TAUM-AVIATION: Its Technical Features and Some Experimental Results

Pierre Isabelle and Laurent Bourbeau

Département de Linguistique
Université de Montréal
C.P. 6128, Succ. A, Montréal, Qué, Canada H3C 3J7

Upon the completion of its highly successful TAUM-METEO machine translation system, the TAUM group undertook the construction of TAUM-AVIATION, an experimental system for English to French translation in the sublanguage of technical maintenance manuals. A detailed description of the resulting prototype is offered. In particular, the paper includes: a) some figures on the size of the system; b) a description of the underlying translation model (indirect approach, analysis/transfer/synthesis scheme); c) a presentation of the basic computational techniques (use of a specialized high-level metalanguage for each linguistic component); and d) some results on the evaluation of the prototype.

1 Background

1.1 Historical Notes

In 1965, with funding from the National Research Council of Canada, the CETADOL research center in computational linguistics was created at l'Université de Montréal. Around 1970, the center narrowed its focus to the problem of machine translation (MT), renaming itself TAUM (Traduction Automatique Université de Montréal). In the next few years, several MT protypes were developed: TAUM-71, TAUM-73, and TAUM-76 (Colmerauer et al. 1971, Kittredge et al. 1973, Kittredge et al. 1976).

Starting in 1973, the Canadian Secretary of State Department (Translation Bureau) assumed responsibility for funding the project, in the hope that tangible results would soon emerge. Between 1974 and 1976, TAUM produced its first practical application: the TAUM-METEO system, for the translation of weather forecasts (Chevalier et al. 1978). Since 1977, this system has been used on a daily basis for the Canadian Environment Department (Chandioux and Guéraud 1981). Its current workload represents an annual volume of 8.5 million words (Bourbeau 1984). In spite of its very narrow scope, TAUM-METEO represents an important breakthrough in MT, since it is the only system that currently produces high quality translation without the need for human revision (although approximately 20% of the input is rejected).

The AVIATION project was undertaken in 1976, even before the on-site implementation of TAUM-METEO was completed. The aim was to develop a system capable of translating aircraft maintenance manuals. Obviously, this was a more difficult challenge than translating weather forecasts. The magnitude of the task necessitated a massive infusion of new personnel and the development of a set of new metalanguages (e.g. LEXTRA, SISIF). A prototype of the TAUM-AVIATION system, restricted to hydraulic system maintenance manuals, was demonstrated in 1979.

The following year, an independent evaluation (Gervais 1980) concluded that it was not possible to envisage immediate cost-effective production using TAUM-AVIATION. This evaluation led the Translation Bureau to stop funding the TAUM-AVIATION project, and to look for a broader funding base for MT research and development in Canada. In the meantime, the TAUM Group had to be disbanded.

1.2 Languages Translated

TAUM-AVIATION is designed in such a way that a core portion of the system is independent of particular language pairs: linguistic descriptions constitute **data** for

0362-613X/85/010018-27$03.00

Table 1. Size of compiled codes and memory requirements.

Component	Size of compiled code (6-bit K/char) (a)	(b)	Runtime CM requirements (60-bit octal K/words)
Pre-processing	60	15	34
Morphological analysis	77	48	40
Source language dictionary	1958	72	40
Syntactic/semantic analysis	205	56	120
Bilingual dictionary	1919	107	40
Syntactic transfer and synthesis	113	63	56
Morphological synthesis	99	64	24
Post-processing	43	15	34

(a) compiled linguistic data; (b) compiled interpreter for linguistic data

Table 2. Metalanguage compiler sizes.

Metalanguage	Used for	Compiler size (compiled code, 6-bit K/char)
SISIF	pre- and post-processing	43
REZO	syntactic/semantic analysis	155
LEXTRA	lexical transfer	130
SYSTEMES-Q	structural transfer; syntactic synthesis	28

the system. However, from a linguistic perspective, the project was exclusively focused on English-to-French translation.

In addition, the linguistic descriptions incorporated into the system are addressed not to general language but to the particular sublanguage of maintenance manuals (Lehrberger 1982). The notion of sublanguage is presented in Kittredge and Lehrberger (1982).

1.3 Project Size

The initial staff of seven researchers in 1976 was rapidly increased to a peak of 20 people during 1979, and then slowly decreased until the project was terminated.

1.4 System Size

TAUM-AVIATION was implemented on a CYBER 173 computer, with the NOS/BE 1.4 operating system, but was designed so as to be practically machine-independent. Most components of the system are based on the following scheme: certain linguistic data (dictionaries, grammars) are compiled into an object code interpreted at run time against the input text. **Table 1** gives an idea of the size of the runtime code, together with typical memory requirements for execution. **Table 2** gives the size of the programs used to compile the linguistic data.

1.5 Size of Dictionaries

The dictionaries list only the base form of the words (roughly speaking, the entry form in a conventional dictionary). In March 1981, the source language (English) dictionary included 4054 entries; these entries represented the core vocabulary of maintenance manuals, plus a portion of the specialized vocabulary of hydraulics. Of these, 3280 had a corresponding entry in the bilingual English-French dictionary.

