CONFERENCE ON MECHANICAL TRANSLATION

The Faculty Club
Sloan Building
50 Memorial Drive
Cambridge, Massachusetts
June 17, 18, 19

Massachusetts Institute of Technology

M. I. T. CONFERENCE ON MECHANICAL TRANSLATION

June 17-20, 1952

A conference on Mechanical Translation was held at the
M. I. T. Faculty Club, Old Lever Building, 15 Memorial Drive,
Cambridge, Massachusetts. The conference began at 9:45 a.m.,
June 17, 1952. Dr. Yehoshua Bar-Hillel, Research Associate
of the Research Laboratory of Electronics for M. I. T. was
the chairman of the meeting.

MR. BAR-HILLEL: Gentlemen, on behalf of the
Committee for Conferences on Mechanical Translation, I wel-
come you here now officially. I hope that not only will this
conference teach us a few things for future research, but
also that you will enjoy the stay here. Due to the gener-
osity of the organization committee and Mr. Sayers, we have,
I think, ideal conditions for our meetings. If we are all
agreed, we can get started on the discussions.

I personally would like to know whether it is worthwhile
to start thinking at this stage on the Translator Machine of
the learning-type which will be able to learn from its own
experience, or whether we should have to wait until the great
problems of the learning-type machine will be somewhat more
advanced. I do not know very much about it but I expected
to learn something more about it here. May I ask one or two
of you to volunteer to tell us something about the machine.
I think Mr. Booth might be able to tell us something in

this line. If not, we will have to skip it, but it will be a pity. I think it is important to know whether we should let it go for now or not.

MR. WIESNER: I have been learning the feed-back machine, but I think Reynolds or Booth could give you an intelligent discussion on that subject. I am not capable of doing it.

MR. BAR-HILLEL: I understand that you have dealt with this subject.

MR. BOOTH: I am prepared to discuss it with any-body that is prepared to do so.

MR. BAR-HILLEL: Well, let us go into it in a lesser form to begin with.

MR. OSWALD: It would be a great advantage to do so on a smaller scale.

MR. BAR-HILLEL: How about you, Reynolds? Can you add anything to the subject?

MR. REYNOLDS: We have not as yet gone into much studying on the learning-machine except from the theoretical aspect. In general, they are not acceptable from our point of view. I seriously thought that they would be applicable to the problem of mechanical translation.

MR. WIESNER: Just consider the machine and its properties, and the advantages we see in them. If we are going to build one, it might be worthwhile.

MR. REYNOLDS: There is one thing that I would like to bring up. There seems to be a misconception as to what a learning-machine is. I am not an authority myself, but there are two particular devices that we know of that are classified as teaching-machines. In both cases, the machine is not learning anything, but operating upon the reflexes that are built into the machine. They are like the human reflexes and that is the end of the mechanical mind in the machine.

MR. WIESNER: Some people have said that the telephone switching system will select routes to a given point when one circuit is busy or not available. This system can remember which system is available or busy and therefore is changing the procedure of the learning-machine. Now, if you want to call this a learning-machine...

MR. BAR HILLEL: I had in mind something else. We shall see letters that the existing machine will translate. We can also translate simple sentences. That is what I had in mind.

MR. WIESNER: Even in that sense, the telephone system operates in that regard. The principle is there.

MR. REYNOLDS: This is getting to be a telephone-learning program, rather than a learning-machine conference.

MR. WIESNER: I think that any machine that would modify the goals we want to reach are the things that I

would characterize. However, let's let it go.

MR. BAR-HILLEL: It seems that it will be worth-while to discuss the problem, but we will have to do it at a later time. We all know that there are huge number of potential problems that exist. Many of us have private problems but it will be good to go into all of them at least in theory. Even if one should be superior in some respects under existing conditions, we should remember that these conditions might well change and cause a reversal in the estimates of which plans are superior to translate a message. A plan which is considered superior to another may not be at all applicable except under bad conditions. With that attitude in mind, our fights over our preferred pets might have to be put into better surroundings.

I think that--in spite of the utter informality of our disucssions, where the only restriction is on our time limit for today--it might be well to have a leader for each one of the discussions. I think it might be well.

MR. REIFLER: May I make the suggestion that after the papers have been read that we attempt to limit all commentary to a maximum of five minutes. It will give us a chance to go all around the table for remarks.

MR. BAR-HILLEL: No, I don't think we will have to do that at this session. If there are no other questions or requests, I think we should get started on our papers.

At this point, Mr. Erwin Reifler read his prepared statement of facts pertaining to mechanical translation. The following remarks are in relation to Mr. Reifler's statements.

MR. REYNOLDS: Did you say that it is necessary to pre-determine a message or do you mean that it is necessary to just get the meaning.

MR. REIFLER: I said it was the ideal possibility whereby the pre-editor would know whether or not it is a question that would reach both the human and the machine.

MR. WIESNER: I don't see what you have accomplished, or just what your idea is.

MR. REIFLER: The moment the pre-editor indicates all the logical meaning from the source, he makes a semantic solution to the tests. There is no problem there.

MR. DOSTERT: He chooses the pertinent meaning, but he does not give them all. He gives them in some cases.

MR. REIFLER: He removes all impurities in the matter of the code language. Anyway, he does not need to know the target language. He can have a dictionary and enter in his native language telling him that in this case there are some problems as far as the target language is concerned.

MR. WIESNER: It is a question of who your customer is; whether it is the creator or the consumer. It would be very easy to make a mechanical dictionary in which a man

could simply call up any word and be given in the language which he is working the possible semantic ambiguities in the language, and also the symbols to indicate which one he has selected. Now there are stilled levels of residual ambiguities.

MR. BOOTH: It depends on the size of the dictionary to correlate to the possible ambiguities.

MR. WIESNER: It presents you with the problem of who is doing the pre-editing, and the proof-editor who knows how to work with this thing. I think it will beat looking it up in the dictionary. I think a man could easily call up the particular word in question. Now, the question I would bring up is in coming back from the target language and more-or-less indicate the choices; or must you have more of the context in the target language before you can dissolve the ambiguities? Can a man having the meaning on the select list accomplish this?

MR. REIFLER: Yes, you don't need to know the context at all. A German word could mean "this" or "that", but it would all come back the same. Now, what do you have in mind? You know, Mr. Bar-Hillel gave a report on some of the peculiarities of the target language. In this case, the German editor could very well say that I mean "this" or "that" or he would know what is meant.

MR. BAR-HILLEL: It is interesting that you need

not get the complete formal definition of the word. All you need to know is whether it is animal, water, or matter. These three solutions would tell him whether he chooses one category or another.

MR. WIESNER: We could take these words, and use the words with four numbers in the dictionary. Now, these dictionaries would be extremely relative to the computing factors and very simple to make. In fact, I wonder why we have not done it before.

MR. BAR-HILLEL: It is done to a certain degree in some dictionaries.

MR. KIEFLER: But, the fact is that we have a dictionary for the human dictionary.

MR. WIESNER: Three years ago I tried to convince our people that we should build such a dictionary. I would think it would be eased greatly in the proof-translator if you didn't have to thumb through the pages.

MR. DOSTERT: Take the medical man, or the army man, or any technical man. If they use a word, they all give the very meaning of the word in different functional terms. They could all be different.

MR. WIESNER: We are proposing now that you can mechanize this in such a way as to have a man spell out his word, and we could easily quote a projection machine so that the word would appear.

MR. BAR HILLEL: These mechanics are completely independent of the final translation. It would be done so that it need not interfere with the actions of the machine. This is a much simpler form, but the effect can be done so that the pre-editor can do it the week before.

MR. BULL: The office could say that I have written to the Mechanical Translation Center, and now we have the machine in the next room so you can type this in. If you don't think this machine would work...

MR. WIESNER: This machine would have usefulness beyond the problems we are talking about, but I think the pages would have to be prepared for mechanical translation. However, it would not come from a different dictionary.

MR. YNGVE: Would it be specifically prepared for the two languages involved so that one couldn't use the same dictionary to translate into another language?

MR. KIEFLER: I think you could get the solution if you have the series of subscripts on your word, so that if you are translating the word "faucet", then the English transcript would be number one, and the Chinese would be number two.

MR. WIESNER: You could make a dictionary of that type. The ambiguities would be very different. It is a point of having them all on one page.

MR. BAR-HILLEL: You will have situations where the

word could be translated into German in the same word. For
example, if you take the work "bank" -- it could be a place
to put your money, or a bank on the river. In German, you
will find that "river" is called by the same nar , but I
understand that in French, all of these would be in two
different terms. In other languages, it would be in diff-
erent terms.

MR. REIFLER: He is thinking in terms of general
mechanical translation. He is thinking in terms of building
a mechanical translator into which you feed one foreign lang-
uage test which is prepared by a pre-editor and that this
foreign language is expected to be fed into the first unit
and then into the next unit, such as Russian or French. It
might be advisable to have a mechanized dictionary to have
all the selections available. But, if you don't think in
these terms, you must be thinking only as far as English is
concerned. Here we have shared semantic translations in
German, so it is not necessary to make any distinction in
the dictionary. They have the English page, which in any
case would have only one equivalent in the other languages.
For instance, in the case of the English "s" and in the case
of the Chinese, you would have an "s" for an animal; but the
"s" for a human being is something else.

MR. BOOTH: Your dictionary of source language
could, in fact, be used by an unskilled person. Supposing

you take an automatic translation of words that have been said. The radio could be a wonderful and useful thing.

MR. WIESNER: I think this is probably something that you would have to do. You may be able to get mono-lingual person, but...

MR. BOOTH: On the other hand, you couldn't have a central bureau where a person of low standards would be able to fit into the picture.

MR. BAR-HILLEL: The first pre-editor would give the indents or put up the suffixes as it would be used. The remainder of the general ambiguities would be used in the first semantics and allow the pre-editor to use the "s" as much as possible.

MR. WIESNER: How does the pre-editor know which words will have ambiguities and which words will not.

MR. BULL: He will not have to know. The machine will know.

MR. WIESNER: But you don't want the pre-editor to examine every word in the test for ambiguities. I would think that you prepare the message for mechanical translation, it goes into the machine which does all the searching, and the word which could be confused in the target language would then come back to the pre-editor. I would guess from such a process, you would reduce the work done by the pre-editor.

MR. DOSTERT: This would amount to post-editing.

It seems to me that on this whole business of editing, that
one thing would appear to be clear. The suggestion that
Mr. Reifler made a moment ago that we have assigned the task
of pre-editing to the consumer and ask him to give us a
transcript that is cleaned up is a good one.

MR. REIFLER: I am only on the initial stage, and
I am describing all the possibilities.

MR. DOSTERT: With the thought that we shall say
to them--or it may be stated to them--that this is one possi-
bility to be considered, it seems to me that you and I do
share Mr. Booth's view on that. You are calling for a degree
of skill from the user which he won't have, and it seems to
me that the solution would be to have a board of expert pre-
editors attached to your mechanical translation centers. I
don't think you will find the skill required. Moreover, if
you pass the burden on to the user, it may come to the point
where it would take longer to pre-edit the test than it
would take to do it by hand translation.

MR. BULL: After you guys get this machine set up,
I can see the Russians sitting back and saying, "Brother,
are we going to foul them up when they try to translate all
of this from us." (Laughter)

MR. REIFLER: We have a possibility where this
brain-child might not be so bad. Namely, if it is the author
or his secretary who wants something translated. They are

the experts and they will be the ones to be told that we want "this" or "that" and they will be in a position to say "this" or "that" or something else. But anyway, I will have it taken care of. This has to be considered.

...Mr. Reifler then continued reading his statements on mechanical translation...

MR. REYNOLDS: I think this might be a convenient point to bring up some points on the record. Essentially, it seems to me that you gentlemen are discussing a language rather than a sub-language. Now I know this is a horrible thing for me to say, but here are part of the problems that are involved in part of these problems. We can have a chemical language that is common to the chemical language; the same is true of the electricians who have a language that is common to the power engineers. Then, we also have the electronic engineers. I would like to point out that in other countries at large, we do have highly specialized dictionaries that pertain to such a field. As such, it is quite possible to propose a machine to build a machine whereby the editing is done within a dictionary. Thereafter, a great deal of this pre-editing that is demanded from Mr. Reifler's paper can be done by merely selecting the proper dictionary. We have to, at the present time, draw the point between the common language or the mental language which we are discussing and the language and its characteristics. On this basis,

the pre-editing can be mechanized to a far greater extent than has been brought out so far. I think it is needed at this point.

MR. OSWALD: I am delighted, because we have come to my private solution of this thing. I find that half the planners share it with me. I call it the micro-glossary. As far as I can see, it is the gimmick to crack the nut.

MR. BULL: May I throw in one thing? There is no such thing as the language.

MR. REIFLER: You are choosing what I call a mechanical translation period. We shall not have to build any of these language machines for a certain language, but for a "certain" type of German; for a "certain" type of person of a "certain" type; and also for a "certain" speech narrow, but not for other things.

MR. BAR-HILLEL: I have some misgivings on this, but I believe we can postpone them until a later time.

...Mr. Reifler continued reading his statements on mechanical translation...

MR. BULL: I am just a little confused. It is not the same form as future perfect.

MR. REIFLER: Yes, future perfect, but as far as a perfect sense is concerned, it is perfect.

MR. BULL: It is perfect, but not as far as we are concerned.

...Mr. Reifler continued reading his statements on mechanical translation...

MR. BOOTH: That does not happen in England. As far as the office secretary is concerned, she is the one who is supposed to do this. As far as the author-secretary is concerned, she is the person to do this process and she won't be able to do this. The English secretary would not have the ability to do this.

MR. REIFLER: Well, fire her and get another secretary. We shall lift the educational level by this action.

MR. BOOTH: That involves a program much wider than mechanical translation.

MR. BULL: I am sure the vast majority of the American technicians cannot consistantly recognize the difference between nouns and verbs.

MR. REIFLER: I have heard many of these things, but I am not going to accept that opinion.

MR. BAR-HILLEL: I am certain that in some cases nobody would accept that opinion.

MR. HELMER: I must insist that we come back to the subject at hand.

...Mr. Reifler continued reading his statements on mechanical translations without further interruption. The final discussions on his statements continued...

MR. DOSTERT: It is now 11:00 a.m. We must have

some discussion on what has been read. We should allow just five minutes per person, because obviously, the time would go much too far. Now, as I see the problem, it might be reworded in somewhat different terms. The basic problem is the non-explicitness of explicitness; either of the language as it goes into the mechanical translator or the output. The question is whether we work to make it explicit, either before we put it into the mechanical translator or after it comes out. Now Dr. Reifler here has said some interesting points. First, the principle of the work; and second, what is to be done by the individual user of the mechanical translator; and third, where the problem of systemization of the pre-editing process throws some form of typographical transmission for mechanical translation. I don't know whether it is the desire of the men to formulate any kind of conclusion on the problem here, but I suggest that for the sake of gravity, we might focus our comments on the three points. Also, we should discuss where the editing is to be done, and where and how we invite brief comments.

MR. BOOTH: I am sorry to differ with you, but I am looking at it from the scientific work that I have been writing. If you are going to have a central office for doing this, I would like to have some things explained.

MR. DOSTERT: I think we are jumping to the three points, but I want to know first if we are all agreed on the

principles of pre-editing in regard to mechanical transla-
tion.

MR. REYNOLDS: The one point that I would like to
make is the great number or the great deal of discussion that
Dr. Reifler has engaged in with respect to problems that
revolve about the question of what the engineers will pro-
vide for the meachnical translation. How, of these, there
are two needs that are necessary to be considered. One
would be the key-board and then the method of scanning the
printed matter. It would also be dependent upon the time
that it would take to complete the operation. It might be
better if we produced a mechanical scanning device. The big
question is whether or not we have to reeducate everyone in
the terms of morphology. The pre-editing could be done by
the secretary by means of merely selecting mechanical preci-
sion instruments, or by symbols on a part of any standard
key-board. She could press the word to indicate a ncun,
verb, adjective, or whatever the word might be. The secre-
tary could select the words that will be needed. This is
independent of the original author's own words. I would
like to read in some of these remarks that show that what is
being done is going to depend, I beleive, on what the engin-
eer is provided with for mechanical translation and on the
transfer, and not on the linguistic characteristics.

MR. REIFLER: But how would you use it in terms of

the implied language, as well as the regular language?

MR. REYNOLDS: As a matter of fact, I have discussed the diacritical problems.

MR. BULL: The point I want to make is this-- and not because I have not had enough sleep lately--I am horribly depressed by the admission of defeat that we are all accepting. In the first place, nobody has asked any questions as to what the machine can actually do. I would like to determine what we have to do as pre-editors and what we have to do as post-editors because of our failure with the machine. We are discussing the logical resolutions of problems before we begin to discuss what the machine will actually do. We have an air of defeatism. There is too much of that type thinking. We said we have to post-edit and also that we have to pre-edit. I refuse to accept this attitude until we demonstrate that the machine can go at all.

MR. BAR-HILLEL: Let me say that I fully agree with him. My talk will refer to the subject. I don't say that it will be the best method. It may not be. But, I will attempt to prove that we can go along with it. I want to say that I believe it can be proved, not only now, but also within the next ten years. However, we will be unable to go along with thecomplete mechanical translation.

MR. BULL: I want to ask immediately, why you say we won't be able to. It is an attitude of defeatism.

MR. BAR-HILLEL: If you want proof, the proof is that I have investigated many methods of eliminating syntactical ambiguities and there is no method available at the moment except the method that I tried the other day. That method is not available now. I believe there are two principles which I believe should be combined. He proposes that a certain message have a method of eliminating these ambiguities. This is the right method which can be discussed for a longer time. His other methods which are available should be used. This is one way, and I believe there is no reason at this moment to discuss it at length at this time. I should object very strongly that this method will help us in the pre-editing stage. The resulting semantical ambiguities from this second proposal advocates the first proposal which we rejected. It won't work in the message. This is not done in capitalization of words or letters, but it is done through what amounts to the use of our dictionary to find out the problems in this case and in other cases. We have the dictionary of other types, but we tried to show yesterday that this would amount to a dictionary of editing. I still maintain that this message is utterly unfeasible, not only for today, but for the next ten years. It may even be 20 years before we will be able to use this in general. It might still be feasible, but not now.

MR. LOCKE: I think you are way out of line. Any

other dictionary does not go along those lines.

MR. BAR-HILLEL: How about the English word "bank"? Distinguish between whether you mean a building or whether you refer to a bank on the side of the river where you sit.

MR. LOCKE: What do you mean by the word?

MR. BAR-HILLEL: Well, you don't need many editors to tell what you mean with respect to every word in a given sentence. Every human being would know what that word would mean. It would have to mean one thing or the other. Every person would know immediately that in a certain sentence the word "bank" is a building. But he also knows that there are other meanings. According to that, the machine would have to have an enormous number of combinations--it would have hundreds of millions of combinations. You see, sometimes you decide on a meaning by one word and sometimes you decide by two or three words and not necessarily just one word. You could have symbols on each side of the word, but I do not think so. I will not say this is how we do it, or that this is the method whereby a human being can decide that this word would mean something. At the moment, the machine won't do it and I don't believe it will for the next ten or twenty years.

MR. REIFLER: I accept your criticism, but I want to consider the machine and how much it will cost. I want to point out that I am fully aware that this machine cannot

solve all the meanings. But, in this case, the topic I have

chosen has a different meaning and I agree that in the case

of the word "bank" I agree that it cannot be done. However,

I can give you a language of numbers and this system should

be considered. I suggest that these words cannot remove a

large number of the problems of the pre-editor; it can lessen

the work of the post-editor.

MR. DOSTERT: I am more optimistic because I have

noticed the results of these interventions. We may, in fact,

be chasing windmills. Let's wait before we invode difficul-

ties which we attribute to the lack of ability of the engin-

eers to build the machine.

MR. WIESNER: I was going to make an operational

suggestion and then reply to a question you had raised in

your statement. Once we get to the end of the program, we

will have a little more freedom. Do not force ourselves into

this schedule. You asked what is the objective of this meet-

ing. I would say that as a group, if we could end up with

certain conclusions, it would be useful. We don't have any

compulsion to do this. On the other hand, there are people

represented here from the Department of Defense, and they

have some very real problems. If this gentleman can be con-

vinced that an investment might yield some useful results,

we might see this thing start taking shape. I think that we,

as a group, ought to make a serious effort to decide just

what our efforts should be. Should it be one that is parti-
ally mechanized, or should we say that we cannot face the
problems of a post-editor's job. In other words, what is
the status of this of our considered judgments? If we can
make any such statements, let's hear them.

MR. OSWALD: I would like to make a statement that
is important at this time. It comes back to a statement
that Mr. Reynolds made. The differences that arise in the
attempts to solve problems of this kind, such as Mr. Bar-
Hillel's dictionary, get us back to the question of whether
we are, or whether we are not trying to translate poetry.
The language of scientific people is, of course, different
from anything discussed here today, with the exception of
the introduction--but who reads introductions. The mechan-
ical solution of the problem of translating scientific dis-
course is different from the problem of trying to solve a
mechanically different language which can be done in the
mechanical sense.

MR. WIESNER: On this business of saying that we
can't do it in five or ten years is out of line, because we
don't know what we can do. You can't place any particular
time on this job. Particularly because of the speed with
which things are done today. I think if you are thinking of
ten or even five years, you have to consider just what it is
that you are trying to do.

MR. REIFLER: This approach would enable us to remove some of the greatest mechanical problems.

MR. OSWALD: I don't want to go through the business of retyping letters. I want to run it right out of the machine. We don't want to worry about the mechanical problems. Let the engineers worry about them. They can do it all up for us.

MR. REYNOLDS: I spent a very interesting session with Mr. Bar-Hillel a couple of evenings ago, and I asked a very pertinent question that has not yet been asked by any of you gentlemen here today. That is, what if any investigation has been made for information in the quite technical theoretical sense. To my amazement, I discovered that there has not been any done. Now, I have a potential solution in terms of switching the applications to this problem. It is along the ideas that Mr. Reifler has proposed and which led to the language dictionary of Mr. Bar-Hillel. We have in this country several companies engaged in the transmission of information. This information is transmitted in phrases. Think of the way that Western Union does the job. I am now referring to the communications and so on. Quite a number of the larger companies use precisely the same method of transmitting messages. They use precisely the same equipment as the Western Union and other groups. They take a four letter group that will stand for a particular phrase

and they are sent out as a four letter phrase. After they reach their destination, they are put back to the original form and then they are given to the public. It would pay to investigate the results of the input of 25 phrases, or in our case, it would be characteristics that represent a particular phrase in English and corresponds to a word that is transmitted through the cable line or through the air. It is translated on the other end by a recipient in another language. This type of switching is quite feasible. This type of switching arrangement can be built in without going into the tremendous dictionaries. To this extent, Mr. Perry, in his talk was very much robbing the engineers in their approach to the problem. It can be conceived as a switching problem or as a reduction of the symbols and a reswitching of the problem into the foreign language. To this extent, we can translate the introductions and a small amount of the poetry. In addition, they can be sent down in the form of a single form. We can apply these to the scientific principles. The answer to your problem is an engineering problem. I am very optimistic on this problem.

MR. OSWALD: How about the boiling down of phrases?

MR. REYNOLDS: The boiling down of phrases is done by a technique of a switching arrangement in which we will have built in the inputs and the outputs. After that, all that has to be done is to switch them.

MR. OSWALD: You have to have an interventionalist before that. It would all have to depend on how this is going to be handled. Where does the linguist come into the picture? Who makes the code book?

MR. REYNOLDS: Well, many of you have been in a far away city around Christmas time and you want to send a message home. You walk into the Western Union and you look over some of the Christmas messages that you want to send to someone back home. Then you pick message "23" which is a nice 15 or 20 word message which they are going to give you for the price of ten words, but, they are going to transmit the 15 or 20 words to your family back home. You are correct. This is a problem for the linguist to solve in terms of setting up the phrase dictionary. So far as the engineers are concerned, the switching techniques are available. I think this point should be considered.

MR. OSWALD: That would take us back into the trillions of combinations.

MR. REYNOLDS: No sir, it does not. There are not trillions involved in radio terms and there are not trillions involved in what J. P. Morgan has to offer.

MR. OSWALD: It would be to translate it into a code number and shoot it out on the machine. Is that right?

MR. REYNOLDS: That is right.

MR. WIESNER: We are using a restricted phrase

dictionary. That is how we find what we regard as the most likely billion phrases. Suppose we claimed that we have that many and that we disregard the others. Suppose that we claim that in our process of translation, the machine is capeable of finding that phrase that is closest to the actual translation of the message.

MR. BULL: The thing that I am postulating is this for you to consider. Apparently there is a small amount of confusion here of the phrase that contains 25 letters. The type induced does not care what is going inside the machine. This corresponds to a certain number of states within the machine. This creates a number of states corresponding to a number of selections that can activate the output of a foreign phrase. This is a storage proposition.

MR. WIESNER: But, when you ask how many possible combinations of phrases can be contained in these 25 letters, you come to a fairly large number.

MR. REYNOLDS: Not really, it would only involve a few million. Now I have something else again. You are talking about the transmission of precise information. The precise semantic contents of what I am proposing involves the transmission of information. Now, from the engineering sense, we realize that any time you transmit information, in some time, the information is degraded.

MR. WIESNER: But suppose you do this technically

The thing that troubles me about the process is when I go to a code book, or I could go to the Western Union, even though Western Union has not used any of the words used in the message, I discover that it costs me over $4 for something that should have cost less than a dollar. Now I could translate what I wanted to say into one of those 25 messages that the Western Union sends and wind up with the same thing. But, whether or not I could make a machine that does this is the thing that I am worried about.

MR. REYNOLDS: The thing that bothers me here is getting complicated. Let us take a 25 letter phrase, but let's keep it as a simple phrase. There are a tremendous number of combinations in any one of the different target languages that involve a four letter combination out to a 65 letter combination. I want the minimum number of switching elements to give me this particular output.

MR. WIESNER: I am misunderstanding something here. Either the machine has the equivalent of every phrase that you can put in so that it recognizes the phrase, or it has to do what I do when I am confused.

MR. BAR-HILLEL: The machine can do exactly what this gentleman is doing with the Stenotype machine. No matter what we say, or how it is said, he has prepared in his brain a large amount of short codes. So far, I don't know how many he uses, but he does have a large number of

code messages which enables him to keep up his speed on his machine--the Stenotype machine--and he can do it very quick. But, I cannot see how you can pick out a number of certain common phrases which will be highly used, or how you will be able to build a machine which will be able to do this.

MR. OSWALD: That is a separate block problem.

