PAC3
at
JALT
2001
Conference
Proceedings

MENU
Text Version
Help & FAQ

International
Conference
Centre
Kitakyushu
JAPAN

November
22–25, 2001

# Machine Translation: A silicon-based solution to the Tower of Babel

***Michael Cribb***
*Kansai Gaidai University*

Machine translation is a rapidly developing alternative to learning English as a foreign language. As computing power and speed increase in line with Moore's law, we can expect machine translation to generate output quality that is acceptable for widespread use in the near future. This will have two effects on language teaching. First, in the short term, it will mean that students will use machine translation for *gisting* and *drafting*, and teaching practices will need to change to accommodate this. Secondly, in the long run, it will lead to a teaching profession which is sharply reduced due to the fact that machine translation will remove the need for people to learn foreign languages.

機械翻訳は外国語としての英語学習の代案として急速に成長している。 モーア (Moore)の法則に従い、計算力と計算速度が増加するにつれ、機械翻訳に近未来にみんなが満足できるほど質が高い出力（機械翻訳）が期待できる。これは語学教授に２つの効果をもたらす。１つ目は、短期的に、学生はジスチング（gisting)とドラフチング（drafting)のため、機械翻訳を使用する。また、このことを受け容れられるように教授法は変わる必要がある。 ２つ目は、 長期的に、 機械翻訳の発展は人々が 外国語を 学ぶ必要をなくすため、急速に語学教授が減少するという点である。

For many scholars, the position of the English language in the world today is such that there appears to be nothing preventing it becoming the de facto global language—if it has not already become so. While regional languages exist which are spoken by large numbers of speakers, none appear to have the widespread, though sometimes reluctant, acceptance English does among the people of the world, and the language finds itself in the fortunate position of rising in prominence when globalization truly is a global phenomena (unlike Latin and French which rose to prominence chiefly in limited regional and social circumstances). Yet while English is unlikely to be challenged by another language, it does, I believe, have a competitor in the form of machine translation which could eventually dislodge it from its precarious position as the world's *lingua franca* and make the need for foreign language learning largely redundant.

Machine translation is the automatic translation of text from one language to another and is rapidly becoming a viable alternative, particularly on the World Wide Web, with an explosion of sites that offer basic, and often free, translation services. While machine translation can only provide output of limited quality at the moment, the potential for improvements over the next two decades is exceedingly good, and we could well see a day when it is used on a regular basis for interlingual communication, thus alleviating the need for people to study foreign languages. This silicon-based solution to the Tower of Babel problem is more parsimonious and "emancipating" (Eastman, 1983), I believe, than the current carbon-based solution, which is seen by many as a form of (chiefly Anglo-American) linguistic "imperialism" (Phillipson, 1992). Before considering this argument, I will introduce the reader to a simple example of machine translation using the AltaVista "BabelFish" translation service.

## An Example of Machine Translation

The AltaVista "BabelFish" translation service can be accessed at http://world.altavista.com/. On entering the site, the user can input a small text (limited to 150 words) and select from a list of eight languages pairs for the translation (currently English to Chinese, French, German, Italian, Japanese, Korean, Portuguese, Spanish and back). After clicking the "Translate" button, the translated text is returned literally within a matter of seconds. A useful exercise is to "cut and paste" the translated text into the text box area and translate it back into English to produce what is known as a "back-translation" which enables the user to compare this with the original in the same language to see how successful the translation was. (An alternative to inputting text is to specify the address of a web page which the system will fetch and translate.)

To demonstrate this system, I wrote a small letter

which I translated first into French and then back into English. The original letter and the back-translation are shown below:

### *Original*

Dear JALT.

I recently lost my job due to the collapse of the language teaching profession in Japan. I do not speak Japanese but I can use the AltaVista machine translation system to communicate with the local people. I hope, therefore, you can find me a job.

Yours sincerely.

Michael Cribb

### *Back-translation*

Dear Jalt.

I recently destroyed my work due to the collapse of the profession of the language teaching in Japan. I do not speak Japanese but I then to employ the automatic translation system of AltaVista to communicate with the local people. I hope, therefore, you can find me a work.

Yours sincerely.

