Martin KAY, President of the ICCL

Ongoing directions in Computational Linguistics

This is the fifteenth International Conference on Computational Linguistics.

It is arguably the fifteenth COLING, although we only adopted the name of a Swedish hobo as our nickname after the third meeting.

Since the first meeting in 1965, this is only the third time that we have returned to a country where we had been before. The first time was in 1984, when we returned to the United States, and the second in 1988, when we returned to Hungary. This year we return to France, where the second conference was held in 1967.

For those who, like me, have been associated with these meetings since the beginning, it is a privilege and an enormous pleasure to return to France, and to a conference organized under the auspices of the University of Grenoble. It is too little recognized how much the field of computational linguistics owes to this country and to that university. My predecessor, the second chairman of the International Committee on Computational Linguistics, the late Professor Bernard Vauquois and the machine translation center that he founded in Grenoble, have done more to shape our field than any other single person or center. They were the only major academic research group to live through the dark ages that followed the ALPAC report and their Ariane system has become the model for the great majority of the commercial machine translation systems that have ever been built. Professor Vauquois, and his students and colleagues have been missionaries for, and tireless teachers of, computational linguistics for thirty years, establishing new research centers as far away as Malaysia.

For a computational linguist, to come here is, in a very real sense, to come home.

In recent years, computational linguistics has been returning to its beginnings in some other ways also.

Much of the driving force in our field comes from the desire to make a translating machine, not just because this was the first problem that we attacked, but also because it is a problem that encompasses all others — it is very hard to imagine any achievement that would count as a contribution to computational linguistics without contributing to machine translation.

But, while it lost none of its motivating force in the intervening years, machine translation received somewhat less attention because the perception has been that the need for machine translation was less than had originally been thought. Now, the need is thought to be greater again, and growing. So, once again, machine translation, machine-aided translation, and machine aids for translators are coming to claim more attention, especially outside the United States.

In the early days of computational linguistics, one of the great opportunities that computers seemed to offer was that of performing massive statistical analyses of running text from which it was hoped that much of the hidden structure of language would emerge. The idea fell into the background because it became clear that, if such a program could indeed be carried through, the amount of data that would have to be considered was still beyond the reach of the machines and techniques that were then available.

he machines are now bigger and faster; orders of magnitude more data is readily available in machine treatable form; and much sharper tools have been developed. Someone entering the field of computational linguistics today will no longer be able to ignore statistics and corpus-based techniques.

But, our return to France, MT, and statistics, does not mean that, to quote Yogi Bear, it is just "déjà vu all over again". The old problems remain unsolved, but the relative naiveté of the fifties and sixties has been replaced by a notion of appropriate technology — of the impact that can be made on practical matters without having solved all the problems necessary for complete automation.

The TAUM-METEO project in Montreal demonstrated clearly and cleanly that we could do useful things with sublanguages that we could not do with unrestricted languages. Machine translation systems all over the world have shown that, when used appropriately, there is value in initial translations of altogether lower quality that would once have been thought interesting.

Interactive methods have shown us how to profit from the complementary skills of people and machines, allowing each to supply the deficiencies of the other. In short, we have learnt to approach practical problems with greater humility and greater realism.

These are some of the reasons that make me especially happy to welcome you all to to France, to Nantes, and a week of excitement at the 15th International Conference on Computational Linguistics.

Palo Alto, Friday, 8 May 1992