

LETTERS TO THE EDITOR

Two years ago at the MT Summit held in Hakone, Japan, Martin Kay raised the question whether the world was not getting ripe for another ALPAC. Kay had a point. Eurotra has been running now for about ten years, roughly the same length of time MT ran in the US before ALPAC reared its head. The same holds true for the MT efforts undertaken in Japan under the Fifth Generation umbrella. Since none of these recent efforts can claim to have finally made the case for MT, one wondered with Kay whether indeed ALPAC II can be very far off. As it turned out, in the case of Japan, Kay's question struck a nerve.

In April of 1989, I attended an international MT conference in Oiso, Japan, where the entire first day and much subsequent discussion were devoted to a consideration of ALPAC and the Japanese experience. It was rather enlightening and bears retelling in these pages, I think, particularly as there were only a score or so of westerners present, and less than a handful from the U.S.

To refresh memories, let me begin by summarizing some of the more telling points of the ALPAC Report (1966). I take the liberty of rearranging the order in which these points were originally made.

First, and rather curiously today, the report does not really bother to argue that MT had failed. The point is made as it were *ipso facto*, by merely including without comment some raw output from these first generation Russian-English systems and letting readers see for themselves. It takes a modern-day reader a moment to realize that post-editing was not a self-understood concomitant of the technology in those days, that Bar Hillel's "fully automatic high-quality translation" criterion (FAHQT) was still the understood measure. But a reader today looking at this raw output cannot reject it out of hand. At least one of the translations included in the report (by the Georgetown system) looks promising, and can easily be visualized as a basis for a useful finished work, given a reasonable amount of post-editing.

Second, the report argues that even allowing for post-editing, these systems have never shown themselves to be cost-effective. This conclusion is based on the Arthur D. Little (ADL) study of the production experience with the Georgetown Russian-English system in the Foreign Technology Division (FTD) at Wright-Patterson AFB, a study clearly unsympathetic in spirit.

Third, the authors seek to persuade that these systems in any case were at a standstill in terms of improvability and that they therefore had no future. This is not so much argued as implied in a kind of sleight-of-hand way by exhibiting a dozen perfectly translated "select" sentences from the much earlier Mark I/II systems of IBM (1954)

and inviting the reader to note how much better these are compared with the raw output of the later systems. We are supposed to conclude that over those ten years the technology has actually retrogressed. They also quote Victor Yngve's more telling acknowledgment that "work in mechanical translation has come up against a semantic barrier" (p. 24).

Fourth, they claim that they were unable to establish a need for the technology in the first place. "The supply of translators greatly exceeds the demand (p. 9)." No government agency could be found, for example, that was prepared to state it was unable to meet current translation needs with the human translation resources available to it. The authors of ALPAC even considered whether there was "a possible excess of translation" (p. 13).

Finally, it was suggested that there were better ways for the federal government to spend these R & D dollars. You can appreciate the tenor and intent of the document when you encounter this last point in the transmittal letters found among the very first pages. The progenitors are concerned lest procurement circles overreact to the report and cut off support for natural language study altogether. To forestall such a consequence, a case is made in the very first pages for funding large-scale computer-based natural language research—research, one understands, that would not pursue the chimera of MT but might conceivably one day make a contribution to its feasibility.

The swift and utterly devastating consequences of this report to the fledgling MT community is a matter of history. Although activity in MT was never entirely extinguished, not even in the U.S., most thinking people took the report to signify that MT was a dead issue. The report went pretty far in that direction: ". . . we do not have useful machine translation. Further, there is no immediate or predictable prospect of useful machine translation" (p. 32). But perhaps, as some claim, the authors did refrain from sealing the verdict once and for all, and even went so far as to quote Yngve as saying he believed high-quality machine translation would be available one day (p. 24). But the report was received otherwise, as a scientific finding that MT was not feasible. It would take almost a generation for that consensus to change.

Today, in academic circles, the nonfeasibility of FAHQT-MT is of course a nonissue and MT with post-editing is increasingly taken seriously whether demonstratively cost effective or not. Many governments too, recognizing need, have set up long-range plans for the development and productive use of MT. The last bastion of ALPAC, as it were, seems to be the U.S. government advisors and planners who are heir to received wisdom and who show

little inclination for second thoughts on the topic. This at least is the impression created when no one from those circles made an appearance in Oiso, at a conference designed at least in part to reconsider ALPAC. But now this too may be changing.

