RATIONALISM AND THE TREATMENT OF REFERENTIAL DEPENDENCIES

Doug Arnold
Louisa Sadler

Department of Language & Linguistics,
University of Essex,
Wivenhoe Park,
Colchester, CO4 3SQ, UK

doug@essex.ac.uk;
louisa@essex.ac.uk

ABSTRACT

This paper describes what we take to be the correct perspective on the 'empiricist' vs 'rationalist' debate which is the dominant theme of this workshop, and describes a technique for dealing with the translation problems that arise where languages differ in the conditions they impose, and the strategies they employ for expressing different kinds of referential or anaphoric dependence.

1. Introduction: Rationalism vs. Empiricism, a Perspective.

This paper has two aims: the first is to describe what we think is the correct perspective on the 'empiricist' vs 'rationalist' debate which is the dominant theme of this workshop. The second is to describe a technique for dealing with the translation problems that arise where languages differ in the conditions they impose, and the strategies they employ for expressing different kinds of referential or anaphoric dependence.

(l) a A qui$_i$ a-t-il donne le livre e$_i$?
b To whom$_i$ has-he given the book e$_i$
(2) a *Qui$_i$ a-t-il donné le livre à e$_i$?
b Who$_i$ has-he given the book to e$_i$
(3) a l'uomo che mi domando chi abbia visto
the man that I wonder who has seen
b the man that I wonder who *he* has seen
(4) a [sinda node] minna ga kanasinda] hito
[died because] everyone SBJ distressed-was] man
a man who everyone one was distressed because *he* died
(5) a Petar$_i$ tried e$_i$ to come on time.
b Petar$_i$ he pokušao da e$_i$ dodje na vreme
Peter AUX tried Comp pro$_i$ come(Pres) on time

(1) and (2) show that one language (French) requires pied-piping (la), and forbids preposition stranding (2a), and another (English) allows both possibilities (1b,2b), in the relation between the focus of a question, and the thematic, or understood position of the focussed item. (3) and (4) are cases where one

language (respectively Italian and Japanese) employs a 'gap' strategy for relative clause formation, and English requires a resumptive pronoun. In (5a), the relation between the subject of *try* and the embedded subject is one of functional control (in the terminology of LFG -- that is, there is token identity, or re-entrance between the coindexed positions), while the corresponding relation in Serbo-Croat (5b) is one of anaphoric binding of an empty (or under some conditions, oven) pronominal (cf. Zec, 1987).

Research on Machine Translation has two related, but not always compatible goals. The first is to provide the basis for better MT systems. One could call this an engineering goal. The second (which is, from a certain perspective, of more lasting importance), one might call the scientific goal. This is to articulate a formal theory of translation -- saying something about the notion of 'possible translation', identifying the different components of this notion (such as the grammatical, common-sensical, stylistic, etc. constraints that determine when, and to what extent expressions A and B are translations), and in a general way addressing the way in which the 'form-function' relation is realized differently in different languages. We believe that this is a crucial goal for MT research, and that both rule-based and hybrid systems (which mix a rule-based core with some statistical or analogical component) have much to contribute to it, but will have little to say about this here.

In terms both of the practical and scientific goal of MT, empirical methods such as the extensive use of bilingual corpora in testing and constructing components of MT systems are extremely valuable, especially when the (translation) facts are unclear (e.g. as in the study of novel sublanguages). Equally important, in our view, is the use of established sources such as translation manuals, bilingual dictionaries, and the expert intuition of professional translators.

With respect to the engineering goal of MT, 'empiricism' denotes a collection of techniques for practical MT. Techniques which are analogical, statistical, or example based are naturally thought of as 'empiricist'. By contrast, a 'rationalist' respect for theory leads naturally to the use of techniques which are 'rule based', or 'constraint based'. Empiricist techniques are attractive as a way of avoiding, or ameliorating at least the following problems of rule based transfer systems:

(i) The very large rule writing effort that is involved in describing transfer pairings;

(ii) The need to explicitly represent all translationally relevant information in one level of linguistic structure (viz. the level which constitutes the organising level for transfer);

(iii) The need to formulate explicitly the information required for disambiguation (for example, to guide lexical selection), and to control the application (and interaction) of transfer rules.