MR. LOCKE: That is not a tremendous problem.

MR. DOSTERT: The army uses codes and they are purely mechanical terms. You can say anything you want and the amount of storage that is done on the coded card is very very small indeed. The only thing is that you would use it from the coded symbols in a different language than that which was originally used. Now, the code office in the embassy does not use the machine at all.

MR. WIESNER: The equivalent of what you are proposing to do in a mechanical translation is a code for the translation of a simple word-level. This would have some great advantages. This is a suggestion which I think would be useful. Let us know something about the machine; what is feasible; and what is impossible to set up as a limit of what we think is a feasible number of storage for this machine. Let us suppose that we can have access to a hundred million phrases. That is a lot but this might be very expensive or a very big machine. Now the question is how much of this do you want to do.

MR. REYNOLDS: I think you are over the amount.

MR. WIESNER: Now you say this is impossible. I am saying that if you gave me the money and allowed me to store the building, I could do it. I could build the machine if I had the money and the storage space.

MR. OSWALD: I am not worried about storing them. I am worried about how many years it would take. In the meanwhile, we have to get out a little old piece of paper and scribble down whatever it is we have to say on the paper. I am confident that it can be done faster.

MR. REIFLER: This would allow the post-editor to limit us and it would lessen the burden on the post-editor for some time to come.

MR. WIESNER: I would like to qualify my number and pull that down.

MR. BULL: Everyone is worried about the multi-million type of thinking. I have been doing some work here. I will try to say this in three minutes..It is something to reduce the factors a great deal and apparently, it has not got into the discussion. I noticed that Mr. Kaplan and Rand did not have it because it is still relatively unexplored. Let's take the phrase, "All those large banks." I would propose that you use the word "banks" from Oswald's vocabulary. In other words, call it by something common. That will eliminate sitting down upon a piece of hard wood

or picking daisies along the bank of the river. Now, if you will notice, we have here six motions in the English language and this is the only word in the English language which can be in this position. These are the only words that can be in this position. If we put a preposition in the sentence, we can tell the machine as soon as you hit the preposition that this is a certain phrase, so we can thereby defend this phrase mechanically. There are only 30 words in the English language that would take this position. There are 500 of the adjectives as used by reasonable men that are available so we have reduced the number whether you are talking about 6 words, 30 words, or 500 words. As a result, this multi-million sort of thing disappears completely.

MR. BAR-HILLEL: Yes, but you have a few hundred or thousands of rules to solve with these prepositions.

MR. BULL: Yes, but let's build your dictionary around these cases.

MR. BAR-HILLEL: I don't think you will save anything in time. But, if you add a few thousands of rules...

MR. BULL: This does not require rules.

MR. BAR-HILLEL: How would you tell that before the English word that is a noun? You will have to determine which of the tens of thousands of adjectives would come before it. This does not tell a thing, because before the word "bank" you could have many different things.

MR. DOSTERT: Do you mean to say that you want the word in the form of a permanent stage?

MR. BULL: Yes, that is what I want.

MR. DOSTERT: But you cannot get it.

MR. BAR-HILLEL: Multiply the words you have there and see how many phrases can have this meaning.

MR. REYNOLDS: You can substitute the total number of phrases.

MR. BULL: Yes, but what I am getting at here is that these words are restricted. They are reduced by this factor to a very small number of categories and you can build up a number of patterns.

MR. REYNOLDS: There is a point here. Mr. Bar-Hillel is quite concerned of the combinations you can get out of this set-up. However, since it is in the engineering end of the problem, we can get a determination from the phrase. The question is whether or not it is an advantage or a disadvantage from the engineering point of view.

MR. BAR-HILLEL: Before "bank" you use 500, but before all the other nouns that appear in commerce you use all the other existing languages.

MR. BULL: That is not true, at least not statistically true. We will establish it later on in the conference and this issue will be settled.

MR. YNGVE: I think the point is centered around

the dictionary and the problem of a system. You want to get at the word that you are after within a reasonable length of time. You can certainly store enough words if you have the time and want to wait for an hour, but the problem is in access. Now, if you can predict by this method that after the word, you have at the most, just five or six possibilities, it could be made to work.

MR. BULL: After the word "all" there are only two sets of possibilities. Either you are going to have the word "bank" or these other words.

MR. YNGVE: But you are now trading one part of · the machine for another part. You are trading the dictionary for something like a predictor.

MR. BAR-HILLEL: You will save in space. You are quite right on that, but you will have an enormously more complicated computor.

MR. DOSTERT: Does anyone care to an exchange in ideas?

MR. LOCKE: Most of the statements last night were to the necessity of treating machine translation whereby the author would not make special material for translation. It seems to me that any system that required that the material be submitted in a special form is "Ipso Facto." It is out of consideration here.

MR. DOSTERT: That is a very important one from

our point of consideration and I would like to throw in a few words on this subject at this time. Yesterday, I tried to broaden the field by mentioning the many agencies and the organizations needed, but Mr. Oswald keeps coming back to the scientific facts. I don't know if we want to continue to narrow the usefulness of the mechanical translations, or to narrow the tests.

MR. OSWALD: We are interested in all special languages when you come down to the last analysis.

MR. BAR HILLEL: I would say that in addition, we have three most important tests which require a high degree of accuracy.

MR. OSWALD: It seems to me that the diplomatic tests have all the emotional charges.

MR. DOSTERT: We have a broad enough scope.

MR. BAR-HILLEL: No, I think it is too broad.

MR. BULL: You are going back to the business of what your failures will be, and that is not good.

MR. OSWALD: Once we decide how to correct it, we will be all set. We can work with the kind of thing that we want to. Right now we are trying to drive the wedge through the wood with the thick end first. Believe me, it is a hard thing to do. It is especially hard on the hands. (Laughter)

MR. BULL: May I ask Mr. Oswald if I have not got a number here that is at least 450 too large.

MR. OSWALD: I have no comments to make on your
figure at this time.

MR. REYNOLDS: I think it is too large.

MR. DOSTERT: Now we are going to the other end
of the process.

MR. BAR-HILLEL: We have to, in principle, support
all possibilities for making the machine. On the other hand,
we have also to learn of something which seems--as far as we
can see at this moment--to be in the not too far future. We
have to give up all limits that would be resolved as far as
fiction is concerned. We also have to have some relative
information so we can find some middle way to the answers
to our problems. This is not too easy to do. We have to do
something which can be done in some future date which is not
too far off and which will not be too fantastic about the
requirements of the storing capacities of the machine. What
I intend to prove is that the mechanical translation between
things can be done without any pre-editing from anything,
but simply from reading by a mechanical device that is to be
given. That can be done. Whether we shall do it or not is
a question we shall all have to work on. It can be done, but
with the price we have to pay there will be no possibility
at all of coming out with the unique translation. There will
be a very simple possibility of it coming out with the multi-
plicity of translation which an intelligent and expert post-

editor will be able to reduce to uniqueness without a great deal of difficulties in general. It will require that a great many things be done that have not been done--even things that can't be done will have to be done--and they will have to be done by linguists. These things should be done by linguists anyway.

So, I shall assume that the machine knows how to transform the original and to do some of the following things. Namely, that the machine will be able to find out things with respect to any form of language; to find out what is what in the sentence with regard to the classification of the subject and so forth; and also to find out what the adverbs are and adjectives and what the mofifiers modify. The system will not be good in this repsect because even if I know it is edited, I still have to use a part of the rules to find out what the adjectives or adverbs modify. Some treatment to this situation will have to be done. I assume that it can be done and I shall try to show how the machine, if given certain instructions, will be able to do this. If the original language is composed of some form of the target language, then by statements and by diverse translation, we will arrive at the possible translation for which a post-editor will be able to find out the correct translation.

MR. OSWALD: How many translations could be done?

MR. BAR-HILLEL: Well, for a sentence of 10 or 20

words, there might be tens of millions or hundreds of millions of possible translations; but, as you know, others that will have seen this, if they are arranged in a reasonable order, they could easily read off the correct translation in most cases.

I will now deal with the cases of those exceptions in which the post-editor will not be able to read off the correct translation. I would like to deal with the questions that have been asked on "idioms". In all public talks, the experts want to know how the machine will be able to recognize all these differenct things. My answer is very difficult for me to give, because I know of so many different answers to give you. The only problem is to decide which answer will be most acceptable.

Now, consider the German word which means, "there is", or, "there are". With respect to this word, you will not find it in any English dictionary of any type that I now know about. You might find it in something which is prepared by others that will tell you that it is the third person singular of the word "giben". Then you can look up the word "giben" in the dictionary and then this whole thing will be translated as "s" to a third person singular in the present tense. Now, if you do this, or if a mechanical translator will do this, then a student, after his first five lectures will come out with something that will make

sense. However, it is something that will make bad sense, because he will come out with the wrong translation. It will be something that is considered good English, but he will not know what happened and I don't think that the ordinary English student who does not know German will be able to find out what happened. This is the situation which we call an "idiom" and I believe it has to be relativized to three things. First of all, it has to be relativized to a dictionary. I will show you how this characterization of this idiom would first appear. It has to be set up according to grammatical rules under certain languages; and it has to be know with respect to its use. You can also talk of the idiom of one language with respect to another language. I don't know any reasonable definition to tell a German that any such words in German is an idiom. There is no linguistic definition of this type, but, there are conventional definitions depending on theuses in the sentence. There are no rigid definitions which will show any German that this "s" is an idiom, except, if you have given certain statistical rules in which this would not fit in. However, in effect, this "s" is an idiom with respect to English, but not an idiom in respect to French. It is not an idiom with respect to Dutch. You would have to find the amount of the translation, and then you would find that you would have to find the exact correlation of this phrase in respect to Dutch

or German or other languages. Whether or not it is still an idiom might be discussed, but it is of no interest at all. We can, therefore, deal with idioms with respect to the implied meaning.

MR. LOCKE: Each idiom is a phrase, the meaning of which cannot be determined just by the accumulation of its partial expressions.

MR. BULL: That definition would lead to a lot of complications.

MR. BAR-HILLEL: I am trying to show you just what the definition leaves out, that is, with respect to the new dictionary and certain other languages. With respect to the phrase "es_gieft", it is shown only with reference to some standard explanation. I will show how it can be easily determined as far as this problem is concerned.

MR. REIFLER: But the moment that you apply a historical word from a certain language period, the idiom will not be known.

MR. BAR-HILLEL: You could look up the meaning in a certain dictionary. In this case you would look up the word in the dictionary and you would see the meanings listed in order. You would know by the meaning whether or not you had a correct translation or not. In this case, it would not. You would find that it would appear as an idiom.

MR. LOCKE: I would like to know whether or not

there is any such thing as an idiom for a native speaker of his own language.

MR. REIFLER: Yes, there are idioms in relation to his own language.

MR. BAR-HILLEL: If he is not educated correctly, there will be idioms.

MR. LOCKE: I believed that with a native speaker of his own tongue, there were no idioms.

MR. DOSTERT: I submit that a man that has never learned a foreign language would never know what an idiom is. You could take a certain person and ask him if he knows what an idiom is; and if he has not had the experience of learning a foreign language, he will not know what an idiom is.

MR. LOCKE: There is no point in knowing what the meaning of an idiom is. It is only when you start to define a foreign language when the meaning is helpful.

MR. OSWALD: I am looking at the overall sense of the word. There are lots of idioms. Let's bring the meaning down to a concise definition.

MR. LOCKE: An idiom is something that you can't translate to any other language and make sense.

MR. OSWALD: So, whether we are right in saying that the whole language is an idiom, is a fantastic thing. Let's take the meaning of the translation.

MR. LOCKE: Then, you may have some questions as

to who you identify an idiom when you start working with the mechanical translation.

MR. OSWALD: But, when you take them over one by one, don't they have similar parts?

MR. LOCKE: It is not the original language that gives you the problem; it is when you take them apart.

MR. DOSTERT: I think you have to go at the idiom as a complete and total factor.

MR. REYNOLDS: Actually, to make this statement, I think we are dealing in fundamental errors here. We have to go with the convictions of languages. We can use languages to describe the things we want. This language you are talking about is a language in itself. We are, at this point, capable of distinguishing idioms within the language. We are capable of distinguishing idioms regardless of other factors. I am afraid that I know only one language reasonably well, and as soon as I switch into the mental language, I make a very precise sentence. I think we must make the distinction here if Mr. Bar-Hillel wishes to speak on this subject of idioms.

MR. BAR-HILLEL: What I intend to show is something by which we can deal with the mechanical translation of idioms. I want to make the idiom disappear completely, so that we will have no problems that will arise for idioms at all. But, the price will be too high. Just add to the

possible translation of "es" all the possible translations
of "gieft", add them all into the English, and you are rid of
the idiom. It is one of the possible translations.

MR. LOCKE: You won't find it in any of the bilin-
gual dictionaries.

MR. BAR-HILLEL: The dictionaries are doing it
differently. You will see that the dictionaries are doing
it the way that I shall show you. If this were done, I
would strongly advise not to use this method. Now you want
to know why I say that. It is because the load on the post-
editor would be found to be too high. In addition to these
translations (depending on the words to be translated) you
will wind up with many other translations which will make
sense. Too many translations make sense. In some cases,
the post-editor might have to decide that he is to choose the
words which mean "he is", "they are", or "she is". Though
this a method which would eliminate all idioms, I would not
advise to use it. You can translate many words into English
by knowing what the best possible meaning could be if you
had some introduction to work on. Any linguist knows how
necessary the introductions are, but there again, we could
eliminate certain things. For instance, in the phrase "haigt
of menner" there would not appear an idiom. In the word
"es gieft" you could find it joined by a hyphen. There may
be other hyphenated words that would appear as one word in

the dictionary, so whether we indicate it by a hyphen or not, does not matter.

Now, you want to know how the machine would react to that? Well, the machine would be told that if a word comes with the "s", you would have to look it up in a special phrase dictionary, whereby the machine would start with the "s" and check back. The machine will be very happy to find a phrase, because, now it will be able to translate a whole bunch of phrases or words immediately. It could find one word and then go on and pick up the other words in the sentence or phrase. The trouble involved in this case would be very slight because even though the definition of "men" and "man" might be different, there would be certain characteristics in the machine to take care of this situation. The machine would provide the correct translation in most cases. In other cases, it might also contain the hyphenated word and also be made to show the hyphen. I do not think that this supposition is a good method here. You would run into some terrible trouble if you have some longer phrases. Most of the phrases might be of more than two words so you might lose the meaning of two word phrases. I am afraid that most of the phrases are more than two word phrases.

My problem is not in determining how many methods there are available. There are many of them. My problem is in determining which combinations should be used. I think we

could use all of them simultaneously, but it would require a great deal of work. Now, the first solutions would have some funny situations. For instance, whenever you have the phrase "es gieft" in German, you would translate it as meaning "she is". Unfortunately, there might be four to six pages of English translations. You could tell the post-editor that in all cases where this phrase appears, that you want it to mean "there is" or "there are" and then he would know the correct translation. The post-editor does not know how it comes out--he perhaps will not know German--he will be instructed to find an English sentence of some sort; or he will know that he should replace it with other words to make the meaning clear. In a slight number of cases, it will have to be translated as "it is" but only in a slight number of cases. You could have a phrase-translator with respect to results of the machine translation from some foreign language. You certainly could have such a thing done by the machine, or you could have it done by the human post-editor.

MR. BULL: Let me ask whether you have considered the case of a man that does not know the languages too well. You might have a man translate "es gieft" to mean "it gives" and you will turn around and say that is not what is meant. I have been told by somebody that this meaning must be replaced, but I do not think it is of too much importance. The message would be completed. It would take only a few

pages for each combination of the different phrases.

. MR. BAR-HILLEL: Did you say it takes only a few pages to give you these instructions? Do you know that I think it might amount to a few hundred thousand phrases.

MR. BULL: May I inject something here that is left out of the discussion? If you were to do the proper linguistic investigation, I am fairly certain--this is just a rough estimate--I am fairly certain--that one thousand idioms will take care of nearly 90 per cent of the problems that you are going to have to take care of. Then you will resolve the problem of the idioms. You could reduce the process enormously.

MR OSWALD: May I suggest one more thing at this point? I don't want to drop the train of thought right here, but we have gone in a full circle and we are now back at the German language again. There is a micro-language--not only of vocabularies--there is also a micro-language of idioms, and the type of idiom that prevents that type of problem. You are just sticking your foot in it. You do not get formal discourse of the type that we are hoping to deal with. Every time I open my mouth, I put my foot in it too, except in the most informal kind of circumstances. I don't know what we can do, but the linguistic users of scientific papers find them extremely primitive. Actually, the more complex and complicated the discourse becomes, the more primitive

your language will get. I believe that many of the problems we are fighting to avoid will not be there at all.

MR. BULL: I want to add something at this point. I have done your technique for the thousand most common words in the English language. All of them are invariabl bound to idioms of some sort. The magnitude that you are going to run into if we don't peel off the thousand most common idioms is this. You set up your sentence and let each one of these signs represent a word. You would then have all the translations set up at a given point and you could use 12 or 15 thousand high-frequency symbols to work with. The machine would produce about 150 pages so it would be impossible for the post-editor to get through all this work. You have to reduce this business by some phrase-substitution technique for the high frequency one or we will go broke just paying for the paper to print this thing on.

MR. DOSTERT: I have one solution to this problem. Take the French words which mean, "That to me is equal". Now the key word is in the word "equal". The most normal translation of that word of that phrase is, "That doesn't matter to me". Now, on the basis of this, I would simply have this key word in the idioms. You could elicit this thing as the number one thing, and the only one thing that is elicited. Then, I would say to the post-editor that the logical thing to do in this event would be to leave it alone.

Now, it is true that you could say, "I don't care," or you
could say, "I don't mind," or even, "It's immaterial to me".
All of those phrases would be possible translation of some-
thing. However, if the phrase, "It does not matter to me",
covers the problem of logicality, you have actually put your
semantic message across and you can stop there. Therefore,
all that you have to do is to tell the post-editor that when
he finds something that does not follow the meaning of the
translation, that he should leave the word or phrase alone.
He should leave the translation alone unless it jogs with
the logical sequence of the factors involved.

MR. HARKIN: May I suggest that we find a more
equal translation of the idiom; something that takes care
of the same thing.

MR. BULL: You could translate both possible
phrases, and then almost any intelligent person would know
how to translate the correct meaning. That to me is "all the
same" and "all the same" can be translated as meaning "any-
thing of no difference."

MR. DOSTERT: The only trouble is that with the
earlier proposal, you don't have to do anything. It could
be translated as meaning "I don't agree with you at all".

MR. BAR-HILLEL: This is only a problem of quan-
tity. How many idioms can be afforded? We might be able to
afford hundreds of thousands of idioms in the dictionary

to mean two or three word phrases. We cannot afford an
indefinite number of them. It would be worthwhile to have
the whole sentence as a whole phrase in the special diction-
ary. If not, the message cannot be decided. I would take
the longest phrases of a sentence--if there are not too many
of them--and I would put them into a special idiom diction-
ary with all the other idiom phrases. I think we have to
arrive at a reasonable estimate. We have to use all of the
methods in the idiom dictionary and arrive at some conclu-
sions. The first method does not lead to conclusions, and I
want you to know that when I said it would take only a few
pa es, I mean it would only amount to a few pages after we
have eliminated a large part through a special phrase dic-
tionary. An intelligent man can handle four or five pages
without any trouble, but he would forget the translator
instructions.

MR. REIFLER: I won't say that we can't reduce
the number of idioms, but when you speak in terms of idioms,
we mean to say that idioms of the language is represented
in another language in a completely different picture. That
does not mean to say that if an idiom of language "a" is
literally translated into the idioms of language "b". It
is not exactly a Chinese language, because the Chinese would
not express such a word or phrase. If you find the literal
translation of this, you would say it is not a very good

example of Chinese because it is a language of these idioms.
We can ignore this example and throw it out.

MR. BAR-HILLEL: I am now afraid that it is not a
very large number, it is a large number of small frequencies
which we will not be able to get rid of. Now, I am stuck.
This is not an idiom for you, but for me it is.

MR. REIFLER: The Chinese would understand the
message if they had a little translation of the contents of
the message.

MR. BULL: I think I can make a rough estimate on
this and probably make some sense. Just take the notion that
a good number of idioms can be translated and make some kind
of sense. Now, I think that this can be done. Let's take
the other point of view that a certain number of idioms in
a message will give you just a lot of jebberish results.
Now let's look at how many idioms of that kind that you
would need to store in your phrase dictionary. I would say
that you can store ten thousand of such idioms and get many
substitution techniques for many of the idioms. You can
get 98 per cent of all the idioms that you would have to con-
tend with, even if you were translating Shakespeare or even
Gothic into the Chinese dictionary. The thing drops with
great speed. You got down to the point where you find these
peculiar idioms occurring once in every thousand pages. You
can afford to stand the risk. I think I am accurate on the

figure of ten thousand, but actually, I think that five thousand would take care of most anything.

MR. REIFLER: But, in the human translation, such idioms are often literally translated because they didn't know of the Chinese equivalent to a particular word or to a particular phrase. There is some danger of confusion in that system.

MR. BAR-HILLEL: It is not what comes out of the jibberish; the danger lies in what comes out when you have the wrong sentence.

MR. BULL: But, we can stand those errors if our percentage of errors is much less than one per cent. We can tolerate that amount of errors.

MR. BOOTH: But, it would be very misleading. I don't mean to say that any of your examples are misleading.

MR. BAR-HILLEL: This would not be the case. You would not be misled, you wouldn't even get it.

MR. BOOTH: I think you would, in a certain sense, be misled..

MR. BAR-HILLEL: It is a problem of equation. You would not be stuck.

MR. DOSTERT: Let's suppose that you have a scientific paper or report of a preceding day which somebody was reading in English. If he used a French term and the person reading the English report didn't know a thing about French,

he would translate a wrong message because he thought that the problems of pre-editorship had been solved. You are going to have problems and difficulties to settle even after the problems of post-editorship is eliminated.

MR. BULL: Actually, I can demonstrate where people would be able to read that message with the greatest ease. You could disregard the prepositions in the sentence and express the rest of the sentence accordingly. The rest of the sentence would not matter because that is the way you would read the foreign language. They pick out the nouns and the verbs and say they don't need the rest of the language. That would solve that problem.

MR. OSWALD: I despair of our ever being able to translate diplomatic language because I can't read it in English anyhow. We would have to have a scanner to go in between the lines to get the meaning of a message to be able to give a true message.

MR. DOSTERT: The Democrats may say that about the linguists. I think that if you are going to restrict the scope of our objectives; if you are going to be so restrictive as to deal only with purely scientific languages; the output being intangible only by highly scientific scientists; you will wind up by not doing anything at all. Now, if it is your target to begin with this as a private project, that is all well and good but it seems to me that you are now

coming back to the question that we are asking.

MR. OSWALD: You made one mis-statement, and that is, if we confine ourselves to scientific discourses, the appeal will be enormous. Also, the scientists have all the money.

MR. BOOTH: I don't know the number of scientists in the United States, but in England, out of a population of approximately five million, there are less than 5,000. There would not be a wide appeal over there. You could take the 5,000 from England and multiply that number with the number of physists in the other countries of the world, and then you would recognize the vast amount of material that is in back of all these physists. There are government agencies and other people who want to buy these things. I would buy the other idea that probably the best way, and the only way to crack this nut, is through scientific discourse. If we then learn how to refine our own race to the point where we can handle the problems of scientific discourse, all well and good.

MR. DOSTERT: Let's leave out the diplomatic discourses in this thing.

MR. WIESNER: We don't know whether Mr. Harkin would consider the military requirements as being primarily in the scientific field or in the general.

MR. HARKIN: I would consider it scientific.

MR. BAR-HILLEL: I would like to stress to you all again that I believe that the problems of idioms can easily be handled in a wider field than just in the scientific discourses. However, it could not be developed in a wide enough field to include slang language. That would run into a very large number and the only results would be to increase the number of words.

MR. BULL: That figure of ten thousand will take care of you.

MR. OSWALD: I think you are right. That figure would take in a lot of idioms.

MR. REIFLER: It has been noted that units for very high frequency are very often irregular. Let's look at it from the formal side, the semantic side, and view the same situation. Words that frequently appear in German and the words we have that appear frequently differ greatly. The target source is very small. However, with "gieft" we have a very large number of those words in the German language that we would have to contend with.

MR. BAR-HILLEL: I have not checked this statistically in reference to the words "there is". I don't know of any Germans that would say there is a special kind of meaning for "this or that".

MR. REIFLER: He would not say that there is a special meaning, but he would understand that there is one.

MR. BAR-HILLEL: I don't know about that, but I do know that all these things obviously could be settled in a reasonable time by a certain pool of linguistic experts. They could tell when a phrase should be treated in the way we would treat idioms. I am sure we should try to show a clear meaning. This can be handled; we have only the problems to solve.

MR. BOOTH: There is a problem of patterns in trying to figure the French language in this way. I think that this has been going on for a hundred years.

MR. DOSTERT: This has been going on since the days of the Revolution.

MR. WIESNER: I want to ask a question in regard to many human translator machines. Is it possible for any given translator to work equally well in two directions?

MR. DOSTERT: It is not to be encouraged. I think it is better to train your people to be skilled in one active language, and skilled in two or three other languages, but he must be native to the language. It is rare that you can find a non-native in another language. There are some, but it is rather exceptional.

MR. OSWALD: John Albert was a resourceful man in the State Department. He had to work under such conditions with English.

MR. REIFLER: I had to translate from English to

Chinese. In the end, I was talking Chinese to the English-
man, and I was talking English to the Chinese. I didn't
realize it, but that was the language I was using.

MR. DOSTERT: I don't know whether or not the plan
that I have last night can, in any way, be of value in look-
ing at the machine itself. Could we think of a process
whereby once your material goes into the machine, that we
could then put on certain codes within the machine by human
operation, so that the message could be directed to one part
of the machine. Could we direct it to one area?

MR. REYNOLDS: This could be accomplished almost
instantaneously.

MR. DOSTERT: Could we shift from one area of the
machine to another?

MR. REYNOLDS: We could work on that problem and
I think we could find an answer.

MR. BULL: I want to make one more point on the
subject of idioms. There is in Spanish a little saying that
says, "He is the kind of a man that makes a road out of a
path". Now, the point is this. To the Spaniard, the inter-
pretation of that is just as difficult as it is to you. The
first Spaniard who said that statement, put to all other
Spaniards the same interpretation problems that you would
have if you set about to interpret that statement into the
English language. Every Spaniard would have to decide if

the man that, "makes a road out of a path", is really doing anything or not.

