7 • 3 Concluding Remarks 2

M.Kay

Xerox Parc, U.S.A.

I have been more impressed than I know how to say by the quite remarkable achievement made by ten or a dozen Japanese companies in the development of obviously very viable machine translation systems. It is a remarkable demonstration of something for which the Japanese have become justly famous the world over, namely that ability to cooperate effectively on major projects directed towards a single, clearly defined goal. Nobody could fail to be impressed by this achievement. Furthermore, they have provided us with some data of a kind that seems to be difficult to get in other places, namely data concerning the effects of scale. With few exceptions, machine translation research in other countries has been conducted on a small scale — with small dictionaries, limited grammars, and in restricted domains of discourse. All the systems we have heard discussed here involve dictionaries of 50,000 words or more. In due course, we can hope to learn something about the effects of scale on modern machine translation systems that would have been otherwise practically unavailable. I am also heartened by the optimistic view of the future that our Japanese colleagues retain after the experience of building these systems.

This introduction is important, because otherwise it would be too easy to construe much of what I have to say in a negative way. But that would be to construe it wrongly. I have very little to say in the way of criticism of what has been done. What concerns me is what has not been done that should have. But if there is anything in what I have to say on this score, it can, and should, be taken as suggestions for the riches that the future holds for workers in this field. There is a great deal to be done, and it is far more diverse and varied than the experience that has been reported on here would lead one to expect.

If I felt any disappointment, it came from two things. One is the remarkable similarity that I perceived among the various projects and systems that were described, and the other was a general lack of excitement in the scientific discoveries and specific technical achievements that must have lain behind them. I was told, you will recall, that the sameness I perceived came from the competition that exists among the parties involved. These are private companies; they are naturally in competition with one another to produce the best system; and therefore, for some reason, the systems are almost all alike. I am beginning to think there must be something to this argument, because I have heard it before. I have heard it made about airlines in the United States as they have been since that industry was deregulated. Deregulation, we are supposed to believe, was intended to remove all impediments to competition among the airline companies, I expected that this would lead these companies to compete for my custom in many and divergent ways, and that this would naturally lead to diversity. But it does not. The only way I can tell which airline operates the plane I am flighting on is to take out my ticket and look at what is printed on it.

The lack of technical excitement disturbs me partly because I am a technical person and am eager to learn what it is that the builders of these systems must have learned. But it disturbs me even more because technical facts and technical excitement seem to me to be the commodities that it would have been most useful to communicate here. I will go home delighted that more has been achieved in the mechanization of translation than I had realized when I came, but frustrated by the knowledge that there is a great deal that I obviously could have learnt, but did not.

The achievements that we heard here were remarkable. For the most part, they were intended to fill specific needs, such as the translation of technical manuals from Japanese into English. But it is by no means clear to me that they are the best way, or even a good way, to meet those needs. Perhaps we can agree on this: Whatever may have been achieved under the heading of "machine translation", it is not translation in the everyday meaning of that term. A reasonable operational definition of translation is surely what translators do. A professional translator is someone who translates and what a translator does is, by definition, translation. Whatever it is that the computer systems we have heard about do, and however useful it may be, it is not translation in this sense. There is not a machine translation system in the world that could take and hold a position as a translator in any company or agency, anywhere. Whatever it is that these systems do, it is at best only a part of what must be done to produce the kind of result to which we usually apply the term "translation". Do not misunderstand me. I am not simply saying that it needs revision, because that is also true of what human translators do. On the other hand, revision of translations is also something that any professional translator can, and does, do. The point is that the process of translation encompasses more than what these systems do, and it is only reasonable that we should judge what they do by the contribution that they make to the enterprise as a whole.

The point that the word "translation" in the phrase "machine translation" means something different, and strictly more restricted, than "translation" in other usages is important for two reasons. The most obvious one is technical — by examining the broader problem, we may be able to reach a more appropriate solution. The other is political, and of potentially far greater importance in the short run. The difference between the two meanings of the term "translation" is, potentially at least, the difference between the expectations of the people who pay for the work we do, and the results we actually deliver to them. The discrepancy between expectations and the reality has been damaging in our field in the past. Specifically, it leads to the convening of the Automatic Language processing committee of the United States and their writing of the infamous ALPAC report in 1967. Since then,

there has been much talk of a possible second ALPAC report. Now, I do not believe that such a report is likely to be written in the United States, if only for the reason the linguistic imperialism of the United States has sapped any interest that they might have had in machine translation. I do not believe that such a report is likely to be written in Japan because I believe that Japanese researchers have been careful to keep the expectations of sponsors and of the public in line with their own. I believe that the report will be written in Europe, translated be traditional methods into the nine official languages of the European Community because there, the gap between expert and lay expectations is perilously large.

