ISRN SICS-T—91/20—SE

Transfer Through Quasi Logical Form – A New Approach to Machine Translation

by
Hiyan Alshawi, David Carter,
Björn Gambäck, Steve Pulman and Manny Rayner

SICS technical report
T91:20
ISSN 1100-3154

ISRN SICS-T—91/20—SE

Transfer Through Quasi Logical Form – A New Approach to Machine Translation

Hiyan Alshawi, David Carter,
Björn Gambäck, Steve Pulman and Manny Rayner

1991

SICS technical report T91:20
ISSN 1100-3154

Swedish Institute of Computer Science
Box 1263
S–164 28 KISTA
Stockholm, Sweden

SRI International
Cambridge Computer Science Research Centre
23 Millers Yard, Cambridge CB2 1RQ
Cambridge, England

Abstract

This document is an introduction to a research project aimed at producing a prototype system for on-line translation of typed dialogues between speakers of different natural languages. The work was carried out jointly by SICS and SRI Cambridge. The resulting prototype system (called Bilingual Conversation Interpreter, or BCI) translates between English and Swedish in both directions.

The major components of the BCI are two copies of the SRI Core Language Engine, equipped with English and Swedish grammars respectively. These are linked by the transfer and disambiguation components. Translation takes place by analyzing the source-language sentence into Quasi Logical Form (QLF), a linguistically motivated logical representation, transferring this into a target-language QLF, and generating a target-language sentence.

We believe that the project was successful in demonstrating the feasibility of using these techniques for interactive translation applications, and provides a sound basis for development of a large scale message translator system. The final section of the paper points to several possible follow-on projects aimed in the direction of practically usable commercial systems.

Acknowledgments

The Bilingual Conversation Interpreter project was funded by the Swedish Institute of Computer Science.[1] The greater part of the work was completed while Manny Rayner was employed at SICS.

Björn Gambäck
`gam@sics.se`
Swedish Institute of Computer Science

Hiyan Alshawi, David Carter, Steve Pulman and Manny Rayner
`{hiyan,dmc,sgp,manny}@cam.sri.com`
SRI International, Cambridge Research Centre

[1] SICS is sponsored by the Swedish National Board for Industrial and Technical Development (NUTEK) and the members of the SICS industrial consortium (Asea Brown Boveri AB, Telefon AB LM Ericsson, Försvarets Materielverk, IBM Svenska AB, NobelTech AB, and Swedish Telecom).

Contents

... The Japanese developers seem to agree ... that the technology they are looking at is dead-ended ... I asked the head of a major Japanese laboratory whether further progress in machine translation would come about from extensions to their systems or from entirely new technology. His answer was so prompt as to be abrupt: "From new technology".

– Bernard E. Scott, Lógos Corporation,
in an open letter to *Computational Linguistics*, 1990.

1 Introduction

For the past year, SRI International and the Swedish Institute of Computer Science have pursued a joint research project intended to investigate the feasibility of constructing a sophisticated semi-automatic machine translation system by configuring and adapting existing state-of-the-art components centered around the SRI Core Language Engine. At the end of the project's first year, this has resulted in a prototype system, the "Bilingual Conversation Interpreter" (BCI), which is capable of interactive translation between English and Swedish, using a vocabulary of about 1000 words and covering a broad range of possible grammatical constructions. This paper will give an overview of the project's design philosophy and the current state of implementation of the prototype BCI; the final section will describe possible follow-on projects aimed in the direction of constructing commercially interesting products.

These include:

- Interactive translation of email messages.
- Semi-automatic speech-to-speech translation.
- Multi-lingual access to knowledge-based system applications.
- Interactive multilingual composition of manuals and other texts.

We will begin by describing the BCI's position within the field of machine translation as a whole. It is at the moment generally accepted among researchers in the field that fully automatic high-quality machine translation is not feasible as a short-term prospect, except within extremely limited subdomains. (The example *par excellence* is weather forecasts, as exemplified in the TAUM-METEO system (Thouin 1982).) Disregarding these, realistic projects must normally compromise, either by accepting low-quality output (which may subsequently be post-edited), or by allowing human interaction during the translation process to supply knowledge not directly available to the system.