MR. BAR-HILLEL: Do you mean to say that the machine which is able to translate one language, should also be able to translate from one language to another?

MR. WIESNER: I wonder if the human translator could do the same.

MR. BAR-HILLEL: That is a very important question for me. This is a serious problem in regard to codes. I wonder if the translator service has to translate from the German language into English, and then be able to translate from English into German. I wonder whether we should build up two machines, or if we should build up a system for a two-way code. The point is there is a certain degree of bi-unique correlation with this problem. Whether or not it is worthwhile or not, I don't know. This is a problem.

MR. WIESNER: This would wind up with us not admitting any difference in the target language or in the source. You want to have the machine take a code language and be able to translate it any way that you want it to be translated. It is complicated enough any way that you try to do the job.

MR. BOOTH: One word could have an equivalent in several languages for a basic idea. If you are called on to translate, you could remember that the ideas of one

foreign language is translated into the ideas of another foreign language, so the output could be one or the other. One particular word is not translated into another idea, it is translated into another word. This means that from the machine point of view, the approach is different. You try to look up a German word in the English dictionary and see the results.

MR. BAR-HILIEL: Let us assume that we have words in German which can be translated into English, and that these same words of English can be translated by one system. Would it not be worse to have these bi-unique words on a special system that would save a lot of money? Then, these fifty thousand bi-unique words could be transported according to the need.

MR. WIESNER: You have your tape, which is not a part of your machine. This may be the way you do it, in which case, I would not think you would try to make it by direct means. If you mean that the storage is extremely high within the machine, you may be able to do it. It is not a question which you want to resolve here. It is largely a question of information.

MR. BULL: The problem of the equivalence of the words would be taken care of in the manner by which you set up your vocabulary.

MR. BAR-HILIEL: That is a mistake. That is a

fallacy. If your certain statement in English can be translated into German in one direction, or German to English, or to some other language, it does not mean that the machine as it is now will be able to translate in both directions. This is not a problem for just one direction of translation.

MR. OSWALD: You raise a very interesting point of the ability of the machine and its mind. I have never given any thought of its reversal in translation. It opens a lot of problems. This is a bilingual wall and we have to find out if the machine can reverse the language of English and German. That brings in a very interesting question. Have we skipped a short-cut? Possibly there is a short-cut to this problem that we have overlooked. This is something which should be given some serious thought.

MR. LOCKE: This has great advantages for the translation of any other language with the use of the universal code.

MR. BOOTH: Instead of having ideas in your language, the translation for any phrase in any language could be done whereby phrases would be the idea of the whole thing.

MR. YNGVE: The next thing that you are going to find out when you try to set up a phrase dictionary, is that you are going to have the same trouble with the phrase dictionary as you have had with words. You are going to have phrases being equal to sentences, but you are going to have

difficulties because this system will not go very well in all the different languages. You will have the same problems all over again.

MR. BULL: Why do we want to go both ways?

MR. WIESNER: I am not proposing that we want to go both ways. My question was largely aimed at trying to understand just a little better, how we would handle the problem. I don't mean to say that we want to make the machine that way.

MR. REYNOLDS: Actually, I think an estimate on the answer is, that it could be done. I think it could be done with one or two machines. The circuits of 1½ machines would meet the required circuits to do the job. Actually, it is a problem from the engineering view point as to whether you want to reverse the functions of the input and output of the machine. It is merely a problem of adding some additional inputs and outputs to build up the switching systems so that we can produce the reversals. In that sense, it is only a machine and a half.

MR. WIESNER: It would require some tremendous storage space.

MR. REYNOLDS: That is right, but I still think it is fully within the estimate.

MR. DOSTERT: I think it would require less space than we realize because there are certain elements that

would not need to be reproduced. You could use a lot of the material that is already built into the machine.

MR. REYNOLDS: You would need certain material to reverse the functions of the machine. From an engineering viewpoint, I think it could be done.

MR. BAR-HILLEL: You should talk about certain components of the machine. The operation of syntax would be as large.

MR. WIESNER: You don't know whether the syntax would be wired in or taped in.

MR. BAR-HILLEL: If it works one way, you can work it both ways.

MR. WIESNER: This means wiping out the memory of the machine and not using it again.

MR. BAR-HILLEL: No, it would require the preparation of extra tapes for the machine.

.....The morning session on Mechanical Translation adjourned at twelve-forty o'clock.....

WEDNESDAY AFTERNOON SESSION

June 18, 1952

The meeting convened at two o'clock. Mr. Reynolds was chosen as moderator for the afternoon session.

At this point, Mr. Victor Oswald read his prepared statements on "Word-by-Word Translation". The following pages contain comments on Mr. Oswald's statements.

MR. BAR-HILLEL: You spoke of the harmless form of the German word "der". Now, there is no reason to assume that this is the translation, but what is the case in the translation of "der" into the English translation meaning "of the woman". I don't know if there are certain artificial meanings to this translation.

MR. OSWALD: If you are going to take a word and shoot it across to all the possible translations, one of them has to mean "of the" or "to the". The only thing which you need to know is very important. You need the basic information telling you that the code has been changed to be translated into meaning the code, "of the woman". It is expected that you would be able to work with these codes.

MR. REYNOLDS: I assume, at this point, that Mr. Oswald is discussing this subject from the views that he discussed with Mr. Huskey. I suggest we call on some of your suggestions to tell what the program was about.

MR. OSWALD: Let's suppose that "der" is number

one in a register. Now, let's assume that "mantle" has gone into the register. It will be number twenty. By whatever device these words come into the machine, without the number, the machine will be instructed to look for them. It will then pick up the form from the register with a number, and then the information it requires which will show that in a combination "der" would mean "the". That is a simple system. Harry said it was; I don't know about it myself.

MR. BAR-HILLEL: You infer that word-by-word translation is impossible. There is not the slightest inclination to this. It has to be done block by block, but this has nothing to do with it at all.

MR. OSWALD: I am translating these words as I go along. I don't know what I am doing if I am not translating these words as I go along. In translating from the machine point of view, the word would come up as "der". Now, the next problem is what to do with these meaning-bearing words, and of course, that is another point. That would get into micro-glossaries and micro-symmatics. It is much too complicated to go into it now.

MR. REIFLER: I am thinking of the languages we are going to work with.

MR. OSWALD: I think Chinese is beautiful, and I think the wisdom of the world would be vastly improved if we could translate all the Chinese language into English. On

the other hand, all the Chinese contributions to atomic research have been rather brief.

MR. PERRY: Do you think you can put all the nouns, adverbs, and adjectives into one category and say that they are meaning-bearing and destinctive, and that they don't have certain syntactive features?

MR. OSWALD: No, as a matter of fact, you have fish and fowl, and then you have something in between them; something like part fish and part fowl. The articles have no meaning, and the demonstrative articles have no meaning. On the other hand, somewhere in between, and carrying most of the load on their backs, the adjectives and possessive adjectives give off with a little of both.

MR. PERRY: Do you mean that the definitions will just sort of slide in there?

MR. OSWALD: Yes, we can lick this one half of the problem by syntactic connections; but we have to lick the other half with a completely different system.

MR. BAR-HILLEL: The basic facts are lined up for the numbers. We have three and a half articles and twenty prepositions. Then, we have thousands of adverbs, and you would have so many verbs and so forth, that the other prepositions would be predictable. Articles have some highly predictable adjectives.

MR. OSWALD: That is one system that I don't know

very much about. I shall assume you are right.

MR. BAR-HILIEL: It is connected with the ability of the machine and the numbers. It is not really the same as some of the others. One is so preponderant, that it will appear in all cases and not by itself.

MR. OSWALD: Mr. Perry has put his foot right where the weak spot is situated. I am how getting quite confident about how to handle nouns and how to handle verbs. But, it is the middle group that I am not so sure about. We will have to solve that problem too. Also, adjectives, and in particular the adverbs, I am afraid of the adverbs.

MR. BULL: Obviously, the word has a meaning, but it is of a different meaning. There is a sign for the class-ification of the generics of speech; and as a classification into the known speeches, there should be other signs. You say that we are talking about words that refer to uniform circumstances and to external relativities in speech. We would have to operate something like this in my vocabulary.

MR. OSWALD: Watch out, because here we go. We are going around the same old circle again. Here we have the words "read" and "red". The language is a good deal simpler than Aristotle's was, but the word gets to be called in a wrong manner. It does not matter much to the man in the street when he refers to the color "red", but that is the level we are working on. I am confident that we could

diagram the prepositions. You see, as long as there is some kind of language-clue whereby the human intelligence can determine what that thing is, there can be a machine to do it too. I think the difficulties in diagraming the situation is so complicated and so space-consuming, that it is better and simpler to feed them through and let the post-editor do the worrying about them.

MR. BAR-HILLEL: It is a global issue, but it has to be determined. Mr. Oswald does not intend to say what he has just said, but he does mean to say that he wants to determine effective ambiguities. We have the machine and we have to consider more than a block. I would say that in general, we have to consider the whole sentence in certain specific cases.

MR. OSWALD: If we are going to be verbal, I said that the smallest thing we could work with was block by block translation.

MR. BAR-HILLEL: But, you do not want to carry out the final translation for the final translation. It is very obvious that Oswald does not intend to have a dictionary like that.

MR. OSWALD: All right, he does not...I do not intend to have a block dictionary.

MR. BAR-HILLEL: And the translation will continue to be a word-by-word translation, preceded by block-by-

block analysis; or as I would say, in general, taking into consideration the whole sentence so that at this point, there is no quarrel necessary. There is still the remainder of the issues which we will hear later because they still believe that the whole preceding block analysis can be reduced to a high degree. However, I completely agree with Mr. Oswald that this cannot be done, in general, by syntactical analysis. They will tell us there is a word-by-word block; and we will, in general, come out with a machine output which will be intelligible to the instructor and to the reader.

MR. REIFLER: You just used the expression "of the meaning-bearing words" in your comments. If so, then Oswald does use the block to determine the meaning of the word, and their meaning of it. He does use the block concepts.

MR. BAR-HILLEL: But of course, he does use the whole sentence to reduce the syntactical meaning of the words in the message.

MR. REIFLER: How about the situations if he does the whole thing in German?

MR. OSWALD: I think Mr. Reifler is still where I was two minutes ago. It is important that there be no misunderstanding. Although the analysis is block-by-block, the feeding from storage is word-by-word. You must also understand that I am not proposing that the feeding from storage

be block-by-block, and not phrase-by-phrase. I misunderstood and it is important we get that clear. The people who are talking about word-wise translation are talking about the method of feeding from storage. I think that the efficient way to apply that system out of storage is word-by-word, but only after the decision to apply it has been made block-by-block. This puts you in the position of being able to do something that nobody but an infant would be able to do.

MR. BOOTH: As far as this particular example is concerned, I found that I could read the translation very easy. I suspect that any student of this university should be able to read it very easily, so I can't understand why your group out in California can't read this also. There must be something wrong with them.

MR. OSWALD: No, it has something to do with what Bill said a little while ago. It was just a case of sitting down and reading the material. I think that unless we can give it to them to read in an acceptable form, that nobody will take the time to sit down and read it. We don't want them to say "take it away" or "I'm busy". It has got to be right.

MR. BOOTH: Suppose that you are a mathematical man, and suppose that someone said to you that a particular book is of vast significance to you. Suppose we say it is a scientific paper and not a book; then, I am sure you would

get a little curious about it. Then, I suppose you would say that it is only in German so you would have to pick up a dictionary and then you would read it because you want to know what is in the paper.

MR. PERRY: I agree with you because I had an experience with just such a situation when a Frenchman was working on the Russian papers. Part of the symbolism was with the fact that he was already able to read. He had a smattering of Russian so all he did was to look up a few of the key words. Thus, he was working in the field of theory or probabilities. What I was able to do for him--with my fair knowledge of Russian--was to give him something that is just about as good as this is here. That is, it was just as good as far as hitting the right technical expression is concerned. Of course, I did this verbally.

MR. BOOTH: It was a set of ideas in his language that enabled you to do it.

MR. PERRY: I gave him something of what you have here. It was really awful. If I could have written out my ideas of what the Russian was, then he would have been able to run ahead and do the same thing that you have done here. He looked at it and said that he was not a scientist. I took the thing back and looked at it for the first time. After looking at it, I decided that he could not figure it out. All his ideas are highly Germestic so I couldn't do it.

It is a highly technical sense and, therefore, it is not
surprising that I couldn't translate it for him. If it
was in my line, I could have done it and he could not have
done it.

MR. OSWALD: This suggests some extremely interest-
ing ideas. I have, by no means, closed my mind on this sub-
ject. It may well be, then, that the real specialist in the
field can read with greater ease than I had thought it could
be done, even if it be "monkey talk". Maybe one of the
things that we had better get around to doing very fast is
to do some of the very broad testing of the intelligibility
of such techniques so that we can find how far we have to go,
and where we have to stop. I went into this because I had
to. I don't have any particular love for it and I feel if
we can throw this system out and have something simpler,
then let's have the simpler method.

MR. PERRY: Yet, I think it has something to do
with it, especially when you start out to learn a language.
When I started out to learn myself, a language that was quite
far from anything I had seen before--for example the Russian
language--I got ahold of a text book in chemistry. My first
attempt at understanding Russian only produced more "monkey
talk". I can see how the Russians expressed their thoughts
because I could understand a little of the Russian language
and a little of the "monkey talk". In other words, a little

knowledge of the subject matter is an immense help if you are confronted with a statement in a foreign language.

MR. BULL: You need approximately five per cent of the running text to arrive at this kind of information.

MR. REYNOLDS: I would like to point out one thing and that is the fact that we are all specialists--each in our own line--but we have a great deal of periodical literature to survey. I think we are all confident in scanning that book. The question is, just how much information do we get in scanning a book? We are not paying any particular attention to syntactical contents. I think you would find very little, and you would also find that the only thing that you have increased is the intenseness of the memory.

MR. LOCKE: You can increase the idea of scanning. That may help.

MR. REYNOLDS: I think you are more correct than you are incorrect.

MR. LOCKE: Someone wrote up some conclusions of a book that I wrote. When I read his conclusions, I had a different view of my own book.

MR. BAR-HILLEL: It is still a question of what has been done and what has to be done. Booth and Reynolds have done more on syntactical resolutions than might appear from what they have been telling us. They have not been doing very much less than you have here, although it does

not appear. We have to be very very careful in this respect. Now, what causes me difficulty when I meet the terms? I am not a linguist. The linguists I know don't use none of the terms. I believe that after two minutes of explanation, I would feel much better at home than I do at this moment. It is not only the question of being an expert in biology, it is also the question of getting some elementary introduction of terms that I don't know a thing about. This is one point of what will come out if you translate something of this sort. It is a problem.

MR. OSWALD: You want to do something, but you only want to do as much or as little as you need to do.

MR. BAR-HILLEL: They have done exactly this. They have translated from French, Russian, and German, into English. They have done it with some kind of syntactical resolution. They have assumed that somehow, we are able to find out that a certain word is something and has a certain function in this sentence. I would like to know a little bit more about some of the things that they agree on. Do they know what is what in this sentence? They have done more experiments on it than anybody else.

MR. HELMER: We have done very little experimentation, mostly just for the fun of it to see what would come out of it. We used three different scientific passages and we simply assumed, or left to the translator the privilege

of using his own volition in adding in the correct inflections or endings to the words. I think I said yesterday that the rule that we followed was to use letters for the first translation given in the standard dictionary. If any of you would like to look at the results, I have a copy that is available for you here. Going from French to English, the translation is very similar.

MR. BULL: I would like to add something to your sentences. If your Russian is saying something that you know about, you can read that sentence. If that sentence is saying what we think it is saying, we know what it says. But, if I am saying something that is new, you don't know what it says. So, you have the ability of reading in a foreign language something that you already know. But, if the man writing the paper surplants what you say, then you are stuck.

MR. PERRY: The point is that I started out with a beginning test in chemistry, but that is not where it stops. I used this because it requires a little research in doing it. You can use that to reduce the structure of the language to some extent.

MR. BULL: If we do it this way, we are doing fine just so long as he remains within our camp. But, if he slips one in on us, we are lost. I think that would apply in the highly scientifical and technical business.

MR. OSWALD: Assuming that all of this is just what you say it is, and considering what we know about genetics, I would like to know if anybody who knows just a few key words would be able to figure out what anybody is talking about.

MR. BOOTH: That is something that I don't know. I could not say for sure.

MR. OSWALD: You got yourself a formula of some sort. Possibly it is an international language.

MR. REYNOLDS: Actually, you have introduced a new concept. So far as this conference is concerned, you are offering your own knowledge which is not being considered. In Mr. Bull's own words, we are looking at a particular phrase and he must supply certain knowledge. This also corresponds to the machine. Things that are built into the machine will come out; but, things that are not built in cannot come out of it. However, it is possible that this need not be so on a word-by-word analysis of syntactical analysis.

MR. DOSTERT: I have taken the first sentence in Mr. Oswald's paper, and I think you can call that something like a general language. Suppose I translated each word by its number one equivalent. Then, suppose I put the alternatives in parentheses. I believe that would add to the confusion of the reader, but the text that comes out in French is intelligible to a reasonably intelligent reader.

MR. BAR-HILLEL: The load on the editors would be
enormous. It is something he would not want to touch. If
he were forced to sit down and sweat it out, he could do it
because we all are so intelligent. But, the load is so high
and he would be confused. I know that the next morning, or
at any minute, I would leave. If you carry out the complete
statistical analysis, and leave all the original material in
the English language, then you will come out of it all right.
This is something I know to be a fact. I know that he will
not be able to handle this if you do not carry out any of
the grammatical analysis.

MR. DOSTERT: The only thing I was going to do
was to reduce the amount of ambiguities, which the alterna-
tives present to the reader, by eliminating the parentheses.
I think the text becomes more intelligible and I have not
introduced any block process, except in two instances.

MR. OSWALD: The reason we chose German to work on
is because it gives you the greatest number of problems of
that kind.

MR. BAR-HILLEL: You have to eliminate all the
conjections. How do you get all the "s" examples?

MR. OSWALD: We thought we had to do it.

MR. BAR-HILLEL: I don't understand it. How can
that be? Here you have far too many "s" in your text.

MR. OSWALD: We thought of having certain supplies.

Now, this is the very beginning of everything we did. We thought we would have a supply of a choice of singular or plural for every form. We were going to teach the machine to peel off all the endings and shoot the words through. We were going to supply, or give the machine the stem only, and ask the machine to supply us with the ethical alternatives in English.

MR. BAR-HILLEL: I do not agree with that. If you can have a dictionary constructed in that way, the first word will also be the most frequent word. He has some very good reasons to tell us that the thing does not make too much sense.

MR. BULL: There is a dictionary now out--it has been out for three years--and it is based on the frequency count of the English words. They set up the problem of listing the first entries as the most frequent words. You cannot resolve the difficulties that you run into on that, because your first entry may have a peripheral meaning and the functions would throw off the person that comes at it from a different point of view. Then, they make a compromise between the two meanings. The next step is that the dictionary was found to be invalid, so they had to go to work to correct some of their errors. In the first place, they didn't know about their mistakes, and they spent a quarter of a million dollars making the dictionary.

MR. BAR-HILLEL: Now, I don't know about that. If
you were to take into consideration all the German-speaking
people, you could make a dictionary of the lowest frequency
words. So, if you combine this with the micro-glossary, if
you combine the first with the micro-glossary; then, you
might come out with something that is feasible. But if not,
and you continue the general dictionary, you might have some-
thing that would work out for two or three sentences, but on
the fourth sentence, you would run into a lot of trouble.

MR. LOCKE: I think you are cutting the dictionary
up into five different fields of interest. One of the best
solutions for this multiple-definitions-problem is the
glossary, and then if it is possible, give the order of the
frequencies in the field.

MR. BULL: That is something that you cannot do.
You cannot do it in the order of frequency. It all depends
on what article you are reading. They have a discrepancy of
fifty per cent and that is just about a s bad as you can get.
It might very well wind up in a higher figure than that.

MR. OSWALD: The assumption of the Booth-Richards
operation is that you take the first dictionary meaning and
carry out the translation on that plan.

MR. BOOTH: No, that is not it. You can use a
micro-glossary.

MR. BAR-HILLEL: It is of very great importance to

find out how much the syntactical analysis would amount to.

Do you have to read all of the original text in order to

find a verb? For instance, why was it so easy to under-

stand my text? The reason is that I have instructions that

tell me to look up the verb at the end of the sentence.

However, I think that an Englishman that does not know any

German will be lost. So, who is he going to call on for

help? It is impossible, of course, but maybe we should

give him another machine to help him out in the conversion.

MR. BULL: It is cheaper to do it in the machine

than to go to all the trouble to educate a human.

MR. OSWALD: May I try to point out a consenus

on this problem? I would like to get awfully clear in my

mind just what I think I think. We are agreed that a cer-

tain amount of statistical flip-flop is necessary, and that

it also depends on the range you are working from and the

language you are working toward. We agree that a micro-

glossary is also necessary. It seems to me that the

thing we want to find out is just how little syntactical

flip-flop we can get away with, and conversely, how big the

micro-glossary has to be.

MR. BAR-HILLEL: There is one thing that disturbs

me. How many fields are you going to investigate? You

have physiology and a special dictionary for brain surgery.

Then you would have a special dictionary for every other

specialized field. If you set it up that way, you are going to have a lot of trouble because you are splitting it up too much. You would have to combine 27 dictionaries into one, so, I think this is not the best way to get the facts. I agree completely with the principle that we have to use the micro-glossary, but it is very difficult.

MR. BULL: If you are interested, I know the figures on that system.

MR. REYNOLDS: This problem of assembling the micro-glossaries is by no means something that is impossible if you are willing to think of it along the lines that we discussed last night.

MR. PERRY: We ran into this sort of thing when we started out to make analysis of certain indexes. It became evident very early in the game that you have certain words which are used by a very broad range of specialists, and these words seem to have the same essential meaning. Then you could consider a word that has two meanings, like the word "solution". That word has at least two meanings. There is a "solution" for a problem in equations, and there is also a "solution" for matter. These two common meanings run right across the board. Then, you could have a word like "cullet" which is peculiar to the glass industry and it would have its own meaning.

MR. PERRY: It seems to us that words like that

are going to have a different treatment. We were concerned with the problem of setting up a micro-glossary. We want it available for indexing purposes of the different words, and to have it arranged according to the different fields of specialization. Under this system, we will know that anyone indexing something under "cullet" is talking about something that is related to the glass industry. What we finally did was to say that here is a category of certain things that have been arranged on a practical basis. You know, each individual and each specialist tend to generate his own special language or jargon.

MR. BULL: It has been counted that there are 80,000 such jargons in the United States. I suppose that these jargons are based on the American dialect, and also on the American society.

MR. PERRY: It will not be easy.

MR. YNGVE: When you get right down to it, you will find that each person has his own dialect.

MR. PERRY: That is true in almost any professional field or industry. You will always find that they have two classes of meanings for a lot of words. One set of meanings is used by the specialists or the professional men, and these words have a special meaning. But, they are not used outside of the industry, if they are used at all. Then, you have a great deal of common and scientific words that are

common to many other industries. There are also some general words that keep coming up under different conditions.

MR. BAR-HILLEL: This micro-glossary system could be used and less and less space would be consumed. You see, you will have the history of certain situations to work on. For example, in biology, even the best scientific people will use an example from a completely different field just to bring out a certain point. Now, if you have an expert who would take care of these few things that the micro-glossary will not be able to do, you might be able to say that this system is all right.

MR. BULL: That would only include about ten per cent of the meaning-bearing words.

MR. OSWALD: Without realizing it, we have come around to discussing my second paper. What happens is that the casual word that strays in is not the one that in any way would affect the context as far as we can see, and there is a perfectly good reason for it. I will draw a curve for you. I am also glad to change terms in the middle of the stream, so I will draw a monotonous curve for you to show you the amount of stray words.

MR. LOCKE: There is one other point. When in some technical work, I find there is a good per cent of new words being used that nobody around here knows anything about. But, if you know the literary language, you can make

a darn good guess as to what the words mean.

MR. OSWALD: Those words will not be on the micro-glossary.

MR. DOSTERT: That would be a glossary of the American terminology.

MR. BAR-HILLEL: Each day, we have to add up the easily expendable kind, not the rigid ones; so that every day, you have all kinds of new terms that will have to be used. Not all dictionaries are flexible.

MR. LOCKE: It seems to me that the minute you get in trouble, you are going to have some method of looking at the complete definition. We are going to have to phrase, and we are going to have to get a complete definition. On some of the words, we should get a blow-up to get the meaning of the word.

MR. REYNOLDS: We are in agreement that a micro-glossary needs to be developed for people who will be the ultimate users. They will be a "particular" customer, or a "particular" user. However, we do have to give some consideration to the syntax which was brought up with the experiments of the Rand Corporation, and was brought out in the papers of Mr. Oswald. I suggest that we postpone the dis-cussions on syntax until after Mr. Bar-Hillel has given his lecture on syntax.

...This concluded comments on Mr. Oswald's lecture...

At this point, Mr. Yehoshua Bar-Hillel gave a lecture on "Operation Syntax". The following pages contain the comments on Mr. Bar-Hillel's lecture.

MR. BAR-HILLEL: Just as a chemist--even if he knows all the proper things of all chemical elements and all other things--even if he knows them by heart, he would still be unable to determine the exact chemical substances of something that had been composed of mixture of four or five substances. It might take him one, or two, or three years for him to find out what he is going to do, and nobody else could tell him what to do. Even with his knowledge of all things, he still has to stop to figure out what to do, and in just what sequence he will do things in. He is told that the first thing to do is to soak it in hot water or sulphuric acid. If everything dissolves, it is all right and he has no problem because he has nothing more to work with. If nothing dissolves, then they continue in another way, and so on until something happens. At the end of each step, he is told exactly what to do next. Of course, sometimes he is told not to do this or that, but he is still told not to do this or that. It does not matter what line he attacks the problem from, no chemist would be able to pull the answer out within any short period of time, even if he knew everything about the substances. There is no system for the poor chemist. But everyone, if he likes, can try to develop

his own system. However, in this case, it would not do him any good and he could not come to any conclusions in less than a year or two because he still would not know what to do. He would be really stuck.

Obviously, the same situation is existing here with us in the language problem. You could give a man a grammar book and some German sentences, but he would still be at a loss because he would not know where to start. What shall I do with this? Even if we gave him the elements of the German language, he would be unable to do anything as far as phonetic analysis is concerned.

MR. REIFLER: I was told exactly what to do first and what to do next. I was also given the third step, and any others that were necessary.

