QUESTIONS AND DISCUSSION

GARVIN: Since I am one of the discussants, let me get the ball rolling. First of all, I want to say that I agree with Dr. Swanson's initial statement. He said that there is hardly anything that can be said about multiple meaning that is not either obvious or wrong. Since none of us want to be wrong, all of us who are working in the area are belaboring the obvious, and I think if we do this intelligently and systematically, something will come of it. It is, for instance, quite obvious that dictionary problems are individual problems in the sense that each lexical unit has to be treated more or less in its own right, and that ordinarily grammatical problems are problems of classes: word classes or syntactical-occurrence pattern classes. Now, the basic question of multiple-meaning research is how to bring classes into the treatment of the dictionary. From the comments of the speakers, there emerged essentially two directions. One is a systemization in terms of categories of determiners, which I would like to call, for lack of a better term, the syntactic approach. On the other end, it would be, what I would like to call, the indexing approach; that is to say, the systemization of areas of application, as our colleague, Dr. Micklesen, has said. I think that ultimately there will have to be an approach from both ends against the middle and that those problems which, for instance, cannot be resolved by merely finding the determiners might be resolved by the indexing approach, and conversely. The example вид was a very nice one where indexing would not, in and of itself, yield a proper result, but a combination of the two That is to say, in biology there are some passages in which might. obviously means something like "manner" or "mode" and other вид passages in which equally obviously, it might mean "species". Then you have the lovely areas of uncertainty where, no matter what you do, you will get the wrong translation. I think that the requirement for an encyclopedic understanding might perhaps, at some time or another, be met by this intersection of indexing and syntax. That is all I have to say, except for a little comment on the recurrent statement of finding out how the human mind works and what to do about it. I think one additional question that ought to be asked is, what human mind?

Do we want to duplicate the way a very ingenious human mind works (such as the minds of some of our fellow investigators) or do we want to duplicate the way an uningenious, or even silly, mind works? It, too, is human, and presumably has the same number of cells in it.

HARPER: I would agree with Paul that this province must be attacked from both ends. I am surprised and disappointed that people have not attacked it from the simpler end. I think this should be done. On the other hand, I am not sure that the attack of the more general semantic or semiological problem may not be productive, particularly because of the contribution it makes to syntactic analysis. I would like to refer to a sentence or phrase which I think Chomsky introduced: "They are flying planes". Three types of ambiguities are here. But, in addition, to "flying planes" being a phraseological unit, the trouble is because you have a transitive and an intransitive verb. But, consider "they are disturbing questions", or "they are sickening questions". Do we The dictionary says "disturb" is both a tranhave an ambiguity here? sitive and an intransitive verb, or least, as a participle, it may be used as an attribute. The fact is that people do not disturb questions or When I mentioned yesterday about Mrs. Rhodes' sicken questions. prediction problem, I think this example, though it may be rather weak, illustrates the difficulty. Who can say what class of nouns objects may follow the word "disturb" when it appears in this particular tense or aspect form in English or the word "sickening"; perhaps the word "sickening" is followed by an animate noun or by a substitute? On the other hand, "sickening results" or "disturbing results" might be either. That is, something may disturb the results of your experiment or disturb something about it. You may have a different class of noun objects It is this kind of thing, rather than the easier type, that may get here. us somewhere, maybe more slowly, but certainly it will get us something more general--something that will provide a generalization leading to immediate usefulness in the analysis of what we may more properly call syntactic problems. I might add here, there is something like the phrase "increasing velocities" which has the same kind of ambiguity in theory: "we are continually increasing velocities of particles" This is a syntactic probversus "the increasing velocity is observed". lem. It does not happen to exist as often in Russian. Since "increase" is both transitive and intransitive, it may be that "velocity"

is not the kind of word that will solve the problem, but there certainly are cases where the noun object would do it. What are these classes of noun objects?

MANLY: I have a very brief comment on this last point. Perhaps this is a third type of approach to semantics other than the two mentioned, or maybe you would call it a syntactical one. When you say "they are flying planes" you start to refer to the previous sentence and the connections between sentences, and this would be generally very easily resolved. I don't think you have included that in the previous two classes, at least not explicitly.