So be it for background.

Let's look at these five points from the Japanese perspective as it was revealed at this conference.

1. The official spokesmen at the conference were all refreshingly candid about where Japan collectively stands today in MT. By their own estimate they are still considerably short of the mark. We should not be misled when we read, for example, that Fujitsu's ATLAS system has been installed in over 300 sites. A speaker from Fujitsu stated that only 10% of the installed base actually use the system. Casual conversation with conference participants revealed that even this number could be misleading. Users who do make use of the system apparently do so by constraining the input. As one of these users explained it to me, the system translates "baby sentences" very well. Another speaker acknowledged openly that the defects of the first generation system cited by ALPAC apply equally to them. Professor Nagao summed up the situation in his opening statement: "Generally speaking, the systems are quite imperfect and are not easily used by ordinary persons."

As it turns out, then, the Japanese systems today (there are about 10 commercial offerings) are producing raw translation that is probably not much better in quality than that which ALPAC looked at and rejected 23 years ago (though clearly their task is more difficult, given the nature of their language). What's different, of course, is the fact that the Japanese prefer to see the glass as half full.

2. Japanese users reported that the current systems are able to reduce E-J/J-E translation costs by roughly 30–40%. This would seem to indicate that the Japanese have come a little farther, even with their more difficult language. And that may be so. By contrast, the ADL study quoted by ALPAC shows virtually no cost advantage to machine translation at FTD. But this difference might simply reflect the fact that the Russian–English operation was government run, and also the fact that the Russian source documents were not pre-selected for their suitability for MT.

What is more interesting about this, however, are the contrasting ways in which practical experience with MT was and is viewed by the two mentalities. The Japanese attitude reflects the national talent for improving on the given, for taking an imperfect work and effecting those "thousand tiny steps" to make it better. If that is a Japanese stereotype, evidence for it abounded at the Conference. For example, after disposing of ALPAC, the rest of the time was spent considering how to effect incremental improvements to productivity. (This was not a linguists' show.) They spoke of obviously possible things that can be done to improve the numbers, like six-month training programs for post-editors, schools for which have already begun to operate. They want to establish post-editing as a

recognized profession and they asked if anyone could suggest a name for it that had more positive connotation. They talked about introducing technical writing courses into the curriculum to propagate a national technical writing style, something that would be new for Japan and that would greatly simplify the problem of machine analysis. They talked about constraining the input manually, by establishing writing standards, and mechanically, by writing automatic and semi-automatic pre-editors. For example, constraints on writing would require that complements be adjacent to the element they modify. A semi-automatic pre-editor, for example, would enable a user to bracket Japanese sentences that were not composed according to this standard. They also spoke of creating a simplified Japanese ("small Japanese") for use in preparing documents destined for machine translation (akin to "Caterpillar English" and the practice at Xerox).

In short, the Japanese do not want to abandon MT, they want to make it work.

In contrast to this Japanese predilection for patient fine-tuning, the ADL report and the use made of it by ALPAC strikes one as American bottom-line thinking at its worst. But that may be unfair. By far the biggest piece of the total translation cost in the FTD operation went for post-editing, and the system administrators could hardly ask Russian scientists to write in "small Russian." Nor would it have been obvious then that translators' workstations were coming and that productivity gains can come about from sources other than linguists' rules. Having said that, one still regrets not finding any sign in ALPAC of the other mentality, which asks how these systems might have been improved and made to work, which sees cost factors as opportunities. Who knows what might have happened had they done so.

3. On the matter of the improvability of core technology, the Japanese developers seem to agree with their ALPAC predecessors that the technology they are looking at is dead-ended. They may not say as much, but knowledgeable people like Harry Somers (UMIST), echoing Yngve, confide, "they know they are up against a wall." I asked the head of a major Japanese development laboratory about this. I asked whether further progress in machine translation would come about from extensions to their present systems or from entirely new technology. His answer was so prompt as to be abrupt: "From new technology."

The Japanese see what the ALPAC authors saw, but the conclusion they draw could not be more different. Why is this? Why are the Japanese not discouraged? Their technology today is vastly more sophisticated, but the bottom line results do not place them much further along than where matters stood in the days of ALPAC.