MR. BAR-HILLEL: First of all, those linguists would be unable to do this and not get any meaning, or find out the structure of the thing. He points to all things which a machine is not able to do. Of course, the language-teacher could teach some rules on what should be done if he runs into a German sentence. Of course, since the student would understand immediately some parts of the language. The student would use some short-cuts because he immediately would understand what is going on. He will not make certain mistakes after a certain time, and he will be able to make

this analysis. Of course, with certain exceptions, even a student will find problems that he is not able to solve. Now the machine has to be constructed in such a way as to carry out the analysis of each and every German sentence. This requires that we give the machine a sequential syntax; a syntax resulting in such a way that the outcome of these and a few more preceding places will be correct. What is the next thing that has to be done? There has to be one system for a given machine for a given language, and a given machine for a different language, and so on. This is a hard thing to do. It requires great linguistic ability because you have a multiplicity of choices. It has an almost infinite number to work from. You will never be able to tell whether this system is extra good, or to find another system that is better. All we want to do is to find one reasonable system that will enable the machine, or a human being, to find out the construction of a sentence in a certain definite time. It is a task which has stood to be accomplished for any language. The proposals have been made to do it, but I still say it is quite a step to take. Many people would have to work for many months to fill this out and be able to do it.

MR. OSWALD: I would say that we quit because we did not realize how formidable the problems would be.

MR. BAR-HILLEL: My statement is that even a good

linguist with two or three students could not finish the job in less than a year or so.

MR. REIFLER: A good linguist and a man of experience could do it very rapidly.

MR. OSWALD: I don't think so, because he wants a series of sequential diagrams. The minute we try to work that out, the steps would ratify and multiply. It would be very difficult for us to pull it into one program.

MR. BULL: You are both wrong. I have been working this problem in Spanish now for about eight years. I have got the operation syntax worked out to approximately 60 per cent of the language. We have sunk into this job, I would guess, if it were all paid for, about 25,000 worth of labor. That is all clerical labor, plus the time I put in on the job for eight years. We had a staff of 15 people, plus a good linguist. I would say it would take four years.

MR. OSWALD: You are talking as though we have not done anything. We don't have to invent the compass to discover America...we are well along.

MR. REIFLER: Try it on Chinese.

MR. BAR-HILLEL: It would be very difficult.

MR. BULL: I mentioned the help that I had on this job. For a year, I had ten people putting in about ten hours a week on the job. I spread it around over that amount of time. Let me put it to you another way. In a

three week period, any bright linguist can take care of 75
per cent of the foreign language if you tell him exactly
what it is that you want to know. Now, it is the remaining
20 per cent that kills you because you cannot give a suffi-
cient number of examples of the material to him and list it.
In other words, we run into many difficult problems. I have
certain problems on this thing which I would like to resolve,
but you cannot find an example for it in less than 845 pages
of material. The point is that I need 500 examples of it--
maybe even a thousand examples--before I can determine the
operational pattern. It takes hours, weeks, and months of
hunting before you can get enough of them analyzed. Now,
once you have found a model...

MR. DOSTERT: Once you have found a model, can't
you create other models?

MR. BULL: We tried, but then you get in ruts.
You simply multiply the same thing.

MR. LOCKE: It looks to me as though a solution
might be to mechanize the 75 per cent and store it in the
machine.

MR. BAR-HILLEL: Exactly, but I don't want to
start with 75 per cent. I want to start with 90 per cent.
I want to run the machine for just six months on an experi-
mental basis.

MR. BULL: If you start with 90 per cent, and run

the machine for a half a year, you will wind up with a percentage of 9.5 per cent.

MR. BAR-HILLEL: We could leave it and start it all over again.

MR. OSWALD: In the meantime, try adding up the expense of the paper.

MR. BULL: Going from 75 per cent to zero is a tough proposition.

MR. LOCKE: You bring up a point that is bothering me. How perfect a translation are we expecting to get out of this machine without the services of someone to smooth it up. You talk about 99 per cent--you're buying a lot of machine when you talk of a machine with 99 per cent perfection. Are we going to be satisfied with a machine which someone could operate satisfactorily at 60 per cent? It is amazing just how good 60 per cent is when we get it. Won't it be something?

MR. REIFLER: We can decide what we expect is enough or important.

MR. OSWALD: Who does your operating?

MR. BAR-HILLEL: I can give you an example.

MR. REYNOLDS: I would like to have some information on Pollard's technique. Could we have a brief definition of what this technique was?

MR. OSWALD: It is called a key to the rapid

translation of German, and it is a text book that will help
the graduate student in particular to be able to get a good
essential grasp of the German language so they can read in
their fields. It is very clever. It is reasonably complete
but the only reason that it is not of immediate use to us is
that it starts from the assumption that the student trans-
lator has had a semester of German or knows the elements, or
knows grammar. In any good course, he would get those facts.
Then Pollard's system would be of some help to him.

MR. BAR-HILLEL: I would say that the compila-
tions of Pollard and Oswald would give us 60 per cent of
this operation syntax, so if we are satisfied with a very
low degree of translation, we could start our work within
a half year from now.

MR. BULL: Yes, but 60 per cent of what? Is it
higher?

MR. BAR-HILLEL: No machine, at this moment, would
be able to do it.

MR. OSWALD: As the "daddy" of those proposals, I
know how many holes there are in them.

MR. WIESNER: I would like to probe the mean-
ing of this percentage example. We can give articulation
tests. You know, I was shocked when I discovered--when I
was working with Bell--that when they tested a certain long
line of circuits in a machine that has a coefficient of 70

per cent, they found it to be satisfactory. They found it
was satisfactory because the redundancy in the set was OK.
But, I wonder, whether we have something of the same sort
here. I would say this is 85 per cent on some scale, but
I would like a little discussion of what this scale is.
Does this mean that we will get something that won't make
any sense to us, or does it mean that we will get something
that won't be accurate 15 per cent of the time? Does it
mean that the machine would break down 15 per cent of the
time? It is not clear to me.

MR. BAR-HILLEL: This means that if somebody would
write German, and have sentences that have one or two commas,
you would run into a certain amount of complexity and wind
up with two sentences. But, as soon as somebody starts to
use sentences of 30 or more words, you would have a different
problem to solve. You could use every simple German sentence
and according to the combined methods of Oswald and others,
you would come out with a unique resolution that would be
too deep for the human-being. German sentences are usually
not written or spoken without the use of commas. Sometimes
a German sentence cannot be completely analyzed, and you
will find cases, from time to time, where the contents of
the sentence is much too large. We do not expect to have a
machine taking contents larger than the sentence into
account, but this is only a small set-back. If you have any

German sentence, I will show you. I can give you the following example to work out.

"Hans dache dass Paul krsuk ist."

In German, it might turn out that the first order to be carried out is to continue the number of combinations.

MR. WIESNER: Do you guarantee that we will also need a pre-editor?

MR. BAR-HILLEL: We have a certain amount of redundancy which would take care of a certain amount of the so-called mistakes.

MR. WIESNER: It seems that you are building a translating machine that will be used only for the German language.

MR. BAR-HILLEL: You could start almost any place. If I say start counting commas, it is because I am certain that this will be a quick method. But, if there are no commas at all, then you have to look up the last word of the sentence.

MR. OSWALD: How does the machine know what the last word of the sentence is?

MR. BAR-HILLEL: The machine would know because the period would be at the end of the sentence. There would also be certain graphic distinctions. If this last word were to start with a capital letter, the machine would know it is a noun so it would do something else. I will not

explain what it will do because it is much too complicated.
If this test turns out negative, then you have to go to the
dictionary, because, under this condition, there is no parti-
cipial construction in the combinations of the sentence. If
a certain test fails, then go ahead with the translation.
You don't have to do anything, because the German word order
and the English word order will coincide..

MR. OSWALD: The last word in that sentence is not
a noun.

MR. BAR-HILLEL: No, it is not a noun. However,
there is a certain test--which I will not explain--for the
occurrence of participial construction which need not contain
a participle at all. If the dictionary tells me that it
might or might not be a noun, then I would have other prob-
lems. This is the same thing that I told you before and it
is extremely complicated. I don't want to go into it. But,
we have solved the problems of all the possible conditions,
because if the last word is a noun, we would have no parti-
cipial construction. If that is the case, we can go ahead
and translate the sentence. If there is a comma, then the
situation would become very complicated. First of all, we
would have to count the commas, and there may be one or more
in the sentence. If there is just one comma, all you have
to do is to look up the word after the comma, and also the
word before the comma. At the moment, I am not sure of the

correct sequence. You could do it this way or the other way.

MR. REIFLER: Did you say after the first comma or before the first comma?

MR. BAR-HILLEL: I don't know if there is more than one comma, so I can't say whether you would start from the left or the right. I don't know. If there is just one comma, look up the word after the comma. Then check to see whether this word falls into a certain list of words which we shall call conjunctions or subordinating words. You would have to find out what type of word it is. In this case you would find you would find out that it is one of the few words among the list that connects the introduction by a subordinate clause. I think that is what it is called. If there is only one combination, then you would have to check the last word. If it is a noun, you will have to translate whatever is before it and then go on. If it is a verb-- and in this case there is only one verb--you will have to translate it right away. But, if it introduces a subordinate sentence, then check to find out what the last word in the sentence is. It usually will be a verb. If it is a finite verb, then I would continue to find out what the so-called subordinate is. This is a very complicated thing and I won't tell you how it is done. The secret of doing these sentences is that once you have found out the subject, you have to transfer this verb which appears at the end of a

subordinate clause. I can only give you rudiments and nothing more. In this case, you only have to transfer the order and then start your translation. This could be done in a relatively quick sequence because this whole sentence has only six words. This is one of the most simple German sentences. It is not the most simple, but as far as doing work on sentences which are more complicated, I am at a complete loss. At the moment, Pollard is not too much help because he has told us that he assumes too much. The machine does not know anything. It does not know the subject or the predicate. All these things have to be determined to a certain degree, and where the subject might very well depend on the fact that it cannot be an object of what is shown, or that it cannot be a predicate or a verb, you can see that we would have some problems to determine just what it is. Sometimes, it might depend on the fact that it could not be a noun. This would give us a mythical subject, and you know how to handle that problem. This can, in principle, be handled, even though it is quite complicated. Such a sequential system will tell you at each step what you have to do next after the commas. It will tell you that if there is only one comma, you will continue in a certain way; but, if there is more than one comma, it will tell you to continue in another way. There will be examples for every situation. However, the whole system is of great complexity.

MR. OSWALD: I have one question in regard to your sentence. How many words will it take, from the machine's point of view, to give you the 90 per cent efficiency that you are looking for?

MR. BAR-HILLEL: It would take an enormous amount. I don't know exactly, but it will take thousands and thousands of words.

MR. OSWALD: How many will kick back?

MR. BAR-HILLEL: Each order will have to be carried out thousands of times; but I am not afraid of this, even though it will be done with thousands of words. However, don't forget that all of these orders are reducible to about four or five elementary orders. The most important operation is in matching up the terms.

MR. DOSTERT: You are going to edit a number of recurring facts?

MR. BAR-HILLEL: Yes, but the complexities of this problem are of a much higher degree than you realize.

MR. BULL: As far as I know, there are only three recurring patterns.

MR. REYNOLDS: That is right. The process is essentially the same and the machine has completed its task regardless of how many elements go to make up the task.

MR. BAR-HILLEL: What will happen to you in the most important orders--if you even come to such a thing--is

that you will know that you have failed in the vast majority of cases. From time to time, there will be cases where you will have to check what you have coming out of the mis-fits of the machine. Now, I don't know how many mis-fits you will have to work with. There are an enormous number of operations that you will have to try because the machine is dumb, unless, you find one system that fits. The machine will not stop unless you come to other examples, in which case, you will have to find other fits.

Take as an example the German sentence, "John thought that Paul lied." This could be the number one sentence. In principle, this sentence is ambiguous. The machine should come out with both resolutions. In one case it is a demonstrative.

MR. REIFIER: But, there is only possibility. You would have two of the "s" signs given in that sentence.

MR. BAR-HILLEL: In English, you could change the meaning of that sentence, so there is really a principle of ambiguities and if you take them into consideration, I am sure you would find that it would almost never occur. In principle, I think the machine could come out with both of the answers. Then, it would be up to the post-editor to throw away one or the other of the possible outcomes and the resolution would fit. I don't think we will have to give any special orders to the machine to stop and then print out all

the possible outcomes.

MR. REIFLER: Is it possible from the point of view of the engineers, to carry out the research along the different lines simultaneously? Otherwise, it would take quite sometime before you get the solution. That is, first you would find out one thing, and the next step would be to do something else. But, soon you will come to a step that will not fit and it will, therefore, carry you a thousand or three thousand stages away from the solution. Could you be able to carry on in a different manner?

MR. BAR-HILLEL: In principle you could, but you would have to have three, or four, or even five machines. However, in principle, it could be done.

MR. WIESNER: I agree, but I would say that the additional time would not change the order of things.

MR. BAR-HILLEL: The machine could work on each of the operations for 15 micro-seconds--I don't know how much time, but it would not be too much--so the thousands of operations would still be done in from five to 15, or maybe 20 millionths of a second.

MR. REIFLER: But, are you finding a speech or a system to show the "ing" words?

MR. BAR-HILLEL: No, you would have to use both ways because no "ing" will test through in the machine. It will not tell whether it is one way or the other, but you

usually will be able to tell. Let me give you a final state-
ment on the testing process. I will give you a very simpli-
fied account of what this is. This is called the manner of
research. Now, this is a refinement of examination so that
the numbers are just arbitrary. The relations between the
numbers don't mean a thing. Now, I would like instead of
using some arbitrary numbers to use an occasional system to
indicate the relation between them. Let me explain it very
shortly. Let us use the classes of proper names and let us
designate the classes of intransitive verbs by $\frac{"s"}{n}$. This
would form a sentence. If you would be more specific, you
could call it the noun to the left, and that would from the
sentence. Now, though it can be a noun, the machine does
not know from nothing, so as far as the machine is concerned,
it could just as well be a noun as a verb. In English, it
could be a noun, and it could also be intransitive. It will
not happen very often, but I believe that almost every one
of the English verbs could be used that way. Let us just
consider it as some sort of a classic terminology. Now, we
would have "$\frac{s}{n}$" or a system of $\frac{"\underline{s}"}{n\ (\bar{n})}$ which would make up
the sentence. For this purpose, it could make some sense.
However, it is complicated and I would rather skip it. You
could have the "n" off to the right-- $\frac{|}{n}$ and it would
form the name. You would have to refine this thing, so that
an adjective, which is something out of a noun, would form a

noun, or a noun phrase. The word following it would form a
noun phrase, so it would be a $(\frac{n}{n})$. Let us assume that
there are no ambiguities, and you would come out of here
with a multiple of syntax and a lot of different analysis.
These will not all work because the machine will not be able
to tell which of these systems will fit. If it does fit,
the whole thing will come out as an exact sentence. Then,
we would know that our analysis, with all of the other prob-
lems, was able to fit into the system. Now, the operation
which is called "quasi" or "quasi cancelling" is very simple.
Whenever you have the sequences of forms, you can cancel the
forms according to the ordinary fraction of words. If this
is something that is out to the left, it can be cancelled to
the left. You cannot cancel to the right. For instance,
let us start from the right and see what happens. If one
"n" cancels out another "n", then all that you would have
left would be an "s". Therefore, I would know that all that
is left is a transitive verb, and as you see, that is corr-
ect. If I had "lied with Paul" in the sentence, it would
not work. It would not work because it would not fit. The
machine will have to carry out all the possible cancella-
tions. I could have started out with "$\overline{n}$" and the results
would be the same. I could continue with the letter "s" and
show you that we have two methods to work the problem with.
We know that the other "s" would fit and we know how it is

worked and operated upon. However, this one "s" is more incompatible more frequently than you would think. This is because of the different types. The one type would be something which would let us call it a subordinate clause or something, which out of the sentence forms a noun. The other is a demonstrative adjective.

MR. REIFLER: I want to suggest something. Now, what I am going to suggest will cause you to be horrified. However, all those clauses--in a sentence of this type--can be put into classes and considered in German, English, and quite a number of other foreign languages. You would still come out with nothing else but the sentence saying, "Paul lied." It would be perfectly intelligible. If you did this in the case of German to English, it would be perfectly clear all the time. It just depends on where you would put the sign on it. You would have "Paul had this" or "Paul had that" and it would be perfectly clear all the time.

MR. BAR-HILLEL: But, this is a singular case.

MR. REIFLER: I can show you that in Chinese.

MR. BAR-HILLEL: For the German language, it would be necessary because we distinguish the artificial. We distinguish "dass" as meaning only "that".

MR. REIFLER: Otherwise, you would have two principles of thought.

MR. BAR-HILLEL: That is right, but this does not

apply to any of the other 15 German introductory statements. This is only for English. For German, it would be different. In addition, the comma tells us the story; in English, you have the choice of it meaning one thing or the other.

MR. REIFLER: I could say that whenever a German double "s" is preceded by a comma, ignore the comma and just consider the other factors. It would simplify the situation.

MR. BAR-HILLEL: Don't be deceived by this message that has to be refined. It has to be refined to a very high degree. The machine will be able to find out just what is a block, and also to find out what is not a block. A block is something of which the exponent is the final index or that the exponent is the "ing" letter index. Everything in which the "ing" letters are indexed are a block. Now, the machine will be able to automatically find out whether the final index of the sequence is, or is not a block. There are only two possible ways that the machine can decide that this is a possible block. It is very obvious. If you take out the sequences of three verbs, we would not know what the results would be. It could be a block or it may not be a block. In the phrase "all the very great banks" the machine would have no trouble to decide whether or not it is a block. It might be a block in certain circumstances. Now, under the word "banks" we could use "n". The word "great" would be "$\frac{n}{n}$" , then, "very" would be something which would

form a "$\dfrac{\frac{n}{(\bar{n})}}{\frac{n}{\bar{n}}}$". In this case, "very" modifies the word "great". So, "very" is all right and "the" is just an adjective. To show this, we would have $\dfrac{"n"}{(\bar{n})}$. The word "all" would be $\dfrac{"n"}{\bar{n}}$. Now, how shall we treat the situation? There are two possibilities of starting to work on this problem. The machine is dumb and you have to try out both of the systems. Now, this is something which is being done. Therefore, the machine will have to work to the left and to the right. By crossing out some of these "n's", we can prove that this sentence was not a block.

Now, if the machine is given a categorical dictionary, a mono-lingual category list that will tell you that the spaces that occur between the verbs in the English language, the machine will be able to tell you to which category it would belong. You would have to have a million and a half entries which tell you what are the possible categories. Then you have to give a complete list of its categories..if you come out with the wrong answer to your problems or if you came out with an ambiguous statement, it only means that something is wrong in the categories. But, if the original was ambiguous, you have another problem.

MR. YNGVE: Do you mean that the word "all" has several possibilities and that the word "the" has several possibilities? Now, with this particular sentence, can you give us an idea of how many different tasks this machine has

to do for this one example?

MR. BAR-HILLEL: You want to know how many different combinations there are? I think there are over 200 of them. I am not sure of just how many more.

MR. REIFLER: This was suggested to me one time. I happened to discuss the problem of subordinate clauses and conjunctions. Now I have heard that all we have to do is to just discuss clauses. As far as the engineering problems of the subordinate conjunctions are concerned, I think it is an isolated problem. The machine will establish the operations for what is to follow, and if you consider the conjunctions by themselves as infinitives, that will solve the problems of clauses. In German, we could have a third, or single clause that would simplify the whole thing. The other suggestions would break up the whole thing into other components. It is a big problem.

MR. BAR-HILLEL: But the rule of commas does not coincide with the rules to tell you to take it out of the sentence.

...The comments on Mr. Bar-Hillel's lecture were concluded at this point...

MR. REYNOLDS: Gentlemen, I hate to call a conclusion to this discussion, but we only have time for one more paper. I would like to summarize for you from this paper that has been handed to me. It gives me a completely

new concept of what we are doing in determining grammatical structure. It is a technique I did not know about. I do think there is some hope that we do not need to regard his promise too seriously, but this is something I think can be done. I find myself at a loss to explain it.

...At this point Professor Stuart C. Dodd joined the conference. He was introduced by Mr. Reynolds...

MR. LOCKE: Since there has been continuity in the program, I assume that Mr. Dodd is going to be heard too.

MR. REYNOLDS: Yes, we shall hear from Mr. Dodd now if he is prepared to discuss his material.

MR. OSWALD: It is getting late but I suggest that we hold over to hear what he has to say.

...At this point Professor Stuart C. Dodd gave a lecture on Mechanical Translation...The following comments are in reference to Professor Dodd's lecture...

MR. REYNOLDS: I wish I had time to indicate the scope of the problems we would run into if we have to build the machine with upper and lower cases. I think we could build in a mark to indicate the letters to be capitalized.

MR. BAR-HILLEL: You will have to have some sort of a capitalizer. It is necessary.

MR. REYNOLDS: I think there are numerous ways of handling the problem, but they are all difficult. Also, I don't think there is any necessity for including the question

marks at the end of the sentences. There is no necessity for keeping the signals in one operation.

MR. BULL: Since you are reading, you won't know what the punctuation will be at the end of the sentence until you come to the end.

MR. WIESNER: During the period when you were filling the machine and waiting to find out what was going on, you might be able to carry out some other operation.

MR. REYNOLDS: To what extend does the machine have to know that there is going to be a question in the target line?

MR. OSWALD: You don't have any good understanding in phrases, except for the questions. All these parts are just a part of the language.

...Mr. Dodd continued his lecture on Mechanical Translation...

MR. BULL: There is one thing that is bothering me. What would you do with a specialist's terminology for all the objects that he has to work with? We could run into enormous examples with the student of bugs. We now have classified nearly a million species. You can't make a million descriptions in perfect English and still get them into a book.

MR. DODD: For the purposes of mechanical translation, you are not limited to a small vocabulary at all.

MR. BULL: So, for mechanical translation, you can take all the nouns you need; but for model English, we just throw them away.

MR. BAR-HILLEL: There are reasons. There is a basic model English and a model English.

MR. LOCKE: Well, first we were aiming at the world language. Then it occurred to me as I discussed it with Mr. Dodd that model English may be applied to the problems of mechanical translation. Even in mechanical translation, we have considered the possibilities of using the effects of model English for mechanical translation because we would not have a limited vocabulary. You would take from the model language whatever is practical.

MR. REYNOLDS: It is getting very late so I think we should conclude this day's meeting. Tomorrow we will really be able to discuss the lecture that was partially given to us by Mr. Dodd. I know we all have a lot to ask him, and we all want to hear what he has to say.

... Upon the suggestion of Mr. Reynolds, the meeting adjourned at six o'clock ...

CONFERENCE ON MECHANICAL TRANSLATION

Thursday Morning
June 19, 1952

MR. BAR-HILLEL: I think we're going to continue
for a few moments with Mr. Dodd's "Model English."

...‚ ‚Mr. Dodd then resumed his address ...

MR. REIFLER: Which mixes American with Chinese.
It could be a literal translation from the Chinese.

MR. DODD: I suppose that would be the result of
this rule of trying to get a language in which the Chinese
are, perhaps, the best example.

MR. WIESNER: Is Chinese of fairly regular ratio?

MR. DODD: Yes, it is one of the words that is
difficult to change ordinarily.

MR. OSWALD: Perhaps, it would be more familiar
if you would use the plural "these" for the singular, and
the plurals "these" and "those." I think it would be more
acceptable.

MR. BAR-HILLEL: This is not necessary at all.
The plural of "this" book and "these" book? "This" is only
when it stands alone.

MR. OSWALD: This be not the ones I want. These
be not the ones I want.

MR. BAR-HILLEL: This is O.K.

MR. LOCKE: You can't tell whether "this be" is

singular or plural.

MR. OSWALD: This be the one I want. You don't know whether it's singular or plural.

MR. BULL: Use "they." and "that's."

MR. OSWALD: So you say, "This is they." or "That's they." "This is they." or "Be that the ones you want?"

MR. BULL: You don't have an indefinite, but as far as singular and plural is concerned, after all, when you're demonstrative, you have got a finger anyhow.

MR. LOCKE: Not when you're writing.

MR. BAR-HILLEL: This is fairly well established.

MR. OSWALD: This is the form in which I want to do it.

MR. BULL: The form in which I want to do it. Everybody knows what you're talking about, and there are many ways.

MR. BAR-HILLEL: I think the simplest solution would still be to add a few more words, to add "these" and "this." The machine can use it much more readily, so it's simpler for the operator on the intake or the output.

MR. REIFLER: We could ignore the distinction between singular and plural because Chinese does it, and it works very well.

MR. OSWALD: If it works, that's the final word.

MR. LOCKE: How about eliminating the gender between "he" and "she" and "it?"

MR. DODD: In the pronouns?

MR. LOCKE: Why not "he," "she" and "it" without distinction?

MR. REIFLER: The Chinese do not distinguish between them, and it works very well.

MR. DODD: (from prepared speech) "To get rid of inflections of person from the English verbs..."

MR. OSWALD: Scandinavian gets by with exactly that.

MR. DODD: (from prepared speech) "The unchanging words "more" and "most" can easily be used..."

MR. BULL: Why not regularize and make it "more more" and "less less?" Then you will eliminate more words.

MR. DODD: But it would be better to make them particles.

MR. BULL: You would just have "more good."

MR. OSWALD: What do you say by "More, more good?"

MR. BAR-HILLEL: You don't save a thing. Instead of saying "very good," you say "more and more." You don't save anything at all. "More" and "most" are two independent verbs with no connection whatsoever.

MR. BULL: What about "good," "better" and "best?"

MR. BAR-HILLEL: Oh, yes, "more good" and "best."

MR. DOOD: Just by reduplication of as much empha-
sis as you want.

MR. REIFLER: I believe you want to say, "He's
very good." This is "good good" in Chinese.

MR. LOCKE: We say, "good and drunk" in this
country, too.

MR. OSWALD: You say the same thing in Yiddish,
don't you?

MR. DODD: (from prepared speech) "Take the pres-
ent tense in the work and let it not be inflected..."

MR. DODD: That little trick is a sweeping trick.

MR. LOCKE: Is teaching and was teaching.

MR. DODD: You have the present participle, "He
did be teaching." and "He also will be teaching."

MR. LOCKE: By eliminating all perfect tenses,
what are you going to do in a compound sentence when one
action takes place after the next? "When you have read the
meeting, you turn the switch." How do you expect to take a
thing like that?

MR. DODD: I intend to keep the perfect tense with
the use of the word "have."

MR. LOCKE: I assumed that the perfect tenses were
all eliminated.