DOSTERT: I hope that, although generally formulated, the specific applicability of some of my comments will be readily sensed by the audience. I should perhaps apologize for having introduced a false metaphor into the discussion of this group--that blank wall business. Of course we know that metaphors are very dangerous things. What I meant to suggest, when I used the term "blank wall", is that in that area which we so clearly define as semantics or semiology, there is a certain lack of clear discernible data. I also meant, by way of example, to suggest that when, in the course of a session like this morning's, we refer to the same concept as multiple meaning, polysemia and multivalence, one could question whether or not our thinking is quite as sharp as it might be. It is conceivable that the three terms are not strictly synonymous. The area of semantics has not been the subject of very thorough study on the part of the linguistic scientist. It would be a mistake, however, to think of it as being amorphous. Semantics, like any other area of language behavior, has structure, and Martin Joos has very clearly manifested the structural character of semantics by saying that, rather than being amorphous, it is crystalline. What I was trying to suggest by the blank wall metaphor was that we had not yet pursued our study to the point where the structure or crystalline character of semantics is readily known to workers in the field of translation or of linguistics generally. Let me make it quite clear that I don't believe that total, absolute, and perfect translation is possible, either by human The best thing that you can hope for is approxibeing or by machine. Therefore, I should like to correct any mation in varying degrees. impression that may have been given erroneously that anyone on the stall of Georgetown University now, or at any time, has claimed that a perfect

translation had been produced in any field, in any measure, on any machine; nor do we expect ever to attain that goal. It seems to me that it would be helpful in dealing with the broad problem of ambiguity in language to make a distinction between the problem of what you might call internal ambiguity and that of transfer ambiguity. I think it will lead to a considerable measure of wasted motions if we overfocus on internal ambiguity. For example, the hackneyed "Flying planes can be dangerous". There is no ambiguity when it is formulated in French. If a Frenchman wishes to say that planes in flight can be dangerous, he will say "Les avions en vol peuvent être dangereux". If he wishes to say that to pilot planes can be dangerous, he will say "Il est dangereux de piloter des aéroplanes". Therefore, to conjure up these imagined difficulties and to spend endless time trying to resolve them is like trying to see how many angels you can get to dance on the point I am not saying that it is not intellectually stimulating, but of a needle. I question the relevancy to problems of machine translation and to problems of transfer ambiguity. I think the tools that we have for the reduction of transfer ambiguity can perhaps be classified as follows: we have the tools of microglossarization or idioglossarization, or as Dr. Hays suggested, topical glossarization. It should be made quite clear that this is, as yet, only in a hypothetical state. We had in mind at Georgetown to try to push the work beyond its present status by trying to establish a Russian microglossary for the field of organic chemistry. In so doing we shall have the guidance of a specialist in the field of chemistry who is also solidly oriented in linguistics and in Russian. We will then try to keypunch a certain length of corpus, say 500, 000 words in another discipline, and then match the new word list against the organic Perhaps we will get an idea of the magnitude of what we have word list. been referring to as idio- or topical or microglossary. I think this will be a rather interesting set of data to present to the community interested Another idea that we have as a possible approach to the reducin MT. tion of transfer ambiguity is that of idiomatization, and it may well be that for the calculatable future some of the non-resolvable difficulties that we could encounter in the semantic area and, indeed, structural area could, as a provisional solution, be dumped into the basket of idiomatization, pending more refined and sophisticated solutions. The third tool that we have is that of structural analysis, and we know that even adequate as the tool may be at the morphological, syntagmatic, and syntactic level, there are still a number of problems which at the present