Michael Cribb

As can been seen, the back-translation, while not perfect, does preserve much of the semantic context of the original and the reader can get the general gist of the message. (All the more impressive given that the text has gone through the translation process twice. In practice, once would be enough.) Astute readers may have noted that the original text was fairly simple in nature with sentence length and complexity held to a minimum in order to produce better results. Further, the language chosen for the back-translation, French, seems to work particularly well with this system, due largely in part to the structural similarities between English and French, while other languages such as German and Japanese are less successful. However, the potential for improvements in machine translation output quality is considerable, and when they come, the position of English as a global language, a topic I will turn to next, could become all the more precarious.

## English as a Global Language

The position of English around the world is often conceptualized using the concentric circle model proposed by Kachru (1985). In this model, an inner circle represents the countries where English is spoken as a native language such as Great Britain and the USA. Surrounding this is an outer circle which represents the countries where English has attained an institutionalized status and is chiefly spoken as a second language (e.g.

India, Singapore). Finally, there is an expanding circle which represents all the countries where English is generally learnt as a foreign language such as Japan, China, etc. Crystal (1997, p.61) estimates that the number of people who have at least a "reasonable competence" in English within the three circles is between 1,200 and 1,500 million; quite a considerable proportion of the world population and indicative of the status and international recognition English has achieved.

Yet this figure belies the relatively fragile foundation that English rests on. The expanding circle represents a large proportion of this total, and for most of the people in this circle, English is nothing more than a means to an end. English gives them access to the global village and the means to survive and function in a world where a single standard language is necessary if productive interlingual communication is to take place. Moreover, it requires governments in these countries to continually allocate time and resources at the secondary and tertiary educational levels to replenish its stock of speakers, unlike countries in the inner circle where English is acquired naturally in early childhood. English as a foreign language in the expanding circle, in reality, is vulnerable to attack by alternatives by virtue of the fact that by definition it is a "foreign" language. And linguistic history provides many examples of languages which, once vulnerable, can rapidly decline and even disappear as soon as the alternative arrives.

## Advantages of Machine Translation

While English does not appear to have any serious competition from the other major languages of the world, it is, I believe, vulnerable to attack from machine translation due to a number of reasons. Firstly, there is Moore's law (Port, 1997) which states that the speed and capacity of computers will double every eighteen months approximately. This is a powerful tenet which essentially implies we cannot accurately predict what algorithms computers will use in the future; tasks which seem impossible today may be routine tomorrow as new algorithms give rise to new solutions.

Secondly, more and more content is beginning to appear in electronic form such as newspapers and magazines on-line, journals and books on CD-ROM, as well as e-mail. As this volume increases, the advantage to be gained from automatic translation of this content into other languages becomes all the more obvious. Imagine, for example, the benefits a web portal such as CNN.com would achieve in increased readership if it could translate its daily on-line newspaper content rapidly and effortlessly into the major regional languages of the world. (Translation from English into Chinese, Arabic and Spanish, for example, would give native access to approximately a third of the world's population, assuming figures from Katzner, 1995)

Finally, machine translation has an advantage when we look at the needs of the new user in the global village with the electronic communications revolution. In the past, anyone who had a need to conduct interlingual communication often had several years of foreign language learning behind them. These days access to the global village can be much more rapid and newcomers often find they don't have the foreign language skills to cope with the task. However, competence in English will not develop for these users as a result of them simply surfing the Web, and we cannot reasonably expect them to embark on time consuming and costly English language studies. Instead, the companies that form the backbone of the Web will need to take the content to the user in the local language.

Of course, the quality of output from machine translation is far from perfect today and users can only really get a "gist" of the original text to see if it is worth following up. But as machine translation improves and new algorithms lead to higher quality output, there may come a time when the consumer will accept its limitations and make adjustments for it to be used on a daily basis. After all, when speakers of English as a foreign language produce output (spoken or written) which is far from perfect, we often praise the speaker and make adjustments in order to achieve our goals. It is surprising, then, that when machines produce output of similar quality, there are so many who are willing to write it off as "junk".