MR. DODD: That can be done, but it could be that
you could keep "had" or some equivalent form for the single

or imperfect tense.

MR. LOCKE: I think that in fact we usually use the present for the future. That gives you the feeling that you don't need the future in English. We don't say it that way in English.

MR. BULL: You've got to have a perfect in English. Even in the most primitive languages with no tenses, you have to always get an aspect of perfection.

MR. LOCKE: This isn't covered here. Have you covered the use of "have?"

MR. BAR-HILLEL: Of course, a few say, "After this is completed." This is evidence that everything else has been done.

MR. OSWALD: How are you going to say it?

MR. BAR-HILLEL: "After he will."

MR. OSWALD: How can you say, "After you have broken down this compound."?

MR. BULL: "After broken down this compound, you go about your business."

MR. REIFLER: The Chinese say, "After broke down the compound."

MR. BULL: In English you've got a suffix of perfection there somewhere.

MR. REIFLER: "That is very right."

MR. BULL: So after it breaks down, "After broken

down."

MR. BAR-HILLEL: There are so many languages.

MR. DODD: When the committee working down this final form gets into difficulty with their rules, Chinese with their experience might be a way out to make a conventionalized ruling.

MR. LOCKE: It seems to me that it's applicable. You can say, "You can do this because of Chinese or Norwegian." But if you can eliminate everything...

MR. REIFLER: You choose that happening?

MR. LOCKE: If you only choose, but if you eliminate from Chinese everything Norwegian doesn't distinguish, this is what we've been doing. This is the way we've been arguing all along.

MR. BAR-HILLEL: But the application, there's a strong play. Otherwise, of course, why shouldn't we do as you propose. You wind up with some symbolic language that works fine.

MR. BULL: In French I don't see why you should push yourself in the position of writing your progressive tenses. It is just as easy to say, "When he talked to her, he played the piano." Your "when," a simple tense, is just as effective.

MR. LOCKE: You have an incompleted action. The imperfect is progressive in most of its uses. It is

imperfect in the sense of the word.

MR. DOSTERT: We will eliminate by cutting over the edges. of the past indefinite, and we will retain the past definite--the simplest of all.

MR. LOCKE: You need something to indicate that something is complete.

MR. DOSTERT: The context is what we rely on for some of our meaning.

MR. DODD: We ought to get a one to one correspondent for each item.

MR. DOSTERT: "While he," or "When he talked to his cousin, his brother played the piano." In french instead of using the imperfect, I would say (A sentence in French was spoken).

MR. REIFLER: I would like to say to Dr. Bar-Hillel that we don't need to consult Chinese, we can consult Pigeon English--and it works very well there. But as far as this so-called wording, it is a matter of fact that it is a substance of the word meaning, a word meaning to complete to perfect. "Thus he came." is in Chinese, "He come."

MR. OSWALD: I was saying that we need the gadget, but not the whole system. There are thousands of ways that it can be done. The German does it exactly like the French. It combines the present tense of the verb "to be" with the

present participle.

MR. DODD: (from prepared speech) "Leading with the verb "be" with the passive participle '-ed..."

MR. OSWALD: You didn't point there that you don't have to have a passive voice.

MR. DODD: No, but again that is a matter of judgment. But if you want to give that much conception to familiarity of the current uses, then, perhaps.

(from prepared speech) "Then for mood the present 'helping verbs,' they might be called..."

MR. REIFLER: In all those languages, at least which I know, are those helping verbs auxiliaries of mood. Each one of these has occupied part of the territory of another one. You have constantly to consider, in English, all the adjectives.

MR. BULL: If you sit around long enough, "can" will do it in English.

MR. BAR-HILLEL: You could cut those by half and not lose anything at all.

MR. DODD: (from prepared speech) "The two very useful participle endings..."

MR. BULL: I fell off the boat right there. You were going to have "lion" and "lionness?"

MR. DODD: If you want to express the idea of a female lion.

MR. BULL: Then you've got two forms, but you haven't got a free form.

MR. WIESNER: You're inventing a new word for "ness."

MR. BULL: You've concocted a new word for "lionness." So all things that haven't got an "s" in front of them are male?

MR. DODD: They would be unidentified as to sex.

MR. DOSTERT: We all have the word "she," as "shedog" and "shelion."

MR. BAR-HILLEL: A "shedog" in model English means her dog.

MR. REIFLER: It's not a female dog?

MR. BAR-HILLEL: It's her dog.

MR. REIFLER: Whose dog? (Laughter)

MR. LOCKE: You eliminated the redundancy, now you can't talk anymore.

MR. REIFLER: That's what it means in model English?

MR. BAR-HILLEL: In model English you say "shedog" or "her dog." So you have to be very careful.

MR. REIFLER: "I saw she-dog."

MR. OSWALD: It sounds like Chinese.

MR. DODD: Of course, you can use "male lion" and "female lion."

MR. REIFLER: You can indicate the feminine gender, it would enrich the English language. You don't know whether it's a male or a female doctor. It would be better to use "professor" or "professoress."

MR. BAR-HILLEL: I think it would be better to use "she" for "her."

MR. BULL: "She-professor."

MR. DODD: (from prepared speech) "For mechanical translation, a model language may be useful at the pre-editing or post-editing stages or both."

MR. BAR-HILLEL: You were speaking about meaning. You mean apparently one meaning, because otherwise it cannot be achieved.

MR. DODD: You can approach it just as dictionaries give the different chief meanings of the word, as you have in conventional use. And there are now semantic dictionaries that give the meanings of phrases as well as the multiple meanings of single words. We could work along that line as far as possible.

MR. BOOTH: I think the time is getting on. I think it would be better if we went on to Professor Locke.

MR. BAR-HILLEL: I think we should discuss it longer, the summary is not enough.

MR. BOOTH: There was a period of silence.

MR. BAR-HILLEL: Those proposals are of great

importance. It is necessary to comment upon at least a few things. There are two things completely independent from the mechanical translation, and I believe we are not interested here in discussing it in this way.

How far could one of the three letters, these three are the host real. You can interpolate and compromise as much as you like. There are three letters, as you remember, where one could accept this language, and I believe you could concentrate on the less radical problems because as soon as you run into it, the situation becomes even more complicated.

The achievement of completely regularizing the grammar, which is English, is pretty easy, but much more important the achievement of regularizing the syntax would be unique. Work up to a certain degree. This is something that would obviously simplify to an enormous degree mechanical translation, as we have seen, or the recognition of the syntactical features of the given sentence in English. You have seen that each noun could be a verb, perhaps, an adjective and many other things.

MR. DODD: There could be different ways to say things, the synonyms, the richness of expression in a language. But actually, in the long run, you can do it by standardizing the word order and having every part of speech become able to be used as all other words of speech.

MR. BAR-HILLEL: I entirely agree with that, but a human being would be able to handle this easily--for a machine it would be harder. If you have this array of six or seven words, each of which can be a noun, can be an adjective, perhaps, sometimes an adverb, and so on; then this fitting in business is what does this function in this context. You will say, "Of course, it depends on the order." The first, but now what's the first? Not the first verb is the subject. It's the only thing you can say is the first group is the subject, where the first group may contain more than one verb. If you have seven words, what is the verb? It's the second or the third, fourth, fifth, and so on and so on. So the uniqueness of the verb order lies easily and very well to word groups. But how does the machine, and to a certain degree a human being, able to find out what form in a group? And there will be complications in ordinary English because each verb would be able to forefill more functions than in ordinary English. It might turn out that there would be complications which we do not see at this time, at this moment.

MR. DODD: You can indicate which is the subject in the text, at least, at some point in the sentence, which modifies the other things.

MR. BAR-HILLEL: A human being doesn't need it; then if he doesn't, then the machine, in principle, doesn't

need it.

MR. REIFLER: I should like to express most strongly from the outset, before we ever thought of this machine, that we should not narrow down all of the thinking to those languages which have not such an absolutely fixed verb order, as Chinese, as proposed as a model language--where the dictionaries indicate that this is a noun and so forth.

Yesterday Oswald pointed out that we are, in the first respect, not interested in the language like the Chinese because we are not interested in translation of the Chinese language. I do not think at all in the mechanical translation of the Chinese language it is possible that you may say that the rating of the Chinese language, as far as sciences are concerned, it's not very great today, but it may be greater later. But what I'm getting at is that there is a lot of information in Chinese literature and so forth which scholars need, not atomic physicists, but there are other branches of human knowledge where enormous information is available in China, and which we would like to have at our disposal.

On the other hand, we cannot be sure that, say, countries in which English is spoken or French or German and will continue for all times to be the main sources of knowledge, of scientific knowledge, we have from the very outset to consider to develop mechanical translation in lines so

that MT may serve to maintain with respect to any sources of available information. And if we do that, we should develop mechanical translation on the lines of regularized language, or a language with a fixed verb order. If it is not possible to mechanize the determination of the grammatical meanings, every word according to its position, we would have to find another means. We would have to have a human arranger, the service of a human agent to do those things.

For our immediate interests, it lies practical to do so, but in the long run it would be much better for us and for mankind.

MR. PERRY: I was going to say something which apparently the Professor said yesterday. That is that model English would make an obvious ideal target language, and the machine can enforce the target language on the consumer. Unfortunately, as Bar-Hillel said, this the smallest of our problems. We would like to start on a model language. In other words, this reduces the problem but does not eliminate our most serious problem of MT.

MR. BULL: I like some of the things here. Supposing you were to take any foreign language and do your durndest to translate it into English and let the machine decide what your model English is going to be. Because if we have limitations with bugs and so forth, and that may produce for you automaticly, in spite of yourself, your model English.

MR. REIFLER: My second paper deals with this problem. I'm thinking of model English not in this form, but as a model target language adjusted to the particularities of the source language. As you say, people leave it to the machine. But we could set down the principles for it. We could study each of the source languages and find out as to what respects we can adjust model English to it, in order to simplify the English problem.

I have already certain definite suggestions for purposes of the world language. This is very good, but for our purposes we may make a different selection. One example is, for instance, the ending "ed," the past tense of all verbs. That would probably be good for a number of languages, but certainly not for German or Chinese because we don't use it there. But we do have something for "ed," therefore, we would use the "ed" to express the past.

But what are you going to do with the past participle? But English has a second form for the past participle.

MR. BULL: We've another alternative that maybe we could say, "He has bringen." and so forth. Thus this is not the only case now as far as German and Chinese are concerned. The model target language which we are going to put into the machine to those pecularities, even our eliminations it still would be legible. It should still be clear. He should be able to understand that type of English, and since we are

choosing between different alternatives...

MR. DODD: I don't see why the consideration of German, that it doesn't have the "ed" form, should make a difference. Because you're going to have to translate it, just translate it in the "ed" form.

MR. OSWALD: The thing that occurs to me is that some of us are thinking primarily in terms of an operation that could be done reasonably soon, and the prerequisite for doing this fairly soon is the turning of the language into monkey talk. I wonder where we are going to get.

MR. DOSTERT: I think it might be well to have that as the first objective of our efforts. When that time comes, to proceed with an experiment that would feed authentic language at the infant stage so that we automaticly eliminate the monkey talk at the beginning. Then it would be my disposition to try to learn by doing and not to learn by guessing. See what comes out of your machine. See the degree of unintelligibility that you have there, and then proceed to work on your output to reduce the area of the unintelligibility by regularizing, if you will, or by defamiliarizing, if you will.

You would do it within pragmatic limits, and then see what you've got. It may be that you have to run several times before you reach a degree of legibility, but it seems to me that if you start first by experimenting, rather than

establishing a limit, which would become a strait jacket, before you know it you are on the right track of the best process.

Many of us are going to set out merely to prove either the monkey talk that goes in, or the monkey talk that goes out. Why should we bring or confer all our prejudices on the machine. The machine comes to us with a virgin mind-- that is well. There are few people that approach any problem with a virgin mind. Here we've got the incredible probability of getting somebody on the machine to do something on languages without any kind of acquired prejudices. I think we would be doing good.

MR. BAR-HILLEL: Unfortunately there would be human readers instead of machine readers.

MR. OSWALD: You can only insist on simon pure virginity.

MR. DODD: You can go even further and set up a school population that could be a standardized population to resolve these questions. So that whenever you ran into any difficulties; such as, "Shall we make it more regular or more familiar, which would it be?" Then you could have this more regular. This use and the alternative uses and find out which one resulted in intelligibility, as determined by those children in the scores they were able to get in spelling that material. It could resolve all your difficult

problems. You could determine from the children whether this was going too far or not. You could come up to a crucial answer to every problem. Therefore, you could accept this ruling or not.

... Mr. Locke then gave his address ...

MR. LOCKE: But the minute they started writing it crept in. Then I suspect that the first writing might not have been an attempt to represent words, but may have been pictorial, and only an attempt to represent things--the special relations between things. But certainly very early we had that attempt to represent words, and then later syllables, and then finally the attempt to represent what people thought of as single speech sounds and what are usually called phonemes today by the linguists.

MR. BULL: That can all be traced in the Mayan. They actually perform a logical procession. They were just on the point even of getting phonetic representation.

MR. LOCKE: This seems to be the logical order, but it stems from what we know from other languages. The Egyptian has a mixture, the Karna in Japanese, and in English we have gone over to attempt to show simple sounds by letters.

The historical development gives us a lot of unnecessary letters and peculiar letters. We do have alphabetical letters that I define as an attempt to represent the sounds.

So to try to sharpen up this distinction between the written
and the spoken language, let me say that things are perfect-
ly obvious that written language as written language is
visual symbols distributed in space according to conventional
patterns. The eye has the ability to resolve these patterns.
I had these visual symbols, and I don't know whether the eye
or the ear is theoretically capable of taking information
faster.

MR. BULL: The eye.

MR. LOCKE: So the written language has an advan-
tage in that it can assemble information faster than any
other.

MR. OSWALD: It depends on whether you are audio-
minded.

MR. WIESNER: There somebody may know the answer,
but I've heard...

MR. BULL: You can read many more words than you
can hear.

MR. DODD: The ear can sometimes take in more
words than the eye.

MR. WIESNER: But you've seen the attempt to
determine whether or not you have the ability for comprehen-
sion audibly or from sight.

MR. BULL: Take a thousand words, a reader will
get through a thousand words at a speed that would be much

than the ear could take in that many words.

MR. LOCKE: Nobody has made a phonograph record clearer by speeding it up.

MR. DOSTERT: You have already modified your statement. You speak of a very fast reader, but there are also very fast hearers. I don't think you can be categorical about this. I would rather say that I can hear faster than I can read one hundred words at a rapid speed. I would get through those hundred words faster than if I used my eyes.

MR. REIFLER: In one respect you can be categorical, when you have something to read, you can skip a lot.

MR. WIESNER: There is some evidence that the eye is faster than the ear.

MR. LOCKE: We are getting a bit mixed up about what the eye can take. When you talk about which is faster, the ear is limited, the eye can take a space pattern.

MR. WIESNER: I don't think anyone has ever been able to determine it.

MR. BULL: I was wrong.

MR. REYNOLDS: In the first place I think the discussion is talking about the speed with which we can resolve the information presented now in terms of actual storage mechanism. There was some rather interesting experiments that were reported by Licklots of M. I. T., and in this particular series of reports you will find an analysis of a

cat's brain. Now the reason I bring the cat out is that, and the reason it was started, the cat accepts information apparently at the same rate through its ears as it does through the eyes. The question was a neuro-physiological question. Just how does the cat scan information received orally and information received visually? Certain ears were located in the brain. These were compared in terms of area and in terms of estimated number of neurons involved in accepting the information and interpretation, and in the cat they turned out to be quite comparable in size and ability to resolve. So far as the human being is concerned, we can identify the same kind of areas in the brain as we can in the cat.

The relative number of nerve centers that are used for the interpretation of the information are considerably less for the ear than for the eye, and they have been patterned out. I believe that it was in the last issue of "Scientific American," there was a report by a great British physiologist where he said that we can pick up more information through the eye, as compared with what we can pick up through the ear.

We are talking about a machine. We should be able to recognize at the outset, here I am harking back just a little small amount to what I said previously, we can condition the input either to accept information by the medium of

sound or by the medium of the printed page. The one is more difficult than the other because we respond intellectually. I am holding in reservation here what I am trying to say here, that there are some techniques which could be discussed at length in the engineering. The rate of input is questionable. I think that the visual scanning is still by far the more rapid of the two. The only thing I wanted to put out here is, to get some factual information these two will give some background in a not too technical sense.

MR. LOCKE: I think I went off the topic of my paper here. I didn't know which was a better way. To come back to the visual symbols of written language, the printed form of language is composed of a combination of these things separated by short or long spaces and punctuation.

Handwritten language has various curving which simulates the printed forms, but rather wide deviations separated by relatively short or longer spaces and some punctuation, probably not as much as in the printed and probably not as logical as in the printed form.

Now opposing these two types of written language is the spoken language. The spoken language is composed of auditory symbols distributed in time. The spoken language is composed of various combinations of notes. I don't know what to call them, so I'll say combination of notes which may be sounded on different parts of the musical scale. But

these cords retain their identity. They are separated by relatively shorter or longer pauses, which usually do not correspond to spaces of the written language. The pauses in the spoken language do not correspond to spaces between words. There are no spaces between words in the spoken language, you have nothing of this kind.

They have separated these combinations by accent and rhythmical features which do correspond in part to punctuation and to word spaces in the written language.

One more type of spoken language which I would like to condense into one word "processed speech." This means speech which in anyway has been transmitted, recorded. This invariably includes some degradation of the symbol patterns analogous, perhaps, to the handwritten case of the printed word. Any telephone conversation usually gives you a degradation of the symbols.

Now for the utilization of written symbols in mechanical translation, I'm not threading on anyones toes, I think, in the papers that have gone before. I hope that I'm not threading on the papers that come after. What I want to do now is to study the progression of operation in mechanical translation as I foresee it.

First, usually I think most of the material which we will want to process with MT has been spoken. This is a question. Maybe you would like to say that most of it was

written without being spoken, but I would not be surprised if half or more was in the form of orally delivered papers or speeches. Anyway a good proportion of it was spoken before being written down.

The first operation then was to write it down. This is starting with the speaker, with the person that originates the paper. First he wrote it down; then he had it typed and edited, because almost invariably a paper gets edited.

There's the second operation. After being edited, it is typed or printed and then proofread. If we have a pre-editor for MT, the pre-editor will take this printed and typed material, and he will pre-edit it. Perhaps, it will be possible that it will be typed again, the pre-editor will take and retype it again. That is the third operation.

Now it depends on the type of input for the machine, but it may have to be typed again on a keyboard on the machine. Maybe we can eliminate a pre-editor.

MR. WIESNER: Maybe the pre-editor could put it into the machine?

MR. LOCKE: I'm not sure whether the pre-editor would want to read it again. Even in editing, I think, after you have edited, you would want to look at the final copy. So in all probability the material will be typed again into the machine or punched or taped in some way, to get it into a form which the machine can use. This form

will be read by the machine, and it will then be processed and go through the syntax and dictionary business which we have discussed. If there is a post-editor, the output of the machine will be post-edited. This means that he will have to read again all the material in its entirety. With all the multiplicity of reading and typing through which this material goes, the time involved here is really very substantial. It has to be typed once more because he will pick it off the machine, edit it, that machine version, and then it will be typed in the clear and then, perhaps, printed.

If the post-editor has another editor, the chief editor, who rereads the stuff--as one often does have to clean it up-- if the post-editor has it cleaned up, it will go through another typing stage; and then finally it will be printed in the clear in the target language.

MR. REIFLER: But if you eliminate the editor and the pre-typing, it will save a lot of time.

MR. LOCKE: Look at the time all these people who have got to type it and read it through will take.

MR. YNGVE: How much does it cost to write a paper in the meantime?

MR. LOCKE: We can leave that out.

MR. WIESNER: I think it is a pertinent question, though. Perhaps, if you can translate the paper, you can avoid doing all the background that it requires--doing it in

your own language. Suppose it writes at $100 a word to
write a technical paper, you can avoid doing that, $10 a
word is worthwhile.

MR. LOCKE: You can argue that it's cheaper to
write it.

MR. WIESNER: We've never figured out the cost of
not translating.

MR. BULL: I think we should think in terms of
what it costs us to hire a translator to do it, and at the
present time, at the present moment, this is in line with...

MR. DOSTERT: Point four, point eight, about
point four eight tenths of a cent for translated stuff. The
commercial translators do it for less than that. They make
a substantial profit.

MR. LOCKE: So much for the utilization of written
symbols. Now for the utilization of spoken symbols. There
are two types of utilization. First, is delayed utilization.
In delayed utilization, and this corresponds to the first
state in which the sound is recorded. It may be recorded as
sound; it may be stenotyped; it may be taken in shorthand.
In the second stage the sound is typed out by a typist, or
we shall, perhaps, in the near future type it out on a speech
typewriter, which will automaticly take the sound waves and
put them into phonetic symbols.

MR. WIESNER: That's the most optimistic symbol

you've made this week.

MR. LOCKE: All right, I'll say five years. This is after all when we've considered the rest of this pattern here. Five years, perhaps, is not unquestionable. Perhaps, we'll have it as soon as the rest of these components are ready.

Well, now once it's been typed by the typist or by a speech typewriter, you then proceed as in step one for the written symbols. So you haven't done anything, except that your output of the machine will principally be in written form. But there is no earthly reason to say that the output can be spoken language. There is no earthly reason to say that we can't synthesize it. If we can get it into the machine phonetically, I can see no reason at all why the machine shouldn't translate into a phonetic form rather than alphabetical forms. I don't think there is any difference at all. I'm talking about writing the output in phonemics rather than alphabetical. So the output, principally in written symbols, can be either written or spoken language.

We know enough about beach sand to identify the individual sounds in terms of phonemes; then we know enough to reverse the process and take the symbols and synthesis them.

MR. WIESNER: We could do the second thing before we could the first.

MR. LOCKE: Some attempts have been made, but none
of them really have been successful or satisfactory in terms
of intelligibility. In principle we should be able to
indentify the sounds and make the speech typewriters, and
quick, too.

MR. BULL: We do know that there is a speech type-
writer.

MR. LOCKE: I don't know.

MR. BULL: There is one in Germany. A man by the
name of Grafus Graf made it. Only he could talk into it, he
was the only one.

MR. WIESNER: He's a Swiss.

MR. BULL: Let me give you my source of informa-
tion.

MR. WIESNER: We too could make a typewriter that
would be fine for one person. But going in the other direc-
tion, we could make a machine today that would work from
coding and make some kind of speech if you wanted to.

MR. LOCKE: It wasn't good, but it wasn't bad.

MR. WIESNER: The Haskins machine, I think if they
took away its monotonic quality, it would be a very intelli-
gible machine. So if there was a real premium for making
synthetic speech from coded language today, we could do it.
But you can't be so optmistic in going in the other direction.

MR. LOCKE: It was not to make tapes, but to make

syllables.

Now the other possibility of utilization of spoken symbols without going through all these eleven stems, in other words, transforming it, is the possibility of simultaneous translation. It is simultaneously processed by the machine. You process it without substituting any spacial distribution for the temporal distribution.

Now I don't know that I'm going to convince you, but I'm going to work hard on it--that this can be done. It seems to me that since the internal workers, computers, essentially involve a temporal distribution of impulses, that one may better be able to process a temporal distribution of symbols.

MR. BAR-HILLEL: I don't see the point, you have to store the whole sentence into a simultaneous thinking before the machine can consider it. The machine will have to store it in a spacial sequence of what you are talking about.

MR. WIESNER: You have both. The machine is not doing one thing at a time. It may or may not have only one pulse active in a machine at one time, but it might. I believe that twenty-five years from now the machine will do more than one thing at one time.

MR. REYNOLDS: But I would seriously question the best of the straight temporal processes, even when I'm

listening to you, or I'm talking here, I'm still using a temporary storage until you have completed the word. Therefore, in my own mind I have to store. Yet in the machine I will have to store in order to get comparable results from the machine.

MR. BULL: Now what you're proposing is that the machine doesn't have to store it in a linear fashion.

MR. REYNOLDS: I think that's correct, but it does have to store it.

MR. BAR-HILLEL: It must be linear. You must know exactly that this word is between such and such a word. The machine will have to do something about it. It will have to have linear order.

MR. LOCKE: Now in the output then the target language speech would be synthesized. This would give us the elimination of the time lag and expense that we've been talking about. But here you don't have any stems, you're talking into a microphone.

MR. BAR-HILLEL: It must be clear that the number of stems doesn't mean anything.

MR. WIESNER: You're assuming that you can crowd into this operation a number of things that you think that we can do in the other machine; namely, that I don't see that the need for a pre-editor has been dispensed with just because it comes into the machine in a spoken form.

MR. LOCKE: It seems to me that the page scanner is going to take longer. What about the producing of the type? You've got all manner of producing material.

MR. WIESNER: We're not going to take the fellow in Russia who has just built a Syncrotone. Most of the material we want to deal with will already be in printed form.

MR. LOCKE: Some of the uses, indeed, involve spoken material.

MR. WIESNER: Very little because even in the U. N. it's something that is going to be translated--it will have to appear in worked over form.

MR. LOCKE: .When you want immediate translation, this is the way you do it.

MR. DOSTERT: I would like first to start by asking a question. Do you visualize? You see I am not a technician so I have to ask something in the elementary stage, if you don't mind? Do you visualize language "A" will come in all right? Then that there will be something in the machine that will pick up the sounds of language "A" and transform them into language "B."

MR. LOCKE: Not on the individual sound level.

MR. WIESNER: It will pick up the sound of language "A," and it will convert it into a machine alphabet, from a machine alphabet into the source language; then it

will go through all the translation processes that we've been talking about.

MR. LOCKE: There is no alphabet here because it's on a different level--it's a phonemic alphabet.

MR. WIESNER: But what you have to do is translate.

MR. DOSTERT: It matters a great deal. You have no idea what could come out at the other end. I could write solemn verses in a very, very slightly different way, and you get an entirely different story.

MR. WIESNER: It's a problem that we, namely, are not able to translate poetry.

MR. LOCKE: I'm about to discard it for the spoken language, too.

MR. BAR-HILLEL: In the machine you have a certain representation as to whether the original is sound or whether it is print. You represent it in one way, and if it comes out in print or spoken language, it doesn't make the slightest difference. It takes exactly the same number of stems in the machine.

MR. WIESNER: There is another very significant point here. I think we are prepared to accept a rather ambiguous but halting translation if we read it, because we can then take the time to analyze it and interpret it. You can't do it when it comes out a spoken language, so I think it will require a better translating ability before you make

speech.

MR. LOCKE: Much better.