2 Application Environment

TAUM-AVIATION remains an experimental system. It is designed to take as input a text that is in a photocomposition-ready format; a pre-processing program stores the formatting codes, which will be reinserted in the translated text. No use is made of manual pre-editing.

The translation process is fully automatic. If desired, it can be interrupted after dictionary lookup to obtain a list of unidentified words, and enter any such words in the dictionary.

Revision of the machine output is normally necessary: the domain is too complex for results comparable to those of TAUM-METEO. The designers of the system decided not to rely heavily on "fail-soft" strategies such as constraint relaxation or partial parses; these strategies make the quality of the output totally unpredictable. Thus, the material passing through the system is translated relatively well (very well by MT standards), and the revisor is less likely to feel overwhelmed by linguistic garbage. The price to be paid is a failure to produce any output for a relatively high proportion of the input sentences (somewhere between 20 and 40 per cent, at the stage of development reached in 1981). For a sample of translations produced by TAUM-AVIATION, see the Appendix.

The development of TAUM-AVIATION has not been taken far enough for a definitive assessment to be made of the linguistic and computational strategies that it embodied: the total system throughput was approximately 100,000 words.

3 General Translation Approach

The TAUM-AVIATION system is based on a typical second generation design (Isabelle et al. 1978, Bourbeau 1981). The translation is produced *indirectly,* by means of an analysis/transfer/synthesis scheme. The internal organization of the major components of the system is based on the notion of linguistic level. Finally, the linguistic data are generally separated from the algorithmic specifications.

3.1 Transfer Model

The overall design of the system is based on the assumption that translation rules should not be applied directly to the input string, but rather to a formal object that represents a structural description of the content of this input. Thus, the source language (SL) text (or successive fragments of it) is mapped onto the representations of an **intermediate language**, (also called **normalized structure**) prior to the application of any target language-dependent rule.

No one knows how to construct a universal, language-independent semantic interlingua. The intermediate language used in the TAUM-AVIATION system is largely language dependent: it consists of semantically annotated deep structures for SL and TL sentences. A certain degree of language independence is attained by the use of a common "base component" (a context-free grammar that enumerates the admissible deep structures) for both SL and TL. But the lexical items are left intact, and a transfer module is used in order to map the lexical items of SL onto those of TL.

3.2 Linguistic Organization

The arrangement of the system into three major modules (analysis, transfer, synthesis) reflects a theoretical model of translation operations: it is claimed that these operations take place at a "deep" level, between language-dependent meaning representations. Moreover, each one of the three modules is arranged internally along the lines of a linguistic theory: the components of these modules correspond to the standard levels of linguistic description (lexicon, morphology, syntax, semantics). This contrasts with older systems, the structure of which frequently had no direct relationship to any definite theory of language and translation.

Figure 1 shows the internal structure of the TAUM-AVIATION system.

