ})$

The lexical rule would also make other minor adjustments to fill out the entry; the details do not concern us here.

4. Restriction and adverbial modifiers

Sentence with adverbial modifiers can also be characterized intuitively as having a flat syntactic structure with hierarchical semantic relations, and restriction can also be used to describe the appropriate structural configurations. The adverbial examples above involve only single sentential adjuncts, but our analysis allows for any number of adverbs with all possible scope ambiguities, and it can easily be extended to handle VP modifiers as well. We start with a c-structure rule that assigns arbitrarily many adverbs to the set-value of the ADJ attribute, consistent with the original LFG account of adjuncts (Kaplan and Bresnan, 1982):

(17) $$\mathrm{S} \rightarrow \underset{(\uparrow \text{SUBJ})=\downarrow}{\mathrm{NP}} \quad \underset{\downarrow\in(\uparrow \text{ADJ})}{\mathrm{ADV}^*} \quad \mathrm{VP}$$

The sentence (18a) will be assigned the f-structure (18b) by virtue of this rule. Our goal is to associate with this f-structure the alternative semantic structures in (19).

(18) (a) John obviously just fell.

(b) $$\begin{bmatrix} \text{PRED} & \text{'fall<[John]>'} \\ \text{SUBJ} & \begin{bmatrix}\text{PRED} & \text{'John'}\end{bmatrix} \\ \text{ADJ} & \{\ [\text{obviously}]\ [\text{just}]\ \} \end{bmatrix}$$

(19) (a) $$\begin{bmatrix} \text{REL} & \text{obviously} \\ \text{ARG1} & \begin{bmatrix} \text{REL} & \text{just} \\ \text{ARG1} & \begin{bmatrix} \text{REL} & \text{fall} \\ \text{ARG1} & \text{John} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(b) $$\begin{bmatrix} \text{REL} & \text{just} \\ \text{ARG1} & \begin{bmatrix} \text{REL} & \text{obviously} \\ \text{ARG1} & \begin{bmatrix} \text{REL} & \text{fall} \\ \text{ARG1} & \text{John} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Intuitively, the innermost proposition in both (19a) and (19b) is based on the f-structure information in (18b) ignoring the adjuncts; the middle proposition in (19a) is exactly what we would expect for a sentence that included *just* but not *obviously*, and the middle proposition in (19b) would be appropriate for a sentence containing only *obviously*. Thus the semantic structures again seem to correspond to subsets of f-structure information, and we begin by completing the definition of the restriction operator. In addition to restricting an f-structure by an attribute, as defined in (12), we also define the restriction of an f-structure by an element of an attribute's set-value:

(20) If f is an f-structure and a is an attribute:

$$f \backslash\langle a\ g\rangle = \begin{cases} f \backslash\ a & \text{if } (f\ a)-\{g\}=\emptyset \\ f \backslash\ a \cup \{\langle a, (f\ a)-\{g\}\rangle\} & \text{otherwise} \end{cases}$$

The restriction of a given f-structure f by a particular member of an attribute a's set-value is the f-structure that results from deleting that member of the set and also deleting the attribute itself if the set would then be empty. The relationships in (21) exemplify the pattern:

(21) $f = \left[\begin{array}{ll} \texttt{PRED} & \texttt{'fall<[John]>'} \\ \texttt{SUBJ} & [\texttt{PRED 'John'}] \\ \texttt{ADJ} & \texttt{\{ [just] \}} \end{array}\right]$ $\qquad g = \texttt{[just]}$

$$f \backslash <\mathrm{ADJ}\ g> = \left[\begin{array}{ll} \texttt{PRED} & \texttt{'fall<[John]>'} \\ \texttt{SUBJ} & [\texttt{PRED 'John'}] \end{array}\right]$$

We note that set-element restriction also has commutative and associative properties:

(22) $[f\backslash<a\ g>]\backslash<a\ h>=[f\backslash<a\ h>]\backslash<a\ g>$
$=f\backslash<a\ g\ h>$