MR. REIFLER: In every country, particularly in this country, there is a number of foreigners or people who are not brought up in this country, and who come from mostly different parts of Europe and Asia, and who speak different forms. Therefore, I am not absolutely convinced that what comes out will not be immediately intelligible.

MR. BULL: You can't say, "Now we will discuss the problem of the House, Department of Commerce Building, Chamber of Deputies." and give the ultimate forms to be picked out. If you do that, you're sunk. You have got to hit the nail on the head.

MR. BAR-HILLEL: You cannot have any post-editor.

MR. LOCKE: We can have post-editors. We can have a pre-editor, but I didn't want to bring him in. We can have very simultaneous translation by the machine method, if you are willing to accept it. On the other hand, of course, you can record the speech, and somebody else can listen to it and rephrase it and put it in with any sort of symbols which you want, which will indicate the parts of speech and so on. This seems to be undesirable and unnecessary. A guy with a nice polished language with earphones will speak it.

MR. BAR-HILLEL: There is one point, you cannot present the post-editor with the simultaneous possibilities

of choice. It's very easy in principle to present him on the screen, to present all the various possibilities of the machine in print, and he reads off the correct translation. He will probably be one and a half sentences behind the original speaker. He can see them all simultaneously, but he can present only one to the ear.

MR. LOCKE: Apparently it's possible to get simultaneous listening and translation. You can put on a pair of earphones, and it's very difficult and requires terrific attention.

MR. BAR-HILLEL: Why does he need a terrific memory? He doesn't remember it, he sees it.

MR. BOOTH: It's impossible.

MR. BAR-HILLEL: Because if he hears ten words at the same time--not because it's difficult, it's impossible.

MR. BULL: If you have a series of alternative forms appearing on the screen, you may well arrive at the end of a fifty word sentence before you discover that the alternative you picked for position number six is wrong, and you've got to rescan the whole thing at some kind of electronic speed. If you haven't got a completely photographic mind, you can't do it.

MR. BAR-HILLEL: In the U. N. they can predict what the word at the end of a German sentence will be.

MR. DOSTERT: They correct themselves if they get the wrong word.

MR. LOCKE: With the machine you won't have to wait for the verb.

MR. BAR-HILLEL: No, you see the machine could not do it because we have not seen it so far. No machine which we have considered brings into consideration frequency and practicability of frequency. Now the human translator sees the word which he is predicting as the highly possible one, but it is not. The machine will have to wait and get the whole sentence.

MR. LOCKE: Undoubtedly it will be a rather sloppy translation which we will get. It might be something which will be intelligible without going through different individuals.

MR. BULL: We can't solve that.

MR. LOCKE: If one were able to do it, as I see it, what you do would be, as your spoken language came into your mind, you would indentify the phonemes. You have a speech typewriter, you can identify phonemics, and you would use a conventional type of spelling of all the languages. It would avoid identity which to us would be a very great advantage. Perhaps, the output of this machine could be a written form, as you suggest. We might very well conceive of a machine which would take a spoken form and which would put

out in a spoken form.

We mentioned the criss-cross possibilities, all the exceptions were possible.

Now I have another half of this paper. I think if you gave me another ten minutes, I could finish it.

MR. WIESNER: Probably there should be one more paper for this morning, and the next one tomorrow morning.

MR. LOCKE: The rest of this is an attempt to show that it ought to be easier to translate spoken language on the phonemic level.

MR. BULL: That's a theory problem.

MR. REYNOLDS: I find myself in severe disagreement with both Professor Locke and Dr. Booth. In order to get information into the machine, you gentlemen have been doing a great deal of work in terms of simplifying language.

MR. LOCKE: You mean the source language by a pre-editor then.

MR. REYNOLDS: Let us just state that simplification has been the greatest occupation of the conferees here.

MR. LOCKE: I thought most of us couldn't do very much without a pre-editor.

MR. REYNOLDS: But the statement that I'd like to make is that the machine has to recognize what it has to do. The machine is not intelligent, and I think that you're going to find that out. Now by spelling I mean including

the consonants, using the asterisk, using the various other forms in order to give instructions to the machine. You mean that some one operator is going to put it into the machine, and the raw material will have to be processed. There will have to be a certain number of instructions that must be given to the machine.

MR. LOCKE: I think we can give those automaticly.

MR. REYNOLDS: But if we try speaking into a microphone, this introduces enormous complications. How many times have you gone three words past, or this should have been capitalized, or you have to check the spelling. These are every day phenomenons of a human operator in the same language merely attempting to get the spoken language down into the written form.

MR. BAR-HILLEL: That could be overcome without any slight interference with the speech.

MR. REIFLER: You are just starting out with the comma. You speak up to the comma, or you could have...

MR. BAR-HILLEL: It is to have what he says responded by somebody according to some standard procedure with standard clauses and so on and so on.

MR. LOCKE: The guy who is respeaking it can just touch a key.

MR. WIESNER: You think that this is easier than typing? I've been involved in the past in these books for

the blind which are turned out in this country, and it's a
lot harder to turn those out probably than to control the
quality of the spoken language, then to get good typewritten
quality--and you don't get it as fast.

MR. OSWALD: The one thing that has fascinated me
about Mr. Locke's proposal is that anybody doodling with the
source language can do it faster that you can say comma. I
always tell my secretary, she doesn't know where to put them.
I may get very cross at this machine in time.

MR. LOCKE: Some machines are very sensitive.

MR. OSWALD: It seems to me that in our haste to
insist upon the superiority of the printed form in the out-
put, we're overlooking the fact that you can do a lot more
efficient interferring with the input than you can any other
way.

MR. BAR-HILLEL: This is on the assumption that
you've got a lot more machinery to do it.

MR. OSWALD: The machine will react to what is
going in, as the normal hearing apparatus.

MR. BAR-HILLEL: No.

MR. OSWALD: More so.

MR. BOOTH: It would have to be more so. I don't
say it will, but it might.

MR. OSWALD: It could even be used.

MR. WIESNER: And we want to have a very quick

translation, what would be better and quicker than to have it typed by some stenotypist, whatever it may be, or to have his talk, which is usually not in standard clauses, respoken by somebody for whom a certain machine reacts. He is very intelligent, has to be, and he will try to preformulate all the additional information in some standard way so that the machine will be able to pick it out.

MR. OSWALD: The German compounds, you can break them up by clauses.

MR. BULL: There is a time lag, or there it's fine and dandy to translate automaticly into your own language, but now if you're going to take a nice great big, long syllabic word and pronounce it in its component parts as you are making the simultaneous translation, the number of pauses in between each component part is going to build up on you; until finally you are way, way behind the original speaker. And now we are brought into the business, can you hang on to that, is your memory going to hold it?

We said, "Let's forget about the simultaneous voice translation." We're talking about the voice as the better way to put input into the machine because it makes the pre-editing job easier.

MR. BAR-HILLEL: At the United Nations simultaneous does not mean that the output is simultaneous with the input. It means that many different outputs are simultaneously a

half sentence or a whole sentence behind the input.

MR. DOSTERT: They are only concurrent with one another. A Chinese translation of a given sentence may take much less than a Russian. The original spoken in French, there is a little lag before the Chinese. He may want to wait until the French is all true, and he comes out with three syllables.

MR. REIFLER: I noticed that he clearly indicated commas and full stops just by the changes in the moderation of his voice. And it has been shown that there are these devices in speech in different languages, and it is possible that the machines are created where you can make use of it.

MR. WIESNER: There's a lot of information in the spoken voice. The possibility of translating and transferring a lot more information in the spoken voice than in the typewriter exists, but this is not necessarily an advantage when you've got the problem of machine interpretation. All I'm saying is that I think we've got a big job ahead of us before we can do this. It's a job of magnitude of the problem in translation that we are not willing to talk about. Namely, going back to conceptional symbols and then coming back to the language.

MR. OSWALD: Can the spoken input be interpreted in principle now as readily by the machine as the input of a character?

MR. BAR-HILLEL: How much of what you talk would you like the machine to pick up?

MR. OSWALD: Everything.

MR. BAR-HILLEL: That would take years.

MR. WIESNER: We have a much more complicated problem when you try to convey meaning in terms of punctuation.

MR. LOCKE: I'm not so sure that it's more complicated.

MR. DOSTERT: There is no doubt but what spoken language is far more expressive of shading annuances. The personality of the writer reflects it only through the stylistic arrangement of the selection of words, but in speech the whole personality enters into it.

MR. WIESNER: He may convey quite a different meaning.

MR. REYNOLDS: On these books for the blind that you were talking about. You said something about conveying messages by means of inflections and so on. Yet it seemed to me that would not a monotonous voice in which inflection was kept at a minimum so the importance could be stressed on the meaning. This done partially for the reason that you are saying that what they attempt to do in these books is to leave the speaker's mood essentially to the attempt there to substitute for the printed page. You needn't do that

but this is what the attempt has been. It is an attempt to take the form of the written material and not necessarily to make it easier to understand.

MR. LOCKE: I've another half which Professor Dostert embarked on a minute ago, but I don't know if I should attempt to go into it. But maybe Friday afternoon I could be given five or ten minutes or twenty minutes to follow this through.

MR. Bar-HILLEL: If you remind us, we will do so.

MR. LOCKE: The title of is is "Expressions of Various Aspects of a Message." And what I go into is the expressing of the cognative aspect, the affective or emotional act or the characteristic aspects of the speaker, in writing and speaking and handwritten, printed, processed speech straight spoken. How we expressed them is exactly what I've been talking about when I said they are conventional requisites in which we can express them.

MR. WIESNER: He has much more confidence.

MR. REYNOLDS: I think that it's been extremely interesting, but, perhaps, in view of the fact that we're not decided on how to solve the much simpler problem of translation from written material, I doubt whether anybody else has.

MR. WIESNER: Our problem, when we talk on Friday afternoon about what a program ought to be like, we ought to

talk about how you learn how to translate and not worry too much about the inputs and outputs. But in these preliminary stages we ought to take the inputs and outputs that are easier to come by.

... Mr. Oswald then read his paper ...

MR. BAR-HILLEL: There's no reason why there should be an overlapping between the technical and the non-technical.

MR. OSWALD: That's right.

MR. WIESNER: I presume that you searched both glosseries in every case?

MR. OSWALD: Yes, yes.

The one thing that I want to say, I shan't read anymore of this, is that I want to remind you all that what I think about this, I know only about the nouns. I see no reason why the verbs should not be similarly structured. It is likely that adjectives will prove to have a similar structure. That is, there will be a large number of pure technical adjectives relating to each particular field, and I think the same thing is going to hold true with the number of adjectives chosen from the general vocabulary--the ones I've called non-technical and smaller. I don't know about the adverbs, I can't predict about them--they disturb me greatly.

MR. LOCKE: Do they have specialized meanings?

MR. OSWALD: No, they are not as closely tied to the technical field, and they are not as readily predictable as the adjectives are, when you come to the question of more conventional speaking.

MR. BULL: You can pick your adverbs from almost anywhere.

MR. BAR-HILLEL: When you're talking about adverbs, do you mean in the morphological sense?

MR. OSWALD: The morphological sense.

MR. BAR-HILLEL: Because adverbs in this sense...

MR. OSWALD: There's no problem, no. I'm talking about the predictability of it.

MR. BULL: You have such a wide source of manufacturing probabilities.

MR. OSWALD: I anticipate a much greater difficulty, but the emphasis ought to be on, "I don't know, and, therefore, can not say anything about adverbs." I would be willing to predict that the verb structures will be the same way.

MR. BAR-HILLEL: We are in a better position in such languages as Chinese.

MR. OSWALD: I keep thinking, and I'm glad Mr. Helmer is here, and it's the same thing that I keep thinking of the infinitely laborious process of putting in these diacritical remarks and then conveying the text. What is it, twenty milli-seconds search time to a a drum vocabulary?

MR. LOCKE: You can have a pack of people getting the stuff ready.

MR. OSWALD: Twenty milli-seconds is what makes me bounce.

MR. BAR-HILLEL: You see from the practical point you have such a situation that national governments will not be interested in having their happenings produced from mechanical translation, but usually the consumer is a man who is interested in the certain thing. Since these are the most practical occasions, we are not interested in the fact that the United States wouldn't want its publications in English presented in such a way that the Russians will be able to translate it into Russian. But we are in most practical cases interested to have the Russians' publication output put out as fast as possible. In such cases we don't have enough people who could do the pre-editing. But for practical reasons in the near future, it seems that post-editing has to get some practical priority over the pre-editing.

MR. REIFLER: Why not have two types of machines, one for those languages where you can do it, and languages like Chinese where you cannot do it. We cannot avoid...

MR. BAR-HILLEL: This I don't see. Is there any language where we can avoid pre-editing?

MR. REIFLER: Yes, in Chinese.

MR. BAR-HILLEL: The machine will be simpler. Why should you romanize it?

MR. REIFLER: Because it would be impossible.

MR. BAR-HILLEL: I cannot see it.

MR. REIFLER: I am sure that you cannot do it to the Chinese language.

MR. BAR-HILLEL: Since the Chinese can read Chinese, I can translate it.

MR. BULL: What Reifler is saying is that in Chinese in order to do what Oswald has done, you would have to count by function--not by form.

MR. BAR-HILLEL: Obvious, of course.

MR. BULL: And the function is not revealed in the fore-written form of Chinese.

MR. BAR-HILLEL: Don't you think that this operational syntax applies to Chinese as it does to any other language?

MR. REIFLER: We have tried in vain to do something like that.

MR. BAR-HILLEL: How does the Chinese understand it? How do you read a Chinese sentence?

MR. REIFLER: I read to a certain point, and then I realize that this couldn't be, or I'm not quite sure; and then I have to go back. You get an expert interpreter, and he will tell you that we can only begin to understand

Chinese only after we have read it very, very much. Do you want something in modern Chinese speech? A most common written form?

MR. BAR-HILLEL: Do they have a period for written endings in modern Chinese?

MR. REIFLER: Yes, now they have. It is not used to indicate any differences in grammatical meaning.

MR. BAR-HILLEL: I mean only the sentence up to what point to pick something up.

MR. REIFLER: You get a whole text.

MR. BAR-HILLEL: The machine will be able to handle sentences. However, I would like to see a sample to see how the machine would be unable to find out what's going on less than a human being.

MR. BULL: Let him take, "Birds of a feather flock together." and do it in Chinese. But take any other one.

MR. BAR-HILLEL: If a human being does it, a machine can do it.

MR. REIFLER: I should, of course, choose of the very larger number of very complicated sentences. This is a simple situation because of this thing here at the end. This is something that tells us the sentence ends here. Now do you want a literal translation here?

MR. BAR-HILLEL: What do you mean literal?

MR. REIFLER: There is nothing to indicate the

grammatical relationship between them. It is one of the possible alternative meanings the word has. I cannot give you the indication of the tones.

MR. BAR-HILLEL: Can it be a noun?

MR. REIFLER: It can be everything. That means that the rest of it has a certain grammatical meaning and function, but if it is not a noun, we will have to try something else. How is the machine going to know that that is a verb? We know because of certain words that occur very frequently, and we know the meaning of the characters. It is a summation. If it wasn't for that, no Chinese would be able to understand a Chinese sentence.

MR. PERRY: In principle you are perfectly right, but our brain is much more complicated.

MR. REIFLER: Our Chinese syntax is much less complicated. It is so simple that you cannot create machines for the next hundred years to determine what we are going to do when we are writing. The modern Chinese are trying to do something here, but it will take a long time for the simple reason the first steps we have to do is to take the literature production of one person, all his publications, and make an analysis; then we get a whole period, and then we may do it, perhaps, but I doubt it.

MR. BAR-HILLEL: Now we come to it.

MR. DOSTERT: Swinging from a language without

endings and so on over to the other extreme, which is Rus-
sian. Here you have a very elaborate system of endings. At
first when you begin to try to read Russian, you are not
fully familiar with all these endings, and yet if you read
something, some simple text, you can get along all right
without knowing the endings. But when you begin to get over
into complicated texts; then if you don't know the endings,
you will quite often be stuck. The point I am trying to make
is this, that in Chinese you don't have these extra clues
that are already worked into Russian. And from the point of
view of machine translation, I would suspect that your end-
ings in Russian would be a little complicated to handle
because you have a variety of endings and variations in end-
ings. But I would think that the Russian grammar is more
complicated and would make it simpler for the machine.

MR. REIFLER: In many cases Chinese scholars are
unable to decide   the meaning of certain contexts until
they have read the whole thing, and even then sometimes they
would have to consult another one. You often see Chinese
scholars going from one text to another.

MR. BULL: This is what Oswald meant by verb
translation.

MR. REIFLER: We can create the romanization which
serves our purposes.

MR. BAR-HILLLL: No machine will be able to do the

romanization.

MR. REIFLER: Yes, of course, but then the human agent will have to come in because we cannot romanize. We cannot put in those marks or capitalization by a machine. Then someone will have to come in and put a mark here and there.

MR. BULL: Let me make a summary point on this in relation to the problem of the machine in translating Chinese, which is comparable to what a Chinaman does when he reads a sentence. The way the Chinaman determines the functions of the various components of the sentence is in terms of density of a variety of other factors, usually semantic. So that he is making a frequency analysis of recognizable potential semantic factors, and out of that summation he decides whether his first word is a noun, an adverb functioning like an adjective, or a verb.

Now I would then assume that the machine would have to scan all of the phonemes of Chinese; then it would have to make a statistical analysis, and it would then have to scan all of the possible semantic values of each sign; then it would go back and determine the function.

MR. BAR-HILLEL: No, it isn't feasible.

MR. BULL: What Mr. Reifler is doing now, putting in those marks, is the summation, you see.

MR. BAR-HILLEL: If he can do it, the machine can

do it.

MR. REIFLER: The moment I do this, you can do it in a very simple way. The machine you build on this basis will be simpler, and it will be cheaper.

The machine is not able to understand. It is only able to understand what has happened, and for obvious reasons that for Chinese there is no grammar in the second stage existing which describes the structure, describes the grammar to such a degree that the machine would be able to handle it. I am quite ready to admit that it might take 100 or more years before scholars will be able to do such a key that a machine will be able to handle it. What would take a year and a half to two years for other languages would take for Chinese a hundred years. Therefore, this is not to be done.

MR. BAR-HILLEL: The pre-editor would have to do the work on the basis of his semantic understanding.

MR. REIFLER: The tones are also very important. But in the Chinese script there is nothing to indicate any relationship.

MR. OSWALD: It is a well-known phenomenon that reading Chinese is very difficult, and speaking it is very easy.

MR. BAR-HILLEL: I do not envy Chinese physicists who have to read the stuff in Chinese.

MR. REIFLER: There is no need to pity them, they have no difficulty whatsoever.

MR. DOSTERT: They are probably no worse off than the Germans.

MR. BAR-HILLEL: They have to put much more brain power to understand a Chinese sentence than an English one.

MR. DOSTERT: In German you have to wait for the verb.

MR. REIFLER: The moment I see the first word here (indicates the blackboard), I know the next word will be behind, and the moment I see that "yo" following after this, I expect the noun behind it.

MR. OSWALD: We had a little exchange here, and I just wanted to wrap up the microsemantic thing with one more statement. That is the extent to which we will go into microglossaries will be tied in very closely with the degree of smoothness and finish that we want to put on the product we turn out. I gather the microglossaries you've been working with are not so micro as this. You've been working in medicine, just in biology you see. For some of these scientific fields there are adequate dictionaries, for others there are not. And the one thing I wanted to get set up before us all is that whether we go to the business of making these, I now think relatively small compact glossaries for special fields, or whether we decide to take larger

domains, let's say we take all of medicine, we'll still have to do a lot of plain lexigraphical work ourselves. It is not available in the books, so the lexicography will have to be done all over again. Whether we do them on a small scale or a large scale, we will have to do them over.

If you allow any kind of dictionary, we have to have an operational syntax. I now see no reason why one can't say that since we can make microglossaries, I see no further problem.

    MR. BAR-HILLEL: There will be.

    MR. OSWALD: Just one reason why I quit the operational syntax one year ago, because they are very closely tied in.

    MR. BULL: In a microglossary you may find that a certain noun in that glossary can only have a noun function, in which case your whole approach on the operational syntax is different.

    MR. YNGVE: I just would like to make a remark that both the operational syntax and the microglossary are really two aspects of the same thing. As I think you have begun to point out in that we're dealing with the reduction of the meaning. Given word, we're trying to get at the meaning by some method. The microglossary is concerned mostly with the technical terms that are for the whole of the language infrequent, and are more frequent in the particular

sense that we have.

MR. OSWALD: Peculiarly frequent.

MR. YNGVE: On the other hand, the syntax has
given you the meanings, the specific meanings, that the word
is carrying at the moment on account of the structure of the
sentence. In other words, we're dealing now with the struc-
ture of the sentence, before we were dealing with a much
bigger structure of the whole of the language.

Now I think there is something in between that is a
problem. There are a vast number of words which are less
specialized than your less specialized microglossary, but
these are not the special words that you have been talking
about, and these carry a large number of meanings. How are
we going to get at that problem?

MR. OSWALD: When you get this 80 per cent off the
top of the curve, the words that drift in carry less meaning.
They nearly become non-meaning words.

MR. REIFLER: A different text may start out with a lot
of explanation and then later only refer or assume a lot,
but not expressing things which have already been mentioned.
To what extent would this interfere with your glossaries?

MR. BULL: We find that the same verb which on
principle can be a noun and a verb when analyzed , turns out
that this verb is also used as a noun. You say that a human
being will go on and predict it. Well, it seems that in

this book it will be used always as a noun, and if it could be put in the machine that way, it would save a lot of time.

MR. BAR-HILLEL: I don't know what it would cost to do that.

MR. REYNOLDS: That's cheap. Of course, you will run into trouble.

MR. OSWALD: I'm doing no syntax, syntax was a pipe dream. The heck with my results, I want Bar-Hillel's system. I think we ought to forget that I ever did a thing called "Proposals." But I would be the first man to admit that these were incomplete, and I think Josh's system is complete. Don't be so modest, Josh.

MR. BAR-HILLEL: This is what, instead of having the machine find out?

MR. BULL: I don't know how much this word is going to be used.

MR. WIESNER: Suppose you're translating a book in which a certain word comes up many times with the same meaning, it would be very simple for the machine to catch on to that after a few pages.

MR. BAR-HILLEL: Would you tell the machine to forget about the other meanings or to do something to help it?

MR. BULL: Let's analyze it in this fashion. You have to do the microglossary to get the semantics to put in

the memory. At the same time the microglossary will not solve for you the frequency of the function of its total vocabulary.

You have a word "X" which may in the general language be in four parts of speech. Your microglossary might not resolve for you completely that this word "X" will be consistently only in one function in the literature. The point is this, that at this stage you could build into the machine, as Reynolds pointed out, a very simple counter. The machine goes along and it finds that after it's translated five books on brain surgery, a certain word always appears as a noun—at which point the machine erases all other things. The machine is building its own function.

MR. BOOTH: This is the feed bank process, teaching the machine.

MR. REIFLER: Glossaries are smaller than dictionaries?

MR. OSWALD: I was thinking in terms of size actually rather than of contents.

MR. REIFLER: A glossary would be enough, but not distinct enough for us. Ideoglossary would be the thing for us.

MR. BAR-HILLEL: Microsemantics means ideosemantics, you treat of the semantics of a certain field.

MR. WIESNER: We can retain microsemantics to mean

what it means.

MR. YNGVE: If we extend this microglossary, after you've made a functional analysis of the sentence, you know from the breakdown that the second word in the sentence coming after it is a noun. Now you look it up in the microglossary of nouns, and it gives its noun meanings.

MR. BAR-HILLEL: But how could you tell whether to start with this or that?

MR. WIESNER: Statistics will still be very useful. You still can go back to your syntactical analysis.

MR. OSWALD: Once we start to doodle, we'll find all kinds of short cuts.

MR. REYNOLDS: You have a statistical frequency, and it has a particular meaning to it.

MR. BAR-HILLEL: How can anybody tell its meaning in an average English sentence? How many different structures could such a sentence have?

MR. BULL: Most English sentences tend to have just one. We could tell the machine the one thing. For Chinese it wouldn't work, no.

MR. REIFLER: An author would have a word that would be grammatically possible in two ways, but the author only has one of the two in mind, and the machine cannot find out, therefore, both would have to be included. This has to be investigated by linguists.

MR. BULL: If you mean by syntactical structure the permutational possibility of syntax blocks, it is about twelve.

MR. BAR-HILLEL: You mean an ordinary English sentence could have twelve syntactical structures?

MR. BULL: No, factorial. "Yesterday at noon when you arrived, I was tired." Now by the process of analysis at which we arrive at syntax blocks, we have this, this, and this--factorially three.

MR. BAR-HILLEL: Can this be analyzed in any other way than what you have done at the moment?

MR. BULL: I don't see any other system.

MR. BAR-HILLEL: You, as an ordinary English speaker, could you understand what you have written now.

MR. WIESNER: The question was, "Can you give it a different interpretation?" ·

MR. BULL: The blocks would be the same.

MR. BAR-HILLEL: There was a possibility of arriving at two different blocks.

MR. BULL: By and large this will come out. You can mechanically arrive at the syntax blocks in 70 or 80 per cent of all.

MR. BAR-HILLEL: If this is so; then we can construct the machine.

MR. OSWALD: I still haven't heard the answer to

the main question. That is, if your factorial digit is one, how many computations does that leave you?

MR. BULL: 3,850,000.

MR. WIESNER: That is close to one.

MR. BAR-HILLEL: This is too close to one.

MR. OSWALD: In the case of the sentence you gave yesterday...

MR. BAR-HILLEL: It is much more complicated to have the machine learn syntax in the meantime.

MR. WIESNER: But how many instructions do you have to give the machine?

MR. BAR-HILLEL: In a machine of such a type it would be a huge amount.

MR. BULL: The position of the adverb is disturbing. In English it is the most disturbing of all. It can come anywhere in a sentence. There's no telling the number of different meanings the machine might give an adverb.

BAR-HILLEL: If we could tell the machine, "When you come to the first possible fit, forget the others...

MR. BULL: The number of little things like that in the English language is very small.

MR. REYNOLDS: But certainly those of us who write papers in this country are continually accused of being very, very poor writers. We are attempting to convey information, and our writing is monotonous because it is regular. An

investigation can be made by anyone who cares to read
scientific literature and attempts to follow the subject-
predicate formation.

MR. WIESNER: The machine could put through a
hundred pages, and you catch the error; and then you could
tell the machine that these are the peculiarities of the
particular meaning.

