Methodological Considerations in the METAL Project

Winfield S. Bennett
Siemens Information Systems, Inc.
and
Linguistics Research Center
University of Texas at Austin

For almost ten years the METAL machine translation project has had something of a split personality: on one hand, it is a research project seeking to create and refine a theoretically sound framework for machine translation; on the other, it is a development project seeking to produce a commercially viable system capable of translating a number of language pairs. Decisions on the project must always consider both the commercial and the theoretical aims. Further, METAL is a multinational project with research and development work being accomplished at four sites in four countries. The consequences of this situation have had significant effects on the approach and accomplishments in METAL research and development.

This paper will examine some of the methodological considerations which affect the work on the METAL project at all its sites. In particular this paper will examine the use of end-user input in the research and development process, the environment for linguistic researchers, and some of the effects of having multinational research sites.

1. Introduction

Research in machine translation has continued at the University of Texas at Austin since 1959, with only a few breaks due to lack of funding. Until the advent of funding from Siemens AG in 1979, the research was conducted on a theoretical basis, using various approaches to the question of how best to translate from one human language to another using computers. The current METAL system finds its roots in this earlier theoretical work, a fact mirrored in the very name of the system which was first used for a earlier model. When Siemens became a sponsor of the work a new dimension was added to the project METAL came to be viewed as a possible commercial MT system. These commercial aspects were added without loss of the earlier theoretical dimension.

Since 1979, then, METAL has had something of a split personality: on one hand, it is a research project seeking to create and refine a theoretically sound framework for machine translation; on the other, it is a development project seeking to produce a commercially viable system capable of translating a number of language pairs. All major, and even some minor, decisions on the project must take into account both theoretical and practical aims. Any solution must answer the question: is this theoretically justifiable within the METAL framework? as well as: does this contribute to the goal of creating a working commercial MT system?

METAL is not unlike other projects which find themselves caught between theoretical and practical considerations; sometimes the dual roles lead to a kind of schizophrenia, an uncertainty whether the theoretical model must be compromised 'to get something done' or the practical needs must be ignored in the search for an elusive optimum theoretical model. METAL's situation is further complicated by the fact that there are now four sites in as many countries at which work is going on in both the theoretical and practical domains. The situation has led to a number of methodological considerations which will be explored in this paper.

2. End-Users in METAL Research

One place where the question of theoretical versus practical considerations becomes highly significant is that of just how much influence should or can end-users have in the process of machine translation research and development. In projects which are concerned chiefly or only with the creation of a theoretical model for machine translation, this is not an issue — nor should it be. For projects which are concerned solely with producing a commercial MT system, the needs of end-users may be the chief consideration. Those projects which have theoretical and practical aims, however, must constantly wrestle with the needs of potential end-users as well as theoretical concerns.

The situation is further complicated by the fact that often researchers and end-users have different and sometimes divergent expectations of what a machine translation system should be. On the basis of my experience it is clear that my concerns and interests as a researcher are often quite different from those of the potential end-users. If end-users and researchers shared the same viewpoints and expectations, there would be no real issue; the reality is that research and development goals are often the product of researchers' decisions without input from end-users; these goals may not accurately reflect the needs of those who will ultimately use the system on a day-to-day basis. Failure to include end-users in the design and development of an MT system which is intended to be commercially viable will result in a less acceptable system. User acceptability is one of the most significant measures of a commercial machine translation system — if not the most significant.

As a commercial application METAL is intended to be a system which is to be used day in, day out by people who are trained as translators, not linguists or computer scientists and whose interest in MT is somewhat removed from the niceties of theoretical models. On the METAL project, any decisions made about the system, thus, must take into account the needs and desires of the end-users.

In a paper I delivered last year at the ATA conference (Bennett 1987), I stated that any MT research and development effort would do well to employ a 'research end-user', a person who is part of the research

effort, but whose function is to represent the viewpoint of the ultimate end-users (a group of people whom I called 'production end-users'). My position is based on the presence of research end-users on the METAL project, even though they were not originally hired as such.

