

PANEL I - DISCUSSION

DOSTERT: I note in your paper, Mr. Joos, at the end, the same note of pessimism which was found in your review of Locke and Booth's collection of articles on MT. I don't know whether I am reading you correctly—or translating you correctly in terms of my concepts—but you seem to think that we are something in the nature of a millenium away from the solution of the problems of the translation of prose by machine processes. Do I interpret you correctly?

JOOS: The millenia have a way of getting shorter these days.

DOSTERT: You wouldn't care to indicate the rate of shrinkage in respect to machine translation, would you?

JOOS: As I remarked in a publication which will be out in a couple of weeks now, the *Readings in Linguistics* volume, the development of linguistics in this country from 1925 to 1955, 30 years—one generation—is about the same in extent as the development of mathematics from Newton to about 1850, or in chemistry from Dalton's atomic hypothesis of 1810 approximately down to nuclear fission, about two generations. So we have skipped from 200 years down to 60 years, and now down to 30. Thus the time scale keeps changing. The culture is developing and technology is developing more rapidly all the time. I do regard machine translation—which I would consider as being adequate as being in some sort of competition for what I can do myself in some of the languages I know—I do regard that as millennial. But considering the very swift progress in technology, possibly my grandchildren may witness it.

DOSTERT: It may be pertinent here in commenting on Professor Joos's remarks, to indicate the orientation of some aspects of our research. We have taken an existing English translation of texts in the field of organic chemistry. We have tried to standardize the English texts. That is to say, to remove from the translation some of the unnecessary, superfluous language and some of the idiosyncrasies which the human translator, who is usually frustrated in his creative processes when he translates, inserts into the translation in order to somehow manifest his personality in what he produces. After standardizing the text we proceed to analyze the material in terms of the transfer process; that is, the transfer of meaning from one set of symbols to the other. We do so by lexical and grammatical analysis. We have deliberately chosen what might be called a corpus in the 'inert' form

of language, because in the scientific treatises we are presented with a corpus of language relieved of many of the more elusive, and therefore more difficult to render in translation, forms of expression.

While I would go along with Professor Joos in saying that the translation of, say, *War and Peace* by mechanical process, it may well be that our grandchildren alone will see that sort of thing. In relatively accessible language, in the descriptive language of the sciences, we may be able to do something mechanically which will be acceptable within a reasonable number of years. In two-three years—we might have something adequate. We will have to retain human intervention. That is to say, the output text will have to be reviewed. I don't mean retranslated, rather edited. We cannot expect the machine, at least in the foreseeable future, to come up with a text that will stand completely on its own feet. Our Russian opposite numbers are saying that they are already doing translation on fairly broad scope today, though they give little precise information on their technique.

GARVIN: I'd like to make just an extremely brief remark on this matter of literary vs. non-literary translation. The linguists in Czechoslovakia who have dealt with literature and translation in this respect have claimed—and I think they are correct—that in literary translation you really deal with more than translation—with a sort of “poetic re-creation”, and obviously this cannot be expected to be done by a logical machine since it involves, in addition to logic, some of the other faculties of the mind. As far as technical translation goes, I think I would say that, as Mr. Dostert said, no more editorial work than is ordinarily required for any human translator would be required for the machine. Most translation agencies edit what they get from their translators and I don't think that if we get a product two or three years from now it will be worse than what the human translator presents from the standpoint of requirements of editing.

JOOS: Concerning this Czech literary theory, I would like to use for their approach the old term “transcendentalism” and I expect such transcendentalism from the Czech school of literary critics. I am, as a more-or-less ordinary American linguist, rather on the “positivist” side against the transcendentalists. I hold that there is nothing ultimately transcendental in any literary document and I have done literary criticism work myself and also glossemses of literary texts and annotating of them. My approach to this I think is adequately summed up in my long concluding paragraph on denotation and connotation which I reduced to a more elaborate programming of the kind that is