However, in evaluating the attractions of empiricist techniques, two things should be borne in mind. First, in fact, neither (i) nor (ii) are necessarily limitations of rationalist approaches *per se,* or at least, there are other ways they can be alleviated. For example (i) is alleviated if grammatical and lexical knowledge can be re-used from application to application, e.g. if formalism and monolingual information are not specifically tailored for MT. Similarly, if the formalism is sufficiently high level and embodies the right kind of linguistic theory, the rule writing effort is correspondingly reduced. (ii) is not a problem in approaches that are 'sign-based', or use the notion of projections (see below). The use of formalisms that have properties of declarativity (i.e. their interpretation is independent of order) and monotonicity helps with control and interaction in (iii), and further eases the rule writing effort mentioned in (i).

Moreover, empiricist techniques have inherent limitations. For example, while it is easy to see how empiricist techniques can be used in lexical or structural selection (e.g. in PP attachment, or in preferring one kind of translation over another), it is quite difficult to see that they currently offer an *alternative* to rule based approaches (rather than a useful adjunct to them), when one looks at translation problems such those in (l)-(5) above, since 'pure' statistical techniques are limited to the treatment

of strictly 'local' phenomena.

The first difficulty with constructions like (l)-(5) is to find a way of making information about the 'antecedent' available in the location of the 'anaphor', specifically when the constraints on various types of anaphora differ between the two languages. Where the basic representation is a tree, this is problematic: copying the information makes *Everyone loves his mother* look the same as *Everyone loves everyone's mother,* which is incorrect; if the antecedent is moved to the anaphor position, so that *Which books does Sam admire* is treated as *Sam admires which books,* one loses information about the antecedent location, and has the problem of reversing the movement in synthesis (cf. Arnold et al, 1988; van Noord et al 1990). This problem is overcome in representations based on attribute-value (A:V) structures, where it is possible to have *re-entrance* -- one single piece of structure occurring in two distinct locations. However, since this sort of dependency is generally unbounded, one cannot, in general, write a transfer rule that mentions both antecedent and anaphor. In practice, this means that 'structural' approaches which use A:V structures -- that is, approaches which have transfer rules that operate over such a structure to build target structures -- typically produce representations where the re-entrance has been replaced by some kind of copying, so that the advantage of using an A:V structure representation is lost[1] Secondly, the problem is compounded in (l)-(5) above by the fact that the conditions on anaphora are not exactly parallel across the each language pair. Dealing with this sort of difference between languages appears to need very complicated, and powerful rules. This is especially worrying, because these constructions are very widespread, and this sort of difference between languages is very common.

We will try to show in this paper how these problems can be overcome in constraint, or description based approaches. What is needed is some sort of negotiation between two rule sets (governing the distribution and behaviour of anaphoric links of various sorts in source and target languages). In the remainder of this paper, we will describe a technique for doing this. The basic idea is that one states constraints over the translation relation which underdetermine the target structure. This is combined with further descriptions from the target grammar to describe a set of possible translations, from which a 'best' solution may be chosen, either by some general non-analogical principle, or by means of some form of corpus-based choice by example or analogy.

2. Constraint Based Translation

In constraint-based approaches to MT (such as Whitelock 1991, Kaplan et al 1989), the focus of transfer is not on the the application of rules to representations, but on defining the constraints which describe such structures. In such approaches, the entire translation process can be seen as a process of collecting, and resolving sets of constraints. *Instead of transfer applying rules to a source structure, constraints on the source-target mapping are collected during analysis.* The partial target description so derived operates as a constraint on synthesis, which can be thought of as a kind of forward inference from a set of premises, the target language constraints, with the target grammar as a source of further hypotheses. One effect of this is that the input to synthesis may be under-specified in various ways -- this allows the target language specific information to be omitted from transfer, reducing the effort of describing transfer (cf. (i) above). Bilingual constraints are in principle no different from monolingual constraints. There is no necessary requirement that the translational constraints are restricted to one dimension (or level) of linguistic description, and thus linguistic information from a wide variety of sources can be brought to bear on the translation process without the need for one organising level (cf.

