

The ALPS computer-assisted translation system in an academic environment

Patrick Corness

*Department of Language Studies, Coventry Lanchester
Polytechnic, Coventry, Warwickshire, UK*

INTRODUCTION

May I begin with a few words about the background to the polytechnic's association with ALP Systems, or ALPS, as it is known to us in Coventry? This began in May 1984, when, searching for software which would help me to compile an electronic dictionary, I contacted, amongst other firms, ALP Systems of Provo, Utah. My colleagues and I in the Department of Language Studies became very interested in the advanced computer-assisted language translation facilities being offered by a number of commercial companies. We believed we saw in the ALPS interactive approach considerable potential for educational applications beyond the electronic dictionary look-up capability which was originally looked for, and beyond the automated translation capability seen purely as an aid to productivity for the busy professional translator. Subsequently, in the spring of 1985, Lanchester Polytechnic in Coventry became the first institution of higher education outside America to install an ALPS computer-assisted translation system.

At Coventry we only began to introduce ALPS to students in October 1985 and the Polytechnic is, for the time being, I understand, the only educational ALPS user in Britain, indeed so far the only one outside America.

I am therefore going to be talking rather more about plans and aspirations than about hard experience. My remarks are predicated on a certain degree of optimism and a good deal of confidence that the ALPS system, while it may in practice turn out to be less than everything one could wish for, is something which represents a significant advance in the application of technology to language translation, and something which

has considerable potential for language learning purposes.

If this confidence in the system as a useful aid for translators is justified, then it follows that advanced language students should be acquainted with its principles at least, and hopefully they will have the opportunity to gain hands-on experience with ALPS.

I should emphasise at the outset that we see the system not only as an advanced tool for the translator, which students of translation can therefore with advantage learn to use, but as an aid to language learning in a much more general sense. It forms, in fact, a new element of the computer-assisted language learning (CALL) facilities which have been developed and introduced over the last three years or so in our language courses. In this development of CALL the principle has been adopted that the use of microcomputers, like the use of the language laboratory, should be fully integrated into the teaching at all levels, from beginners to final honours. The Department is continuing to develop this integrated approach, and intends that ALPS shall play its everyday part in an overall teaching scheme. The ALPS installation will not be a curiosity item, nor something exploited by a limited number of machine translation buffs, but a tool which all students will gradually learn to use for appropriate routine tasks.

This paper is not intended to be an advertisement for ALPS. The fact of the acquisition of this system by the Polytechnic obviously reflects confidence in its potential for academic purposes, but a full and objective assessment of the system will have to wait until there is extensive experience of its use in practice. Of course, users' needs vary, and ALPS offers different levels of assistance, designed to match these. As far as the applications of ALPS in the academic environment are concerned, it will be vital to demonstrate that they represent a real educational advantage and that ALPS, unlike Everest, is not simply being used 'because it is there'.

COMPUTER-ASSISTED LANGUAGE LEARNING (CALL)

Now a few words about computer-assisted language learning (CALL) and the prospects for its development, in order to consider the introduction of ALPS at Coventry Polytechnic in this particular educational context. Then follows a brief outline of the facilities ALPS offers, before some talk about the ways in which these computer-assisted facilities are being used at the moment for language teaching, and about the prospects for the fairly near future.

Today, students in some polytechnics and universities are beginning to use computers as an aid to language learning as naturally as they would a tape recorder. Indeed the tape recorder and the microcomputer are now

being linked in a new development called audio-enhanced computer-assisted language learning, or AECALL.

I am not able to cover the latest research work in CALL in other polytechnics and universities. The CALL packages which have so far been developed in my own Department¹ – and which are, if I may mention it, also being widely used in other institutions – enable many kinds of self-correcting language-learning exercises to be produced. Such exercises are now extensively used on our languages degree courses; indeed this development has stimulated a different approach to the teaching of languages at Coventry Polytechnic. It would not be appropriate for me to go into the detail of the teaching method here: can I just re-emphasise the point that it is fundamental to our approach to CALL that work on the computers is fully integrated into the teaching scheme, rather than being an optional extra, or a peripheral activity.

In CALL, the programming of language-learning exercises can allow for a variety of acceptable responses to questions posed to students by the computer: it does not have to demand, completely unthinkingly, one response only, to the exclusion of other responses which may be reasonable alternatives. However, in self-testing CALL exercises, the number of acceptable alternative answers has to be finite. In other words, the fundamental problem of the language-computer interface is that while human language is infinitely creative, the computer can only accept what it has been told to accept, to the letter.