We have chosen the second alternative, for the following reasons: firstly, there is a large class of applications where two monolingual humans can achieve a goal by carrying out a dialogue in real time; for the sake of concreteness, we have during the project focussed on a hypothetical application, where the BCI is being used by a Swedish car-hire firm in order to communicate with an English customer.[2] Secondly, such an architecture allows practical systems to be built at the level of the current state-of-the-art, while providing a smooth development path for future improvements. As the basis of Natural Language Processing technology improves, less human interaction is required.

2 Translation by Quasi Logical Form transfer

The central technical idea in the BCI is the concept of Quasi Logical Form transfer. Here, we have attempted to create an intelligent compromise between the opposing paradigms of "syntactic transfer" and "knowledge-based interlingua", which we will first briefly summarize.

In the syntactic transfer approach, which is in practice by far the more common one,[3] translation is carried out in three stages: the source-language text is transformed into a syntactic representation (most commonly some kind of tree-like structure), which is then *transferred* into a target-language counterpart. Finally, the target-language text is generated from its syntactic representation, as in Figure 1.

[2] A similar, though less sophisticated, system for translation between Japanese and English is reported in (Miike *et al* 1988).

[3] Syntactic transfer is e.g. used in Siemens' METAL system (White 1985) and the European community project EUROTRA (King 1985).

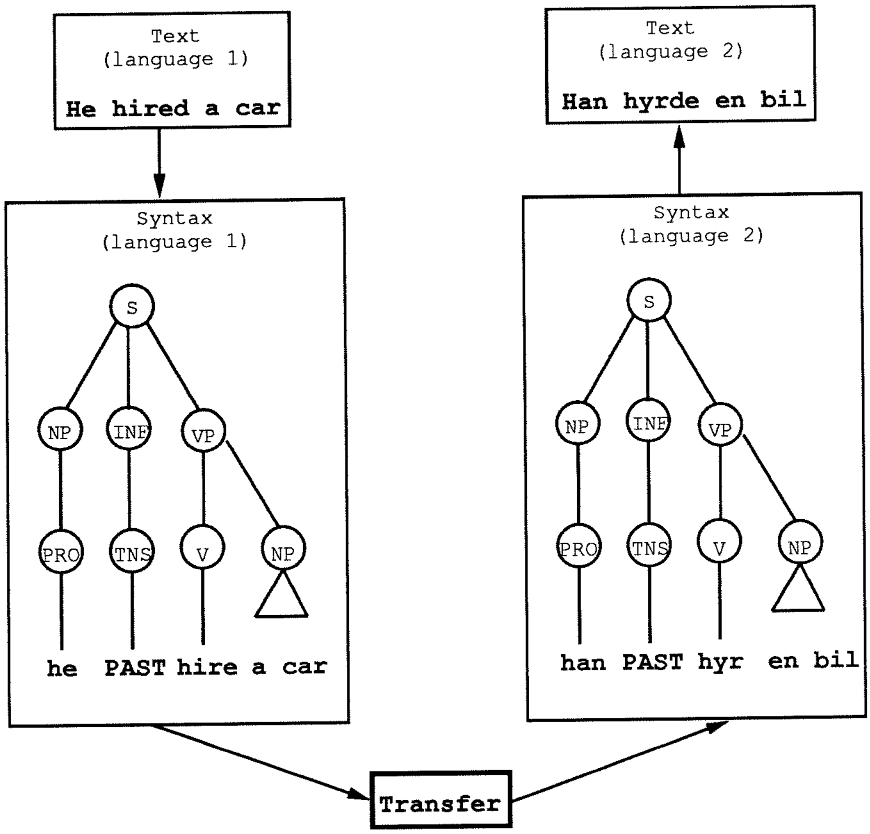

Figure 1: Syntactic transfer

The notation used in the syntax tree:
S = sentence, NP = noun phrase, VP = verb phrase, V = verb
PRO = pronoun, INF = inflection, TNS = tense, △ = subtree

Knowledge-based interlingua-based systems, in contrast, perform translation in two stages: the source text is reduced to a language-independent intermediate representation, and the target text is generated from this representation directly. The ideal picture is shown in Figure 2.