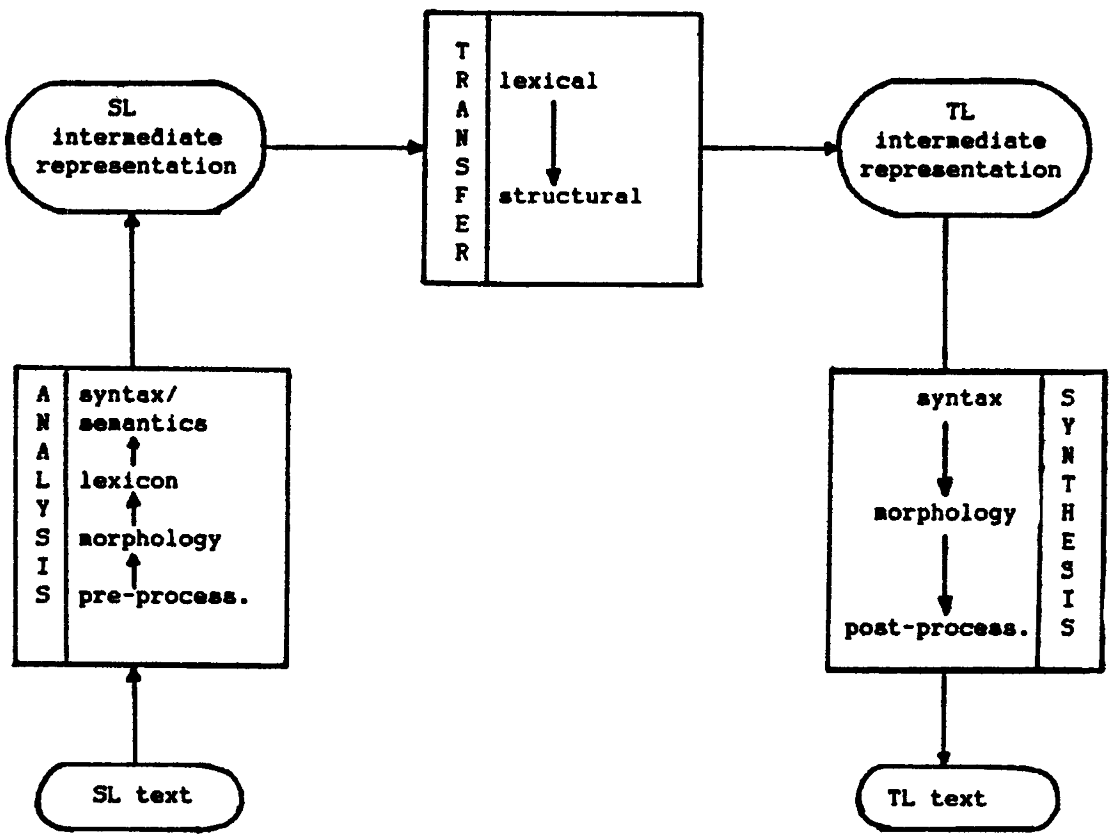

Figure 1. Overall Structure of the TAUM-AVIATION System.

4 Linguistic Techniques

4.1 Processing Units

It is well known that some translation problems can only be solved trough textual, as opposed to sentential, processing. However, we still know too little about discourse analysis techniques to use them effectively in large-scale systems. Thus, the processing unit in TAUM-AVIATION is the sentence.

Fortunately, anaphoric pronouns are quite rare in technical manuals. A more frequent problem is the use of anaphoric definite noun phrases. Consider for example the text fragment in (1):

(1) Remove hydraulic filter bypass valve. This valve is located below accumulator No. 1.

A word like *valve* cannot be translated correctly in isolation. Depending on the type of valve, French will use *clapet, robinet, soupape,* etc. In the second sentence of (1), the word *valve* is used anaphorically. To translate it correctly, one has to refer to its antecedent: the modifier *bypass* determines a specific French equivalent.

TAUM-AVIATION cannot solve problems of this type. The designers of the system preferred to concentrate their efforts on the best possible sentential analysis. And in fact, in spite of a relatively advanced sentence analyzer, translation failures due to weaknesses in sentential processing (e.g. scoping problems for conjunctions, nominal compounds, etc.) turned out to be *much more frequent* than failures due to anaphor problems, as evidenced by the error compilations of Lehrberger (1981).

4.2 The Analysis Module: Ambiguity Problems

Ambiguity is a language-internal phenomenon and it is the responsibility of the analysis module to resolve it. Sometimes, it is possible to ignore certain ambiguities, in the hope that the same ambiguities will carry over in translation. This is particularly true in systems like TAUM-AVIATION that deal with only one pair of closely related languages. The difficult problem of prepositional phrase attachment, for example, is frequently bypassed in this way. Generally speaking, however, analysis is aimed at producing an unambiguous intermediate representation.

The analysis module comprises four components: preprocessing, morphology, lexicon and syntactic/semantic analysis (see Figure 1). The pre-processing component segments the input text into successive words and into processing units. In this latter function, it can be seen as a degenerate text grammar. Because this is carried out deterministically, without interaction from the other components, segmentation problems occasionally arise.

Morphological analysis includes complete rules and exception lists for English inflectional morphology, category assignment rules for numbers and rules for dealing with unknown words. No rules are provided for derivational morphology. The system handles some types of compositional morphology, but this is done in the syntactic component, since compounds frequently exhibit properties that are otherwise thought of as syntactic; for example, internal conjunction is possible (e.g. *four- and six-cell batteries*).

Syntactic and semantic analysis are very tightly integrated in the TAUM-AVIATION system. First, both of them are implemented using the same metalanguage, a particular version of Wood's ATNs (see section 5, below). Second, both components interact freely during analysis. It is nevertheless convenient to describe them separately.

4.2.1 Syntax

The TAUM-AVIATION system includes a large-scale grammar of English capable of handling most constructions that occur with some frequency in the sublanguage of maintenance manuals (Lehrberger 1982). The rules are based on an extensive lexical subcategorization scheme: 12 standard categories are further subclassified using more than 75 features (excluding morpho-syntactic features). This is in addition to the use of lexical "strict subcategorization" frames comparable to those of transformational grammar.

Since the intermediate representation used for transfer is a type of semantically annotated "deep structure", and since maintenance manuals make use of a very complex syntax, it was necessary to provide the parser with a rich transformational component. Thus, the inverses of several transformations from standard transformational theory are used: passive, extraposition, raising, etc.

In dealing with texts as complex as technical manuals, the parser is faced with difficult ambiguity problems. Ambiguities are already present in the input to the parser, at the lexical level. These ambiguities may concern the syntactic properties of the lexical element (e.g. *light* is a noun, a verb, or an adjective); or they may concern primarily its semantic properties: pure homographs like the two nouns *lead* or polysemous items like the noun *line*.

The parser will as a side effect eliminate some lexical ambiguities; for example, if *Check valve* is to be taken as a sentence, syntax tells us that *check* must be a verb. However, the parser will itself introduce *structural* ambiguities, owing to the existence of syntactically undetermined choice points in the application of grammar rules. Two examples of structural ambiguity are adjective scope as in (3), and conjunction scope, as in (4).

