And after the Second Generation ...

JOHN HUTCHINS
University of East Anglia, University Plain, Norwich NR4 7TJ, U.K.

In discussion of the future development of research in machine translation (MT) it is probably useful to make a preliminary, if somewhat simplistic, distinction between two basic types of activity. On the one hand, there are research programmes dedicated to the exploration of methods, techniques and programming languages which are untried in the field. These embrace the small-scale projects which explore the application of a linguistic theory, such as LFG or GB, to some problematic area of MT, the projects which investigate example-based or statistics-based methods, the projects which explore the feasibility of a particular programming environment, such as Prolog, connectionism, etc. The essence of this type of research ('exploratory research') is to push back the frontiers, to open up the field to innovative approaches.

This mode of research may be contrasted with efforts to build prototype MT systems, which may or may not form the basis for operational MT systems. The objective in these projects - typically on a larger scale and funded for longer periods - is the development of systems which undertake the whole translation process (with or without human intervention). This type of research ('prototype research') is inevitably less innovative. The basic aim is to cover a wide range of linguistic phenomena and to establish the framework for large dictionaries. It has to be based on a relatively solid well-established foundation of reliable and tested methods and techniques. It cannot change radically in response to the latest advances. Instead, this type of MT research must attempt to provide a flexible but stable framework; its conservatism is a strength rather than a weakness.

It should be remembered that any prototype system with pretensions to eventual commercial viability must work reliably and efficiently, and this means they must have been well tested. Bringing a MT system out of the laboratory into the marketplace can take many years; Siemens estimated that they would need 15 years to convert the LRC METAL prototype into a commercial system – in fact they took less time, but nevertheless this, the most sophisticated commercial MT system, is based on linguistic and computational techniques now ten or more years old. It ought to be no criticism of commercial systems that from the perspective of current research they are outmoded. The criteria for their evaluation should not be innovativeness but the quality of output and whether actual performance matches the claims of makers.

In the late 1970s there was a general consensus that the model characterised by Harold Somers as the 'second generation' type was the most suitable basis for the development of systems which promised higher quality output than most pre-ALPAC designs. It was linguistics-based (essentially syntax-based), non-interactive (batch processing with post-editing), stratificational. During the 1980s the range of options broadened and it is now obvious that this is not the only possible design for MT 'research prototypes'.

The evident passing of a long-standing orthodoxy as far as MT prototyping is concerned raises inevitably questions about the foundations of MT research itself. What is often raised is the lack in MT research of an agreed theoretical basis. Researchers borrow much of the theoretical baggage and apparatus of related disciplines, notably artificial intelligence and linguistics. But MT research should not be regarded as simply the application or testing of theories from outside – even if, legitimately, it can be regarded as an excellent testbed for formal linguistic theories, since the output can be evaluated independently of knowledge about the theoretical basis of the system. Whereas AI research may be seen as the development of computational models which simulate human behaviour in some way, MT research cannot be defined as the development of computer programs which simulate human translation. Rather, MT research aims to develop programs which actually translate; and which are certainly not simulations of human translation. More crudely, while AI may be a theory of intelligent behaviour, MT is not a theory of translation; it is translation.

What kind of theory would be relevant for MT research? If MT is seen as essentially an 'applied' science (the application of computational linguistics, etc. to translation) then a theory could be defined in practical terms as the embodiment of those techniques, methods and approaches which are known to give desired results. On the basis of such a theory it should be possible for any researcher to predict the effect of using a particular technique in a given circumstance, e.g. using a particular parsing method to produce a particular desired representation. There may be a considerable body of knowledge which is relevant, both implicit and explicit, and which may be derived from the experience of MT researchers, but what is lacking is the appropriate abstraction. A theoretical basis of this kind could serve, for example, as a corpus of theses or hypotheses for refutation or for confirmation in experimental research projects.

However, such a conception of MT theory is unlikely to be either satisfactory or sufficient, since MT research is not generally seen as a field of 'engineering'. What is intuitively desired is a theory based on principles independent of particular practices. The Eurotra researchers developed a theory which defined translation as a series of transformations of linguistic objects (representations of texts) from a source to a target language. The Carnegie-Mellon University researchers define translation (broadly) as the