[From: Makoto Nagao (ed.) *Machine translation summit*. Tokyo: Ohmsha, 1989]

2 • 3 Machines and People in Translation

M. Kay
Xerox Parc, U.S.A.

It is useful to distinguish a narrower and a wider use for the term "machine translation". The narrow sense is the more usual one. In this sense, the term refers to a batch process in which a text is given over to a machine from which, some time later, a result is collected which we think of as the output of the machine translation process. When we use the term in the wider sense, it includes all the process required to obtain final translation output on paper. In particular, the wider usage allows for the possibility of an interactive process involving people and machines.

Machine translation, narrowly conceived, is not appropriate for achieving engineering objectives. Machine translation, narrowly conceived, provides an extremely rich framework within which to conduct research on theoretical and computational linguistics, on cognitive modeling and, indeed, a variety of scientific problems. I believe that it provides the best view we can get of human cognitive performance, without introducing a perceptual component. When we learn more about vision, or other perceptual modalities, this situation may change. Machine translation, narrowly conceived, requires a solution to be found to almost every imaginable linguistic problem, and the solutions must be coherent with one another, so that it is a very demanding framework in which to work.

During the last twenty years — the period we are focusing on — there have been essentially no advances in the field of linguistics of sufficient size or significance as to affect our ability to build working machine translation systems. We remain today at essentially the same place we were in at the time when the ALPAC report was written. Furthermore, I doubt whether many professional linguists would be disposed to contest this. If this claim is right, then we have no reason to expect to be able to build significantly better machine translation systems today than we could then. This is construing the term narrowly. But, if we construe the term widely, we can hope to do better, thanks to improvements that have occurred in computer science and related fields because, while it is clear that fully automatic machine translation can only be as successful as the linguistic theory on which it is based, semi-automatic methods rest on a much wider set of factors. In the wide conception, the problem is not so much to build a machine that can translate as to bring about as good an impedance match as possible between man and machine when they are working jointly on translation.

Seeking this impedance match brings up a great many important questions, few of which have been addressed. They are not questions of morphological analysis, of building chart parsers, or transfer components and pivot languages. They are a new set of problems including such questions as how best to involve monolingual people in the course of the total translation process — how to use people that know a lot about the subject of the text being translated, but only one of the languages involved. How can we make a system which, when faced with ambiguities or vague formulations, knows how to present questions to a person in a natural way so as to get them resolved? It should be possible to put the questions in the same kind of language that one person would naturally use , and not the one used inside the system. How can we provide to a translator, or to a person who is collaborating with a machine on translation, access method to the linguistic and other information that will make the job easier? How, for example, can we best give him access to other translations that have been made with the same kind of technical terminology as in the current document?

How can we give him access to the encyclopedic information about the subject matter of documents? What can we do to put him in touch with other people that could provide assistance of one kind or another? What can we do to insure that, once he has negotiated with the machine over the proper rendering of a particular technical phrase, that he will not have to repeat that negotiation when the same phrase crops up again? In what ways can we benefit from situations in which the same document has to be translated into several languages? In the EEC, it frequently happens that a document has to be translated into eight languages, but I have yet to hear the suggestion that the French translation might profitably be taken into account when preparing the German version. Might we not at least provide more intelligent text editing facilities to translators so that they could call for every instance of the word “kid” to be replaced by “child”, and automatically have “kids” to be replaced by “children”.

Above all, everything that the human collaborator is called upon to do must be such as to honor his sense of professionalism and his intelligence. We must by all means never put a professional translator in the position of clearing up after an incompetent machine day after day, because, if we do that, we will rapidly achieve the situation where the number of professional translators, already very small, is reduced even further.