Now let me return to my claim that what we have tacitly agreed to refer to as "machine translation" is an inappropriate response to the need we see. There is, and will likely be for a long time, a part to be played by people in the production of all translations of more than the most basic quality, and of texts that cover any but an extremely limited domain of discourse. The first question we must therefore ask is how the labor of producing translations can best be apportioned between man and machine. Behind machine translation, as commonly understood, is the largely unquestioned assumption that the machine should produce the best translation it can and that this should then be transformed into an acceptable translation by a process called "revision" by translators, and "post-editing" in the machine translation community. Though I think I understand something of why this view of the division of labor is so widespread, I also believe that it is quite wrong if the aims we have in mind are indeed practical ones. I think the view persists because it makes excellent sense from the standpoint of research in linguistics, broadly construed. I have always believed that no problem provides a better matrix for the study of language, and of computational linguistics in particular than machine translation. I surely need not spend time on saying why. Roughly, a machine translation must enshrine solutions to essentially all the problems that natural language presents (except possibly that of learning), and the embodiments of the solutions to all of them must be coherent in the sense of fitting into a single operating machine. People who are attracted by the problem of machine translation see it

in this way, and this is the attraction it has for them.

However, I believe that machine translation, in the usual sense, is grossly inappropriate as a solution to the supposedly pressing practical problem of producing usable translations; so much so that I am sometimes lead to the belief that those problems are manufactured, or at least magnified, by those intent on doing the research. It is clear to the most naive observer that a machine can do some of the jobs involved in producing a translation extremely well, and others extremely badly. It can format and help reinsert the figures, it can find words in glossaries, it can search the text for other instances of a word or phrase, and it can check spelling. It cannot find the referents of pronouns and definite noun phrases, distinguish subtle differences of meaning, recognize or reproduce irony, or replace a metaphor by a culturally more appropriate one. Its performance on other things would be less easy for a lay person to assess and could change more rapidly as a result of research. An appropriate technology would presumably be one which consigned to the machine those parts of the job for which it was best suited and to the man those for which his special abilities are still crucial. Perhaps an appropriate technology would also be one that gave a role to several different people, each with special skills. Certainly, it would be one that provided a smooth interface between the various contributors, human and computational, and this is itself a question on which recent advances in our understanding of interactive systems bears.

There is good reason to be encouraged that the kind of appropriate technology could be devised, and good reason to believe that it would look very different from what we have been calling "machine translation". Notice that, generally speaking, the parts of the job that people find especially onerous and difficult tend to be those that would benefit most from mechanical contributions. They are the rarely occurring phenomena — words and phrases that are used in a special way in certain fields; questions of maintaining terminological consistency over a long text or in the work of different translators, replacing every occurrence of this by an occurrence of that. Unfortunately for the proponents of machine translation, machines are worst at the most frequently occurring things — deciding whether a pro-

noun refers to this or that, or what is being conjoined by a given occurrence of "and", deciding which of several meanings of a word is in play, and the like. Most of these commonly occurring problems do not appear as problems to a human translator, or any ordinary speaker. The moral should surely be clear. We should neither take the valuable time of a professional translator solving routine problems that would be much better given to a computer, nor give over to a machine problems that can be solved quickly and easily by a person. We should seek an appropriate compromise.

If men and machines work together in a fairly closely coupled system, then tasks can sometimes be profitably given to the machine even if the expectation of a good outcome is quite low. This is because, in closely coupled system, a small error that is recognized by a human member of the team does not give rise to further errors, but is caught before it has had a chance to produce a cascading effect.