The restriction operator can now be used to describe the semantic correspondences for adverbial sentences. If f designates the f-structure in (18 b) and j and o designate the f-structures corresponding to the adverbs *just* and *obviously*, then the constraints (23 a,b) describe the outermost REL and ARG1 configuration in (19 a) and (23 c,d) describe the next level of semantic embedding:

(23) (a) $(\sigma f\ \mathrm{REL})=(\sigma o\ \mathrm{REL})$
(b) $(\sigma f\ \mathrm{ARG1})=\sigma[f\backslash<\mathrm{ADJ}\ o>]$
(c) $(\sigma[f\backslash<\mathrm{ADJ}\ o>]\ \mathrm{REL})=(\sigma j\ \mathrm{REL})$
(d) $(\sigma[f\backslash<\mathrm{ADJ}\ o>]\ \mathrm{ARG1})$
$=\sigma[f\backslash<\mathrm{ADJ}\ o\ j>]$
$=\sigma[f\backslash\mathrm{ADJ}]$

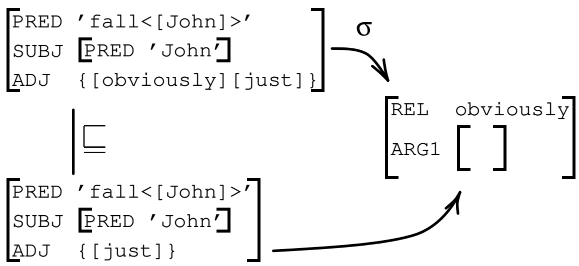
The innermost proposition can be described by interpreting (or redundantly rewriting) $\sigma\uparrow$ in the codescription equations in the *fall* lexical entry as $\sigma[\uparrow\backslash\mathrm{ADJ}]$, that is, by interpreting the σ specifications for all basic predicates as characterizing the semantics of an unmodified f-structure.

The restriction constraints in (23) are sufficient to map f-structure subsumption relations into the desired hierarchical semantic structures, but the number of such constraints depends on the size of the initial adjunct set; indeed, (23 d) suggests that the size of individual constraints used in this construction also grows in proportion to the number of modifiers. Constraints of this general form cannot be produced by the normal recursive analysis of the c-structure because the c-structure itself, by linguistic argumentation, does not have the degree of embedding that these constraints would require. The restriction operator can encode the intuitively desirable constraints, but an additional recursive process is needed to generate those constraints. This recursion can arise from an explicit traversal of the f-structure in the style of Halvorsen's (1983) semantic interpretation procedure. It can also come from lexical expressions of inside-out functional uncertainty; this formal device was introduced by Kaplan (1988) and has been applied to problems of quantifier scope (Halvorsen and Kaplan, 1988) and anaphoric dependencies (Dalrymple, 1993). Here we explore only the description-by-analysis approach, using the following analysis rule to generate codescriptive assertions:

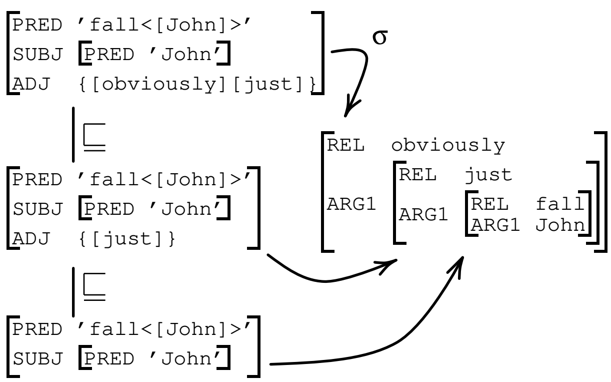
(24) For f an f-structure, $g \in (f\ \mathrm{ADJ})$, and g a sentence adverb,
$\sigma f = \sigma g$ and
$(\sigma f\ \mathrm{ARG1}) = \sigma[f\backslash<\mathrm{ADJ}\ g>]$

According to this rule, a single element is chosen (nondeterministically) from an adjunct set to contribute the relation for the semantic structure of the enclosing f-structure, and the semantic structure corresponding to the f-structure without the chosen element becomes that relation's argument. One application of the rule gives rise to additional structures to which the rule might also apply, and this recursion generates the appropriate set of constraints. For example, suppose f is the f-structure in (18b) and the *obviously* f-structure o is chosen as the instantiation of g in the rule. This produces the equations (23a,b) which define the configuration shown in (25). The rule then matches again against the lower f-structure, thereby completing the picture (26).