MR. BULL: Couldn't you wire your machine in such
a fashion so you've got your statistical counter imbeded
in its operation. At such a point you discover by observa-
tion that that is the wrong operation, you push a button,
and that stops it from operating on that thing.

MR. REYNOLDS: These are all additional things
that you are talking about.

MR. BAR-HILLEL: The machine will wind up with
something that makes sense and makes wrong sense.

MR. WIESNER: You may do just that and end up with
a bad meaning.

MR. OSWALD: I wonder if I can come back to this.
Josh's scheme is one in which the machine would analyze
the syntax components of the sentence. I was wondering if
the analysis wouldn't be simpler. I was thinking this time
of a microglossary. We're thinking of putting together some
small components into an ideoglossary, in the long run it
would still be simpler to lable them when they go into the

machine. It wouldn't have to work out whether it is a noun or a verb.

MR. BULL: That's what I was proposing. In other words, run in all brain surgery meanings; then in the glossary mark a particular word as a verb, and then in that microglossary the machine translates it only as a verb.

MR. WIESNER: You've made the assumption that you're not going to give the machine any additional instructions, and you can.

MR. BULL: You didn't let me finish. You punch the little gadget, and it will run as a noun. You now put a little counter on it, and I suppose we could count every one of them in some way or another; and then, if after five hundred pages of counting...

MR. WIESNER: But there's another point. I predict it would be true. That is, when you have established these microglossaries, which are in a sense a technical vocabulary, you'll find that he excluded the common things in his writings. So I predict that that won't happen.

MR. REYNOLDS: Let's take a couple of terms that were quite prominent in nuclear physics recently. How many meanings does the word "barn" have? How many meanings does the word "crocodile" have? In this country we use alligator for one, and the other has eight different meanings.

MR. BULL: This is your payoff on the thing, we've

plotted the meanings of a thousand words based on the
Thorndyke frequency, and they plot in a decreasing monotonic
curve. Now that is not an accurate business, but it cer-
tainly points to the general conclusion that that's what
you're going to get. And as you consistently get this type
of curve for the meanings as well as the frequency of the
words; then there has to be some correlation.

... At this point the Thursday morning session
came to an end ...

CONFERENCE ON MECHANICAL TRANSLATION
Friday Morning Session
June 19, 1952

MR. BAR-HILLEL: Before we start today I would like
to make a few short remarks. First of all, the schedule is
pretty tight today, and I would like to shorten, if possible,
the discussions as much as possible. Second, we shall have
a business talk this afternoon, and I would like, if possible,
everyone of you to start thinking of comments he can make
on the compilation of our work. Third, I think some form of
publishing this meeting should be done, and I would again
like you to think which form it should take, whether a type
of almost verbatim report or only the major talks with the
summary of the discussions following them, or else even only
a kind of general report on what was going on. Now there is
no discussions this morning, I would like you only to think
about it.

MR. BULL: Before I begin reading my paper, I want
to mention what I've got up there on the screen. This is
the breakdown in terms of form and function that we made in
order to get the statistics. Now what we did was to take
very arbitrarily a word and say, "By our feeling that word
has the form of a noun." Then we looked at the function of
the word in its context.

Now we had to begin arbitrarily and say, "Well, this

word, let's say like *house* has a form of a noun." Now
when we found "house" in a context, we decided whether its
function was that of a noun. If it had an adjective in front
of it and was the subject of the verb, we classified it as
noun one--what we call a common noun. Then we broke the
rest of the nouns into proper nouns which dealt with persons;
then proper nouns which dealt with geographical cities and
things like that. We were exploring whether or not this
would be significant.

In terms of mechanical translation obviously there is
a great difference between nouns which name persons and
places, which you don't have to translate and which would
be silly to try to translate, and nouns which would carry
meanings. That was one of the reasons for breaking this up.

We went through all the parts of speech and made this
recording. The form of the word and the function of the
word, where the form and function were identical, we put the
classification as one. Where they were not, we put it some-
where else. So we had nouns operating as adjectives, and
adverbs operating as proper nouns and so on.

Now where we got our data is as follows. We took a
stencil, a regular stencil, divided it up into four--three
by five cards. We typed that stencil nearly solid with
material. Now on the left hand side of the chart up there,
you see four columns. Each column represents one of those

four cards. We then mimeographed enough of those stencils
to have enough for every word that appeared on the stencil.
That usually took about 500 words; then we underlined the
word as it appeared in the context, wrote it at the top of
the card and put its code number, whether it was noun one,
five or so on afterwards, and then filed the stuff. When
we got through we then counted every card, what was the
frequency of each part of speech on it, and you see there,
card No. 1 had 168 nouns on it..

Now the first column then gave us our raw data. On the
right hand side we had what we call recaps. Obviously you
can't deal with form and function where they are not identi-
cal in the same fashion. So on the right hand side of the
center there, you have nouns, pronouns, adverbs and inter-
jections, where we find out of the various other classes
functions which were not the same as the form. So at the
bottom, at the right, we now have a set of figures which
gives us all of the words which we found only in terms of
their function. This would be what you have to do, Mr. Reif-
ler, for Chinese because the form isn't going to do you much
good, so you make a good word count on a functional basis
in order to make any sense. This required enormous amounts
of labor, so as a result we have only analyzed about 32,000
words. We had sixty sets of stencils, which gives us our
distribution factor and a total of approximately 32,000

words. That is the date. Now I'll go ahead and read the paper.

MR. OSWALD: You shouldn't exaggerate too much because otherwise you have to split up Southern California into three regions.

MR. BULL: You finally end up in studying the frequency of the words.

MR. OSWALD: The fact is that no matter how you count it, and no matter in what area you count this kind of thing, you might just as well be counting the number of petals that grow on dandelions.

MR. BAR-HILLEL: To count together the income of a street cleaner and a millionaire you have to be very obvious. For instance, the average alone, if you learn what's the average, you know we're learning, but in addition to the average you know something. The same is for this. If you tell only so that it tells, "This is the average frequency." it tells something, it tells very little. If Thorndyke in addition would have told us a little about all kinds of sigmas, I quite agree.

MR. OSWALD: If you count this significantly, you would have a significant result. You can't count it significantly.

MR. BULL: There is a fallacy in your argument. An average of the amount of income of various people in the

United States is a very useful piece of information if you
are trying to rate individuals above or below the average.
But the average of vocabulary is of no use to you whatsoever
because you are going to use the vocabulary. You are not
going to find the average money.

Now the problem is essentially this. You want to
actually say a word or put a word into a context or to put
it into the machine or you want to find it in a particular
context as it goes into the machine. The average has no
relation to that because that word in one page cannot be
rated as above the average or below the average. It is
uniquely right there standing in front of you.

MR. BAR-HILLEL: Every insurance company has the
same problem. If you don't know anything about the fellow,
you deal with averages. Since you can't get all the data you
would like to have in certain cases, you have to be satis-
fied with averages, and this is the same here. The only
thing which I completely agree with is that in such cases
the gross average is of very little value. I would say it
is a nice point, but it has no value, has very little value.

MR. BULL: I'm willing to agree with that, but it
reminds me of the Arkansas hillbilly who has paid a hundred
bucks for a load of corn whiskey, but he has paid in one
dollar bills. The man who sold the whiskey counted just so
many bills, and then he stopped counting. The man who had

paid the money said, "Hey, why don't you count the rest of it?" The reply was, "I counted clear up to over half of them, and it was right that far, and I reckoned it would be for the rest of the way." (Laughter)

MR. BULL: (from prepared speech) "If you ask me what on the average will be the first word in the language, we can arrive at that piece of information..."

MR. REIFLER: You could predict a certain general vocabulary. You couldn't predict any specific vocabulary.

MR. BAR-HILLEL: It has to be quantified. You can't predict more or less. It's more or less. Don't put the yes or no--it's more or less.

MR. OSWALD: You can predict to a very low degree the average vocabulary for general language, and it's of no use to anybody.

MR. BAR-HILLEL: Just as you can predict if you know somebody is a graduate of Harvard University, you can predict his income, but you can't predict the income of someone who isn't a graduate.

MR. OSWALD: It serves no useful purpose--that's the important point.

MR. BULL: Why don't we wait until I get through showing the slides, and then there won't be room for so much argument.

(from prepared speech) "There does not exist, nor can

there be devised..."

MR. BULL: Even in writing you get it. I will demonstrate it out of this talk before I get through.

(from prepared speech) "It has been demonstrated by numerous word counts..."

MR. REIFLER: When you say 3.3 per cent of the volume, you should explain...

MR. BULL: You can have 80 per cent of the volume and 3.3 per cent of the meaning by this philosophy--and that isn't getting you very far.

MR. REIFLER: What about different combinations to express the different meanings of ideas?

MR. BULL: If we were to make a semantic count and assign every word that had a different meaning an independent character; then the volume of the 80 per cent would decrease enormously, and most of the words would plot out on a relatively flat curve. You just complicate the problem worse.

MR. OSWALD: The number of one things coming in on the monotonic curve diminish.

MR. REIFLER: He was using the word in a different sense than I was taking it to be. When he said that that many words conveyed only 3.3 per cent of the meaning, not 3.3 per cent of the meaning in the text.

MR. BULL: But the same thousand gives you 80 per

cent of the volume.

MR. REIFLER: It gives you 80 per cent of the
volume, and that may come somewheres being near 80 per cent
of everything he wanted to say.

MR. BAR-HILLEL: It may. Your second number, that
is 3.3, doesn't make any sense. Whether it is 80 per cent
of the meaning or 100 per cent, it has nothing to do with it.
This has to do only with the amount of predictability of the
remaining words. What you are going to say now will be the
important factor. If all the remaining words are predict-
able, these 80 per cent tell us everything. If the remain-
ing words are completely unpredictable; then they tell us
nothing about that. So it depends on the predictability of
what is left out. You will now come to the correct conclu-
sion that this predictability is very low.

MR. BULL: Thirty-three per cent of that vocabu-
lary is made up of it, and these do not carry meaning. So
that 33 per cent of that total volume has no meaning at all.

MR. BAR-HILLEL: Then they carry in so far as
meaning as to what can be replaced into this. It can in
this case be only a noun.

MR. YNGVE: They also carry meaning in the fact
that you say, "The President."

MR. BULL: Oh, sure. I tried to define this word
"meaning" a couple of sessions ago, and Josh got off on

the metaphysics on me, so I will try once more.

MR. REIFLER: Don't say, "meaning bearing words," but say something else.

MR. BULL: When I say content bearing words, I mean by and large the type of word that has a reference in relation. So that if I used the word "book" or "paper," we can go out and find something that we're referring to. If I use it, we have to take it up from his point of view. It has structural meaning, structural significance, but it does not tell us what we are talking about.

MR. BAR-HILLEL: But what comes out if you are saying, "Book on the table," it's a clause. If you are saying, "The book on the table." assuming that there's just one; then there's a difference.

MR. BULL: What I am trying to say is simply this. What we have been trying to do all the way through this conference is say, "Look, even though you don't get the machine to translate a lot of the little gadgets, everyone of the scientists here has been saying, 'Well, look I can read things in Russian and in other languages, and what do I do, I pick out the meaningful words, and the meaningful words are the ones that tell you what this article is talking about.'"

MR. REIFLER: The English word "not" you would consider as a non-meaning word?

MR. BULL: In the sense that it does not tell you what you are talking about.   It gives you a comment.

MR. REIFLER: Then you would know that you are talking about garages or cancer.

MR. BAR-HILLEL: This talking about is one of the most dangerous misleading elocutions by the way.

MR. BULL: If you send somebody a telegram, you leave out all of the little words, and you have perfect communication.   If you code, you leave out a lot of things, and you have perfect communication.   Now when I am talking about content bearing words, I am talking about the stuff you put in telegrams, and the stuff you don't put in telegrams, I am saying, are the non-content bearing words.   It is just a practical definition.

MR. BAR-HILLEL: You never omit in the telegram the word "not."   "I am arriving tomorrow." or "I am not arriving tomorrow."   I believe you would never forget that word.

MR. OSWALD: Nevertheless, you know what the telegram is about.   Once again we're drifting off the base.

MR. REIFLER: Our problem is one of terminology.

MR. BULL: I don't see any point in arguing over the semantics.   If you know what I'm talking about; then I'll go ahead and talk.

(from prepared speech) "My preliminary conclusion on

this manner of thinking relative to mechanical translation..."

MR. BAR-HILLEL: I think that would be a very bad way of taking it. You have to keep the ideosemantics and not what's the type of discourse matters.

MR. BULL: You'd better decide that it's poetry or prose, and that the kind of vocabulary problems you have will be unique to poetry and unique to poetry in major features. Or if it's dialogue and prose, you're going to have the completely different problems of dialogue than you would with prose.

MR. BAR-HILLEL: I wonder. This composition is highly predictable, but, of course, it is very valuable to have this confirmation. But I wonder what difference this would make in mechanical translation?

MR. BULL: If you were to make a frequency count of the record of this conference, you would be amazed to find, to discover, that the number of technical words per hundred words in that account is way below the number of technical words in any one of the articles that anyone of you would write on your specialties.

In other words, in dialogue the "you," the "we," and so forth, the pronouns take up the place of the technical vocabulary so that you have by volume then a distribution which is completely different than if you are writing on the subject.

MR. BAR-HILLEL: That is completely correct, but you come to the conclusion that this has some relevance on organization of work for mechanical translation. I do not see this.

MR. OSWALD: It tells us what we have been guessing, that we don't need to fret about pronouns if our discourse is this kind of thing.

MR. BAR-HILLEL: What do you mean, we don't have to worry? We will have to have "I," "you," or "we." Whether they appear on a page or two hundred won't make a difference.

MR. OSWALD: But you don't need the pronoun "she." I'm just pointing out that this tells you why.

MR. BAR-HILLEL: Then I would continue with Bull and say, "Well, no, this is not in the lap of significant pronouns, it should be sub-divided into "I," "you," "she," or "it," which will occur millions of times in every scientific dictionary. The only pronoun which will be used would be "it," and you can go on.

MR. BULL: In order to prepare your vocabulary for your specialties, you have got to scan some kind of material. If you were to gather material in dialogue to be scanned, you would have, by volume, to take ten times as much material to get your technical vocabulary.

You say this is wonderful. We did 100,000 words, and we found that 750 words represented all that these people

ever said about their technical subjects. So now then you
say, "We can relax, we've got a sufficiently large vocabu-
lary to take care of these fields." You haven't because the
frequency of dialogue is so low that you have to go on and on
and on to get enough volume to pick up what they're saying.

When we are talking we're displacing the content bear-
ing words by other words, building up a sufficient number of
repetitions at the speed at which we can in prose.

MR. BAR-HILLEL: I agree. In our talks in this
discussion all of the technical terms each one has used in
his presented material, everyone of them has occurred or
will occur. And so the negative part is very clear. We
should not judge from the facts that these terms will occur
with lesser frequency. Whether any terms will occur in a
discourse, we cannot afford to have it eliminated.

MR. HARKIN: I did not see the distinction about
what was said on both slides.

MR. BULL: The slide we have up there now repre-
sents the vocabulary, represents its predicted function.
All the things that are called nouns up there may be actual-
ly verbs, adjectives or adverbs. If they function as nouns
on this one, the form is the only means of classification.
In this one the form and the function are identical, so up
at the top where we have "nouns," it must be not only a non-
morphological but it acts as a noun.

(from prepared speech) "The speed and efficiency of mechanical translation depends, vocabularywise..."

MR. BULL: This referrs to the number of words which appear only once in each sample of about 500 words.

(from prepared speech) "Now the suggestion that words which appear only once may be ignored in MT does not appear to be profitable..."

MR. BAR-HILLEL: In a hyperbole if you plot it on a logarithmic scale, it gives you this straight line.

MR. HUSKEY: The point is that these are not straight lines. You can't judge from these samples.

MR. BULL: What I want you to observe is how the different parts of speech behave. The high line, the one going up the farthest, represents the total frequency for the part of speech marked there in sixty samples. The low line represents the number of times that class, that part of speech appeared in our sixty samples. We actually in this case plotted it out on the basis of the three by five cards and used the number 240, so that we could get the distribution showing up better.

Now in these three parts of speech the two lines are relatively close together. In other words, the distribution pretty well predicts the pattern of the frequency, or the frequency predicts the pattern of distribution. However, we're going to have a decreasing amount of correlation.

Now this here, on that number three, the pronoun is indifferent to the subject of discourse, and as a result the distribution and the frequency is practically identical for whatever type of speech you have. In other words then, the degree of deviation between the distribution and the frequency is represented to the content bearing value of the word in terms of defining what you're talking about.

(from prepared speech) "Now going on to the adverbs you notice that we now get a more irregular pattern."

At the bottom you have the specific words that have the frequency and distribution where the lines cross on the graph. It's a case where the form is an adverb form morphologically, and the function is likewise an adverb.

Now we go on to erbs, and here we have the microvocabulary beginning to show up. Notice the tremendous drop of that one form along there in the middle of the slide. It has a high frequency in very few places. Then with the verbs very sharply we are getting classification by the type of subject matter that you're talking about. The jaggedness is more marked on the lower line.

MR. REYNOLDS: I have a suggestion here for future presentation. Wouldn't it be a little bit better to show these slides as bar graphs? It's a purely technical vocabulary with you. It would be better that way instead of this counted line. That is just a suggestion for a future

presentation.

MR. BULL: That would be better. When we present it we will have a detailed description and print the word so you can find out which is the word in each position. But it's a good suggestion.

MR. REIFLER: In such a discourse no other tape is possible.

MR. BULL: Actually I was just taking the word "tape." We would then instruct our machine that "tape" in this context means only this. It doesn't mean tapeworm; it doesn't mean what you put on a cut. Ignore all those other possible meanings that we've got stored in your brain. Essentially that is what we're trying to get at.

(from prepared speech) "Then this finding is the principle of what I've been talking."

MR. BAR-HILLEL: I don't catch, "Everything is so clear?" Why?

MR. BULL: You find, for example, if you take the data that we have here, that the number of words which appear only once is in ratio greater than the number of words which appear only once in brain surgery, because this material is picked from hither and yon.

MR. BAR-HILLEL: What's the consequence?

MR. BULL: You move from brain surgery to medicine, to social science--you increase the availability, you

-182-

increase the total available vocabulary that you may use. Therefore, the rare words in your discourse increase in importance to your communication.

MR. OSWALD: I don't agree with that. What Bill is saying is exactly the opposite of what I've been saying. This is the fundamental disagreement.

MR. BULL: Wait until I've given my demonstration.

MR. WIESNER: What I think you're saying is that you picked the wrong word out of the vocabulary. I can imagine the consequence of this in brain surgery, which I have seen has a small vocabulary.

MR. BULL: (from prepared speech) "The present paper, for example, deals with a highly technical field in linguistics..."

MR. BAR-HILLEL: These words should not be translated by a mechanical translator. These words should have appeared in codes, and a mechanical translation should have to get special rules.

MR. BULL: What do you mean by rules?

MR. BAR-HILLEL: That such and such appears in quotes.

MR. OSWALD: It's indefinite to our context. You can put anything in there that you like. The frequency of the word "sanitube," for example, is simply an example of the veneral disease rate of these speakers. (Laughter)

MR. BAR-HILLEL: I would only try to say that from certain things which would seem not to the point, these might diminish in importance to what you are going to say.

MR. BULL: Let me throw that back at Vic. If you put it in there, and you put it in there, you're not making sense.

MR. OSWALD: These are sort of consonants, as it were.

MR. BAR-HILLEL: On quotations one has to be careful because we alter the speech.

MR. OSWALD: It's the kind of thing which is either indefinite or which the machine would produce as "tilt."

MR. BAR-HILLEL: If you find quotations; then whatever is in the quotation, it will tell that "red" in quotes has three letters. If you translate it into French, it would come out "rouge" and so on. In this case this falsifies the whole situation.

MR. OSWALD: If we get aspirin, let's say that we don't get headache, but we have aspirin.

MR. BULL: O.K., I'm sorry that the English language doesn't communicate as well. Let me point out that that's fine in this particular thing, but if I say the frequency of the word "house" or the frequency of the word "metaphysics," you no longer have a physical stimulus that

is the determining frequency. In other words, how many times I want to say it--it is not a part of my natural activity and consequently is not determined by something outside of me. It is determined subjectively. So if I happen to be a neurotic or I'm a physicist, and I become psychopathic, and I write something, the distribution of metaphysics in my particular piece of writing may be totally out of proportion to what it should be. That is one of the ways you can determine psychopathics.

Consequently we have two problems here. One is that there is a natural factor in our environment and our biology that determines certain frequencies. This is a subjective factor. On the one I give Vic here, you can have blank-blank, and you come out all right. But if I were talking about the other one instead of this one, you don't know what I'm talking about.

MR. OSWALD: Which one?

MR. BULL: Metaphysics, for example, is simply something about these speakers.

MR. OSWALD: I will not know what it is.

MR. BULL: I analyze this paper. I wasn't talking about psychopaths.

MR. REYNOLDS: Isn't the point here that in these particular sentences and phrases from which you've lifted this set of nouns, you are no longer discussing linguistics,

you are giving an example that was supposed to point up what was previously said on that basis. As long as there is a previous context; then I agree with Vic that you could leave two blanks there.

Now if you can show from some of your statements on linguistics that by taking the nouns out that you have distorted the meaning; then I will agree with you. But these selections are unfortunate in that they are not pertinent in our main argument.

MR. OSWALD: You can't take this kind of thing except from this sort of discourse.

MR. REYNOLDS: That is why I said that if he could, he would have demonstrated it.

MR. BULL: I just want to make one more point. It just is unfortunate that there are people who write like this, and picking their example from God knows where. And you simply cannot deny that factor, and so your machine is going along very happily in translating biology; and then suddenly, wham, the biology knows something about geology.

MR. BAR-HILLEL: The use of ideosemantics requires a bi-lingual...

MR. BULL: But now the major point is that the less frequent a word occurs, the more crucial its meaning may be. If a person were in here during all our talks, and he doesn't know technical terms, if the term "pre-editor"

came up once, he would be less able to understand what is
going on than if he hears the term "pre-editor" five or
seven times. Then he will somehow from the context be able
to find out, even if he doesn't understand a single word.
If it only occurred once, he might not know what it is.

MR. OSWALD: The one he is only going to hear once
is Sanitube, and that is irrevalent.

MR. BAR-HILLEL: Assume somebody writes a chemical
abstract, and you understand everything perfectly until you
come to the very minor point where the man is using a term
that is used very rarely in a lifetime. Now in a medical
dictionary you don't have it. You don't know what to do
about it because you missed one little word that is so
infrequent that you will not be able to predict it. Just
for this reason you will not be able to get it out. So even
though these words seldom occur, they might be in certain
context, and they might carry all of the information. So
that the post-editor could handle this if the machine could
translate 99.9 per cent of the thing.

MR. PERRY: I've come across something that I think
is pertinent to this argument. You'll be reading along in a
text book in Russian, and all of a sudden you'll be reading
about the concentration of matter or energy, and then if you
read a paragraph, it begins to set forth that in science
there is a conservation of energy or matter, and that that

constitutes the truth of the non-existence of God. You've hopped from chemistery over into philosophy, and this is a very simple thing. When this happens to you as a human translator, you are all right because you see what has happened. But now this happens in a sort of a concealed form, and you're reading along something, perhaps, about genetics, and it suddenly dawns on you that the words that he is using here in speaking of evolution of species and plants are based on a different scientific background, and that the Marxism business has crept in here.

MR. WIESNER: All that proves is that you have the wrong dictionary.

MR. BAR-HILLEL: It doesn't.

MR. PERRY: The only thing that I'm saying is that it isn't always easy to hit on the right dictionary. This could go all the way from a chance remark that doesn't matter, or something that is woven right into the bottom of it. If you don't know it, you will misunderstand it.

MR. WIESNER: Suppose that Jim Perry is translating from the Russian, and the man chooses to use a very rare molecule that he hasn't had the occasion to see in his lifetime, he's going to be equally confused if he doesn't look it up. All we're saying is that we have to have the machine, or he hollers for the post-editor, which we will all agree on.

MR. BULL: That's what I'm trying to establish.

MR. BAR-HILLEL: But it is the fallacy that has been popping along for years, that if we can get 95 per cent of the text; then we will be able to predict the rest. This might be correct on a great majority of the cases, but to build upon this, it might lead us wrong.

MR. OSWALD: We cannot build a machine that is going to read chemistry better than Mr. Perry can read it. We shouldn't even begin to try. We should get one that will make it easy for him so that he can spot the points and get ahold of it.

MR. BAR-HILLEL: It's Perry plus dictionary.

MR. OSWALD: We have to do the same thing.

MR. YNGVE: On this point suppose you are translating an article in chemistry, and it costs a great deal to build a dictionary that includes all the chemical terminology because you have to look them all up, you have to have a huge staff and it takes you a year; so what do you do? Translate it 80 or 90 per cent, and then you tell the machine that when it comes to a word that it doesn't have in the dictionary to put the original word in the reading?

MR. BAR-HILLEL: That's right.

MR. REIFLER: How does leaving the blank in these cases affect the meaning of what follows?

MR. BAR-HILLEL: The machine might tell us that

it's a noun; and then we can get through.

MR. WIESNER: You may wreck the translation of an entire sentence.

MR. BAR-HILLEL: Of course.

MR. OSWALD: The context will tell you whether it's relevant or irrelevant.

MR. BAR-HILLEL: What might happen is that if one word isn't clear; then you might not be able to translate the whole sentence because the machine will not be able to go ahead with the operational syntax. The syntactical category of each word should be given.

MR. REIFLER: Whether or not you know what it is? Even if you know it's a noun?

MR. BAR-HILLEL: You don't know. Whether the machine can come up with a real solution is a problem.

MR. YNGVE: A rare word cannot tell you in itself what part of speech it is.

MR. BAR-HILLEL: In some languages obviously, yes.

MR. YNGVE: If it's rare, you have to indicate the rest of the sentence.

MR. BAR-HILLEL: It depends on the language.

MR. YNGVE: All right, the affect, you can recognize the affect.

MR. WIESNER: Conceivably if you are going to examine this from the point of view of operational syntax,

you can say, "I have already discovered a noun, or I have not." If I have, it's very unlikely that it is a noun. I've already discovered a verb, and it can be given instructions as to assuming; and then you can see whether it can solve the syntactical problem.

MR. BAR-HILLEL: It might, that is true. But we might wind up with multiple syntactical resolutions.

MR. REIFLER: It's exactly the same problem for the human translator as it is for the machine.