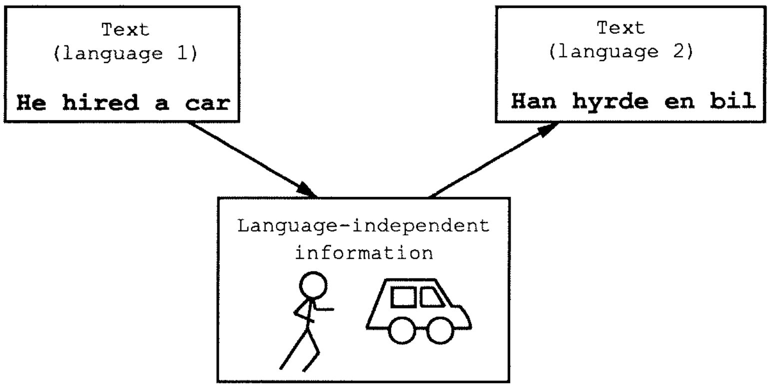

Figure 2: Knowledge-based interlingua

Very few systems are of course completely pure examples of either approach, but to describe the various intermediate positions that have been experimented with would take us rather too far afield. We will restrict ourselves to indicating in general terms the inherent strengths and weaknesses of each approach.

On the positive side, syntactic transfer is the easier alternative to implement, since the techniques of syntactic analysis (and to a lesser extent generation) are well-understood and relatively straight forward. However, the fact that different languages use widely different syntactic forms places a great burden on the transfer component, which becomes correspondingly more complex and harder to understand. To take an example from the English-Swedish language-pair: although the structures of *He hired a car* and the corresponding *Han hyrde en bil* are identical, transforming the sentence into a question already creates non-trivial problems. The Swedish inverts the word-order (*Hyrde han en bil?*), while the English introduces an auxiliary (*Did he hire a car?*), necessitating an extra rule (see Figure 3).

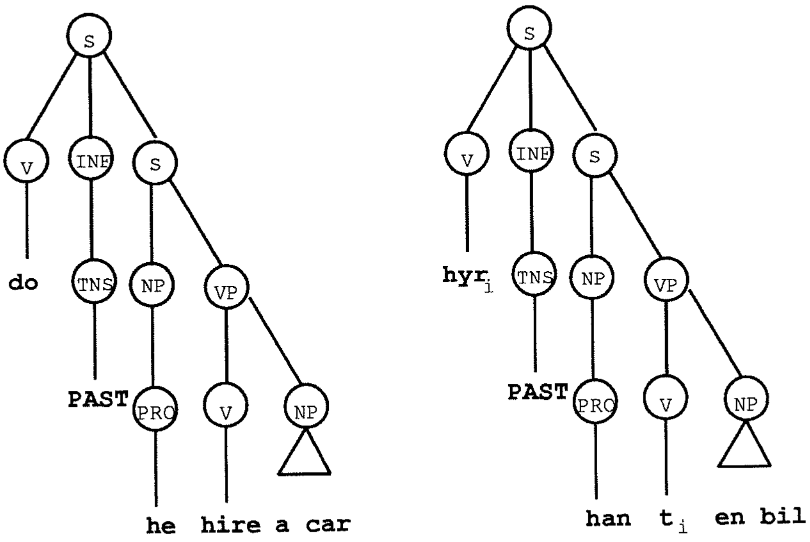

Figure 3: Problems with syntactic transfer

The problem is that the representation is too shallow to "factor out" each language's own way of forming questions. There tend to be many phenomena of this kind, which interact to form an exponentially growing set of complex transfer rules.

On the other hand, pure interlingua systems do not suffer from these problems, since the intermediate representation is not tied to any particular language. The difficulty is rather that too little is as yet known about formal knowledge-representation techniques to make it feasible to specify a robust interlingua for more than a small subset of natural language; moreover, even if the theoretical apparatus were present, transformation to a language-independent form in general requires access to vast quantities of implicit "common-sense" knowledge, the formalization of which is a Herculean task. Although interesting experimental systems have been developed,[4] it seems unlikely that they can be turned into robust products in the short- or medium-term.

Our architecture is half-way between the two positions outlined above.

[4] For example, at Carnegie-Mellon University's World Center for Machine Translation (Goodman & Nirenburg 1989).