(25)



(26)



The alternative in which *just* has wide scope (the semantic structure (19b)) results from nondeterministically choosing the j f-structure in the first rule instantiation. We note without discussion that the rule in (24) is appropriate for sentence adverbs, which take complete propositions as arguments, but a different semantic structure is required for adverbs that only modify the meaning of the basic relation, such as the manner adverb in (27a). The relational embedding in (27b) gives a better account of the meaning of this sentence.

(27) (a) John walked slowly.

(b)
$$\left[\begin{array}{ll} \mathrm{REL} & \left[\begin{array}{ll} \mathrm{REL} & \mathrm{slowly} \\ \mathrm{ARG1} & \left[\mathrm{REL}\ \mathrm{walk}\right] \end{array}\right] \\ \mathrm{ARG1} & \mathrm{John} \end{array}\right]$$

Assuming some suitable marking of the differences among adverbs, perhaps based on the semantic typing discussed by Wedekind and Kaplan (1993), this structure is defined by the additional description-by-analysis rule (28):

(28) For f an f-structure, $g \in (f\ \text{ADJ})$, and g a VP adverb,

$$(\sigma f\ \text{REL}) = \sigma g$$
$$(\sigma f\ \text{REL}\ \ \text{ARG1}) = \sigma[f\backslash\langle \text{ADJ}\ g\rangle\ \text{REL}]$$
$$\sigma[f\backslash\langle \text{ADJ}\ g\rangle] \backslash \text{REL} = \sigma[f\backslash\text{REL}]$$

5. Head-switching and translation

The restriction operator and the description-by-analysis rule (24) provide an account of adverbial modification that is motivated purely on the basis of monolingual linguistic argumentation. The net effect, however, is to provide a natural hierarchical structure to serve as a source of codescriptive constraints in correspondence-based translation. The adjunct translation constraints for a sentence such as

(29) I think that the baby just fell.

can now be associated quite directly with the adverbial lexical entries using the τ' correspondence between source and target semantic structures. For example, the lexical entry for *just* would include the simple equations in (30):

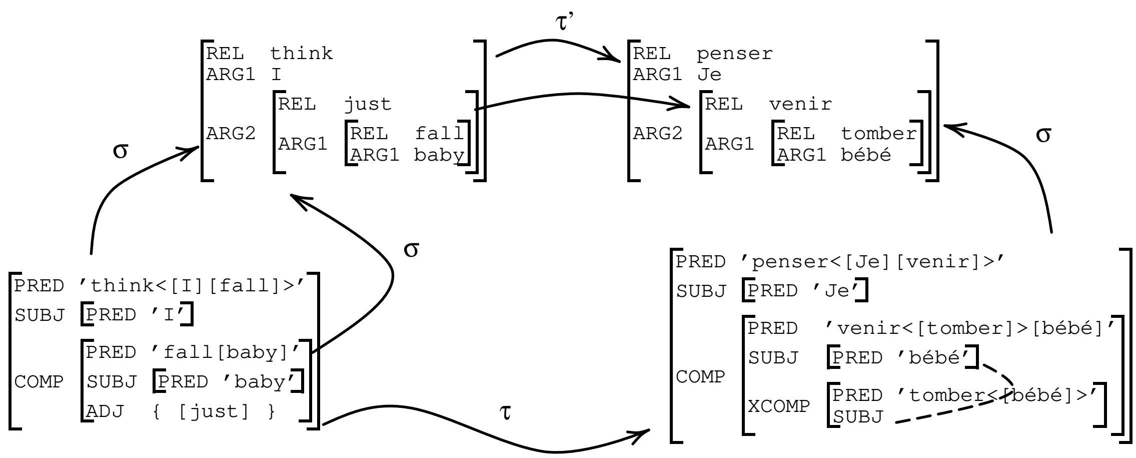
(30) $(\sigma\uparrow\ \text{REL})=\text{just}$
$(\tau'\sigma\uparrow\ \text{REL})=\text{venir}$