In practice, of course, the number of acceptable alternative responses in many language-learning exercises can reasonably be fairly limited. Other exercises, such as translation as a case in point, are, by contrast, very open-ended and creative. An obvious misgiving felt by sceptics who view CALL and computer-assisted translation with suspicion, and hesitate to jump, as they see it, onto the microprocessor bandwagon, is that because language is so fundamentally human a phenomenon, so rich, so unpredictable, it can only suffer if forced into a strait jacket by a machine which is very powerful but fundamentally brainless. The short answer to these misgivings is, I believe, that the machine is not supposed to be the brains: it is supposed to help the human translator to concentrate on the intellectual and creative aspects of the task.

The old conception of what machine translation means, that is to say fully automatic translation by computer, entails the false assumption that the computer can, indeed, become the brains of the operation. But the ALPS approach is based on a more realistic view of the relationship between language and the computer. The fact that the human translator retains control makes ALPS viable as an open-ended and creative CALL facility as well as a translation aid as such.

However, not everyone is convinced that computers of any kind are

potentially an effective aid to language learning, and not simply a gimmick. The teacher's job is to assist students to increase their factual knowledge, to develop their intellectual understanding, to enhance their practical skills and to stimulate their enthusiasm and interest. The question is whether technological aids help the teacher to achieve these aims more rapidly or more effectively. In faculties of science and engineering it is taken for granted that technical facilities, including of course computers, are essential, but in the teaching of languages and the teaching of translation the relevance of the computer is still considered unproven by some.

The development of AECALL was mentioned earlier. This is an attempt to progress beyond the limitations of the first generation of computer-assisted language-learning exercises, in which the written language predominates. The purpose of AECALL is to incorporate the spoken language in computer-assisted learning exercises. It is of course axiomatic that spoken language is primary, and written language secondary. Audio-enhanced CALL is, therefore, one important direction in which CALL research must move. The use of speech recognition and speech synthesis techniques is not relevant here, by the way, or not yet. What is being developed in AECALL is computerised control of human voice recordings for language-learning purposes.

A second, no less important, aspect of CALL research is the exploitation of information retrieval techniques as interactive computer-assisted learning resources. These include dictionary look-up, concordance look-up and other online help files. It is in this vital area of the application of computerised information retrieval to language learning that we are looking at the ALPS system to see what it can offer.

OUTLINE OF THE ALPS SYSTEM

There are three levels of ALPS translator assistance, called respectively:

- SDL (selective dictionary look-up)
- ADL (automatic dictionary look-up)
- CTS (computer translation system).

These terms represent a range of facilities, adaptable to different needs. The different levels are appropriate, for example, for differing degrees of specialisation of subject matter, differing length of text etc. The processing capabilities include the following:

- multilingual word processing
- dictionary building

- dictionary look-up
- keywords in context (KWIC) look-up
- vocabulary frequency analysis
- grammatical and syntactical analysis
- interactive translation.

Once a source text has been entered, using the word processor or an optical character reader, the next step is to provide a purpose-built dictionary, called a document dictionary, for that text. The document dictionary will contain all the vocabulary items found in the source text, and, depending on the level of automatic assistance required, may contain the grammatical and syntactic information codes which the translation program will need to examine.

The first step in building the document dictionary is a process called dextraction, which scans the source text and scans the appropriate permanent dictionaries, as instructed by the user, and then creates a bilingual dictionary containing all the required vocabulary found in the dictionaries searched.

Dextraction also lists, if required, all the source text vocabulary not found in any of the dictionaries searched. The user can then use this Words Not Found (WNF) list to complete the document dictionary. Further assistance to the translator in editing dictionaries is provided by the KeyWords In Context (KWIC) list facility and the word frequency count facility which are also available. When the dictionary is ready, the translator can call up whichever level of assistance is required.

Using SDL, a translation may be typed on the screen using the word processor, to be subsequently saved, printed out etc. During this process the dictionary may be accessed as necessary, to find suggested translation equivalents for particular items.

In ADL, the translation is built up on the VDU screen, sentence by sentence, in the Work Area window. The translation equivalents for all the words or phrases are automatically displayed in the Target Area window, and are numbered so that they can be entered into the target text in the appropriate order, and subsequently rearranged if necessary, by one or two keystrokes. Typing is thus reduced to a minimum, but a translation equivalent offered can always be rejected in favour of a different word or phrase actually typed in by the translator, and the target text can be edited subsequently if necessary.

In CTS, grammatical and syntactical processing is also done by the computer. A draft translation of the text is offered automatically, sentence by sentence, but first of all the computer interrogates the user in order to clarify structural or semantic ambiguities, and to supply missing vocabulary. Editing of the translation offered by the computer can be carried out rapidly if necessary.