[1] Estival et al (1990), and van Noord et al (1990) are two such approaches. Sadler and Arnold (1992) discusses this issue in more detail.

(ii), above).

In the approach of Kaplan et al (1989), LFG style equations and projections are employed to state constraints on the translation mapping. Projections are linguistically relevant mappings between levels of linguistic description. They may be direct, for example the ϕ projection from elements of c-structure (nodes) to elements of f-structure (attributes and values), or the σ projection from f-structure to semantic structures, or may involve the composition of projections.[2] Kaplan et al introduce the projections τ (between f-structures) and τ' (between semantic structures). Achieving translation involves specifying and resolving a set of constraints on target structures, by means of these and other projections. For example, (6)[3] which might appear in the lexical entry for many verbs, or as a default, composes τ and ϕ, identifying the τ of the (source) SUBJ f-structure with the SUBJ attribute of the τ of the f-structure associated with some node (the value of $\uparrow$). That is, the translation of the value of the SUBJ slot in a source f-structure fills the SUBJ slot in the f-structure which is the translation of that source f-structure. (It may help to point out that the SUBJ that appears on the left-hand side is an attribute in the source structure, while that on the right-hand side belongs to the target language).

(6) $\tau(\uparrow SUBJ) = (\tau\uparrow SUBJ)$

Notice that these equations do not form a separate transfer grammar,[4] and there is no separate rule application or recursion through a source structure corresponding to transfer.

3. Differences in Conditions on Referential Dependencies (1)

To deal with the mismatches in examples like (l)-(3), where there is a re-entrance between a focussed position, and a thematic position (i.e. a 'gap'), the basic idea is to state bilingual constraints for only the thematic ('bottom') end of the re-entrance.[5] For the sake of discussion, we will assume the French f-structure corresponding to (la) is along the lines of (8), produced by a rule like (7).[6]

(7) S' → XP S
(↑FOCUS)=↓ ↑=↓
(↑FOCUS)=(↑{COMP, XCOMP}* GF)
(↑ FOCUS{OBJ, POSS}* WH)=$_c$ +
(where GF = {SUBJ.OBJ.OBL$_{go}$,… })

[1] σ is sometimes taken to be a projection from c-structure to semantic structures; ϕ is normally notated with $\downarrow$ and $\dagger$. Where C denotes a c-structure node, and $m(c)$ is its mother, $\downarrow$ on a node denotes $\phi(c)$, i.e. the f-structure associated with the node, and $\dagger$ denotes $\phi(m(c))$.

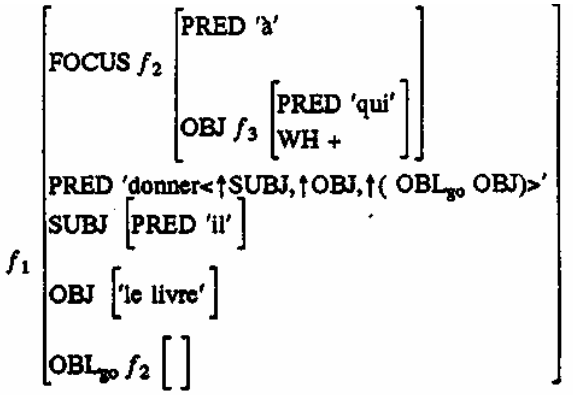
[3] The notation is left-associative, so the right hand side of (32) is equivalent to ((τ↑) SUBJ).

[4] But cf. Sadler et al 1990, for some arguments that one might want to state some aspects of the translation mapping separately in a bilingual lexicon.

[5] Kaplan et al suggest a treatment which translates both the 'top' and 'bottom' positions (both the focussed, and thematic positions). As pointed out in Sadler and Arnold (1992), this will only work if source and target languages observe the same constraints. Specifically, it will not work in cases such as (l)-(3). For example, it will give the ungrammatical (2a), with a stranded preposition as the translation of (2b). Kaplan et al propose dealing with cases of raising (e.g. the translation of *John is likely to see Mary* -- *Il est probable que Jean verra Marie* ('It is probable that John will see Mary.') by translating only the 'bottom', but they do not consider any wider application of the technique.