The source text is analyzed into a representation ("Quasi Logical Form", or QLF), which has been carefully designed so as to represent exactly the aspects of linguistic meaning which do *not* involve context or "common-sense" knowledge. The source QLF representation is transferred into a target counterpart, from which target-language generation is used to produce the target text. This is illustrated in Figure 4.

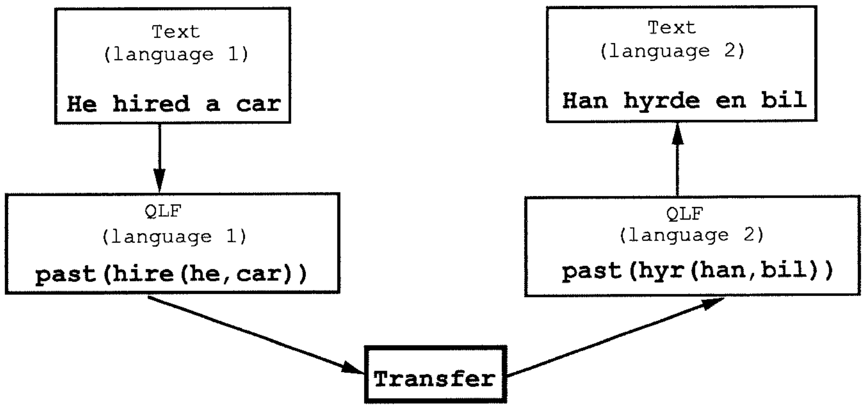

Figure 4: Transfer through Quasi-Logical Form

In other applications, such as NL query interfaces to databases, the QLF representation would be subjected to a further phase of processing which adds contextually determined factors, such as the referents of pronouns (Alshawi 1990). However, our hypothesis has been that a useful translation can be obtained by performing transfer directly on QLFs, when necessary dealing with problems of contextual interpretation by querying the user. These questions are phrased in such a way as to assume no knowledge on the source-user's part of either linguistics or the target language.

Our judgement, based on the experience gained during the first year of the project, is that QLF-based transfer successfully circumvents many of the difficulties that arise using pure transfer or interlingua methods; it manages to factor out the problems caused by linguistic differences, which are reasonably tractable, and leaves those caused by knowledge, which are not. The result is a robust and modular architecture, which can be debugged and expanded with a relatively low expenditure of effort.

In the next section, we describe the BCI's architecture in more detail: a full description is available in (Alshawi *et al* 1991b).

3 Current status of the BCI prototype

The main components of the BCI are two copies of the SRI Core Language Engine (CLE), a state-of-the-art general-purpose tool for natural-language analysis and generation, equipped with English and Swedish grammars respectively. The basic system software and the English grammar and lexicon were written at SRI Cambridge Research Centre between 1986 and 1989, with an expenditure of about fourteen man-years of work. Adaptation of the English-language components to Swedish was done at SICS during 1990-91, and took about 16 man-months. The two copies of the CLE are linked by the transfer and interaction components, which are comparatively small pieces of software; the transfer component consists of an interpreter and a set of declarative transfer rules which can be extended in a modular way.