(3) a) (liquid oxygen) tanks
a′) ?? liquid (oxygen tanks)
b) correct (oil level)
b′) ?? (correct oil) level

(4) a) (pressure and return) lines
a′) ?? pressure and (return lines)
b) jack and (jacking adapter)
b′) ?? (jack and jacking) adapter

These examples show that with ADJ NOUN NOUN sequences and NOUN CONJ NOUN NOUN sequences, two different syntactic groupings are possible. But only one of them is semantically acceptable and results in a correct translation.

Moreover, some lexical ambiguities, instead of being eliminated in the parsing process, will constitute a further source of structural ambiguity, each reading of the relevant lexical item being compatible with a different syntactic structure. In example (5), *drain* can be taken either as a noun or as a verb, when appropriate adjustments are made to the surrounding syntactic structure.

(5) Remove dust cap and drain plug.

Thus by itself, a syntactic parser produces a highly ambiguous output, and further constraints are needed in a practical MT system.

4.2.2 SEMANTICS

Semantic processing in the TAUM-AVIATION system performs two related tasks: a) it filters the syntactic structures, eliminating as many ambiguities as possible; and b) it associates with each node of the tree a set of semantic features which will be used by transfer rules.

Most semantic features originate in the dictionary, where lexical items are described in terms of some 35 features that form a **tangled hierarchy.** Predicative lexical items (verbs, adjectives, certain prepositions) are assigned selectional restrictions on their possible arguments in terms of these semantic features.

Selectional restrictions constitute the main semantic mechanism used by the system to eliminate ambiguities of two types:

a. structural ambiguities introduced by syntactic rules; thus the spurious structure proposed by the parser for (5) is eliminated because the verb *drain* does not accept as direct object something in the semantic category of *plug.*
b. lexical ambiguity in the semantic properties of certain lexical items; polysemous words like the noun *line* (which can denote either an abstract geometrical object, or physical objects such as conductors) are frequently disambiguated by selectional restrictions; for example, in *Flush the line,* the concrete sense is selected.

In order for selectional restrictions to work properly and for trees to be correctly annotated, it is necessary to apply **semantic projection rules** which assign sets of features to tree nodes. In TAUM-AVIATION, the semantic rules work in a compositional fashion, raising selectively certain features from daughter nodes to their mothers (Isabelle 1985). Rules such as the following are used:

- all of the semantic features of a headnoun are raised onto the dominating NP node;
- the intersection of the features of two conjoined NP nodes is raised onto the dominating NP node; and
- when the headnoun is a partitive noun (e.g. *portion*), and the NP has an *of NP* complement, the features of this complement are raised onto the dominating NP node.

The system also makes use of standard **control rules** for subjectless infinitives and gerundives, and of some pronoun/antecedent rules, in order to enforce semantic constraints wherever possible.

Semantic ambiguity, whether real homography (e.g. the two nouns *lead*) or polysemy (e.g. the various senses of the noun *line*), is not handled by creating multiple entries in the source language dictionary. Rather, in its single entry, the word is assigned a number of *semantically incompatible* features. The semantic rules seek to filter out some of these features, so that no incompatibility remains. This strategy prevents the redundant syntactic search that results from a multiple-entry strategy.

4.3 THE TRANSFER MODULE

In principle, transfer rules state correspondences between two sets of unambiguous structural descriptions. Their most obvious task is to relate the lexical items of SL to those of TL. Even if the rules are applied to unambiguous lexical elements, the correspondences are by no means one-to-one: the lexical system of each natural language reflects a specific way of breaking down the conceptual universe. For this reason, equivalences have to be stated in terms of structural patterns rather than in terms of words or strings of words.

To take an example, there is no language-internal evidence that *hard* is ambiguous in English; however, depending on the context, it is translated into French as *difficile, dur,* etc. The French equivalents have more restricted collocations. In all those cases, transfer rules are needed to select the contextually appropriate equivalent.

Moreover, very frequently, these **lexical transfer rules** cannot simply substitute lexical items, leaving the tree structure unaffected. Since SL and TL lexical items frequently have different contextual requirements (i.e. subcategorization frames), translation rules have to establish correspondences between a source and a target structural pattern, as illustrated by the examples in (6).

(6) a. check x against y → comparer x à y
b. supply x with y → fournir y à x
c. cantilever x → monter x en porte-à-faux
d. bond x electrostatically → métalliser x
e. service x → faire l'entretien de x

It is clear that lexical transfer rules must include powerful transformational mechanisms. This basic fact has not so far received the attention it deserves in the MT community. The TAUM-AVIATION system provides for full transformational power at the level of lexical transfer (Chevalier et al. 1981).

The transfer component also involves rules for **structural transfer,** that is, rules that deal with linguistic contrasts not tied to any specific lexical item. Since the same base rules are used for SL and TL, this sub-component is kept to a minimum. Nevertheless, a number of structural differences between SL and TL have to be accounted for by means of contrastive rules. For example, because the intermediate language does not provide for "universal semantic tenses", the tense systems of SL and TL have to be explicitly contrasted by a set of rules.