The benefits of input from such people on our project has been clear; the system is aimed at being acceptable to the non-researchers who will use it. From the standpoint of end-user acceptability we would have saved time had we had at least one research end-user on the project from the very beginning.

An example will illustrate why our methodology includes using end-user input. A typical METAL lexical entry consists of a LISP list with a canonical form ('dictionary entry') and a series of feature/value pairs which provide morphological, syntactic, and semantic information; the canonical form is the CAR of the list; features appear as atoms and values as members of lists in the CDR of the list. A verb such as *arrive* gives a simple example of what a METAL lexical entry looks like:

```
(arrive        CAT (VST)
   ALO         (arriv)
   SNS         (1)
   PLC         (NF WI)
   PRF         (NIL)
   ARGS        ((($SUBJ N1 (TYPE P1)) OPT ($POBJ N1 (TYPE LOC TMP))))
   PV          (NIL)
   TT          (I)
   PX          (NIL)
   ON          (VO)
   CL          (G-ING I-E P-ED PA-ED PR-ES1)
)
```

Virtually everything about this entry (and the thousands of other METAL lexical entries) may be termed 'end-user hostile'. From a researcher's standpoint there is nothing theoretically problematic about the shape of METAL lexical entries, quite the contrary; the problem is that this shape is difficult for end-users to work with, even after months of exposure.

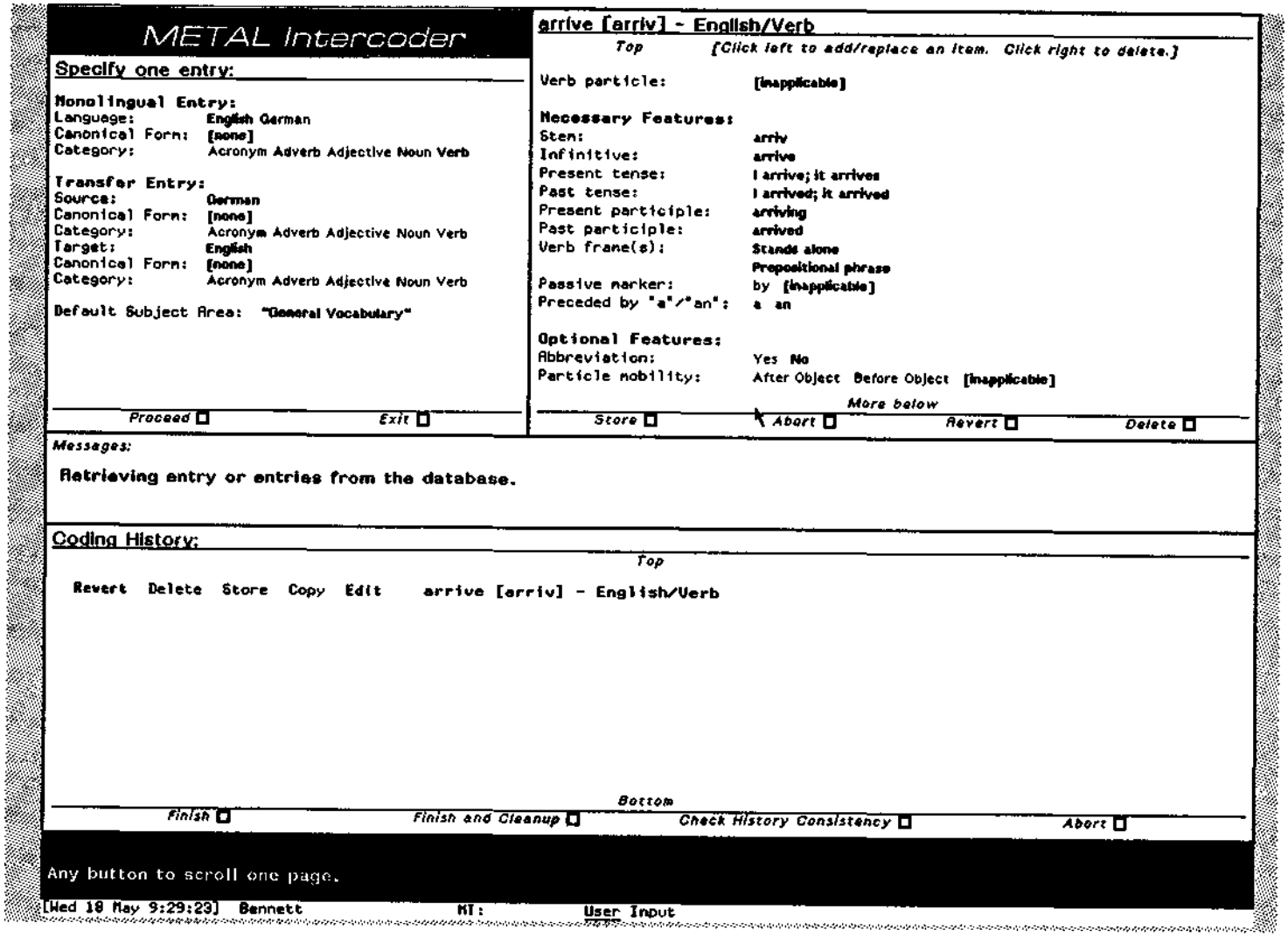
The solution taken on the METAL project was to create the INTERCODER, an environment in which end-users could code and modify lexical items without having to deal with the actual forms found in the lexicon. The INTERCODER employs windows and menus to guide the coder through the process of creating or editing lexical entries. A comparison of the entry, above, with the screen dump of an INTER-CODER window for the same verb, on the next page, illustrates significant differences.
Experience with lexicographers has shown that the INTERCODER is a faster, more efficient means of coding lexical entries for METAL than other methods, as well as being decidedly more user friendly.