[6] The following discussion involves a number of simplifying, or questionable assumptions about the precise status of some items (especially the prepositions *a* and *to),* and the formulation of rules which are not germane to the discussion.

(8)

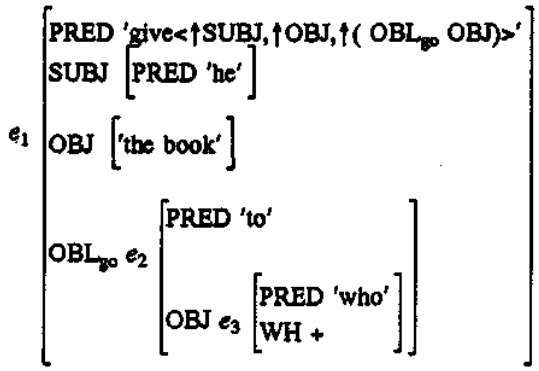


This rule makes the f-structure of a pre-sentential XP into the value of FOCUS, and identifies it (via the functional uncertainty equation (Kaplan and Zaenen, 1989) with the value of a SUBJ, OBJ, OBL_{go} that is embedded inside zero or more COMPs or XCOMPs. It requires a +WH somewhere inside the FOCUS (this ensures that only phrases containing WH words are focussed in questions). Notice that no τ equation is given for the XP -- the content of the XP will be translated in its thematic position (OBJ of *à*).[7]

The entry for *donner* will include information such as (9), stating that the 'Function Name' of the PRED in the English structure will be 'give', and that the translation of the OBL_{go} will be the value of OBL_{go} in the corresponding English structure. Evaluating these constraints (with constraints gathered from other lexical entries, such as the entry for *à,* which will indicate that τ(↑OBJ)=(τ↑OBJ), for example) gives an underspecified English structure, which can be filled out with information from the English grammar and lexicon, to give an f-structure like (10), corresponding roughly to the echo question *He gave the book to who.*

(9) (τ↑PRED FN) = give
τ(↑OBL_{go}) = (τ↑OBL_{go})

(10)



In order to produce a correct f-structure, we must ensure that a FOCUS attribute with an appropriate value is obtained for the English. There are three sources of information which can be used.

[7] No τ equation is needed on the S. Because its f-structure will be identical to that of the VP it contains, their translation are also identical.

(i) A source language constraint -- the solution to the functional uncertainty equation which established the source language re-entrance - is available. In this case it is (f_1 FOCUS)=(f_1 OBL_{go}). By the application of a general schema, we can derive a T equation from this,[8] which in this case is (11), which (again in this case) is equivalent to (12).

(11) (τf_1 FOCUS) = τ ($\uparrow OBL_{go}$)
(12) (τf_1 FOCUS) = ($\tau\uparrow$ OBL_{go})
(13) (e_1 FOCUS) = (e_1 OBL_{go})

If we simply add this to the other τ equations, we will produce a structure parallel to the French, with pied-piping. This is an acceptable result in this case, but not in general, since (as noted) one cannot always preserve stranding or pied-piping of prepositions in translation.

(ii) The monolingual English grammar will contain a constraint, similar to that in (13) to ensure that, if FOCUS is present, some path within the FOCUS attribute contains the attribute value pair WH=+. This might be as in (14).

(14) ($\uparrow$ FOCUS{OBJ, POSS}* WH) $=_c$ +

(iii) The English monolingual grammar itself contains a functional uncertainty equation for establishing a relation between FOCUS and some within-clause function, along the lines of (15). Notice that this is different from the French equation, in allowing the 'bottom' GF to be 'inside' another GF (as in preposition stranding).