stage are left unresolved. I was very glad to see resurrected, in a sense of the word that I used in 1953, the word "predeterminer" or "postdeterminer" in arriving at the determination of lexical meaning. Incidentally, the very concept of pre- and postdeterminers indicated no fixation as to going from left to right or right to left. The "pre" is referred to in sense of the item under examination, the preceding word being the predeterminer and going from right to left. Now, I will conclude these remarks by indicating what we have in mind to do in the next year or so. We are going to work in the field of microglossarization for organic chemistry and a second discipline, and in the former we will have the assistance of Dr. Summers. We are going to work in the field of syntactic ambiguity with the assistance of George Trager, and in the field of semantic ambiguity with the assistance of Martin Joos. I feel that the participation of those three eminently qualified people together with a staff which is reasonably well trained will enable us, when we meet again, to submit to you in perhaps more explicit form than we have been able to do heretofore. I could obviously give one or two examples of approaches to the resolution of semantic ambiguity or semistructural or semisemantic ambiguity by using a sentence in which the preposition "on" was contained. You can say in French, "Put your book on the table, and your sentence on the board". But the first "on" comes out as sur and the second one as au. If you look at the total context it would appear at first glimpse that there is no clue to the choice of sur versus the prepositional article au. Then, if you examine the word that follows sur, you find, in due course, that it is a noun of horizontal surface. Then you find that the noun which follows <u>au</u>, as in the case of tableau, is also a noun of surface, but it is a noun of vertical surface, and therein lies the clue to your choice of <u>sur</u> versus <u>au</u>. In other words, in French, with nouns of horizontal surface the preposition used is <u>sur</u>, and the vertical surface is <u>au</u>. It is not ours to reason why--it just is. The thing to do, it seems to me, is to try to find the clue that will enable you to decide how to translate "on" when you are faced with that particular type of complexity. Another example to be given is the behavior of the particle en in French. As you know, it has In about 80% of its use the clues can be dia multiplicity of meanings. vided structurally--in the remaining 20% they cannot, and that balance is rather important because that particle is a very frequently used item

in the French language. From these two brief examples we come to the establishment of a sort of intermediate class between what you might call strictly structural meaning determination and strictly semantic meaning determination, and we get at a sort of semantic-structural in-between level where both clues of a semantic and of a structural character would help resolve the problem of ambiguity.

APPLEGATE: In the talks which we have heard, there seemed to be the point made that translation is simply a one-way decoding process and we should focus our attention on improving simply the output once we get an output. I think some of the comments made by the speakers indicate that others perhaps feel the same way I do, that it is only when language is considered as a two-way process that the possibility of using syntactic analysis to resolve problems of semantic ambiguity is possible.

OETTINGER: I would like to address a question to Professor Dostert. He said that the translation would not be perfect but an approximation in varying degrees, and in outlining the work for the next year he indicated some very interesting plans, but made no further mention of the plans for production. Since the approximation in varying degrees leaves a great deal of room for variation, I would like to have his opinion regarding, first the quality of the raw output that he estimates to obtain, its reliability, the extent of postediting necessary on this output, and the number and caliber of training and other qualifications of the posteditors, if any are required to turn this output into reliable material, if it is not so already when it comes out of the machine.

DOSTERT: An ad hoc norm for acceptability which we formulated last spring is the following: If your machine brings out the description of an experiment in the field of organic chemistry in such a manner that a chemist reading the output can effectively duplicate the experiment therein described, this would be a practical criterion of acceptability. It would seem to indicate that the meaning content of your source text has been, at least in a practical sense, adequately transmitted since the experiment in question can be duplicated effectively. The amount of editing or revising that will be required will depend upon whether we are going to be concerned with elegance and stylistic perfection or whether we will put up with a certain measure of awkwardness, but retain accurate meaning content. We have in mind--as to the revisers-to take experienced translators who are, of necessity, bilingual, and

who have had experience in the field of revising human translation. We plan to conduct a seminar this summer for eight weeks with them, and take some of the material that has been printed out in the field of organic chemistry and try to develop a technique to show them how we would like for them to check any inaccuracies against the source item and to revise the text to make it more readable to the user. In that connection, I was most interested in what Dr. Swanson was saying about the technique which Ramo-Wooldridge is developing for training of output revisers, and I should be most grateful for any help he could give us with respect to the training of revisers for better performance on machine output.

OETTINGER: The question of revising is one that, to my mind, requires very careful definition. We have conducted a number of experiments on the revision of the word-for-word dictionary output and the conclusions are rather sketchy. At one extreme we have found some people knowing no Russian whatsoever, but who are fairly expert in the subject matter, able to make some sense out of what was going on, particularly since they were warned by the fact that there were half a dozen alternatives for each word and that they could not take anything on faith. Some of them turned out fairly decent work. At the other extreme, we found that very competent bilingual people far from being helped by this material, were actually hindered, and reported to me that they wished they could throw this thing away because they could do the job faster by working with the raw text. I think one might fall into the danger of using competent bilingual people by supplying them, not with a crutch but with a leash that reduces their efficiency. In view of this very broad spectrum, I am not prepared to say where and what the right thing to do I wonder if your experience has been along this line--where in the is. spectrum would you consider things to work out?