The concept and design of the INTERCODER came from Lesley Whiffin Jezierny, the project's chief lexicographer who also functions as a research end-user, the latest form of the INTERCODER represents considerable effort on the part of Jennifer Oppenheim who has created an environment which is not only friendly to coders but which can be customized for any of the research groups with menus and windows in the language of their choice. Without Lesley's ideas we researchers might still be assuming that the form of the lexicon was 'something they [= end-users] just have to live with.'

A related facility in the METAL system is a series of lexical database management tools which assist the lexicographers in maintaining data integrity. Clearly this is an important issue for end-users who must acquire and maintain large lexical databases; researchers, on the other hand, do not require large lexicons for grammar research. The database management software available in the METAL package is again the result of input from project end-users, such as Lesley Whiffin Jezierny.

The issues of inputting lexical entries and maintaining the large system lexicon are only two of many issues which have required consideration of end-users' viewpoints in METAL system development (see Whiffin & Slocum 1985 and Slocum *et al.* 1987 for further discussion of problems and solutions).

In the two major lexicon issues outlined above and in a number of other issues, running the gamut from


METAL Intercoder
Specify one entry:
Monolingual Entry:
Language: English German
Canonical Form: [none]
Category: Acronym Adverb Adjective Noun Verb
Transfer Entry:
Source: German
Canonical Form: [none]
Category: Acronym Adverb Adjective Noun Verb
Target: English
Canonical Form: [none]
Category: Acronym Adverb Adjective Noun Verb
Default Subject Area: "General Vocabulary"
Proceed
Exit
arrive [arriv] - English/Verb
Top
[Click left to add/replace an item. Click right to delete.]
Verb particle: [inapplicable]
Necessary Features:
Stem: arriv
Infinitive: arrive
Present tense: I arrive; it arrives
Past tense: I arrived; it arrived
Present participle: arriving
Past participle: arrived
Verb frame(s): Stands alone
Prepositional phrase
Passive marker: by [inapplicable]
Preceded by "a"/"an": a an
Optional Features:
Abbreviation: Yes No
Particle mobility: After Object Before Object [inapplicable]
More below
Store
Abort
Revert
Delete
Messages:
Retrieving entry or entries from the database.
Coding History:
Top
Revert Delete Store Copy Edit arrive [arriv] - English/Verb
Bottom
Finish
Finish and Cleanup
Check History Consistency
Abort
Any button to scroll one page.
[Wed 18 May 9:29:29] Bennett MT: User Input

dealing with varying source language text formats to the best facilities for post-editing, the METAL project has actively used co-workers who represent the production end-users' viewpoint.