## A Look Into The Future

Given that machine translation will start to have an impact on our lives, how will this affect language teachers and their profession? First of all, in the short term, I would suggest that the use of machine translation for *gisting* (obtaining the gist of a lengthy article to see if it contains content worth following up) and *drafting* (using machine translation to translate a written piece from L1 to L2 before the student "polishes" it up) will become more frequent by students. Some teachers may be horrified at the notion of their students using anything other than their own mental processing power to read and write in English. However, there was a time when calculators were not allowed in mathematics classes because it was believed that students would not learn how to use slide rules properly or do mental arithmetic. Nowadays, though, slide rules are obsolete and nobody would suggest that calculators do anything but augment a student's mathematical capabilities. In the same way, I believe that machine translation, if used as a tool in the right way, can only help our students to be better at interlingual communication. (In a way, electronic dictionaries, which are often used by students in class, are the beginnings of machine translation usage.)

Of course, if a student submits the raw output

from a machine translation system as an end-of-term essay without any re-writing and post-editing, then, as teachers, we need to reject this. In other words, standards need to be maintained. But so long as students come to use and understand machine translation simply as a tool and they continue to develop their own mental abilities, then I believe they should have the best of technology at their disposal and be confident and skilled in putting it to use in an appropriate manner.

Secondly, in the long term, I would suggest that machine translation will eventually remove the need for people to learn foreign languages on a large scale. As machine translation permeates the Web, users will gain access to content in their own language, even though it was originally written in a foreign language. The large web portals which supply this content will no doubt invest heavily in these systems, since the benefits they can gain from the vastly increased readership will be enormous. Eventually, machine translation might find its way onto the desktop, perhaps becoming as common as the spellchecker is today, and allow anyone to conduct interlingual communication without having foreign language expertise. A Japanese salesman in Tokyo, for example, could correspond with a French customer in Paris without either of them having knowledge of the other's language or of English.

Following on from this, the automatic translation of spoken language could become a reality. While this may prove to be more of a challenge than the translation of written text, it is not an impossible feat. Already, speech recognition and speech synthesis, two key components for the realization of spoken machine translation, have started to appear in commercial applications, and only time is needed before they gain widespread acceptance. A more challenging component will be the filtering out of false starts, repetitions, and extraneous noise present in spoken language which is not a feature of written language, but again this problem can be minimized by users adjusting their output to accommodate (something every native speaker does today when talking with non-native speakers).

As machine translation becomes more accessible and ubiquitous, this will inevitably cut into the number of people around the world who choose to study English as a foreign language. With a silicon alternative in machine translation, the appeal of spending ten to twenty years in foreign language study (a reality for many) may not seem so attractive or economical for most people. Governments could start to roll back their massive foreign language learning programs at secondary and tertiary levels leaving just a small core of elite language learners for specialist applications. Of course, there would always be bilingual contact situations where people continue to learn foreign languages rather than rely on machines, but it seems like the number of people around the world who would need to invest in the

pursuit of foreign language learning would be drastically cut, along with a sharp decline in the language teaching profession.

For a while, English as a global language and machine translation may exist side by side, but it seems unlikely that this situation could remain stable for long, since the development curve for machine translation appears to be accelerating while the growth rate for English as a foreign language appears to be reaching a saturation point as the time and effort required to learn it appear to outweigh the benefits for many. Of course, English would not disappear from the global arena altogether and would still be very important in the inner and outer circle countries, but its role would be vastly diminished from that suggested by some scholars as the global *lingua franca*. Strangely enough, the electronic revolution in communications that has helped catapult English into its dominant position around the world could be the very revolution that leads to its downfall.

## References

Crystal, D. (1997). *English as a global language.* Cambridge: Cambridge University Press.

Eastman, C. M. (1983). Language planning for special purposes. In *Language planning: An introduction* (pp. 96-102). San Francisco: Chandler & Sharp.

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge: Cambridge University Press.

Katzner, K. (1995). *The languages of the world*. London: Routledge.

Phillipson, R. (1992). *Linguistic imperialism.* Oxford: Oxford University Press.

Port, O. (1997). Essay: Gordon Moore's crystal ball. *Business Week International Edition.* Available: <http://www.businessweek.com/1997/25/b353225.htm> (2001, December)