

HOMUNCULUS LOQUENS

In the invitation to the electronic colloquium the topic was described in the following way:

A labour-saving robot, a man-made monster and/or a self-portrait? What are we computational linguists really working at?

The stunningly successful development in recent decades of "intelligent" man-made systems have focussed attention on limits of what machinery could ever do, and on the fundamental issue how machine-like we are ourselves. Whenever we can mechanize something which seems deeply human, we gather urgent, often painful, knowledge about ourselves. Whenever we fail, we may learn even more; it is not only in thermodynamics that the great failures mark the great advances. And since nothing appears to us more human than our human language, what engineering triumph could be greater than machines which manipulate and comprehend language? And what could be more disquieting? Computational linguistics is about some very serious questions.

These questions are not all new. Mathematicians studied automata long before they had programmable machines, and Descartes was not the first philosopher to design mechanistic models of the mind. The human effigies which come alive in lore and fiction speak and listen to human speech - and fail at crucial moments to do exactly that, turning unresponsive and frightening if they do not get the right synonym.

How ancient are, at bottom, our ideas and concerns? We shall try to identify some origins of hopes, fears and beliefs underlying modern attempts to make machines read and write, talk and listen.

Comments on this topic were invited to homunculus@com.qz.se. The following is a brief account of these comments which will be elaborated upon in a panel during Coling 90 and later revised and published in bookform. The invitation to participate in the electronic colloquium therefore remains valid!

The computer and the modern methods for computation have introduced something radically new. As Martin Kay once put it, until recently there existed only one symbol-manipulating system in the world; now there exist two. We recognize that we have only seen the primitive beginnings of the consequences - intellectual, economic, social.

And yet: Is it true that our ideas are so new, our situation so unique? Just because our work is future-oriented, we have an extra responsibility to look back now and then. Which occasion could be more natural to widen our perspective along the time axis than when we now meet at a European university?

The ambivalent attitude towards the tools we have been given and are enhancing - delight and awe at the power we may let loose and the secrets which might surface - is shared by others who work with and reflect on computers and computerization in a wider perspective. But for us linguists the conflict stands out even more sharply than for other makers of 'intelligent' artefacts: if we believe that exactly language is what makes us differ from other animals and things then our very success if frightening, a threat to our self-esteem.

The idea of an automaton which can be our intellectual superior in several aspects has a century-long history in mathematics and logics, where the limits of what such a machine could possibly do has been discussed in successively more precise terms long before the first computer could be manufactured. The conflict between delight in algorithmization and distress at too successful reductionism is not new (Jürgen Kunze, Berlin). Leibniz, who dreamt of the Logical Piano, focussed much of his thinking exactly on language (Lenders). Unlike many unsuspecting information engineers of today it seems Leibniz recognized, in his later years, that his grand idea of the Universal Language was too simplistic.

To the question how ancient are the attempts to make a mechanic model of human language Esa

Itkonen, University of Turku (pekka@kontu.utu.fi) gives the answer "Not more ancient than Panini (about 350 BC)". Panini's grammar, or rather the grammar-cum-grammarian as Panini conceived it, can be seen as machine. It is a device for converting meanings into form. First ('vertical') 'expression rules' replace meanings by abstract forms. Second, ('horizontal') 'combination rules' put together abstract sentence-structures. Third, ('vertical') 'substitution rules' produce concrete sentences in a step-by-step fashion; this last type of rules is closely analogous, says Itkonen, to context-sensitive rewriting rules. It is the purpose of Panini's grammar to generate all and only the correct sentences of Sanskrit. Itkonen elaborates the analogy between Panini's rules, meta-rules, rule ordering etc and their modern counterparts.

After a detailed comment on the idea of the talking computer as a selfportrait Damjan Bojadziev, Lubljana (damjan@ijs.ac.mail.yu), summarizes: Yes, if it has its own self-portrait, to which it is related in a way similar to whatever way it is in which we are related to ours. This observation links to Goethe's Homunculus whose existential problem, on which the "Faust" hinges, is that he wants to become existing: he has clearly no linguistic handicap in expressing himself (karlgren@com.qz.se).

The concept of man as a mere machine has provoked fierce conflicts in the past, the most wellknown being those between the cartesian and the Church. Rolf Lindborg, University of Lund, reminds that Descartes did not say that the human mind was a machine; in fact, he insisted that it was not.

The kind of machinery to which we have compared ourselves has changed over the centuries: the generalized meaning of 'mechanical' - as used in, say, Mechanical Translation - reflects a technology which produced such clever devices as astronomical clocks, piano-playing dolls and machines for intricate weaving the term 'machine' goes back to antiquity when impressive machinery was used in theatres to portray men and gods. Other phantasies, like Goethe's Homunculus and Frankenstein's monster, are extrapolations of what chemistry can achieve. Other myths leave the transition from inanimate to animate unaccounted for, as magic, but they may still reflect attitudes towards what we would call technology. - In each period, the high-tech, what-ever it then was, has been frightening as well as discouraging; no colloquium participant has ventured an answer on how long this situation has subsisted, but it was suggested that the fear that the Thing might take over goes back on earlier experience from taming strong animate slaves - animal or human.

A few comments reacted against the view that the over-smart machine is round the corner. Everyone with some programming experience would know that the ceiling is very low. The risk, if any that machines outwit us is very remote indeed and neither practically or morally a concern of the living generation. Computational linguistics, even more than other computation, has taught us the complexity of the human mind. The remaining distance seems longer now than in Leibniz' days.

For some artefacts which spring alive in the old myths the ability to speak and listen is often taken for granted, just as naturally as animals and plants talk in fairytales. (The equally mythical robots of early industrialism were not supposed to talk but to keep on working; they frightened by not listening.) But surprisingly often, and surprisingly early, the artificial men show their subhuman character exactly by their linguistic deficiency: they start and stop on hearing specific commands but are too dumb to understand synonyms, stonewalling in a way which is only too familiar to users of many computerized systems. Like the giants, the artificial men do dull work better than we but they do not possess the human language. That is the comforting message of many of these old tales.

And of their modern counterparts. The dream of the unreflecting robot is the employers' delight (Rolf Lindborg). For many others, the happy end of the story is when the monsters fail, as every monster ultimately must, to use language in a sensible, that is human, manner.

Hans Karlgren