3. Researchers in METAL Research

In the same vein as the preceding section, one might also ask: just how much should an MT project take the needs of the researchers into account? On the surface this question sounds facetious, but it is meant to be taken seriously. Stated another way the question becomes: how much time and energy should be put into developing facilities (software tools and research environments) to assist the researchers in their efforts? In projects whose aims are solely theoretical, I suspect the answer is that the needs of the researchers should be uppermost; expenditure of time and energy for developing such facilities is part of the entire theoretical approach. On the other hand, I assume that projects whose aims are entirely commercial would be less prone to develop any more facilities than were absolutely necessary to achieve the goals of the project; like it or not, commercial MT projects have production goals to meet which limit the amount of time the project can devote to developing research facilities. The question may come down to one of whether the time saved by the researchers in system development is worth the time necessary to develop elaborate tools and environments for that research. Projects which combine commercial and theoretical goals must decide just which goal developing software tools and research environments contributes to. Of course the best case is when both goals are met in such development, but this situation is not always the case.

Some years ago the METAL project began a concerted effort to create and refine tools which could be used by the linguistic researchers in grammar research and other linguistic work. From some modest, but highly useful, grammar and transformation testing capabilities developed by John Bear in the early 80's, the research tools and environment have been considerably expanded (see White 1987 for a discussion of some of the earlier facilities). One of the most notable results of this effort is the so-called 'METAL Window', first designed and implemented by Oliver Gajek and subsequently developed by Roland Polzer and other system programmers.

The illustration on the following page provides an example of the METAL Window in the grammar mode. The METAL Window provides an environment in which the research linguists' work is facilitated by means of means and commands. Creation of this environment has greatly enhanced the research effort.