The architectures that a complete translation system might have lie on a continuum with traditional methods at one end and machine translation at the other. At one end, there is only the translator with his background, his glossaries, and his network of friends. At the other is the machine, with its algorithms, heuristics, and data bases. In the middle are a variety of arrangements, each giving a different proportion of the initiative to the man or the machine, and apportioning the responsibility for the tasks variously between them. One family of architectures gives the initiative mainly to the person. He makes the translation, calling on the machine only occasionally, and when he sees fit. The members of this family are often referred to as "machine aided translation". But this term is also used for a minor variant of standard machine translation in which the machine produces a complete text — either all at once, or piece by piece — with results that are acknowledged to be so bad that they are to be though of aids to the translator rather than translations in their own right. At another point on the continuum, is a family of systems in which the machine retains much of the initiative, and is primarily responsible for preparing the translation, but with the ability to turn for help to one or more human consultants from time to time.

It is not clear at what point, or points, on this continuum the most efficient systems would be found, and is not clear how the points should be expected to move in the near future. It is entirely clear that these questions are receiving essentially no attention. It is clear, to me at least, that our enterprise and the communities that presumably require our products would be best served by a broad attack on the problem at various different points along he continuum. If this view were to take hold, then I would presumably no longer experience at the next Machine Translation Summit the impression of sameness and technological conservation that has characterized this meeting for me.

In the absence of any strong reason to do otherwise, I would recommend directing our first attack to a point on the continuum fairly close to the traditional translator. I have two main reasons. First, this is quite clearly the best way to get the best return on investment in the short run. I have little doubt that the productivity of translators working on technical texts could be dramatically improved — perhaps by a factor of two — by importing simple but appropriate technologies into their work. The cost would be low and the payoff extremely high. Second, a small amount of machinery in the translator's work place could not only improve his productivity, it could also serve as a means of gathering data on what a translator actually does when he is translating a document. If one fact about machine translation research is more surprising than all others is that it has never, even for the briefest moment, awakened any interest in the way translations are done by the only devices we have that are able to do them at all well. We have absolutely no data, as far as I know, on the human activity of translation. Without this, it is hard to see on what we can base our claim to expertise.

If you agree with me that we will see best where technology might be most profitably brought to bear on the translation problem if we take as wide a view as possible of the problem, then suggestions flow in from all directions. Mr. Rolling made some remarks yesterday that suggested to me a potentially fascinating line of enquiry possibly leading to profitable results. He pointed out, you will recall, that there are very great discrepancies in the amount of text that gets translated into different languages — far more into French than into Greek. If an English text must be translated into Greek, then the chances are that it will also have to be translated into French and several other languages, whereas the converse is not true. It follows that, when a text must be translated into Greek, we can usually expect to have versions of it in other languages, notably French, on hand. Now I ask you, should it not be easier to find the information required to construct a Greek version of this text, given inputs in English and French, rather than simply on the basis of the English text. Are we not entitled to expect some, if not many, of the vaguenesses and ambiguities of one of these sources to be supplied by the other? If we are really driven by the need to produce translations of difficult texts into little known languages, should we not therefore be considering how information from multiple versions of a text can be combined?

Let us suppose, doubtless somewhat simplistically, that the number of users of the IBM PC in a country, or among the speakers of a language, is roughly proportional to the size of the population, so that we should expect there to be twenty times as many users in France as in Norway. We should also expect that new releases to be a great deal more expensive in Norway than in France because the cost of translating the manuals, though presumably just as costly when the language is Norwegian as when it is French, has to be spread across a much smaller user community. But it is clearly not commercially realistic to charge orders of magnitude more for each new release of MS-DOS in Norway as is charged in France. The point is that the economics of translating these manuals is quite different, depending on the target language. But, once again, it is reasonable to assume that we could have the French version of the manual, and possibly the German and Italian ones also, when the translation into Norwegian was undertaken. I therefore claim that, research into machine translation, human aided machine translation, and machine aided translation from multiple sources could fill an obvious need but, like every departure, however slight, from the original paradigm of fully automatic, autonomous, machine translation.

There are other situations in which it may be possible to supply to a translation system more information than is contained in a single text. Consider the situation where I send a letter to a foreign place asking for some information. I receive a reply in another language which must be translated. Consider, now, a system that could be supplied with the letter to translated together with a copy of my original request on the supposition that the letter should be translated as a response to the request.

My claim, then, is that machine translation, as usually understood, provides a fine framework for scientific investigations, but that it is only one of a great many ways to approach the problem that presumably besets us, of producing more translations of higher quality more rapidly. Furthermore, of the limitless variety of approaches to that problem, it is the one that has consistently shown least promise.