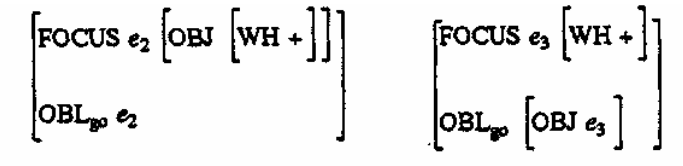
(15) ($\uparrow$ FOCUS)=($\uparrow${COMP, XCOMP}* (GF) GF)
where GF abbreviates {SUBJ, OBJ, OBL_{go}, … })

Intuitively, the source-derived equation (12) can be used to provide the information that there should be a FOCUS attribute in the target f-structure. The idea is that this constraint can be interpreted defeasibly in combination with the target information to find the solutions consistent with the target grammar. In the case of wh-questions at least, there are two target functional uncertainty equations which can be solved in different ways - the first requires FOCUS to contain a WH=+ path (14), the second states a re-entrance between the value of FOCUS and the value of some within clause function (15). In this case, there are two possible solutions for these constraints (16), and (17), corresponding to the partial f-structures in (18), and the strings in (19) and (20).

(16) (e_1 OBL_{go}) = (e_1 FOCUS)
(e_1 FOCUS OBJ WH) = +
(17) (e_1 OBL_{go} OBJ) = (e_1 FOCUS)
(e_1 FOCUS WH) = +

[8] The method for doing this involves taking the functional uncertainty, namely $\uparrow$FOCUS=$\uparrow${XCOMP,COMP}*GF, and adding ($\tau\uparrow$(FOCUS)=(τ($\uparrow a$)) for every solution a of the uncertainty on the right-hand side. This gives (11), equivalent to (13), as one solution.

(18)



(19) To whom$_i$ did he give the book [ ]$_i$?

(20) Who$_i$ did he give the book to [ ]$_i$

In this case, both solutions are acceptable, since English allows, but does not require preposition stranding and pied-piping. Of course, there will not always be such a choice, for example, there is no choice in going from English *Who did he give the book to,* as above in (2), because the uncertainty equations for French do not allow preposition stranding *(*Qui a-t-il donné le livre à),* so the only solution will be one where the WH-item is pied-piped *(A qui a-t-il donné le livre).* Sadler and Arnold (1992) demonstrate how the technique works in these cases.

When there is a choice, however, and in a practical setting, one would like to have a procedure for making a choice. The obvious method for making this choice is to do define some preference regarding the relative 'lengths' of the solutions to the uncertainty equations (e.g. giving priority to short solutions to (14) will produce solutions where the WH+ is less deeply embedded, that is, it will lead to stranding of prepositions, in general). However, there is a clear role for analogical techniques here, based on the frequency of different constructions in actual corpora.

4. Differences in Conditions on Referential Dependencies (2)

The approach just described does not extend immediately to cases like (3)-(5), where one sees a difference of 'strategy' -- where one language encodes a referential dependency by means of re-entrance, and another uses some kind of anaphoric binding (notice that the antecedent and anaphoric positions differ in at least the value of PRED, so cannot be the same f-structure object). In the case of anaphoric binding, we assume (following Dalrymple (1990) and Dalrymple *et al* (1990)), that referential dependence involves pieces of f-structure sharing the same semantics (that is, for f-structure objects f_1 and f_2, $(\sigma f_1)=(\sigma f_2)$) Constraints stating conditions anaphoric binding may be positive (requiring sharing of semantics), or negative (forbidding sharing of semantics), and are associated with each anaphoric element in the lexicon. The equations delineate possible relations between the f-structures of a pronoun and that of its antecedent by means of 'Inside Out' functional uncertainty.

A positive binding requirement schema might be of the form (21), indicating that the semantics of the f-structure associated with the mother of an item is the same as the semantics associated with some AntecedentPath within some containing f-structure (specified by DomainPath), cf. Dalrymple (1990). In the example equation (22), the DomainPath is XCOMP, GF, and the AntecedentPath is SUBJ. This would be appropriate for an anaphoric item which required its antecedent to be a SUBJ, which could be any number of XCOMPs 'higher up'.

(21) $\sigma((\text{DomainPatht})\ \text{AntecedentPath}) =_c \sigma\uparrow$

(22) $\sigma((\text{XCOMP}^*\ \text{GF}\uparrow)\ \text{SUBJ}) =_c \sigma\uparrow$

We will focus on the case of (5), repeated here, though it should be clear that the same problem arises wherever a source language re-entrance is realized in the target language by means of some kind of anaphoric binding (e.g. (3) and (4)), and the solution described here will work equally in those cases.