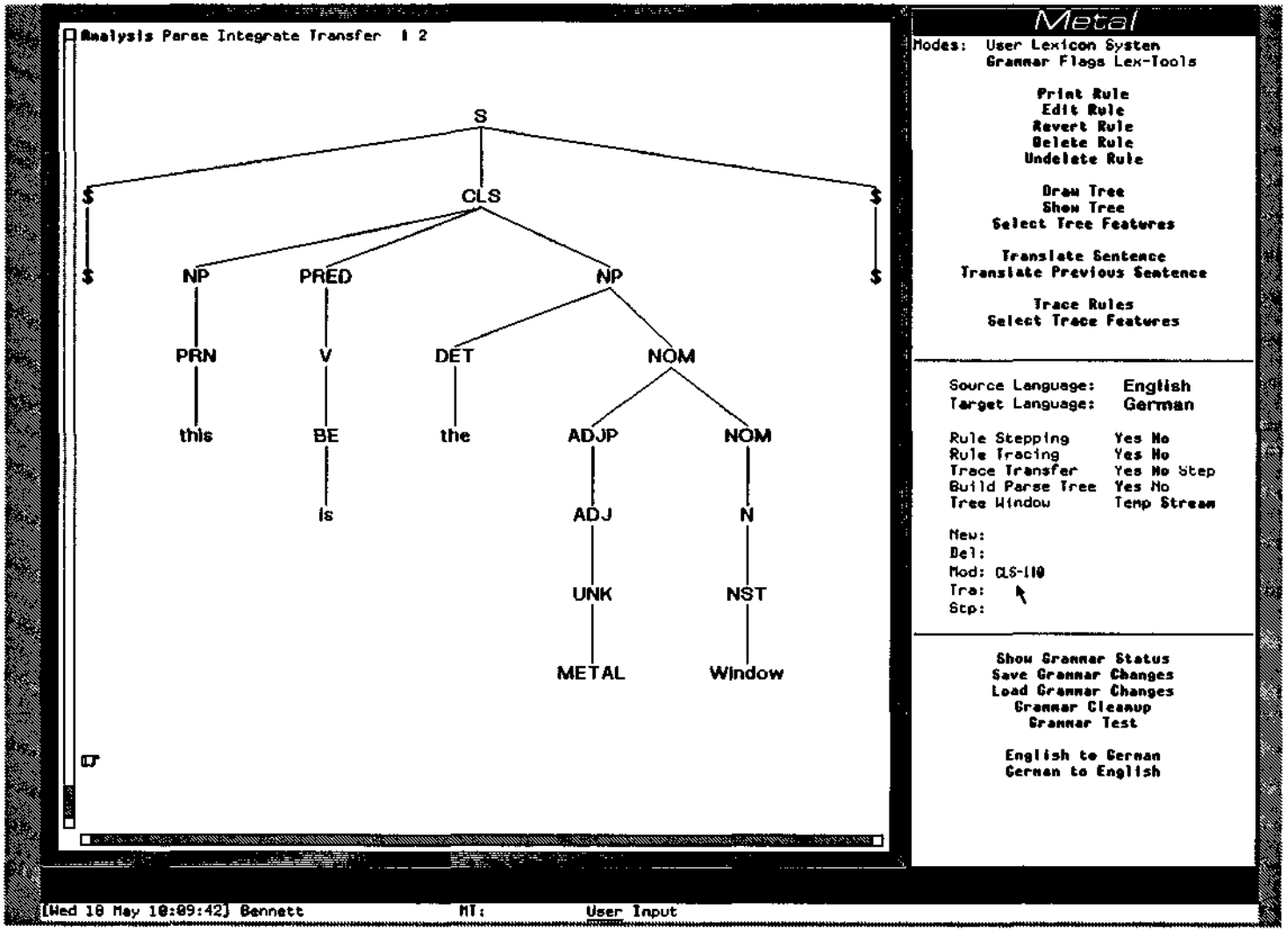
Within the METAL Window environment, there are any number of tools to aid the linguists in grammar development. The collection of tools (called 'METAL Shop') includes facilities for tracing the operation of part or all of the grammar in analysis and/or synthesis, for drawing phrase structure trees, for displaying the contents of selected nodes, and for comparing phrase structure trees, among other things. An example of the tree drawing capability is shown on the picture of the METAL Window, above. The design and implementation of METAL Shop has always been in response to the stated needs of the research linguists. The combination of the METAL Window and METAL Shop has significantly enhanced the entire research effort.

From the standpoint of METAL, devoting considerable time and energy to the development of research facilities was and is an important methodological consideration.

4. Multinational Research Groups

METAL is a multinational project with research and development sites in Munich, Germany; Barcelona, Spain; and Leuven, Belgium, as well as in Austin, Texas. Each site is responsible for development of particular language pairs, but each is also responsible for the theoretical development of the METAL framework. While the situation for METAL is hardly as complex as that of EUROTRA, the fact that research and development is being accomplished at four sites creates certain potential problems. Several methodological considerations for METAL, then, come from its fairly unique situation.

Each site is essentially autonomous from the other sites as far as the work on particular language pairs is


Analysis Parse Integrate Transfer 1 2
S
$
CLS
$
$
NP
PRED
NP
$
PRN
V
DET
NOM
this
BE
the
ADJP
NOM
is
ADJ
N
UNK
NST
METAL
Window
Metal
Modes: User Lexicon Systen
Grammar Flags Lex-Tools
Print Rule
Edit Rule
Revert Rule
Delete Rule
Undelete Rule
Draw Tree
Show Tree
Select Tree Features
Translate Sentence
Translate Previous Sentence
Trace Rules
Select Trace Features
Source Language: English
Target Language: German
Rule Stepping Yes No
Rule Tracing Yes No
Trace Transfer Yes No Step
Build Parse Tree Yes No
Tree Window Temp Stream
New:
Del:
Mod: CLS-110
Tra:
Stp:
Show Grammar Status
Save Grammar Changes
Load Grammar Changes
Grammar Cleanup
Grammar Test
English to German
German to English
[Wed 10 May 10:09:42] Bennett
MT:
User Input

