

DEVELOPMENTS IN VOICE AND LANGUAGE TECHNOLOGY

Authors: Helen McCready and Françoise Moreau-Johnson

IBM European Language Business Unit

This paper is intended for all those who provide, and use, translation services, and who are looking for ways to increase productivity, reduce cost, and improve their working environment. Current developments in voice and language technology address the set of ways by which users can interact better and benefit from their use of computers. The IBM European Language Business Unit is focusing on the use of natural language interfaces in a multi-lingual environment. The paper describes two practical applications of this technology, namely managing the translation process and text entry by voice. Although there will be further developments in these areas, the technology available today represents a major breakthrough.

A MOVE TO HUMAN-CENTRIC COMPUTING

Computer systems have found their way into all areas of daily life: in the work environment, for creating, manipulating or retrieving information; at home, for education or recreation purposes; on the high street, in the form of cash dispensing machines.

The wealth of information now stored on computer is formidable - for many organisations it is their most valuable asset.

The range of software available has become overwhelming. The sheer variety can create its own problems when information needs to be exchanged with others.

In the past 20-30 years we have seen a number of interfaces into these systems, ranging from punch cards and teletypes through to mouse-driven Graphical User Interfaces.

More and more we are looking for better ways to get effective use of the technology. This is the goal of human-centric computing. The technology should suit the way the human wants to work - not the other way around.

With easier access to information, and better interfaces for interaction with the computer, we can expect to see an increase in the use of Information Technology in areas where it has previously been excluded.

Current technology now offers practical applications for text manipulation for translation projects and text entry through voice input.

For those in the translation business, the benefits can be considerable.

MANAGING THE TRANSLATION PROCESS (IBM TRANSLATIONMANAGER)

Competitive business pressure worldwide is increasing translation workload while squeezing costs. Currently Machine Translation technology is unable to address the challenges of large scale translation projects.

The challenges are those of process management rather than translation itself. As customers become more demanding in the standards they require, this gives rise to a set of questions which need to be answered.

How do you ensure consistency of terminology in a project spread across multiple translators?

How do you work efficiently with source text provided in a wide range of word processing and text processing formats?

How can the project be centrally controlled and audited to ensure quality output on a timely basis?

How can you cut costs dramatically by using previously translated text?

All of these issues are becoming increasingly critical to large buyers of translation services. Research carried out earlier this year in the US indicates that about 70% of large translation projects consists of material previously translated. (Source: LISA Forum Newsletter Feb.1994). The potential cost impact alone of using previously translated material is enormous.

The technology now available through products like IBM TranslationManager addresses these management challenges, by providing the following:

Comprehensive terminology management functions ensure consistency across multiple translators and projects.

The ability to work with a wide range of source document input formats without having to use the relevant word processing packages.

Project control to administer, monitor and cost the translation process.

Translation memory -- "automatic translation" -- so translators can detect and reuse previously translated material for updates.

IBM TranslationManager supports 19 source languages. Target languages can be any which are supported by the Operating System, including Russian, Finnish, Greek, Japanese, Korean and Chinese.

User experiences to date show time and cost savings of between 25-50% depending on the nature of the project. As use of this technology becomes more widespread, the potential impact on business efficiency in international markets is enormous.

TranslationManager makes translating a lot easier. It presents a productive working environment and easy-to-use utilities for all the tasks of a translation project.

Of course further productivity gains can be achieved by also speeding the process of text entry.

TEXT ENTRY BY VOICE (IBM PERSONAL DICTATION SYSTEM)

One of the main aspects of human-centric computing is getting away from the keyboard as the main text input method. The keyboard is not a natural interface: we learn to talk first, then write (the quality of which degrades over time) and some of us learn to type. However, as only 15% of people who have a keyboard in front of them are touch-typists (the rest can only type at most 30-40 words a minute), but all of us know how to speak, voice seems to be the best text entry method. IBM has been conducting research in the area of speech recognition for over 20 years in order to transfer the spoken word into text, and can now provide a solution for users to enter text by voice.

This dictated text can be saved to a file, printed, formatted, edited or integrated with other applications, including TranslationManager. IBM Personal Dictation System is a highly accurate and intelligent speech recognition system that analyses and converts spoken words into text. It has an active vocabulary of 32,000 words, which means that you no longer have to use a keyboard to enter translated text.

The system needs to be trained to the user's voice (speaker dependent) and for the speaker to use isolated word mode speech (i.e. leaving a pause between words) as opposed to continuous speech. After the required training (less than 1 hour of the user's time), the user can expect an accuracy

level of around 95%, and can dictate at speeds of up to 100 words/minute.

The use of speech technology is particularly beneficial to professions and environments where a hands-free task is performed (e.g. radiology, medical, scientific). Furthermore, because this technology offers a play-back facility for corrections, it means the user does not have to verify the text as it appears on the screen during the dictation, but instead, the user can be looking at the source material (required in the legal, journalistic and translation professions).

This application, which also offers voice control of the desktop and applications, is currently available in 5 European languages (English, French, German, Italian and Spanish).

Early users have found productivity gains of around 25%.

CONCLUSION

There is still a lot that needs to be accomplished in the future as the technology develops.

In the area of Speech Recognition: speaker independence and continuous speech; recognising many voices at the same time, for example, writing proceedings of an interview, minutes of a meeting, or even recording the debates in the Houses of Parliament (HANSARD).

In the area of Translation: integration of Translation Management tools with Machine Translation technology as it matures .

Undoubtedly, the greatest advances will be achieved through the integration of different technologies, many of which are available today.

The applications described in this paper will not do the job for you - but they will help you get it done.