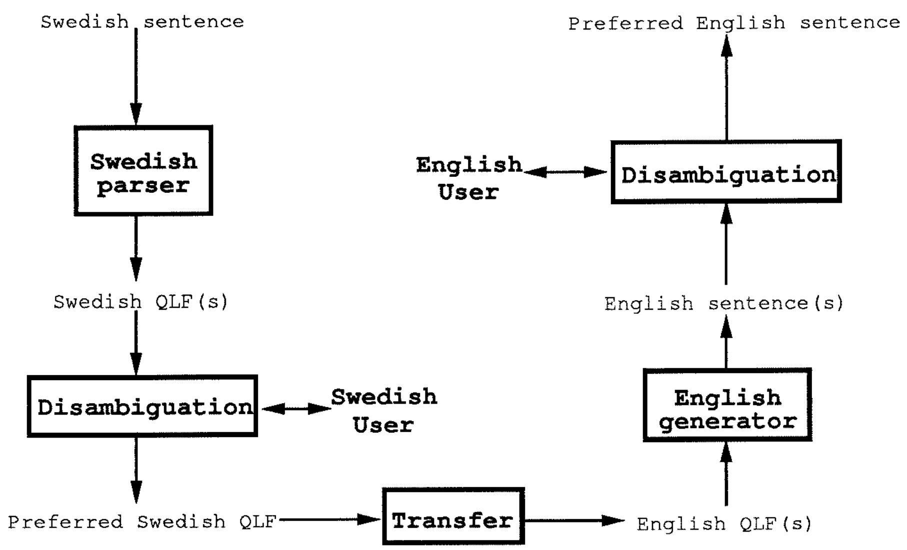

Figure 5: The overall architecture of the BCI

The system is normally run on a pair of Sun SPARCstations under either Quintus or SICStus Prolog. The overall architecture of the system is shown in Figure 5 on Page 7.

The CLE is capable of running both in analysis and generation modes, using a single grammar which is compiled in different ways for the two tasks; generation is performed using the Semantic Head-Driven algorithm (Shieber *et al* 1990). Analysis turns sentences into QLF representations, while generation works in the opposite direction. Intermediate stages include processing of morphology and syntax (grammar).

The QLF notation is a conservatively extended version of first-order logic, and is perhaps best described here by illustration. Continuing the example from the first section, *He hired a car* gives the QLF

```
[past,
 [hire_3p,
  qterm(ex,sing,A,[event,A]),
  a_term(ref(pro,he,sing,l([])),B,
         [and,[male,B],[personal,B]]),
  qterm(a,sing,C,[car1,C])]]
```

while *Did he hire a car?* is

```
[ynq,
 [past,
  [hire_3p,
   qterm(ex,sing,A,[event,A]),
   a_term(ref(pro,he,sing,l([])),B,
          [and,[male,B],[personal,B]]),
   qterm(a,sing,C,[car1,C])]]]
```

The main point to notice is that the representation of the second sentence differs from that of the first only by having the operator `ynq` ("yes-no question") wrapped round it. This principle of cleanly separating out distinct aspects of the sentence in its QLF representation is adhered to consistently, and as explained above greatly simplifies the transfer process.

The English and Swedish grammars are both fairly large and cover most of the common constructions in their respective languages, including questions (YN- and WH-), topicalized clauses, imperatives, passives, relative clauses, negation, cleft constructions, conjunction, noun-phrase and verb-phrase modification by preposition-phrases, adjectives and adverbs, various kinds of complex determiners, proper names, dates and times, possessive constructions and about fifty different kinds of complements to verbs and adjectives. There is a good treatment of inflectional morphology, which for Swedish covers all main inflectional classes of nouns, verbs and adjectives.

The function-word lexicon for each language contains about 400 words, including most pronouns, conjunctions, prepositions, determiners, particles and "special" verbs. In addition, there is a "core" content-word lexicon with common nouns, verbs and adjectives: the English one contains about 900 more entries, while the Swedish version, which is still under development, has about 350. New lexicon entries can be added by users using a tool developed for the purpose.

The BCI has been debugged using a test-set of about 500 sentences (copies of which are available on request), mainly taken from the hypothetical car-hire domain. The current level of robustness of the prototype is that of a fairly good "demo" system. For short sentences typical of the car hire domain, median total processing times for analysis, transfer and generation are around ten seconds when running under Quintus Prolog on a Sun SPARCstation 2.

A short sample dialogue follows on pages 10–11: "Client" refers to the English-language user, while "Agent" is his Swedish-language counterpart. Glosses for the Swedish text are in italics, and some of the less interesting disambiguation questions asked by the system have been omitted for the sake of brevity.