But instead of discouragement, Japanese planners are busy laying the ground for a fresh coordinated attempt at a new generation of MT, to be realized in 10 years, around the turn of the century. Part of this plan is already in motion. A large electronic dictionary, for example, has been under construction for several years in a project

jointly sponsored by government and industry. The work is being carried out by Japan Electronic Dictionary Research Institute (EDR) in cooperation with many other organizations. In addition to the traditional terminology bank (Japanese and English, initially), it will include such things as semantic networks and thesaurus-type orderings. There will also be a co-occurrence dictionary. The work will go on for another six years. There were no details about the MT superstructure to be built on this platform, and one had the impression that it was as yet largely unspecified. But there was no doubt of the broad commitment to this new undertaking and of the importance being attached to it.

4. The reasons for Japan's commitment to MT are not hard to find. The situation their language puts them in vis à vis the rest of the world could be cause enough. I won't belabor the obvious, but the numbers are interesting. We heard that over 200 million pages of translation are done annually and that this number will likely double in two to three years. Not surprisingly, the preponderance is Japanese-English, no longer the other way around, in keeping with the trade picture. Japan represents a six-billion-dollar annual translation market, growing at the rate of two billion a year. For Japan, translation and foreign trade are inseparable realities. Europe too is no stranger to this, and 1992 can only make translation more of moment there. As for the ALPAC authors, we may excuse them for not having made this connection themselves, given the circumstances of the time. The annual U.S. government outlay for translation in those days was 13.07 million dollars.

But the conference revealed deeper reasons for Japan's commitment to MT, reasons that could hardly have occurred to the ALPAC authors in their day and that might surprise some of us today. One of these is the long-range contribution machine translation is expected to make to information technology generally. Like many, the Japanese see information technology as *the* technology of the 21st century. Like no one else, however, they have assigned machine translation the role as "test bed" for developing and proving the critical engine for this technology—language understanding, that will give information-handling circuits of the future their intelligence, and that will enable machines to perform complex conceptual tasks and engage in humanlike communication.

There's more. They spoke also of a kind of geo-economic motivation for MT. We were told about an Asian multilingual project being developed by the Center of the International Cooperation for Computerization (CICC), which on the surface looks like an Asian Eurotra. The CICC project is a joint development program involving Japan, China, Indonesia, Malaysia and Thailand. The system will translate from and to the five respective languages and, when completed, will serve as the basis for a multinational information net, allowing for regional economic planning, cooperation, and data sharing. There were very few details except that Japan has committed 50 million dollars over

the next six years. (About the cost for the engines of one Stealth bomber.)

One wonders what the authors of ALPAC would say on seeing lowly machine translation, with all its warts, being given so lofty a role as that to which the world's leading economy has assigned it. MT is being asked to help secure that economy's place in the vanguard of global technology, and as if that were not enough, MT is being prepared as infrastructure for an entire region's economic future. One wonders whether we will soon find ourselves looking to Japan for "smart" circuits the way we do today for automobiles and VCRs. One wonders whether the ALPAC authors, if they had had forevision of all this, would have changed their recommendation. Perhaps not.

5. The ALPAC authors, of course, recommended that practical efforts at MT be abandoned and that funds be diverted to more theoretical studies and to basic research in computational linguistics. Breakthroughs in the handling of natural languages they believed would only come from such sources, not from slogging it out on the front lines of MT. The Japanese naturally agree on the need for basic research and for breakthroughs that might come from such research, but to their minds the real world involvement of MT provides the best environment in which to incubate such breakthroughs, to verify them and then to bring them to maturity as usable technology.

This is an interesting question. The ALPAC authors acknowledged that MT had made an important contribution to the advent of computational linguistics (p. 30f). But in the future the shoe would be on the other foot. Has this been the case? Were the ALPAC authors correct in assuming that progress in MT would only come about from research, albeit large-scale and computer-based? Or is the development of MT in the final analysis an inductive nitty-gritty business where you learn as you go, where theory comes after the fact (as in much of science) and where the closer you stay to the real world of language the sounder your work will be?

If ALPAC had not seen the glass as half empty 23 years ago, might it have been almost full today? An interesting question. Certainly the way the Japanese have answered Martin Kay's ALPAC challenge must give pause for thought among those whose thinking is still influenced by this legendary document.