These involve the codescriptive composition of τ' and σ together with the φ and M that are implicit in ↑. Along with all the other constraints from the lexicon, grammar, and description- by-analysis rules, these define the translation configuration outlined in Figure 4. The restriction f-structure that the `fall` semantic substructure corresponds to is not shown in the figure, but as we have seen, it is crucial to the declarative, order-free construction of the embedded head-switching translation. This arrangement shifts the translation burden for adverbial modifiers from the f-structure to the semantic structure and from the τ correspondence to τ'. However, despite its somewhat diminished role with respect to modifiers, the τ correspondence is still important in controlling the translation of simpler grammatical function assignments and other more superficial grammatical properties.

6. Conclusion

We have suggested in this paper that certain problems of correspondence-based translation, in particular the difficulty of adverb and verb interchanges, are more properly considered the result of defective monolingual analyses. We have proposed a new description-language operator, restriction, to permit a succinct formal encoding of the informal intuition that semantic units sometimes correspond to subsets of functional information. This operator, in conjunction with an additional recursion provided by a description-by-analysis rule, is the basis of a more adequate account of head-switching phenomena.

Figure 4. Translation based on monolingual semantic embedding



References

[Butt *et al.*, 1990] Miriam Butt, Michio Isoda, and Peter Sells. Complex predicates in LFG. Unpublished manuscript, Department of Linguistics, Stanford University.

[Dalrymple, 1993] Mary Dalrymple. *The syntax of anaphoric binding*. Chicago: CSLI/The University of Chicago Press, to appear.

[Halvorsen, 1983] Per-Kristian Halvorsen. Semantics for lexical-functional grammars. *Linguistic Inquiry* 14, 567-613.

[Halvorsen and Kaplan, 1988] Per-Kristian Halvorsen and Ronald Kaplan. Projection and semantic description in Lexical-Functional Grammar. *Proceedings of the International Conference on Fifth Generation Computer Systems*, Tokyo, 1116-1122.

[Kaplan, 1987] Ronald Kaplan. Three seductions of computational psycholinguistics. In P. Whitelock, M. M. Woods, H.L. Somers, R. Johnson, & P. Bennett (eds.), *Linguistic theory and computer applications*. London: Academic Press, 149-188.

[Kaplan, 1988] Ronald Kaplan. Correspondences and their inverses. Paper presented to the Workshop on Unification Formalisms: Syntax, Semantics, and Implementation, Tititsee, Federal Republic of German.

[Kaplan and Bresnan, 1982] Ronald Kaplan and Joan Bresnan. Lexical-functional grammar: A formal system for grammatical representation. In J. Bresnan (ed.), *The mental representation of grammatical relations.* Cambridge: MIT Press, 1982, 173–281.

[Kaplan *et al*., 1989] Ronald Kaplan, Klaus Netter, Jürgen Wedekind, and Annie Zaenen. Translation by structural correspondence. *Proceedings of the Fourth European Conference of the Association for Computational Linguistics*, UMIST Manchester, 272-281.

[Sadler *et al*., 1989] Louisa Sadler, Ian Crookston, and Andy Way. Co-description, projection and "difficult" translation. Working Papers in Language Processing No. 8, Department of Language and Linguistics, University of Essex.

[Sadler *et al*., 1990] Louisa Sadler, Ian Crookston, Doug Arnold, and Andy Way. LFG and translation. *Third International Conference on Theoretical and Methodological Issues in Machine Translation,* Linguistic Research Center, University of Texas at Austin, 11-13.

[Sadler and Thompson, 1991] Louisa Sadler and Henry Thompson. Structural non-correspondence in translation. *Proceedings of the Fifth European Conference of the Association for Computational Linguistics,* Berlin, 293-298.

[Wedekind and Kaplan, 1993] Jürgen Wedekind and Ronald Kaplan. Type-driven semantic interpretation of f-structures. *Proceedings of the Sixth European Conference of the Association for Computational Linguistics,* Utrecht.