It should be apparent from the dialogue that the BCI's output, while perhaps not perfectly idiomatic, is nevertheless of a good standard and is readily comprehensible. Note also that non-trivial translation is possible, as evidenced for example in the seventh sentence, where the expression *Det går att* (literally: "It goes to") is rendered as *It is possible to.*

1 Client:

```
I want to hire a car.
```

System to agent:

```
"Jag vill hyra en bil"
```
(I want hire a car)

2 Agent:

```
Skulle du vilja ha en Volvo?
```
(Would you want have a Volvo?)

System to agent:

```
What does "en" mean?
1 någon (some)
2 precis en (one)
Enter a number:  1
```

System to Client:

```
"Would you like to have a volvo"
```

3 Client:

```
I don't like Volvoes.
```

System to agent:

```
"Jag tycker inte om Volvoar"
```
(I think not of Volvoes)

4 Agent:

```
Vi har en Saab och en liten Audi.
```

System to Client:

```
"We have a saab and a small audi"
```

5 Client:

```
Can I pick up a Saab at the airport at 20:45?
```

System to Client:

```
Consider the following bracketings...
1:  can i {pick up a saab} at the airport at 20:45
2:  can i pick up {a saab at the airport} at 20:45
Enter the number of one you like, or minus the number
of one you don't like:  1
```

System to agent:
`"Kan jag hämta en saab vid flygplatsen 20:45?"`
(Can I collect a Saab at the-airport 20:45?)

6 Agent:
`Jag tror att kontoret där stänger klockan 19:00.`
(I think that the-office there closes o'clock 19:00)

System to Client:
`"I think the office there is closing at 19:00"`

7 Agent:
`Det går att hämta bilen från kontoret i staden`
(It goes to collect the-car from the-office in the-city)

System to Client:
`"It is possible to pick the car up`
`from the office in the city"`

8 Client:
`OK, where is it?`

System to agent:
`"Okay var är den/det"`
(OK where is it (common)/it (neuter)?)

9 Agent:
`Det är vid Slussen.`

System to Client:
`"That is at Slussen"`

10 Client:
`Thankyou!`

System to agent:
`"Tack"`

4 Future developments of the BCI architecture

In this section, we will describe five ways in which the existing BCI prototype could be extended into the basis for a commercially viable system. The first of these is simply a more advanced version of the existing BCI; the second is a variant on the basic idea, used to translate email messages.

The third is a speech-to-speech translation system built by equipping the BCI with a speech-analysis front-end and a speech-synthesis back-end. This is perhaps the most exciting of the possible ways in which the project can be continued, and is by no means as impractical as one might immediately be inclined to suppose.

The last two ideas are more speculative: the fourth is a "multi-lingual workstation", which uses QLF transfer as a representation to facilitate information exchange between applications, including access in several languages to a knowledge-based application. Finally, we look at the possibility of a system for multi-lingual composition of business or technical texts.

4.1 A second-stage BCI

A more advanced version of the BCI could include either an improved coverage of the existing English-Swedish language pair, or the introduction of new languages; based on our experience in adapting the CLE to Swedish, we consider it likely that similar adaptations to most European languages could be carried out with comparable investments of effort. SRI and SICS can quickly access requisite expertise for Danish, Dutch, French, German, Norwegian and Spanish.

The English-Swedish BCI could be improved in several ways. The prime determiner of performance is good monolingual coverage, since this is where most of the work gets done; the system design means that the transfer rules are easy to write, as long as the various grammars correctly encode the relationship between surface strings and their QLF representations.

A tentative plan would include the following points:

- *Upgrading of grammar*: Tests made at the end of the project's first year indicate that certain parts of the grammar (in particular, the treatment of tense and aspect) could profitably be rewritten. Following this, the grammar would be improved by empirically determining shortcomings in its coverage and extending it appropriately.

- *Completion of Swedish core lexicon*: The core content-word lexicon should be expanded to give coverage equivalent to the English one. This involves adding something between 400 and 800 new entries.

- *Construction of domain lexicon*: If the BCI were to be used with a specific real application in mind, a large domain lexicon would need to be constructed. This would involve frequency analyses of real texts, and construction of 2000-5000 entries using the lexicon acquisition tool.

- *Interfacing to large machine-readable lexica*: Better fail-soft performance could be provided by giving the system facilities to allow it to access existing bilingual lexica when unknown words are encountered; experiments have already been carried out in which the English CLE has been interfaced to a monolingual dictionary, and a fair proportion of the software support needed for this task is consequently already available.

The result of such an effort would be a large-coverage general-purpose system which could easily be adapted to a variety of dialogue-based applications.