Bernard E. Scott
Lógos Corporation
Mt. Arlington, NJ 07856

I am delighted that *Computational Linguistics* is increasing the number of books on speech recognition that it is reviewing. Speech recognition is an area of particular interest to me and I look forward to seeing many more reviews.

I am writing because I found Joan Bachenko's review of *Prosody and speech recognition (Computational Linguistics 16(1), p. 46)* by Alex Waibel disturbingly out of date.

By 1988, when that book was published, there were already commercial speaker-independent systems that exceeded the descriptions she provided (e.g., speaker-independent systems can handle “one to five phonetically distinct words”). By 1987, one speaker-independent, continuous-speech system, for example, already had 20,000 words—considerably larger than the 1,000 words Ms. Bachenko cites as the maximum for large-vocabulary systems.

In 1990, there are commercial speaker-adaptive systems that recognize 30,000 or more words. Those systems operate with all of their vocabulary potentially active (which may be what Ms. Bachenko meant by “on-line”). There are also speaker-independent systems that recognize up to 120 words; speaker-independent systems that recognize 16 or more words over the telephone; continuous speech systems that recognize 7,000 words; and the speaker-independent, continuous-speech system I mentioned earlier now has more than 40,000 words.

Speech recognition is a dynamic, rapidly changing field. It is far more advanced than Ms. Bachenko (and, perhaps, Mr. Waibel as well) appears to think. It disturbs me that your readers are being given an inaccurate view of it.

Judith A. Markowitz
J. Markowitz Consultants
8439 West Catherine, Suite 225
Chicago, IL 60656

Ms. Markowitz is correct in noting that, as results presented at last April’s Speech Tech conference showed, my assessment of present-day speaker-independent and speaker-trained systems was too pessimistic. However, I feel that, equally, they show that her picture of recognition technology is too optimistic. With so many people waiting so eagerly for a speech-to-text system, I am uncomfortable with any stance that encourages unrealistic expectations from this technology. I personally know of a situation in which handicapped people have been misled by vendors whose enthusiasm for new markets overcame their better judgment.

Joan Bachenko
AT&T Bell Laboratories, 3D462
Murray Hill, NJ 07974

The brief notice by G. H. [Graeme Hirst] of my *Syntactic phenomena of English* (*Computational Linguistics* 16(1), p. 56) contains the baseless accusation that I “attempt to be relatively theory-neutral.” In my book I in fact take positions on a large number of issues of syntactic theory and semantic theory and arrive at conclusions that are sure to offend those advocates of other approaches to syntax who take the trouble to read what I am saying. Perhaps G. H. has fallen into the common error of using *theory* only with reference to the better-known name-brand packages of theory and metalanguage, in which case it would have been

more accurate for him to say *nondenominational* or *non-aligned*. If my book is deficient in explicit vilification of the name-brand approaches, it is only because I regard their advocates as potential customers; I attempt to write books that those with whom I disagree on matters of linguistic theory will find evil but indispensable.

James D. McCawley
Department of Linguistics
University of Chicago
Chicago, IL 60637

Professor McCawley is playing prescriptive linguistics. However, the descriptive data in the *OED* suggest that a *theory* may indeed be a package (or system, or body of principles), and a *theory-neutral* text is one that is not expressed in terms whose use presupposes the acceptance of some theory. The preface to Professor McCawley’s book seems to be fairly explicit in stating his preference for describing syntactic constructions in simple English rather than “notational systems . . . embodying grossly inaccurate presuppositions about what factors play a role in syntactic phenomena” (p. xii). Notwithstanding this, I am happy to withdraw, without reservation, any suggestion that the relative theory-neutrality of his book is the result of any intent on his part.

Graeme Hirst
Book Review Editor
Computational Linguistics

The following reference was omitted from my paper in *Computational Linguistics* 16(1):

Katz, S. M. (1987). “Estimation of probabilities from sparse data for the language model component of a speech recognizer.” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. ASSP-35, 400–401.

I wish to apologize to the IBM speech group for this unfortunate mistake. It was certainly not my intention to overlook their contribution to statistical language modeling techniques.

Kenneth W. Church
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

This was indeed an editorial oversight for which we apologize.

James Allen
Editor
Computational Linguistics