The ALPS computer-assisted translation system at the CTS level is not intended to replace the human translator and it cannot translate entirely

automatically a text fed into it. However it can accelerate the translation process.

It is important to realise that the analysis is limited to one sentence at a time. This means that the computer has no knowledge of the world beyond the purely grammatical structures it is programmed to recognise and beyond the meanings of words and phrases it is programmed to recognise. In other words, it cannot resolve for itself the ambiguities contained in a sentence such as, 'Coventry Polytechnic is a big computer user'. The intervention of the human translator is needed here: knowledge of Coventry Polytechnic's actual use of computers is needed to resolve the ambiguity as to whether the Polytechnic is being described as a 'big user of computers' or as a 'user of big computers'. Even such background knowledge may not resolve the ambiguity, since both interpretations may be valid. The human translator may be able to tell which meaning is intended from the context, since the rest of the text may refer consistently to users of big computers or consistently to big users of computers. However, the computer programs which are analysing a sentence have no knowledge of the context beyond that sentence. The system can be programmed so that 'big computer user' is always translated as 'big user of computers' or always translated as 'user of big computers', but in principle this is unsatisfactory, since it limits the usefulness of the program to one particular text. The translation might as well have been done by hand in the first place.

What is needed is an automatic system which is independent of particular texts and particular vocabulary items, a system which enables all such ambiguities to be resolved as they occur. What the ALPS computer-assisted translation system does is to ask the human translator to resolve the ambiguity, by displaying on the screen a question such as:

Is the text speaking of
(a) big computers
(b) big users?

The translator presses either *a* or *b* on the keyboard to tell the ALPS computer what the correct choice is in this case. If the computer's dictionary contains more than one possible translation for a given expression, the translator will again be asked to make a choice by pressing *a*, *b*, *c* etc.

The translation which is offered by the computer after the ambiguities have been resolved may not be satisfactory. Some editing will usually be needed to produce an acceptable final version. Much depends on the dictionary which has been built up in advance by the translator.

At Coventry there is presently a two-workstation installation, incorporating all three levels of ALPS translation assistance (SDL, ADL and

CTS) from English into French, into German, into Spanish and into Italian, running on a Data General minicomputer, the Model 20 Desktop. We are awaiting the software for translation into English, and also we hope soon to have ADL facilities for Russian.

ALPS AS A DEVELOPMENT OF CALL: PRINCIPLES

One of the problems of language learning, including translation training, is that of heightening awareness of contrasts in structure between the native language and the foreign language, awareness of the variety of stylistic expression available in a given language, and awareness of the many factors influencing the choice of particular forms and structures in a given case.

One can see in ALPS new opportunities to demonstrate these contrasts, giving students – literally at their fingertips – a display of a variety of possible translation equivalents for a given word or phrase, showing examples of their occurrences in a variety of contexts and meanings. As they proceed with their translation, students either make choices from a number of suggested translation equivalents shown on the screen, or else decide on equivalents of their own. In the latter case they then update the dictionary by adding the new equivalents as additional entries for future reference. Naturally this work has to take place under the supervision of the tutor, but it results in a translation which is the product of a series of conscious decisions which the computer obliges the user to take in a systematic way.

There might be a suspicion that working with pen and paper is really more creative than doing what the computer apparently tells you to do, but pen and paper translation would normally involve reference to dictionaries too: one difference between the traditional method and working with assistance from ALPS is in the speed and convenience of access to helpful information. ALPS users are not banned from looking at traditional reference works, by the way! And the translator is not obliged to accept the suggestions on the screen any more than the suggestions in a traditional dictionary.

As I see it, one principal advantage of ALPS as a language-learning aid consists in the power of the computer to make rapidly available a wide range of online look-up facilities during the translation process, as required. We are planning in Coventry to build up further, specifically education-oriented, look-up facilities in addition to the dictionaries and KWIC lists. These will include a set of help files for convenient reference to information, for assistance on problems of grammar, for example. The extent of the available help can also be controlled by the tutor, depending on the nature of the exercise involved. Access to certain files can be

temporarily restricted for testing purposes, for example.

The fact that students can be made aware that a dictionary does not represent tablets of stone, but is rather the object of creative development work, to which they themselves can contribute, learning by doing, also seems to me to offer a distinct pedagogical advantage.

A further significant feature of ALPS as a learning aid is the labour-saving facility. Since words and phrases are selected from the screen menu of translation equivalents and entered into the target text by a very small number of keystrokes, the time and effort spent on sheer writing or typing is considerably reduced, and so the system should enable the student to concentrate on the actual intellectual activity of translating, of making intellectual decisions. Writing or typing everything out by hand does have its own pedagogical advantages, of course. It helps the learner to remember what was written. But the potential advantages of more intensive and extensive exposure to the languages and their points of similarity and contrast which the language student will eventually have as an ALPS user do seem to be considerable. Besides, students will not do all their work on an ALPS terminal. They will still do plenty of writing in longhand.