DOSTERT: I agree with you, Dr. Oettinger, that if your machine brings out a nondescript dictionary output, your reviser will probably tell you that he would rather tackle the thing initially than to try to see which of six possibilities will fit the contiguous environment. I assume that the machine output will be of such quality that the bilingual professional reviser instead of producing 5, 000 words in one day, will produce 10, 000. If we can double the productivity in terms of finalized text of a given man by the use of the machine during a transition period, I think this is a

significant attainment. I have had enough experience in the field of human translators to say to you that it is not only in respect to the machine output that a reviser sometimes will say, "Let me do it from As a matter of fact, I remember once having a reviser on scratch". my staff in Nuremberg to whom I gave a piece of paper which I attributed to a person who had taken a test to join our staff. He went over it and he said the person was not good enough. It happened to be a translation which he had done about six weeks before. There are a good many translators who are afraid that machine translation is going to displace them from their job, and therefore it is difficult to avoid a bias among some of them who will not see that the purpose of machine translation is to liberate them from the drudgery that is involved in translation and to focus human talent on human problems in translation. Many of them, in their sense of insecurity, and in their desire to damn the machine--they think of the machine as their enemy rather than as their liberator--will tell you that they would rather start from zero. The thing to do, it seems to me, is to approach them calmly and objectively and to wait for experience to tell you what to do rather than to prejudge the situation. If we find that among 40 revisers there are 2 or 3 of them who make a shambles of the thing and keep saying that nothing can be done with it, then we will send them back to their desks. If there are others who are more adaptable and receptive and whose performance is more adequate, then we will try to refine their training for better performance. It is not fair to ask for a firm and categorical statement as to just precisely what you will tell your reviser to do and how to do it. As you know, I happen to be something of an empiricist and am willing to be guided by my experience rather than an a priori position, and this is my view with respect to the problem of what we are going to do with the problem of revising machine output. I know that there are some among us who feel that we should not try to disseminate the output of the machine until it can do without revision. That day will undoubtedly come, but I think that an interim period during which we shall harness human talent for the improvement of the machine output is not without meaning to our total objective.

RHODES: I am reminded of the story of the very wealthy and glorious lady who went to Fritz Kreisler and asked if she were to hire him to play for her guests how much he would charge. He said he would charge

\$5,000. "Oh", she said, "\$5,000 is a lot of money, but I must warn you, sir, that you cannot eat with my guests." "Well", he said, "if that is the case, I will charge you only \$1,000." This is what happened to me when a very important chief of a translation section came to me with a big manuscript on mathematics and asked what I would charge to do the postediting. I said, "Where is the original?" "Oh no, not from the original, just from this thing." I said, "Five cents a word." "Five cents, why the translator only charges me one and one-half cents." "Oh, if I am to translate, then one-half a penny", I said. Does that show you what I mean? It takes 10 times as long to do postediting.

CANNON: We seem to agree that by restricting attempts to scientific literature we have made our problem easier with respect to polysemia. I would like to say that, in one respect, we have made it more difficult, because we are not in that situation with respect to scientific literature, where there is a possibility of repeating experiments and where there could be a correct interpretation and yet the results be on the wrong I would like to tell a story which has bearing side of a critical area. upon our session. I was a second-year graduate student and very much interested in analysis while working for my Ph. D degree. Dr. Aurel Wintner, who died a few years ago, patted me on the shoulder one day and said, "Cannon, you report to the mathematics department seminar on this article". Now, we had to report twice a year on technical articles and the objective, we all knew, was for the faculty to demonstrate to us that we had quite a bit more to learn. So we took this matter quite seri-I replied, "Dr. Wintner, I cannot do it; ously during the preparation. what language is it?" "It is a short article in Danish. You use dictionaries; it has a liberal sprinkling of equations; you know German and French--you will have no difficulty." I spent about a week in the library I had all the dictionaries and I had some capacity, preparing for this. which I think probably equals or transcends the capacity of man-machine combinations as we have developed them to this point, to select the proper target word from among multiple choices. I began to deliver my speech on this article, and about halfway through Professor Wintner began to redden a bit and to squirm, but he did not say anything. Immediately upon my completion, before I could sit down, up he jumped, saying, "Cannon, it is all wrong". Just as quickly, two other teachers arose and said, "It is correct". A very interesting discussion on my presentations ensued among the teachers. It turned out that my translation was