(5)a Petar$_i$ tried e$_i$ to come on time,

b Petar$_i$ he pokušao da e$_i$ dodje na vreme
Peter AUX tried Comp pro$_i$ come(Pres) on time

In English, the relation between the subject of *try,* and the complement subject is one of functional control, i.e. there is a re-entrance, they have the same f-structure. On the other hand, Zec (1987) argues that the complements of certain predicates in Serbo-Croat (the *da$_2$*-complements, which include the complement of the translation of *try)* must be analysed as containing an anaphorically bound null pronominal (Serbo-Croat is a pro-drop language, under certain conditions an emphatic pronominal is possible in the controlled position).

Taking English as source, on the approach described above, the τ equations for *try* and *come* would produce an f-structure in which *Petar* appeared as the SUBJ PRED of the embedded complement (as well as the PRED of the matrix SUBJ):

(23) try:
(τ↑PRED FN)=pokušati
τ(↑XCOMP)=(τ↑COMP)

(24) come:
τ(↑SUBJ)=(τ↑SUBJ)

The target lexical entry for the translation of *try, pokušati,* will indicate an obligatory anaphoric binding between (↑SUBJ) and (↑COMP SUBJ), which latter should, accordingly, contain a pronominal PRED value. The problem is that τ equations will have assigned the translation of *Petar* as the value of the embedded SUBJ attribute. There is a contradiction between these sets of requirements -- the translation that is naturally proposed is incompatible with the target grammar.

The solution we propose is to alter the translation information for nominals, exploiting the possibility of using translation constraints at more than one level. For simplicity, we will suppose that the semantics of *Peter* is an A:V representation of something like **peter'(x)**, and the semantics of *Petar* is something like **petar'(x)**, as in (25). (This is simplistic, but nothing here precludes nominals having more complex, and adequate, semantic values).

(25)

$$\begin{bmatrix} \text{REL} & \text{peter'} \\ \text{ARG} & \text{x} \end{bmatrix} \qquad \begin{bmatrix} \text{REL} & \text{petar'} \\ \text{ARG} & \text{X} \end{bmatrix}$$

Briefly, where the lexical entry for a nominal is currently as in (26), our proposal is to replace it with something like (27).

(26) (↑PRED) = Peter
(σ ↑ REL) = peter'
(τ ↑ PRED FN) = Petar

(27) (↑PRED) = Peter
(σ ↑ REL) = peter'
(σ (τ ↑) REL) = petar'

The last constraint states that the function name of the REL slot in the semantics associated with the target f-structure is peter'. Rather than assigning a PRED to the target f-structure directly, as in

(26), the entry in (27) implies only that whatever this PRED is, its semantics must be compatible with having **peter'** as its REL. This will be consistent with having (inter alia) either 'Petar', or 'PRO' as the PRED (that is, there being a re-entrance, or a pronominal). In the case in hand, the target language constraints will mean that only the pronominal will be permitted.

More generally, the revised style of entry will leave the target grammar with an underspecification. If a re-entrancy from the source side can be established in the target side, it will be, and the REL requirement will be satisfied. If it cannot be, then the target grammar must look for a suitable anaphor that will be consistent with this semantics. This should be reasonably efficient, since there will typically be very few such pronominals (at most as many as there different pronominals in the language). As with the indeterminism discussed in the previous section, one might also try to use some kind of analogical or statistical techniques so that the most plausible choices are tried first.

5. Conclusion

This discussion is intended to illustrate the way constraint based approaches contribute to the 'rationalist' enterprise, by indicating how one can alleviate some of the problems that face traditional rule based systems. The intuition it reflects is some limited kind of negotiation to resolve constrains that arise from source language grammar, target language grammar, and translation mapping. The particular proposal involves weakening the constraints on translation (leaving the translation of the FOCUS undefined, e.g.), and permitting remaining constraints on translation to interact defeasibly with monolingual constraints. This seems to us an interesting, and not at all unintuitive view of what translation 'really' involves.

