

Session 4: METHODOLOGY

QUESTIONS AND DISCUSSION

HAYES: Are there questions from the floor?

WALLACE: Semantic problems that have been dealt with here peripherally, I think, may be due to dealing traditionally only with the written language and the written language is only a part of any tongue. I think it fair to say that there is no linguistic analysis of any modern tongue that is completely satisfactory, and you attempt to deal with two of them at a time usually. You have dealt with a very restricted area of text which is a bare shadow of the larger language. This perhaps is a methodological problem in the large -- that is the conflict between those who wish to deal with a much larger area and those who wish to deal with a very small area in a cyclic fashion. So by little bits and pieces you may assemble something out of these very biased samples. There is also the question of variable redundancy in content of different text in the written and the spoken tongue. These two things in the semantic area I think are real, and you have not dealt with them so far as I have heard.

EDMUNDSON: We purposely have not dealt with the spoken mode. This has been of interest to scientists for some time; as you know, it is of obvious interest to the U. N. We feel that we should crawl before we walk, and we feel that the written mode should receive attention first.

WALLACE: I do not believe that you can really satisfactorily deal with a linguistic problem in the small by dealing solely with the written tongue. I think that your truly horrendous problems are going to be enhanced by the fact that you confine yourself this way and that you are not sufficiently concerned with linguistics, per se, in the large.

JOSSELYN: As a linguist, I will make the statement that we do not confine our efforts solely to the examination of written language, it is just the other way around. We consider that written expression is

Session 4: METHODOLOGY

just a very small portion of language structure. It is simply that in machine translation we deal with those little black and white marks. That is all, that doesn't mean we are ignoring the problem.

OETTINGER: As a matter of fact, it has taken a lot of beatings over the heads of linguists to get them away from speech to realize that in machine translation all you have to start with is the printed page, and there is nothing else available and that is that.

LAMB: There is another point involved in this, too. Once you get to the morphemic level and you are dealing with grammar, syntax and semantics, it doesn't really matter whether the expression was in the form of speech or writing.

GARVIN: I would like to address myself to the initial three categories that Dr. Edmundson mentioned. First, I think it would be useful for us to differentiate between, on the one hand, a cyclic research procedure which I believe the panel has pointed out is an operationally trivial distinction, because it is only a matter of the size of the cycle, and on the other hand, a particular iterative-type computer program as over another computer program which is not necessarily iterative. By that I mean a program which involves carrying out pretty well the same operation several times on the same sentence versus a program which carries out one single set of operations per sentence. I think this would be a more suitable differentiation than cyclic versus non-cyclic. Next I would like to get into the other question of theory versus empiricism. I think there the problem is the extent of theory that is then subjected to empirical verification. There are some groups which felt that it is desirable to start out with a minimum of theory and derive a theory from a tabulation of empirical results, whereas other groups have felt that we should have a reasonably extensive theory (which in turn may have been derived from the empirical experience of straight analytic linguists, but from the standpoint of MT itself, is a theory) and then put this into a program and test it. I think in a sense this gives you small versus large cycles. Finally, there is the matter of the quality of the product. I wonder how sensible it is to worry about a mathematical measure of the

Session 4: METHODOLOGY

quality of the product since there are so many facets involved in this which at this point are not as yet sensibly measurable. You could for instance think of one measure; you could tabulate how many inversions of order are necessary in the sentences of a given text. That is to say, given a Russian article, how often does the object precede the verb and the subject follow the verb, and then gauge the quality of the syntax program by the number of times this inversion has actually been effected. This is one of the few instances in which a measure is possible. Some programs have a built-in measure in the sense that either a chain number is increased for those cases in which this has not been done, or a notice of some kind if printed out on the information tape if this could not be done because the program has not yet reached a state of completion where this is inherently possible. I think the problem of the quality of the product, aside from the degree to which specific problems have been formulated and tabulations can be made in terms of specific problems, is largely an impressionistic decision. One very obvious answer to whether or not one should go into production is, as was very properly pointed out, up to the customer. If he will take the product and if you feel that you have no further intellectual interest in the matter, then by all means let us go into production. If on the other hand, the customer is not satisfied with the product, or if you feel that your primary intellectual interest is not a first crude product but an ultimate finer product, then let's not go into production. Let somebody else do it and we will attempt to shoot for higher quality.