being worked on at present. It is particularly interesting to me to see how well Mr. Garvin's paper and mine meshed, specifically with his example of the two translations of эфир as *ether* and *ester*. Where he finds that 75% of the cues are adjacent, 22% are local (by which I mean in the same sentence), only 3% can be called ambiguous in any way, and I think that the programming of the kind of machine-computer that they are working on now—a computer that learns by experience—has to include a pleasure principle, or a reward-and-punishment in the machine. If you include that, the 3% will shrink to the level of the carelessness of the original author. And then again, concerning the original author, when editing is spoken of here we generally mean post-editing, eliminating pre-editing, but I do not think that technical translation by machine on a large scale can afford to do without a certain modest kind of pre-editing which can be done by a monolingual editor—an editor who knows nothing about problems of translation, but operates quite mechanically. Such a person would have certain specific jobs to do. One, he must reduce the original text to sentences no more than twenty words long each. By hook or crook, they've got to be broken up in sentences each one of which is not more than twenty words long. Then this pre-editor is to *introduce* few stylistic simplifications. (That's not essential, but I think such an editor can be taught to introduce enough stylistic simplifications so that the machine can handle it more easily, and then this pre-editor will have to standardize abbreviations, expand some, expand all abbreviations that are not on the standard list. These three things can be done by a monolingual, and then the machine that I envisage for technical translation might be complete, I think, in a generation.

GARVIN: Would it be the job of the pre-editor to correct misprints?

JOOS: Oh, I assume that. This is a secretary's job anyhow. We expect the secretary to straighten out simple misspellings—simple misprints which can be called misspellings.

ALO RAUN (Indiana): In connection with connotation and denotation, would it not be possible to establish a linguistic *norm* like *se-memes* as opposed to *allophemes*?

JOOS: I think those terms can be defined usefully. I have not been tempted to use them myself. In my reading where these terms are found, most authors have used them for begging various questions, so I don't use them myself. But I might have a use for them. I think

I could define them in a useful way. More or less as I attack the question of defining the two words denotation and connotation.

ZARECHNAK (Georgetown University): Denotation and connotation are not both properly the subject of manipulation by machine. My experience has been that the machine does not have to handle connotation, but if I have a good dictionary with the meanings in Russian I can translate adequately from Russian into English and from English into Russian. A Russian mathematician would understand terms which I would not. A machine has to translate the structures of the source language into the target.

JOOS: I agree that the normal or average reader, let us say the reader of the *Reader's Digest*, does not get a very large fraction of what I call connotations, and if you want to make your translation for a normal reader, then you don't need to worry particularly about connotations. I introduced the term as leading up to literary translations, and literature is so far away from complete grasp by any single reader that of course the problem of interpreting literature by college professors is a profession, at times lucrative.

CARLTON HODGE (Foreign Service Institute): I would like to ask if, granted that it will take some time to get anything that would translate literature, but in the process what linguistic byproducts can we expect from the research?

JOOS: The history of linguistic science in the last generation has proceeded in a way you might say from the small to the large. Phonology was well under control at the beginning of the second world war. Then, morphology and syntax, and we are now moving evidently in the direction of a linguistic semantics; surely the linguists' semantics will profit greatly from intelligent MT research (in meaning) transfer problems.

PAPER: (Michigan): I would like to add just one comment on what Mr. Joos just said. I think that in the areas he calls "inside meaning" since we now have at our disposal these complex and rapid data-processing machines, we can now get back to detailed study of the statistical frequencies of form classes, of words, of morphs, a project which we could not even dream of undertaking before because we didn't have the equipment at our disposal.

JOOS: Perfectly right. But when I start thinking about it, it seems to me that there would have to occur simultaneously substantially the same text in two languages, that is a good translation from one to the other. Feed them both in together in order that the machine may be informed, then study context distributions to see what semantic differences these context distributions are correlated with. I don't see quite how the machine can do it in one language, but if I retire into silence for a while, I may be able to figure that out later.

PAPER: Of course, as a linguist, I would be perfectly happy to get the statistical frequency information about the occurrence of morphs, for a particular language, regardless of the eventual feasibility of MT or not.

JOOS: As a linguist, or a semanticist, or literary critic, I compare the text with the imaginary picture of its reference as I build it up. I read a literary text and I build up a picture of a situation. That serves me in place of that other language. The machine needs it in the other language, or in Mr. Newman's type of analysis. For comparison, I use instead the real world for comparison of the text frequencies. Now, my concerns here have been mostly in the field of literature. For example, in Middle High German I find that the words for "eye", and "heart" tend to occur in context with each other, not necessarily in the same sentence. But if you take the same paragraph they are in context with each other better than three-quarters of the time, and that is certainly worth knowing for semantic research.