References

D. Arnold, S. Krauwer, L. des Tombe, and L.G. Sadler, "Relaxed Compositionality in MT," *Proceedings of the Second International Conference on Theoretical and Methodological Issues in Machine Translation,* Pittsburgh, 1988. also, Working Papers in Language Processing 5, Dept. of Language and Linguistics, University of Essex.

M. E. Dalrymple, *Syntactic Constraints on Anaphoric Binding,* Unpublished PhD thesis, University of Stanford, Stanford, Calif., 1990.

M. E. Dalrymple, J. Maxwell, and A. Zaenen, "Modelling Syntactic Constraints on Anaphoric Binding," *Proceedings of COLING-90,* vol. 2, pp. 72-6, Helsinki, Finland, 1990.

Dominique Estival, Afzal Ballim, Graham Russell, and Susan Warwick, "A Syntax and Semantics for Feature-Structure Transfer," *Third International Conference on Theoretical and Methodological Issues in Machine Translation 11-13 June 1990,* Linguistics Research Center, Austin, Texas, 1990. (Page numbers not integrated)

P-K. Halvorsen and R.M. Kaplan, "Projections and Semantic Description in Lexical-Functional Grammar," *International Conference on Fifth Generation Computer Systems,* Tokyo, Japan, 1988.

R. M. Kaplan and A. Zaenen, "Long-Distance Dependencies, Constituent Structure, and Functional Uncertainty," in *Alternative Conceptions of Phrase Structure,* ed. M. R. Baltin and A. S. Kroch, pp. 17-42, Chicago University Press, Chicago, 1989.

Ronald Kaplan, Klaus Netter, Jurgen Wedekind, and Annie Zaenen, "Translation by Structural Correspondences," *Fourth Conference of the European Chapter of the Association for Computational Linguistics,* pp. 272-81, Manchester, 1989.

Ronald M. Kaplan, "Three seductions of computational psycho linguistics," in *Linguistic Theory and Computer Applications,* ed. Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod L. Johnson and Paul Bennett, pp. 149-188, Academic Press, London, 1987.

S.G. Pulman (ed), *EUROTRA ET6/1: Rule Formalism and Virtual Machine Design Study,* Commission of the European Communities, Luxembourg, 1991.

Louisa Sadler, Ian Crookston, Doug Arnold, and Andy Way, "LFG and Translation," *Third International Conference on Theoretical and Methodological Issues in Machine Translation 11-13 June 1990,* Linguistics Research Center, Austin, Texas, 1990. (Page numbers not integrated)

Louisa Sadler and Henry S. Thompson, "Structural Non-Correspondence in Translation," *Proceedings of EACL-91,* pp. 293-8, Berlin, 1991.

Louisa Sadler, "Structural Transfer and Unification Formalisms," *Applied Computer Translation,* vol. 1, no. 4, pp. 1-22, 1992.

Louisa Sadler and Doug Arnold, "A Constraint Based Approach to Translating Anaphoric Dependencies," *Proceedings of COLING-92,* Nantes, 1992.

G. van Noord, J. Dorrepaal, D. Arnold, S. Krauwer, L. Sadler, and L des Tombe, "An Approach to Sentence-Level Anaphora in Machine Translation," *EACL-4,* pp. 299-307,1989.

Gertjan van Noord, Joke Dorrepaal, Pim van der Eijk, Maria Florenza, and Louis des Tombe, "The MiMo2 Research System," *Third International Conference on Theoretical and Methodological Issues in Machine Translation 11-13 June 1990,* Linguistics Research Center, Austin, Texas, 1990. (Page numbers not integrated)

J. Wedekind, "Generation as structure driven derivation," *Proceedings of COLING-88,* pp. 732-737, Budapest, 1988.

Pete Whitelock, "Shake and Bake Translation," Unpublished Ms., Sharp Laboratories of Europe Ltd., April 1991.

D. Zec, "On Obligatory Control in Causal Complements," in *Working Papers in Grammatical Theory and Discourse Structure,* ed. M. Iida, S. Wechsler & D. Zec, pp. 139-68, Chicago University Press, Chicago, 1987.