Another task left to structural transfer is to deal with observable contrasts concerning the use of **optional movement transformations.** In all likelihood, the use of these transformations is governed by discourse phenomena that the system does not attempt to analyze. The strategy used in TAUM-AVIATION is to take advantage of the frequent parallelisms between SL and TL regarding these aspects of surface structure organization. Thus, the intermediate representation retains "traces" from SL surface structure used by the synthesis component to maintain a certain parallelism with SL. However, in some cases we know that the two languages exhibit systematic differences in their use of certain movement transformations. The structural transfer grammar describes these facts. For instance, TAUM-AVIATION includes complex rules for translating English passives with various French constructions, as illustrated in the following examples:

(7) Quick-disconnect fittings should not be removed.
→ Ne pas enlever les raccords à démontage rapide.

(8) Ensure that pump and lines are bled.
→ S'assurer qu'on a purgé la pompe et les canalisations.

(9) The flaps are operated by hydraulic system no. 1.
→ Le circuit hydraulique no. 1 actionne les volets.

4.4 THE SYNTHESIS MODULE

Synthesis of the TL text involves three steps: syntactic synthesis, morphological synthesis, and post-processing.

Syntactic synthesis is carried out on the basis of a large-scale transformational grammar of French. Since the input to the synthesis component is normally a well-formed unambiguous sentential deep structure, synthesis here is much simpler than analysis. This is not to say that synthesis of natural language texts is generally easy. Generating a coherent text from an abstract discourse representation is certainly a very difficult problem. But in TAUM-AVIATION, synthesis can only be achieved on a sentential basis. Therefore, no attempt can be made to describe the complex discourse factors that influence sentence generation (e.g. application of "optional" movement transformations). As mentioned in the previous section, the strategy adopted is to try to preserve a certain parallelism with the SL sentences, since both languages have relatively similar means of expressing discourse cohesion.

Syntactic synthesis produces a string of lexical items annotated with all the information required to inflect them correctly. The morphological synthesis component then determines the final form of each word. This is done on the basis of an exhaustive description of the rules of French inflection (together with their exceptions). Post-processing reformats the TL text, making use, wherever possible, of the formatting codes of the SL text.

5 COMPUTATIONAL TECHNIQUES

From the computational point of view, the TAUM-AVIATION system is more complex than TAUM-METEO, which is entirely written in the Q-SYSTEMS metalanguage (Colmerauer 1971). One of the ideas underlying TAUM-AVIATION is to make use of specialized tools for different tasks in the interests of increased efficiency, though somewhat at the expense of overall simplicity.

In the implementation, the actual modules closely match the components of the linguistic model presented in Figure 1. They are applied sequentially and communication between components is achieved by means of a **chart structure** (a type of loop-free graph). The arcs of these charts are labelled with tree structures whose nodes are labelled with complex symbols: a categorial label plus a set of features.

Most components are based on the following scheme. Certain linguistic data are described with a high-level metalanguage; in this metalanguage, the linguist expresses facts about tree structures. These descriptions are compiled into an abstract formal structure interpreted at run time against the material to be translated. Most of these compilers and interpreters are written in PASCAL.

5.1 PRE- AND POST-PROCESSING

These relatively simple components, which map character strings onto sequences of chart structures and vice-versa, are implemented as sets of rules in a metalanguage called SISIF; a set of SISIF rules amounts to a deterministic finite-state automaton. These rules are compiled into list structures, which are interpreted against the input text at run time.

5.2 INFLECTIONAL MORPHOLOGY

Since it was possible to exhaustively describe the inflectional morphology of both French and English, there was no compelling reason to use a very high-level formalism. Consequently, in the interests of efficiency, two PASCAL programs were written for morphological analysis of English and morphological synthesis of French.

5.3 DICTIONARIES

5.3.1 SOURCE LANGUAGE DICTIONARY

A dictionary system called SYDICAN enables the linguist to write lexical rules that associate a complex of lexical information with a string of base forms, forming a path in an input chart. Two types of rules are provided: a) rules that simply add a new path (labelled with the complex of lexical information) to the chart; and b) rules that, in addition, have the effect of taking precedence over shorter matches in the chart (cf. "longest-match" strategy).

The rules are compiled into list structures. At run time, they are retrieved from an arbitrarily large lexical data base and applied to the chart. The lexical database system includes some maintenance facilities, such as integrity constraints on its contents, and facilities for retrieving entries through arbitrarily complex requests on their contents.

5.3.2 BILINGUAL DICTIONARY

We saw in 4.3 that lexical transfer involves rules that perform complex transformations on tree structures. The LEXTRA metalanguage makes it possible to associate with any lexical item an arbitrarily complex set of tree transformations. These transformations describe a pattern (anchored in the relevant lexical item), which is to be matched against the tree structure at run time. When a match is found, a series of associated actions specifying structural changes is performed.

An important idea embodied in LEXTRA is that a transfer component should have an explicit description of the intermediate language. In the TAUM approach this intermediate language is partially defined by a set of context-free rules that describe a common base component for SL and TL. LEXTRA takes as **data** this context-free grammar and guarantees that any manipulated tree structure corresponds to a permissible derivation in terms of that context-free grammar. This notion is to be related to computational formulations of transformational grammar such as Petrick (1973), where the deep structures produced by the inverse transformations are checked against the rules of the base component. No equivalent check is performed with parsing systems like ATNs.