wrong, but the mathematics were correct, and I had missed the point completely. I feel that I knew the subject matter; as I said, I spent almost two years in analysis; I feel that I had the equivalent of idioglossaries available; and I feel that my presentation was smooth; and yet I had reached the wrong results. But, you see, if we hand translation to an expert in the field, from the standpoint of content, he is not always able to determine whether we gave him a faithful image of It is one thing to state a criterion, it is another to state the original. how it can be used and whether it is practicable. Depending upon his nature, he cannot separate the nature of the criterion from his nature, so he will either be more strongly founded in the field or in the language. No matter which, I feel if he has tended to produce a large quantity of output, or at least pass it by as acceptable, he will be forced to depend upon smoothness of rendition, and I am not sure that will bear the closest relationship to faithfulness of the transfer of content which we desire. With respect to the word "postediting", it has been used throughout the session and never defined. From the standpoint of the experimental approach, I do not think postediting, as I see it, is enough. It would probably be better to have 20, 000 words of text and a translation by a man who knows both languages and the field, and then to have a group comparing the machine output with that output. We might call that operation postediting. I think this might be more conducive to acceleration of the convergence of the method to something we would consider worthwhile.

SEE: I would like to amplify a few remarks along the lines of Dr. Cannon's comment about reading the output. Some statements were made by Professor Dostert about the fact that a chemist who knew Russian had evaluated material that was produced by a computer and found there was a meaning transfer. I would like to open a small avenue of discussion about criteria. One criterion is that in judging the output of a translation process we must consider the relation to the output to the original. Consider that we have a Russian chemist or someone who knows Russian and is a chemist, and we hand him an article in Russian on chemistry. Clearly he can read it. Suppose the chemist also knows English and we hand him a word-for-word translation of the I contend, although it is not clear that he can read it, Russian article. that some people in this position will be able to absorb a considerable Many have, in fact, done so, but how is the amount of the content.

I would like to offer one hypothesis as to how one might article read? read such an article. One can look at the sequence of English meaning that is produced, introspect the possible source in Russian of these meanings, recover the original Russian text, translate the text, and thereby be well informed as to the content of the article. This, in fact, is what I have done when I have studied machine translation outputs. I have come to a difficult passage in English, sometimes I have to rely on dictionaries because I am a dictionary translator of Russian only, but I can stumble along. I meditate on the source of these words, recognize a Russian idiom that perhaps had not come through the procedure, realize what the Russian was, and translate that. I offer that this is roughly equivalent to reading Russian. I propose that most output looks more like Russian than it does English, although the words are actually English. I would like to offer the hypothesis that the use of a bilingual to judge the output of a translation system is invalid as a test to decide whether or not a conversion has taken place from the original language to English because he understands both. I contend that it is impossible to tell what he understands. Therefore, I would invalidate myself in any case as a positive judge of the quality of translation. If somebody gave me an excellent mechanical translation output I would be unable to say that it is excellent.

HAYS: The discussion has ranged pretty far from semantic problems, but as long as it stands where it does I would like to express my point of view on this. First of all, the postediting process that we are using at RAND is not intended, by any means, as a source of publishable material, lthough we are beginning to hope that it is worth collecting publishable translations as a byproduct at only slightly greater expense than the production of the same translations from the original material. I don't believe that anyone has stated the only question that is of interest to a sponsor who wants to buy translation. He does not care how the posteditor works, and he does not care what the machine output looks like. All he cares about is the cost of getting a certain amount of material translated within his fixed time requirements and with the I believe that a case could be made for machine personnel available. translation as a useful step in a purely human translation process that requires less competence in the staff and less total cost. Presumably it would go like this: First, find the cheapest possible way of getting the text into the machine. Second, use a reasonably cheap processing