ALPS AS A DEVELOPMENT OF CALL: THE EXPERIMENT SO FAR

We are experimenting at the moment with some fairly general texts for translation from English into German, using ADL, with first year students. The subject matter of the texts is modern German history and the ADL level is being used in this way to provide exercises in what we call 'guided translation' into the foreign language.

In this context I should mention that the Department's microcomputer network has been used during the last three years or so to give practice, amongst other things, in what is called 'text reconstruction'. In this exercise, an original German text, say, previously translated into English by the students, has to be faithfully reconstructed by them on the computer from memory using the English translation as a guide. This is, of course, not a translation exercise as such, but is used to reinforce knowledge of target language structures in context. The idea behind this is not too far removed from the hallowed principle of learning poetry by heart to reinforce knowledge of a language.

Now that the ALPS system is installed, we are introducing, experimentally, exercises in a more open-ended form of guided translation, using ADL. In these exercises, students are expected to produce their own translations with varying degrees of assistance from the computer, and a model translation is made available for purposes of comparison only. The student's version and the model version can be displayed and printed out

in parallel columns for easy checking and comparison.

One such type of exercise using ADL is very closely guided translation practice, which is really more of a grammar test than a translation task. In this case the appropriate vocabulary is defined in advance by the lecturer, who builds the dictionary. The student is then automatically provided with the vocabulary to be used, and is required to compose the translation using the words offered, in their correct grammatical forms, building a correct sentence structure.

Certain grammatical information is available in the ALPS dictionaries, and can be looked up as needed. This grammar look-up facility is being extended (something not normally needed by a professional translator) for language-learning purposes. It is planned to make such help available on ALPS on a fairly comprehensive basis ultimately.

Other ADL assignments, designed to be more like translation exercises, involve more interaction. A relatively wide range of suggested translation equivalents is displayed, and the student is required to translate the text with the help of this display and the other online look-up facilities available. Working with a normal dictionary, learners might not bother to look a word up, but assume that they knew the best translation for it. ADL can ensure that students consider a range of possibilities, including some they may well not have thought of. We will be developing look-up facilities which will show the use of vocabulary in various contexts, to assist in finding the best translation, and enhance students' awareness of the problems involved.

Professional translators using ADL would design their dictionaries in such a way as to minimise interaction and save time, but for language-learning purposes it may be appropriate to deliberately maximise interaction.

CTS is being used to demonstrate some of the problems of translation arising out of contrasting structures. My colleague Gwyn Howells is experimenting in this area, systematically testing the capability of CTS to cope with the translation of specific structures from English into Spanish.

CONCLUSION: ALPS AS A LANGUAGE-TEACHING AID

For language students, the three different levels of ALPS systems (SDL, ADL and CTS) offer the benefit of hands-on experience with a range of information retrieval and information-processing techniques, from word processing and database handling to computer-assisted translation. This provides them with basic skills in these areas, which they are likely to need if they become professional linguists.

The ALPS computer can also be used as a broader CALL facility. The ALPS company² has provided some education-oriented software, so that

the system can be more readily adapted at all levels for language-teaching purposes. This software helps teaching staff to prepare controlled translation assignments, which students can carry out without the need to learn complicated procedures and without the risk of crashing the system. This is a big advantage, enabling all students to be given some experience as ALPS users with only a small installation.

The information storage and retrieval facilities are equally useful for purposes of accelerating the translation process and for purposes of accessing information for the language learner. They offer an open-ended, creative CALL facility which complements the self-correcting testing programs.

The vocabulary analysis – both alphabetical and frequency lists, as well as KWIC lists – can be used to assist the teacher to design graded teaching materials.

Last but not least, the CTS system and how it works, with its limitations just as much as its positive capabilities, presents a rich field of enquiry for the theoretical linguist.

To conclude, we believe that experience with ALPS will help students become more aware of the nature of language in general and of the problems of translation in particular.

REFERENCES

1. UNITEXT (for roman-alphabet languages) and RUSTEXT (cyrillic) suites of authoring and testing programs, written by R.M. Benwell of the Department of Applied Physical Sciences, Coventry Lanchester Polytechnic, in co-operation with the Department of Language Studies.
2. European Office: Automated Language Processing Systems, Route de Boudry 14, CH2016 Cortaillod/NE, Switzerland.

AUTHOR

Patrick Corness, Department of Language Studies (CALL Consultancy Group), Coventry Lanchester Polytechnic, Coventry CV1 5FB, UK.