LEXTRA rules are compiled into list structures. It was found that some of the constraints on admissible tree structures could be enforced at compile time (Gérin-Lajoie 1980). This mechanism is very useful in the complex task of dictionary development. It helps validate the work of the lexicographer. At run time, the LEXTRA interpreter searches the tree structure for SL lexical items, retrieves the associated lexical rules and applies them to the tree.

5.4 SYNTAX AND SEMANTICS

5.4.1 ANALYSIS

The English grammar for syntactic analysis is written in REZO (Stewart 1975, 1978), TAUM's version of augmented transition networks (ATNs). The REZO metalanguage is different from Wood's ATNs (Woods 1970) in several respects. Some of the differences are:

- REZO does not support morphological analysis, which is performed in a separate component;
- tree nodes are complex symbols that include sets of features on which boolean operations can be performed;
- REZO includes a number of primitives to perform pattern matching over tree structures;
- in addition to regular ATN states where all transitions are tried, REZO includes "deterministic" states where only the first transition whose test is met is followed;
- REZO accords special status to the states to which a recursive call can be made, so that the resulting grammar is a collection of sub-networks.

The REZO grammar is compiled into a set of instructions for a virtual machine, which is simulated by the runtime interpreter. Parsing is done in the usual top-down, depth-first, left-to-right, serial manner. The interpreter can either work in an all-paths or in a first-path mode. One important difference from Wood's ATN interpreter is that REZO takes as input a chart structure in which lexical ambiguities are encoded and applies the grammar in parallel to all the paths of this chart. The result is also a chart structure: REZO is thus a chart-to-chart transducer.

In 4.2, it was mentioned that syntactic rules create structural ambiguities, and that semantic processing can eliminate some of these. Serial parsing provides another means of selecting a particular reading. Since the transitions of the REZO networks are followed in a fixed order, the grammar can be made to produce the most likely reading first. In TAUM-AVIATION's analysis grammar, the ordering of the transitions reflects:

- general parsing principles such as those discussed in human performance studies (e.g. Kimball 1973); and
- sublanguage-specific statistical tendencies.

5.4.2 STRUCTURAL TRANSFER AND SYNTHESIS

Structural transfer and syntactic synthesis are implemented in the well-known Q-SYSTEMS, which we will not describe here. This introduces some heterogeneity into TAUM-AVIATION, since: a) unlike the other metalanguages, Q-SYSTEMS do not support trees with complex symbols as node labels; and b) the compiler and the interpreter are written in FORTRAN while PASCAL is used for the other metalanguages.

In fact, the original design of the system included provisions for a new metalanguage well suited to synthesis; but time constraints precluded its development.

6 Experimental Results

6.1 COST-BENEFIT EVALUATION

In 1981, the sponsor submitted TAUM-AVIATION to a cost-benefit evaluation, in order to determine if the system was usable in a production environment. This evaluation made by an independent consultant is reported in Gervais (1980), and we will only summarize the main conclusions.

Raw machine output was deemed to have a degree of intelligibility, fidelity, and style that reaches 80% of unrevised human translations (HT).

Table 3. MT/HT: Compared costs per word ($CDN).

Tasks	MT	HT
Preparation/input	$0.014 (8%)	–
Translation	$0.079 (43%)	$0.100 (69%)
Human Revision	$0.068 (37%)	$0.030 (21%)
Transcription/Proofreading	$0.022 (12%)	$0.015 (10%)
TOTAL	$0.183	$0.145

Revised MT and revised HT have a comparable degree of quality, but revision costs are twice as high for MT; thus, globally, revised MT turns out to be more expensive than revised HT as shown in **Table 3.** However, it is noted in the evaluator's report that MT reduces by half the human time required in the translation/revision process.

The direct costs of MT could probably be reduced to an acceptable level, for example by interfacing the system with a suitable word-processing environment and by reducing the percentage of sentences for which no translation is produced (Isabelle 1981).

Cost-effective production would require the system to be applicable to at least 6 million words per year. In order to reach that target, the system would have to be extended to translate domains other than hydraulics. But the indirect costs involved in these extensions (e.g. dictionary development) are very high. Gervais concludes that it is impossible to assert that translation using TAUM-AVIATION would be globally cheaper than human translation.

6.2 TECHNICAL EVALUATION OF PERFORMANCE

Cost-benefit evaluations are certainly necessary, but a single evaluation of this type tells one very little about how the system can be expected to perform on different texts, or after further investment. TAUM developed a methodology for analyzing the performance of an MT system through a systematic examination of its translation errors.

The first step is to collect all the errors in the translation of the sample text; translators/revisors then have the responsibility of deciding what is to be counted as an error. A classification scheme for translation errors will include headings such as the following: incorrect TL equivalent for a word, incorrect word order, lack of an article, etc.