OETTINGER: I would like to take issue with you, Dr. Garvin, on the question of letting the customer decide. I regard this to a certain extent as an abdication of professional responsibility. Caveat emptor may be a valuable maxim in the marketplace, but to let the buyer decide by himself whether he wants a translation scheme that might turn out wrong translations is not sound professional practice.

REIFLER: I should like to say something on what Dr. Oettinger just said. In fact in actual life it isn't a proposition of either -- or. In our particular case we were satisfied with very little, and in the course of the process we convinced him that we might be able to do

Session 4: METHODOLOGY

more. The customer, of course, was only too happy that we could do more.

EDMUNDSON: My appeal for devising these notions of a mathematical measure is because I feel that whether we like the output is like buying a car. I would like to factor it into; do we like the color, do we like the horsepower, do we like four wheels. I think by specifying some of these factors as criteria we may be able to understand why people like a Volkswagen and others like a Cadillac. I also feel that the customer if given a Volkswagen may in a few years want a Cadillac.

SEE: In relation to Dr. Edmundson's comment about the Volkswagen and the Cadillac. It appears to me that in discussions of grey it is difficult to draw a line, but in some cases I think we can. I will state that at present I don't think we have a car of any kind let alone a Volkswagen.

SWANSON: I would like to make a brief comment on the caveat emptor remark in which I would take issue with Dr. Oettinger taking issue with Dr. Garvin. I think there are merits on both sides of the question. Perhaps it is not trivial to point out that a machine has a lot more integrity than does a human being. A human translator probably cannot be programmed to warn the customer when he is turning out a bad translation, and is more likely to do everything he can to conceal the fact that it is bad and to turn out a smooth product. Whether or not all machines are being programmed to translate in just the opposite way I don't know, but certainly they can be. Certainly we could deliver to the customer a product with a good explanation of what could go wrong with it and to program the machine to be a good deal more fail-safe, than can be done with a human being.

DOSTERT: I want to correct what may have been an erroneous impression. If we do succeed within a year or a year and a half in producing usable text in one discipline, surely everyone should understand that we are not going to stop with that rather rough and partly inadequate material. After all, before the 1960 Fords came out some

Session 4: METHODOLOGY

of us remember that there were Model T's. True they could not travel as fast nor anywhere near as elegantly but they traveled, and in this I disagree with Mr. See that we cannot see a first clumsy pre-Model T in one discipline from one language to another within a reasonable period of time. I should like to point out that I am not the only one with claims of limited modest production, but that our friend and colleague Dr. King made much wider promises than anything that we ever did.

KING: Well, in answer to that, I have customers -- paying customers.

RHODES: I was amused watching everybody sitting here wondering whether we should cooperate or we shouldn't cooperate, just as if they had all the money in the world and all the will to do what they want to do. That isn't true. We are being paid and we are being paid by the government, by the taxpayers, and we owe them something. We owe them a great deal. It isn't up to us to make the decision. We have a certain duty to perform and I don't think we can sit here and say, "yes we will cooperate or no we won't cooperate".

SHERRY: We have been discussing production here lately and the question comes up whether products are usable or not. I believe we have two possibly usable products; we have IBM and Georgetown. These are the two that so far have come to light. In my mind at least I think I have a clear picture of how usable the IBM product is. On the other hand, I don't quite have the same feeling with Georgetown. I cannot tell at this point whether or not the output that will be produced, will involve the use of posteditors. I believe from Dr. Zarechnak's statement that since he expects to do more research, there will be the need of posteditors since the output will not be quite good enough. Now, if this is true how much need will there be? Will it be equivalent to just getting a translator perhaps? If you need posteditors what kind of training is involved? I think answers to these questions would clear up some of the comments that have been going back and forth where nothing has been answered.

Session 4: METHODOLOGY

LAMB: Dr. Swanson made an excellent remark just a minute ago. I think maybe it didn't get across quite well enough and I would like to amplify it a little bit. If there is an interest in getting early production, and I don't see anything wrong with such an interest, then the attitude that should be adopted by those who want early production, should be to try to make a distinction between what is known and what isn't known. Then in that area where it isn't known, instead of programming the machine to make a guess, have the machine offer all the alternatives which are possible.

G. BROWN: There are at least 10 languages of major importance from the world today; English is one. Therefore there are 9 source languages to be worked on. Well, if people's interest is primarily theoretical, why do they all choose Russian?

HAYS: Just one word on this choice of Russian. I think that a wise over-all view of the utilitarian needs of the country would suggest that whether your interests are pure or applied, everyone's choosing Russian is not the best thing for the country as a whole. There are some other things to be worked on.