4.2 Interactive translation of electronic mail

A variant on the basic theme of the BCI, which may well be of greater immediate interest, would be to integrate it into an existing electronic mail system to allow interactive composition of messages intended for foreign-language addressees. Messages would include lists of source-language QLFs as well as the actual source-language text, these QLFs being deciphered at the receiver end by target-language versions of the system. This configuration makes it possible to produce output in multiple languages transparently to the source-language user.

Our impression is that the quality of the output from the BCI is of a level well-suited to this kind of application. Although it can sometimes be unwieldy or stilted in style, it is rare for it to be incomprehensible, or actually incorrect, in the sense of expressing something radically different from what the source-user intended. Consequently, both users can communicate through the system with a fair degree of confidence; this last point is of great practical importance, since it would normally be impossible either to check the translation or to post-edit it.

On the basis of our experience in the project so far we think it likely that an application of this kind, which was sufficiently robust to be of real use in practical contexts, could be created without doing more than systematically extending the system in the ways indicated in the previous section.

4.3 Speech-to-speech translation

Although automatic speech-to-speech translation was not long ago regarded as being well beyond the state of the art, it appears that progress in this area has been more rapid than was generally anticipated; although the time-scales involved are certainly longer than for applications of the kinds described in the first two sections, there is already talk about products appearing on the market by the beginning of the next decade. NEC recently announced a prototype system which could carry out automatic speaker-independent translation between English and Japanese with a 500 word vocabulary in a circumscribed domain, and it is well-known that ATR have made large research investments in similar systems. Reports of this work have sparked off general interest in the field, not least of course due to the enormous commercial opportunities it presents.

Our impression is that it would be quite feasible to consider converting the BCI to work in a speech-to-speech mode; there are in particular two factors that weigh heavily in its favour. Firstly, the high quality of the translation output becomes doubly important in this context, since post-editing is for obvious reasons impractical. Secondly, there is already a large body of experience on using the the CLE for speech applications; one project, at SRI Menlo Park, has been in existence since early 1988, and a second one, a collaboration between SRI Cambridge and RSRE, Malvern, is nearing the end of its first year. In both cases, the CLE has shown itself to be well-suited to the speech environment. Given that speech front-ends to the English-language version of the CLE and Swedish-language text-to-speech systems already exist, it is likely that a research prototype of an English to Swedish system could be built in a relatively short time.

If the system were to be used within a limited application domain, it would also be possible to consider incorporating machine-learning techniques recently developed at SICS, which use examples of typical sentences to "tune" a grammar to a specific application (Rayner & Samuelsson 1990; Samuelsson & Rayner 1991). This would have the effect of automatically cutting down the range of possibilities for the speech recognizer, increasing its accuracy for domain sentences, and also simultaneously speeding up the analysis phase by a substantial factor.

4.4 The "multi-lingual workstation"

In a wider perspective, it would be possible to use the idea of QLF transfer to create what might be called a "multi-lingual workstation". There is already a growing interest in the idea of using natural language as a standard for accessing knowledge-based applications: for example, (Bobrow *et al* 1990) describe a system in which English is used as a standard query-language to access a set of knowledge-based tools totaling over 800 different functionalities. The user is not obliged to state which tool or tools he wishes to use, the system deciding this automatically. Broadening the coverage of such interfaces to encompass two or more natural languages has obvious advantages in terms of making them available to a wider audience.

To do this efficiently, it is extremely important to choose an architecture which will allow the main kernel of the system to be shared between the different languages covered; although the peripheral parts, such as the grammar and lexicon, may be quite language-specific, the reverse will tend to be true of the components encoding the central functions of reasoning, database query construction and answer generation. QLF-transfer appears to provide a natural solution to the problem. Instead of using transferred QLFs as input to the generator, they would be fed into the knowledge-based query application; similarly, output QLFs encoding the system's answers would be bypassed through transfer before being sent to the selected output language's generator. This would have the effect of transforming a single-language system to a multi-lingual one by doing little more than configuring existing components.

One experiment in this direction that could be undertaken in the near future is a synthesis of the BCI with CLARE, a second-generation natural language interface prototype currently being developed at SRI. CLARE, which includes a version of the CLE as one of its components, is capable of advanced types of linguistic processing, including the interpretation of anaphora, tense and ellipsis. In the proposed experiment, CLARE would be supplied as a "black box" in an object-code version.