In itself, an absolute number of such errors for a given text is not very revealing; but a comparison of the ratio of errors to word tokens in different texts, or at various stages of development of the system is an initial source of useful information.

Still, from the point of view of system development, these "surface" translation problems are merely symptoms for problems in some component of the system. To provide an answer to questions such as:

- how many of these problems have a known solution?
- how long would it take to correct them?
- how much better would the performance of the system be after n person/months of work?
- what should the priorities be?

it is necessary to identify, for each surface problem, one or several causes in the functioning of the system. For example, the fact that, in the translation of a given sentence, a French adjective is incorrectly inflected could be caused by one or more of the following factors:

- incorrect marking in a dictionary entry;
- mistake in TL syntactic rules for agreement;
- incorrect scoping in SL analysis (e.g. give the wrong bracketing to an ADJ NOUN NOUN sequence); or
- absence of relevant marking in SL (e.g. when translating *federal and provincial governments* into French, should one pluralize the adjectives?).

A sophisticated error classification grid was developed, so that the sources of translation errors could be investigated in a coherent and meaningful way. Basically, this error grid reflects the internal organization of the system, so that translation errors can be assigned a precise cause in the operation of the system.

Once a coherent scheme is available, one can proceed with the classification of the translation errors found in the sample text. This classification process is difficult and tedious, but it is crucial that it be done with accuracy and consistency. Frequently, one has to follow "execution traces" to discover the exact source of a given error.

The final step is to look at the possible remedies for the problems that have been identified. A careful examination of each problem source will reveal whether or not there is a known way of eliminating it, and if so, what amount of effort is needed.

If this type of technical evaluation can be carried out at successive stages of development (with both old and new texts), one gets a clear picture of the evolution of the system. The figures obtained will reveal whether or not:

- there has been substantial progress compared to previous stages;
- an asymptote has been reached in the curve investment/improvement.

The same figures will also help determine development and research priorities.

This sort of technical evaluation was applied twice to TAUM-AVIATION in the final year of the project; only a few person-months of development had been invested in the system between the two tests. The main goal was to see how well a system developed on the basis of corpora from the domain of hydraulics would fare in the domain of electronics. Some results were as follows:

- In both tests, more than 70% of the failures were classified as having a known solution; the vast majority of these could be corrected in 12 person/months of work.
- From a syntactic point of view, there is no notable difference between hydraulics and electronics. In fact, as a result of a minimal effort in correcting some problems discovered in the test on hydraulics, the overall performance of the parser turned out to be better in the electronics test.
- As expected, there was a major dictionary problem in going from one domain to the other. Selectional restrictions as assigned for hydraulics worked so poorly that they did more harm than good to the final result.

A definitive assessment of the linguistic and computational techniques on which TAUM-AVIATION is based would have required a few more applications of this evaluation/correction cycle.

7 Future Directions

Before TAUM was disbanded, Isabelle 1981 voiced the views of the group on a possible course for short- and long-term machine translation R&D activities.

The difficulties encountered in the AVIATION project convinced the Translation Bureau that a more permanent and broader R&D base would be required for MT to be viable in Canada. In 1983, the Translation Bureau in conjunction with the Canadian Department of Communications, funded a large-scale study to review natural language processing technologies and examine opportunities for Canada in this field.

The consultants have submitted their report and the two Departments involved now have to determine the best way to implement the recommendations that are made therein. In the area of MT, there would appear to be three fronts on which R&D could be pursued:

- the development and integration of various computer aids to human translation within a translator's workstation;
- the application of second-generation MT technology to promising sublanguages (Kittredge 1983);
- research on third-generation MT technology.

References

Bourbeau, Laurent 1981 Linguistic Documentation of the TAUM-AVIATION Translation System. Groupe TAUM, Université de Montréal, Montréal, Canada.

Bourbeau, Laurent 1984 Transfert du système METEO sur micro-ordinateur: étude de faisabilité. Bureau des traductions, Secrétariat d'Etat, Ottawa.

Chandioux, John and Guéraud, Marie-France 1981 METEO: un système l'épreuve du temps. *META* 26(1): 18-22.

Chevalier, Monique; Dansereau, Jules; and Poulin, Guy 1978 TAUM-METEO: description du système. Groupe TAUM, Université de Montréal, Montréal, Canada.

Chevalier, Monique; Isabelle, Pierre; Labelle, François; and Lainé, Claude 1981 La traductologie appliquée à la traduction automatique. *META* 26(1): 35-47.

Colmerauer, Alain 1971 Les SYSTEMES-Q: un formalisme pour analyser et synthétiser des phrases sur ordinateur. Groupe TAUM, Université de Montréal, Montréal, Canada.

Colmerauer, Alain. et al. 1971 TAUM-71. Groupe TAUM, Université de Montréal, Montréal, Canada.

Gervais, Antoni 1980 Evaluation of the TAUM-AVIATION Machine Translation Pilot System. Translation Bureau, Secretary of State, Ottawa, Canada.

Gérin-Lajoie, Robert 1980 Vérification des manipulations d'arbres en LEXTRA. M.Sc. thesis, Université de Montréal, Montréal, Canada.