technique; most of these processing techniques that we have been discussing are cheap. Third, design some human reprocessing technique for the output which produces good translations and keeps the cost and the competence level required as low as possible. This has not been said, and this is really what sponsors care about; not about whether the product of the machine is good or bad, useful or not, by what criterion it is to be judged. It is not to be judged at all; what is to be judged is the whole man-machine system and its final product. I could suggest some ways of setting up human reprocessing techniques that might be successful and cheap. For example, find out at what level of training a professional man picks up adequate knowledge of the language of his field. Do not try to hire bilinguals to do the whole reprocessing. Hire monolinguals in the target language who have just exactly enough training in the subject matter to be able to handle the job and provide them with supervisors. Give them the machine product in parallel column format so that the source and the target material is all there at once, let the monolingual go through this, getting the final product, as nearly as possible and as cheaply as possible, maybe with a dictating machine. Let him call on the bilingual for help when he feels he needs it, let him call on a more advanced specialist in the field of the subject field when he requires that, and you may well save a good deal of money and skilled personnel.

LAMB: I would like to suggest a way in which we might be able to do something that would make it unnecessary for us to carry on discussions of this type in the future. This is an experiment which could be performed very easily, and I suggest that it should be performed. Get 20 experienced translators who translate from Russian into English and, in a completely random way, divide them into two groups of 10 each, then take 10,000 words of new Russian text in a particular field and assign 1, 000 words to each of the 10 people in each of the two groups. Have one group translate their 10, 000 words and let outside impartial Run the same 10, 000 words of text through, observers time them. let us say, the Georgetown automatic translation system, and take the other group of 10 and have them revise so that it also comes out to a good translation, and time them. See which takes the longer time. This would be something that would be very nice to know. I think the same experiment might be done also with regard to the Harvard automatic dictionary, and perhaps it also could be done with the output of

433

our automatic text analyzer after it gets into operation.

SWANSON: I doubt the productiveness of that sort of an experiment. I think a good translator can translate as fast as he can dictate if he is an expert in the language and in the subject matter, and I would like to see any posteditor go that fast. I think Dr. Hays brought up the more important aspect of it, which has to do with the degree of skill required in both the question of whether or not you can use the less skilled per-My other remark has to do with the description son as a posteditor. that Dr. Dostert gave earlier of submitting the machine translation product to a bilingual expert in the subject matter. I would like to carry Mr. See's remark one step further by claiming that even the expert in the subject matter does not necessarily constitute a good evaluator of a machine translation. It takes very few clues for someone intimately familiar with the subject to understand an article sufficiently to be able to duplicate an experiment. I would submit that if a pure transliteration with a few endings stripped off and accompanied by all of the equations and all of the text, were submitted to an expert with enough time to do the job of translating, then you would find that it would be evaluated as a machine translation that transfers a good deal of meaning. The question might be how much time it would take him to do that. I think that is a very relevant point.

REIFLER: Much has been said today about a human posteditor, but I did not hear anything about an automatic posteditor. Before I go into this I would like to say something about another problem that has been brought up here; namely, the use of bilinguals to judge the output. I do not think that everybody working in machine translation is thinking of creating such machines for bilinguals, and if we were only concerned with bilinguals we would, of course, not think of wasting our time. Now. the whole problem of translation quality is only a problem of how you look at it. It looks different from the point of view of sponsors and it looks different perhaps to the people in universities who have an academic interest in this. We could consider the output of the first translation systems, not as a final translation product after which we would do nothing more, but we could consider this not as a translation at all but as an input for another device which could by then--by considering a number of criteria to be established--try to perhaps improve in gradual stages. There could be another intermediate output, which again is taken as an input to another device, and so forth. I want to point

out that this is the approach we thought of from the very outset at the University of Washington. All members of the staff, present and past, have always thought of the output of the translation system we are working for as an intermediate output, not actually to be published. If, in the meantime, some sponsors consider this kind of output to be useful, all the better.

OSWALD: I have been delighted to hear so many remarks about output. The one thing no-one has said is that this nation is in dire and urgent need of mechanical translation. We need it in many fields; we need it in intelligence, we need it in science; and the sooner we arrive at output, the better we will be and the better the nation will be. I have the impression that the sooner machine translation groups arrive at useful output, the sooner we will arrive at a happy situation where even more ample government finances will be made available to carry on these and other linguistic projects.

435