It is worth pointing out explicitly that the sort of application sketched here is only possible with an translation architecture of the kind used in the BCI, in which the output of semantic analysis can be input both to the transfer component *and* to the understanding and reasoning processes.

4.5 Interactive multilingual composition of texts

Looking ahead, it may be possible to improve the system to the point where it could be used for interactive multilingual composition of formal documents such as business letters and technical manuals, although this places rather heavier demands on the quality of the output text; existing systems for machine translation of this type of document all rely either on extensive post-editing, or on a highly constrained input language. To achieve the necessary fluency, it would probably be necessary to use a number of new techniques, including generation with reference to statistical collocation information.

5 Summary

Comparing the various proposed follow-on projects, one can summarize them roughly as follows. The "second-stage BCI" and email translator projects are essentially just extensions of the current system, and could reasonably be expected to yield serious prototypes in a short-term perspective. The main problems are increasing efficiency and robustness, but there are well-understood methods that could be applied with good hope of success. The speech-to-speech application is interesting in a somewhat longer-term view. Construction of anything approaching a product would undoubtedly require the solution of a number of fairly hard research problems, but many of the basic building blocks are already there, and the potential pay-off is very substantial. The multi-lingual workstation is a straightforward project that could be carried out with a relatively small investment, as it requires little more than the reconfiguring of existing components and the adaptation of the BCI to a new domain.

Readers who wish to know more about the technical aspects of the BCI can find a short (eight-page) discussion in (Alshawi *et al* 1991a), which focuses in particular on the design decisions behind the idea of QLF transfer and the scalability aspects of the system. The full project report is available as (Alshawi *et al* 1991b).

References

Alshawi, H. (1990). "Resolving Quasi Logical Forms". *Computational Linguistics* **16**:133–144.

Alshawi, H., ed. (1992, to appear). *The Core Language Engine.* Cambridge, Massachusetts: The MIT Press.

Alshawi, H., D. Carter, B. Gambäck and M. Rayner (1991a). "Translation by Quasi Logical Form Transfer", *the 29th Annual Meeting of the Association for Computational Linguistics*, University of California, Berkeley, California, pp. 161–168.

Alshawi, H., C. Brown, D. Carter, B. Gambäck, S.G. Pulman and M. Rayner (1991b). *Bilingual Conversation Interpreter: A Prototype Message Translator. Final Report*, joint SICS (R91011) and SRI (CCSRC-018) Research Report, Stockholm, Sweden and Cambridge, England.

Bobrow, R.J., P. Resnik and R.M. Weischedel (1990). "Multiple Underlying Systems: Translating User Requests into Programs to Produce Answers", *the 28th Annual Meeting of the Association for Computational Linguistics*, University of Pittsburgh, Pennsylvania, pp. 227–234.

Goodman, K. and S. Nirenburg, ed. (1989). "KBMT-89 Project Report", Carnegie Mellon University Center for Machine Translation.

Miike, S., K. Hasebe, H. Somers and S. Amano (1988). "Experiences with an on-line translating dialogue system", *the 26th Annual Meeting of the Association for Computational Linguistics*, State University of New York at Buffalo, Buffalo, New York, pp. 155–162.

Rayner, M. and C. Samuelsson (1990). "Using Explanation-Based Learning to Increase Performance in a Large-Scale NL Query System", *the DARPA Speech and Natural Language Workshop*, Hidden Valley, Pennsylvania, pp. 251–256.

Samuelsson, C. and M. Rayner (1991). "Quantitative Evaluation of Explanation Based Learning as an Optimization Tool for a Large-Scale Natural Language System", *the 12th International Joint Conference on Artifical Intelligence*, Sydney, Australia, pp. 609–615.

Shieber, S.M., G. van Noord, F.C.N. Pereira and R.C. Moore (1990). "Semantic Head-Driven Generation". *Computational Linguistics* **16**:30–43.

Thouin, B. (1982). "The Meteo System", in V. Lawson (ed.) *Practical Experience of Machine Translation.* Amsterdam, Holland: North-Holland, pp. 39–44.

White, J. (1985). "Characteristics of the METAL machine translation system at production stage", it the Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages, Colgate University, Hamilton, New York, pp. 359–369.