Isabelle, Pierre; Bourbeau, Laurent; Chevalier, Monique; and Lepage, Suzanne 1978 TAUM-AVIATION: description d'un système de traduction automatisée de manuels d'entretien en aéronautique. COLING-78, Bergen, Norway.

Isabelle, Pierre 1981 Some Views on the Future of the TAUM Group and the TAUM-AVIATION System. Groupe TAUM, Université de Montréal, Montréal, Canada.

Isabelle, Pierre 1985 Machine Translation at the TAUM Group, In: King, Margaret, Ed., *Machine Translation: the State of the Art.* Edinburgh University Press, Edinburgh.

Kimball, John 1973 Six or Seven Principles of Surface Structure Parsing in Natural Language, Indiana University Linguistics Club, Bloomington.

Kittredge, Richard 1983 Sublanguage-Specific Computer Aids to Translation: A Survey of the Most Promising Application Areas. Translation Bureau, Secretary of State, Ottawa, Canada.

Kittredge, Richard et al. 1973 TAUM-73. Groupe TAUM, Université de Montréal, Montréal, Canada.

Kittredge, Richard; Bourbeau, Laurent; and Isabelle, Pierre 1976 Design and Implementation of a French Transfer Grammar. COLING-76, Ottawa, Canada.

Kittredge, Richard and Lehrberger, John 1982 *Sublanguage: Studies of Language in Restricted Semantic Domains.* Walter de Gruyter, New York.

Lehrberger, John 1981 Possibilités d'extension du système TAUM-AVIATION. Groupe TAUM, Université de Montréal, Montréal, Canada.

Lehrberger, John 1982 Automatic Translation and the Concept of Sublanguage. In: Kittredge and Lehrberger (1982): 81-106.

Petrick, Stanley 1973 Transformational Analysis. In: Rustin, R. Ed., *Natural Language Processing.* Algorithmic Press: 27-41.

Stewart, Gilles 1975 Le langage de programmation REZO. M.Sc. thesis, Université de Montréal, Montréal, Canada.

Stewart, Gilles 1978 Spécialisation et compilation des ATN: REZO. COLING-78, Bergen, Norway.

Woods, William A. 1970 Transition Network Grammars for Natural Language Analysis. *Communications of the ACM* 13(10): 591-606.

APPENDIX

INPUT TEXT:

HYDRAULIC PRESSURE IN-LINE RELIEF VALVE

(See figures 2-4 and 2-5.)
30 Identical interchangeable hydraulic pressure in-line relief valves (in-line relief valve) are provided for each ac hydraulic pump and for the dc hydraulic pump. The in-line relief valves are located in the hydraulic service center. Those for the No. 1, No. 1A, and No. 1B ac and dc hydraulic pumps are on the left side next to the No. 1 service centre assembly. The in-line relief valve for the No. 2 ac hydraulic pump is incorporated in the No. 2 service centre assembly.

31 The in-line relief valves are poppet-type, spring-loaded to the closed position. A pressure of 3450 psi impinging on the poppet is sufficient to overcome the opposing spring force, and the poppet will move from its knife-edge seat.

RAW MACHINE OUTPUT:

CLAPET DE DECHARGE INCORPORE DE PRESSION HYDRAULIQUE
(Voir les figures 2-4 et 2-5.)
30 Les clapets de décharge incorporés interchangeables identiques de pression hydraulique (clapets de décharge incorporés) sont prévus pour chaque pompe hydraulique ca et pour la pompe hydraulique cc. Les clapets de décharge incorporés sont situés dans le compartiment hydraulique. Ceux pour les pompes hydrauliques ca et cc no 1, no 1A et no 1B sont du côté gauche à côté du bloc collecteur no 1. Le clapet de décharge incorporé pour la pompe hydraulique ca no 2 est intégré au bloc collecteur no 2.

31 Les clapets de décharge incorporés sont champignon, sont rappelés par ressort à la position fermée. Une pression de 3450 psi s'exerçant sur le clapet-champignon est suffisante pour vaincre la force de rappel du ressort et le clapet-champignon se déplacera de son siège en couteau.

HUMAN REVISION:

CLAPET DE DECHARGE INCORPORE DE PRESSION HYDRAULIQUE
(Voir les figures 2-4 et 2-5.)
30 *Des* clapets de décharge incorporés interchangeables identiques de pression hydraulique (clapets de décharge incorporés) sont prévus pour chaque pompe hydraulique ca et pour la pompe hydraulique cc. Les clapets de décharge incorporés sont situés dans le compartiment hydraulique. Ceux *des* pompes hydrauliques ca et cc no 1, no 1A et no 1B sont du côté gauche à côté du bloc collecteur no 1. Le clapet de décharge incorporé pour la pompe hydraulique ca no 2 est intégré au bloc collecteur no 2.

31 Les clapets de décharge incorporés, *du type* champignon, sont rappelés par ressort à la position fermée. Une pression de 3450 psi s'exerçant sur le clapet-champignon est suffisante pour vaincre la force de rappel du ressort et le clapet-champignon *s'écartera* de son siège en couteau.