concerned. One restriction on autonomy, however, is that every analysis module must be able to be fit with every synthesis module and vice versa, since the model is designed to have interchangeable modules which will permit the analysis modules from German - English and from Dutch - French to be used with the corresponding synthesis modules to produce German - French and Dutch - English pairs, in addition to the original two pairs. This point does not mean that internal workings of the actual analysis and synthesis phases for a given pair must be exactly like those for another pair; rather, the points of interface, i.e., the output from the analysis phases and the input to the synthesis phases, must be compatible. Each site, then, must work within this provision.

The fact that all sites share the same software and METAL framework also imposes some constraints on the methodology for METAL research and development. Two obvious consequences of this situation are that modifications to the software and to the framework must be reported to every site and that software must be released simultaneously to each group. Typically, groups work on new software or modifications to existing software or to the framework on their own, and release the results once they are satisfied with them. So long as communications between the sites is reasonably effective, this method of operation is successful. A clear advantage to approach is that software and framework development can be divided up among the groups.

The presence of four groups on the METAL project has led to significant 'cross fertilization' within the overall project. As with many projects, METAL is always in some danger of getting into a rut, failing to see possible innovations. The addition of other groups and other language pairs has helped us avoid such problems. A specific example is the framing mechanism used originally in the German - English grammar to block ungrammatical clauses and to assign grammatical functions to the arguments of the verbal predicate. The original mechanism was very German-dependent and therefore unusable for analyzing either Dutch or French. This situation was remedied by the Belgian site who developed a much more language-independent framing mechanism for use in the analysis and synthesis of all the language pairs (see Gebruers 1988). The new framing mechanism is only one of many improvements to the overall METAL framework which can be cited as resulting from the collaboration among the sites; new ideas continue to arise at all the sites (see, for example, Meya & Vidal forthcoming and Alonso forthcoming).

The existence of four METAL project sites on two continents has both positive and negative consequences, which, in turn, have influenced the methods in which research is conducted in each group. Having multinational research sites is an area in which our joint approach to research and development efforts will continue to evolve.

5. Conclusions

Work on the METAL machine translation system represents both a research effort, seeking to develop a theoretically sound model for MT, and a development project, seeking to produce a commercially useful MT system. This dual role, coupled with the fact that work on both the framework and on particular language pairs is being carried out in four countries, has led to a number of specific methodological considerations, some of which have been discussed above.

Any discussion of methodological considerations in METAL, as in other projects, is open-ended, since such considerations must evolve with changes in the composition and thrust of the project as a whole and with ideas from other projects. Although I have focused on METAL, my hope is that the experiences from our project may contribute to more general discussions about methodological considerations in machine translation research and development

REFERENCES

Alonso, Juan A. forthcoming. A model for transfer control in the METAL system. Proceedings of COLING 88.

Bennett, Winfield S. 1987. The proper place of end-users in MT research. Proceedings of the 28th Annual Conference of the American Translators Association, 425-427.

Gebruers, Rudi. 1988. Valency and MT: Recent developments in the METAL system. Proceedings of the Second Conference on Applied Natural Language Processing, 168-175.

Margaret King (ed.). 1987. Machine translation today: The state of the art. Proceedings of the Third Lugano Tutorial. EDITS 2. Edinburgh: Edinburgh University Press.

Meya, Montserrat, and J. Vidal. forthcoming. An integrated model for the treatment of time in MT systems. Proceedings of COLING 88.

Nirenburg, Sergei (ed.) 1987. **Machine translation: Theoretical and methodological issues**. Cambridge: Cambridge University Press.

Slocum, Jonathan, *et al.* 1987. METAL: the LRC Machine Translation System. In King (1987), 319-350.

Whiffin, Lesley, and Jonathan Slocum. 1985. Machine Translation: Viewpoints from Both Sides. Working Paper LRC-85-3. Austin TX: Linguistics Research Center.

White, John S. 1987. The research environment in the METAL project. In Nirenburg (1987), 225-246.