MR. BULL: Could I summarize the point I was going to make and then go on? What I was getting at in relation to Oswald's technique of getting a microvocabulary, I was trying to establish the limitations of it. Now I know the literature in this frequency rather well. The words "coordinate," "linear," "population," "well," and "correspond," which are not portable words in this discourse as far as I can tell, do not appear in the previous literature on frequency. Consequently if we were to build up a microvocabulary going through all of the existing literature on frequency up to this moment, a machine can translate this paper, and I'm pointing out that that is hard because this paper is not a micropaper. I am able to bring in here information from a variety of fields, I am introducing many things together at once. As soon as we move in that type of discourse, we now have an entirely different problem in organizing our

vocabulary. We have a microvocabulary of linguistics, which is in this paper, but we also have a vocabulary that is free of geography, population, and linear.

The people who have done studies in word frequency have been professors in French and English literature and haven't learned about statistics.

Now the other point that I would like to make is this. As you recall yesterday, in one of my bad moods, I summarized the discussion here by saying, "What we are interested in this alimentary canal was the stomach." Now what does the machine do when it's translating a highly technical conference on machine translation, and it comes to alimentary canal and stomach? You simply could never predict that I was going to say it. I didn't know that I was going to say it, but it summarized the discussion. It was critical, and that's the problem that I'm trying to point out.

MR. YNGVE: Maybe it wasn't critical?

MR. BAR-HILLEL: It might have been. So who will decide whether it is or it is not? I'm not sure that the final reader will know whether you could predict that a certain passage that you don't understand is critical.

MR. OSWALD: I think it's the kind of thing that I think we will all agree, to get back to the basic assumption, it's the kind of thing which is increasingly rare as you proceed into the sciences like chemistry and physics and

mathematics.

    MR. BULL: I'm saying that as you expand your subject, the likelihood of this stuff creeping in increases. I'm just establishing that fact.

    MR. REIFLER: I must say that your example was badly chosen, because I think there is multiple meaning in that. We have a literal meaning in Japanese and Chinese.

    MR. BULL: Yes, but it wouldn't be in the machine.

    MR. BAR-HILLEL: Stomach is not sufficiently frequent in this.

    MR. OSWALD: Who is going to determine gastroendoitis?

    MR. WIESNER: Your general glossary will be as big as a normal dictionary, won't it?

    MR. BAR-HILLEL: Why should we? In the general dictionary there might be words that don't come up at all in a thousand times.

    MR. OSWALD: That, I think, is what Bill is saying. That there is no such thing as a general vocabulary. We can't get it.

    MR. BULL: You can only have a total vocabulary.

    MR. BAR-HILLEL: All the vocabulary on educated people? I don't know.

    MR. WIESNER: What do you mean by your general glossary? I have misunderstood you for three days then.

You've always insisted that your microsemantics would be
supplemented by a general glossary.

MR. BAR-HILLEL: But in spite of general...

MR. OSWALD: I'm talking about a technical glos-
sary and a non-technical glossary. The non-technical is
also in a certain sense a technical glossary. It's the
words that the guys keep using to express themselves.

MR. WIESNER: Your non-technical glossary would
not be a general glossary?

MR. OSWALD: That covers 90 per cent of the run-
ning items. Stomach is in the other 10 per cent.

MR. WIESNER: Suppose that your general glossary
was larger than apparently you predicted, that it was a
general glossary itself rather than a glossary that was
associated with your technical microglossary, and your
machine was instructed to always regard the microglossary
first; and then if it doesn't find the thing you are looking
for, let it go to the general glossary. What difficulties
will you get into by having a larger general glossary?

MR. OSWALD: Storage of excess type?

MR. WIESNER: No, semantic trouble.

MR. OSWALD: No, no.

MR. BULL: What I'm trying to say, however, is,
if you will let me get up on the blackboard for a minute.

MR. OSWALD: Presumably the machine can give it

to you. I'm not making such a problem. I'm trying to untangle it.

MR. BULL: Let this up here represent the total potential vocabulary, and it is presumably infinite because you can keep adding to it forever and ever and ever. Everyday we invent new words or compound new words, and we go on and on and on. So we have to deal with this as though it were never coming to an end. Whether or not it is true depends on whether earth is truely infinite--but let's not argue that.

Now on the opposite of this we will have what we're calling "a microsubject" and "a microvocabulary." In other words, from here on down we will agree particularly pertains to one special field of activity. It can be the anatomy of the human body, or the anatomy of an elephant, or it can be brain surgery, or the construction of a screwdriver; but we know from the activity that it is not making love to a blond in the park, in that sense.

Now we move up to a group of less technical words which tend to cluster around these in the position of any special activity. This is what Oswald is calling the non-technical vocabulary in brain surgery. It belongs more to the total language than this down here, but it has very special meanings attached to it in association with this. So if we say, "Cut the head." in surgery, we're not meaning what Mr. Booth

says to his assistant when he says, "Disconnect the power to the head of the drum."

MR. BOOTH: I wouldn't say such a thing.

MR. WIESNER: Why do you distinguish between these pairs of microglossaries?

MR. OSWALD: The bottom first.

MR. BAR-HILLEL: No rigid distinction?

MR. OSWALD: This is the kind of stuff which is so specificly brain surgery that it can't very well be anything else. You take a word that is used in brain surgery. It has a tremendously high frequency rate as long as you talk about brain surgery, but I don't predict that this would be a high frequency word in any other way.

MR. WIESNER: I don't think it matters whether you have to have these or one.

MR. BAR-HILLEL: What is the degree of reaction if you reduce four to three? I would not put it. This is a pragmatic question, when it pays to do it.

MR. WIESNER: I wasn't proposing that either Reynolds or I are capable of making the machine today which would use the large vocabulary or the other. I was trying to arrive at why you are ingnoring it. Whether it was a logical problem that you couldn't deal with, or whether you were worrying about other problems?

MR. OSWALD: This takes care of a large part of the

discourse, and since the other stuff is totally unimportant...

MR. WIESNER: You're talking about excess time. You're talking about a very infrequent thing, so maybe you are prepared to stall and have the long excess time to go back to the auxiliary memory?

MR. OSWALD: If you can determine the non-critical words--maybe you don't care. This is where we have to let the human being take over.

MR. WIESNER: Still it's a lot faster than if the post-editor has to do it.

MR. BULL: Sure it is.

MR. BAR-HILLEL: I think the source of the general glossary should be... The technicals are so obvious that it's not even necessary to discuss it.

MR. BULL: I can't probably get through the remainder of the material in five minutes. I've got ten more slides to show you.

MR. BAR-HILLEL: We will have to skip it.

MR. BULL: I'll go back to the rest of it.

(from prepared speech) "What I was trying to get at is that the microvocabulary appears feasible..."

MR. BAR-HILLEL: Do you have to split up into high frequency or low frequency?

MR. BULL: In this vocabulary on brain surgery the word "brain" appears with very high frequency, and that tells

us that we're talking about the brain. That does not tell us what we're saying about the brain, and what we're saying about the brain appears in a much lower frequency even in his own counts. That's all I'm trying to get across.

(from prepared speech) "The middle parabola..."

MR. BAR-HILLEL: We are in a bad position. We have only. thirty minutes left for today's morning session. We have to eat lunch today in here, and we have, therefore, to leave the room at 11:30.

... Mr. Bull then finished his talk without any further discussion on it ...

... Mr. Huskey then gave his speech ...

MR. HUSKEY: First of all to clear up any misunderstanding, maybe you know this anyway, but I would like to tell you anyhow. I am not a linguist nor an engineer. In fact, I am a mathematician, and this brings up a question of why I am here. Well, that's maybe a question that I had better not answer.

This discussion of the basic operation that are necessary in machine translation, the first thing that I would like to start out with is the idea of what sort of machines are in existence at the present time. Then we can look into the question of where do you go from here to accomplish what you need to get the machine translation.

Now I would like to put down this table here. (on

blackboard) In terms of memories of the memory computer, and the rate at which you can obtain information from this memory. In the figures that I'm putting here, they don't represent any particular machine that is in existence, they represent a combined amount of figures, and I think there will be a machine in a few months or in a year with all these present.

I'm talking about a head speed memory of a thousand numbers, and I'm using the term "numbers" instead of words. Whereas if I was talking to computing machine people, I would use "words" instead of "numbers."

This is a specific amount of information, so I am dealing in numbers which like this can be coordinated to represent alphabetical information or anything else. Suppose we stick to this amount, the rate of access is 460,000 numbers per second. If you want to talk about bits, you multiply by thirty or forty if you like.

Now this sort of thing is not possible in all computing machines that we're talking about here. It's the sort of access that is present in high-speed electrostatic memories or in certain methods of using   other types of memories. So this represents a high-speed memory. No machines are in operation at the present time, but more than this amount of high-speed memory, and quite a number in operation were considered with less than this.

Now in the in between area we're talking about a size
of 10,000 numbers, and the access is on the order of 125,000
numbers per second; and then finally you go up to memories
on the order of 100,000 numbers again. The access thing
here, I'll put down this figure, 0.005. These are dependent
on just how you set up the ideas. For example, if I talk
about this many numbers in this memory, you can certainly
have a second unit and a third unit. People aren't prepared
to face all the engineering difficulties of multiplying
these units or to face the cost, and the figure here is some-
thing on this order. Of course, I mean that this was an
electrostatic memory. The 10,000 word size would be a
magnetic grouping, and the larger one would be a magnetic
tape. If you happen to be in the middle of the tape, you
can reach the number in just...

MR. WIESNER: Is that number consistent with that
factor ten? I can imagine a tape which would take three
minutes, but I would think in going from a group to a tape
you would improve it.

MR. HUSKEY: This is very possible.

MR. WIESNER: I think that if you have a factor
of ten, you wouldn't find the access time changed very much.

MR. HUSKEY: I would like to take this figure with
somewhat of a grain of salt (0.005). I won't argue with you
here, but at least this gives us a pessimistic view here.

Now the question is, "What can you accomplish in machine translation with something of this sort?" And again I'm talking about things here, well, those are composite machines with these characteristics. There are machines under way here that will soon be in operation with all these features and promise somewhat more. Now the question is, "What can you accomplish in the language translation from here?" Again this is a question which we can't answer from our point of view, at least we are not quite sure what the linquist would like to do. In other words, how many words would you have to have in this category?

        MR. BAR-HILLEL: Do you intend to give the answer?

        MR. HUSKEY: I don't think the answer is known.

Now let's think about this in the following fashion. I specificly called these numbers. They are at fixed lengths, and the reason I am insisting on this is that in the present computing machines nobody has thought about having machines where the pieces of information are variable.

        MR. REYNOLDS: One of them is practically working.

        MR. WIESNER: If I understand Mr. Huskey's proposal, you are not prepared to say what you think the linguists want. Maybe at some stages we ought to converge with the linquists.

        MR. HUSKEY: They have to do a lot more studying.

        MR. BAR-HILLEL: I can give you certain requirements

or a few alternatives.

MR. HUSKEY: From my point of view, you don't know what you want.

In talking about these numbers of fixed lengths, the organization of a computer for numerical purposes at least has been generally built up on this sort of a premise. This means that if you are going to handle alphabetical information or words, you run into problems of the variety of information you want to handle. This has to be taken of by changes in the design or something of the sort. Of course, I'm talking about things that these are machines with each one of these seperate categories in operation and in satisfactory operation. As far as I know, I don't know of any machines where the numbers are of variable lengths which are in satisfactory operation.

The point is that in dealing with numerical problems, people have been satisfied by having a fixed length that they deal with. To tie this into alphabetical information, the length of these numbers probably corresponds to something like seven alphabetic characters. In terms of complete words, this represents much fewer words, generally speaking, if you are dealing with some language like German which tends to be longer due to the combination of the various groups of words.

Now so far we're talking about the internal

characteristics of this computer. The other, or the next category that is of some importance here, is the idea of what you can do in the way of input and output speeds. The real problem in the input speed is back in the preparation of the material to feed into the computer. There has been an operation at the present time for the transfer of information from a slow-speed storage to the high-speed storage at rates on the order of a thousand numbers per second. So that once you have information on a storage medium of this character; then it can be transferred into the other memories at any reasonable speed. That is, that speed is commensurable with the rate at which you can handle it there.

Now the actual putting of the information onto this, this input medium, of course, is probably a bottleneck here. Either you build a machine that reads a printed text automaticly, or somebody types on some sort of typewriting machine. That brings up a question of economics that you can hire a translator that is as fast as the person who types. We all agree that anybody can translate faster than any of these present machines can translate--at least the cost per word. So principally we are interested in an idea here that maybe in a few years we'll figure out a way to make a machine that will be able to do it better than a human being can.

Now what do you do about input? As long as somebody

has to type the information onto an input tape, it's probably cheaper for a human being to do thework. It has been worked on machines to read texts, and I don't know about it. But it's probably something that somebody should get interested in if you are going to follow throuh with this proposition.

The printing is, perhaps, in somewhat better shape because there is all the equipment that will print alphabetic information at quite a number of lines per minute. I understand that there many machines that will put out something on the order of a little less than 200 characters per minute, and there/under way machines that will do better than this. Whether they are really considered satisfactory machines for application here, I am somewhat doubtful. But I think they will come quicker than the other parts will come.

So that's the first thing in terms of dollar cost. I don't know whether these figures are worth putting down, but certainly a machine of this composite character here will cost you on the order of a half-million dollars or more. A human being is at most five thousand dollars say, so the cost ratio of a machine to a man is quite markedly different. The machine will cost you a couple of hundred thousand dollars to run per year, and the initial cost can be written off

in five years or thereabouts.

MR. WIESNER: Is it fair to equate an individual and a machine?

MR. HUSKEY: Yes. These are conditions that you are going to have to face in the future before it is worth-while.

MR. WIESNER: I don't know that you've proven that the machine is faster or slower than the human being?

MR. BAR-HILLEL: This machine is incomparably slower than the human being.

MR. HUSKEY: Suppose we go ahead here. Now let's look for a moment at just what you need to store in this machine, for each one of those words that you're going to have to consider. Now I picked out of Dr. Oswald's paper one of the bad examples, I picked this German word "ausfellerscheimungen." This has twenty characters in it, and consequently occupies quite a number of positions. If you don't store this word; then you must complicate the input operation to compensate for this. So you have to store this word; you have to store the pre-editorial information about where you are going to split it--or you don't split it--and you have to store more words in your memory. You have to store whatever corresponds to the English equivalent to this, and the point about the English equivalent is that they have to be stored some place in reasonably complete form. But

ultimately the machine has to print out the English word,
and presumably you store it some place in the memory where
it's accessible at a rate that corresponds to your printing
rate at least.

But at least it appears here that you're going to have
to use quite a number of memory positions in this sense to
store information that is required here. The other informa-
tion you have to store here beside the equivalent part of
speech that this represents, if it's a noun or pronoun or
participle or what-have-you. Because the rearrangement of
the sentence in the proper English word order or into an
English word order will depend on the parts of speech here.
So this has to be a code number in the sense of this mechani-
cal resolution. It tells you what to do about it from the
grammatical point of view. This is not much storage space,
though. Certainly the machine storage space is a problem,
and the question of whether you have to store all these
words, or whether you can chop them off at some point in
your vocabulary remains. So the point is that machines of
this character are not going to handle a great many words
in the vocabulary. So this means that you have to restrict
this at some point.

So let's think for a moment about the procedure in
language in this operation. Let's think in the terms of a
typed word clause as a starting point. The procedure that

the machine will have to use in working with this informa-
tion, and here I'm going to present it from the point of
view of Dr. Oswald's paper. That essentially you want to
rearrange this order of these words--and this is particu-
larly true in German--so that you come out with a desired
word order in English. It is the best idea I have seen so
far as to that point. You can have certain fixed word orders
in English so that you can minimize the number of word orders
in English that you desire. This is not producing an ele-
gant translation. Anyway that will lead up to another point
in a moment. So the first thing your machine has to do is
to have at least ten words that will appear and pick out from
the memory from wherever they are stored the information
about these parts of speech. So you can proceed to do the
rearrangement. Perhaps, while this picking out is being
done, you may be able to pick out meanings or you may not.
The actual selections of meaning may depend on what happens
in the rearrangement, of course.

Let's suppose that you can pick out the information
while you're going about it. Now to be able to arrive at
any sort of figures about what time is involved here, we have
to search in these various places of memory. If we say off-
hand that on the average you can store two thousand words
in a memory, I don't think this is optimistic in any sense.
You will not store very much here because this has to contain

your operating instructions, which will have to be available
at high speed. Here again it's back to the question of how
elegant a job you want to do in the rearrangement of the
information; so maybe you can store two thousand numbers
here. Then the question goes back to the word count in the
microglossary. What percentage of the terms does this cover?
Does it cover enough of it to give you the gist of the infor-
mation? At this point I think it may be fair to say that
you've got 80 per cent of the words. So offhand you can
handle eight of these words now. This is again not the
proper way to draw the conclusion.

I think we are getting out 125 words a second on this
memory, so it takes in the order of a tenth of a second to
get this information; so the other two words will have to
be forgotten and not translated--which may not be satisfac-
tory. Although again this is not a question of what you
want to do in the translation. If you want to search in
this type of a memory for them, this is going to take a
reasonably long time. Where you have to stand through the
point where you have time to run back to the middle, after
any particular search, it's going to take on the order of
minutes. If you take the 0.005 figure here, it represents
maybe a three minute run from the middle of the tape.

MR. HELMER: Human translator will have to handle
more than 80 per cent of the words that he indicates. If he

has to make a dictionary reference, the estimated time is less than twenty seconds.

MR. OSWALD: Too low.

MR. HUSKEY: It takes more than twenty seconds to find a word?

MR. OSWALD: It all depends. You mean find the word or find the meaning for it?

MR. HUSKEY: He looks at the first meaning, and this meaning reminds him of what he wants to know. At the present time the human being is going to be faster than the machine, as long as you have to depend on memories of this sort.

Now there is work going on in the country of memories that do not involve this search medium. Where you have a space selection rather than the business of running through the whole thing. This may cut down the research time for memories of this order. But this still means that there is a good deal of work to be done here. So maybe by the time everybody has figured out what he wants from the linquists' point of view, they will have better equipment.

The machine after it's picked up this information will rearrange it in more detail than it has, and then it will point out the final information here. I don't think there is any point in going into that part of it here, but I would like to propose the following idea. Out in Los Angeles we

are not proposing to do any general translation. In other words, we fully realize that this is a very big problem compared to the sort of equipment that can be built at the present. But what we're interested in is the idea of giving an approximate translation where you omit some of these infrequent words, and on occasion, of course, you may lose the meaning of full sentences so that maybe you will lose the meaning of the whole article. But in contrast to this the goal is not really a complete translation of this thing, but only a sufficient translation so that the person who is reading this article will know whether he really wants to go to the dictionary and look up these missing words. We pointed out earlier that the machine may print the words in a foreign language the words that it doesn't want to handle.

Now the point of view that I've been taking in this respect is that one application for this sort of thing is to translate scientific articles. And in this case I would like to point out that I think you could dispense with the post-editor, and the people who are interested in reading these articles will serve as their own post-editors.

So to sort of restate this goal then I would be interested in seeing experimental translations tried on the restricted fields, such as brain surgery or certain aspects of mathematics, to devise a microvocabulary to divise that information and see what sort of information you'll get out

here. And possibly you won't even try to handle all the possible grammatical constructions that you would run into. I think this is a good goal to have as a starting point. We have to do something of this sort before we can prove that it's worthwhile to really go ahead with it on a larger scale. I think I am ready to quit.

REIFLER: In a limited way we have the same situation not only in the human translation, but in human untranslated words where there are mistakes. It's mistakes originating with the author, mistakes in typing or printing.

MR. BAR-HILLEL: I think this talk was of tremendous importance for me, and maybe for others, to keep us down to earth and to see just what exactly could be done more or less at this very moment. For me certain consequences are very clear, and I would like to discuss them later. It would be very good to digest this information before talking about it. We will go and have lunch, and after lunch we will discuss this.

For me it takes sometime to imagine what this means in terms of present-day tests of possible translations, and also to compare what I would like to have in the future.

MR. OSWALD: Would you allow just one or two questions, Josh, to which the answers are either "yes" or "no" and numbers?

MR. BAR-HILLEL: All right.

MR. OSWALD: Is it in the opinion of the computer people at the present time that the capacity of the high-speed memory of the electrostactic machine is great enough to handle the syntactic rearrangement we ask for? Will the number of spaces available in the high-speed memory enable us to do, however crude, this syntactic regrouping? ·

MR. HUSKEY: It may be 50 per cent that this size memory would handle of what you propose to do in the mechanical resolution.

MR. OSWALD: It would do 50 per cent of what I proposed?

MR. HUSKEY: Within plus 50 per cent as to what you propose to do in your article on mechanical resolution.

MR. BAR-HILLEL: That high-speed memory is going to be ten thousand instead of one thousand.

MR. OSWALD: Would anyone in existence be able to handle it?

MR. BOOTH: Some, not all.

MR. OSWALD: What bothers me there is, if you think in terms of ··that little proposal paper, it doesn't propose nearly enough. Now question number two. How many of these ten thousand unit drums can you get hold of?

MR. BOOTH: As many as you want.

MR. HUSKEY: Not only does the dollar cost go up, but the maintenance cost goes up, too.

MR. OSWALD: I want to know how many magnetic drums can you give me to play with tomorrow?

MR. WIESNER: How many could you put in a room this size?

MR. OSWALD: Really, you've got them?

MR. WIESNER: They're ready to go.

MR. OSWALD: I want to know when you're going to give them to a linguist to play with? Who's got a machine with drums attached to play with?

MR. BOOTH: I have them.

MR. RICHARDS: I have them, too.

MR. OSWALD: There are machines at the present moment with a sufficiently high-speed memory in the electrostatic machine?

MR. BOOTH: I'm not sure what you want to store anyway.

MR. WIESNER: Actually I think you're really wanting to throw away the tape for a moment. Let's assume that you can afford to do that. Certainly if all you're going to get from the tape is the factor of ten; then you can certainly afford to have ten drums and not run into any maintenance problems.

MR. RICHARDS: We are going to exceed that in four or five months, and IBM will have ten machines which will exceed those by a factor of two in all departments.

MR. OSWALD: There will be machines available to do it?

MR. WIESNER: You can place an order for an IBM.

MR. OSWALD: We know that.

MR. RICHARDS: It will cost him $18,000 a month rental.

MR. OSWALD: How soon?

MR. WIESNER: Richards and Booth are sitting with a corner on the market.

MR. REYNOLDS: I would just like to make one statement here. So far as the program for this machine is concerned, all that we're talking about right now is high-speed. If you wish to program a regular line of IBM sorters, correlaters and punchers, this problem can be solved but not in a short length of time. We can set up a machine system that would translate for you. In theory I can show in programming that this is feasible.

MR. WIESNER: If you Victor or somebody else said, "I want to do the following job." and it was very interesting, IBM would probably do it for you for free.

MR. BULL: That's what we want to know.

MR. WIESNER: There's a high probability that it might happen.

MR. REYNOLDS: I would say with a high degree of probability that Jerry is correct.

MR. BAR-HILLEL: If we would like to have a general translator without a pre-editor, we would need not two and not five, but a thousand to five thousand times as much combined speed and storage. But I will give you the exact figures later.

... At this point the Friday morning session came to an end ...

THURSDAY AFTERNOON SESSION

... Mr. Reifler read his prepared speech ...

. . MR. BULL: Let me see if I can give a resume of these points. First, Mr. Reifler points out that there exists a great many universals which we're going to have to isolate very soon in all problems of mechanical translation. Secondly, there exists a number of pseudo-universals which may not be too important in mechanical translation. Third, when we do set up...

MR. BAR-HILLEL: Are not important linguisticly, but important mechanically.

MR. BULL: The last point is simply that we should completely ignore the traditional analysis of language when we are thinking of comparing the two and say, "Can we find a gadget in one language which performs the same function, regardless if one is a noun and the other is a particle?" And that these correlations of gadgets exist as a necessary consequence of universality, and so that primarily what we ought to do is study more linguistics before we have a machine.

MR. BAR-HILLEL: All this is shop talk of our own topic, and we should, not only for mechanical translation, but even in general, put too much weight on the sense that the linguist will tell you, "Well, inherently we don't have

in Chinese such and such a form. This inherently I don't
like." I know you wouldn't agree to that.

MR. REIFLER: This doesn't concern us here.

MR. BULL: It tends to say that this is a beauti-
ful philosophical discussion, but it isn't getting us any-
where. It is not getting our program forward. I love it,
but it isn't getting us anywhere.

MR. HARKIN: "This" and "is" are equivalent.
"Those" and "that" are what? Is there any similar parallel-
ism?

MR. REIFLER: Both languages permit it, not only
for "this" and "is," but "those" and "that."

MR. HARKIN: What's the equivalent of "that-those?"

MR. REIFLER: The Chinese word for "that" and the
Chinese word for "those," I'm sorry, this is not a good
example.

MR. OSWALD: You don't mean seriously that the
English "is" is in the framework of English?

MR. REIFLER: Historically that goes back to early
European times. It makes us open-minded.

MR. OSWALD: The third singular qualifier is the
same thing as the demonstrative?

MR. BAR-HILLEL: Not "is," "was."

MR. REIFLER: Now in English it just happens that
only "is" has survived there, but in German it is there, and

in Latin it is right through.

MR. HARKIN: I still don't have the answer to my question.

MR. REIFLER: We cannot do it in this way that we say, "All right, now we have just found out that there is something between 'this' and 'is.'" Let us say, "What is the difference between 'that' and something else?" You just make a very broad research of our problems, and here and there you find something which happens to be of value for us.

MR. HARKIN: What does the Chinese say for "that" as distinguished for "this?"

MR. REIFLER: I want to point out that the Chinese have quite a variety of words meaning "this," and quite a variety meaning "that." And I have to scan each one of those. In this particular case it just happens that we are concerned with two types of grammatical meanings represented by the same speech. In other cases it may not be two types of grammatical meanings. In Chinese that has the following meanings: The meanings of "that," which corresponds to our "what" or "which."

MR. HARKIN: Can you tear the parts down?

MR. REIFLER: By origin, of course, we can. Well, I do not think I can do that. I will give it to you private-ly afterwards.

MR. REYNOLDS: How many more universals exist to

your present knowledge, and is it necessary to prove these
in the sense of comparing between highly numerous languages
for the linguist's benefit, or do enough of them exist that
the engineer could immediately make use of them without
regard of whether this linguistically is of value or not?

MR. REIFLER: I can only answer if you give me a
hundred specialized linguists for a hundred years. But as
far as the Chinese language and English, or Chinese or Ger-
man and Chinese and French is concerned, I have already quite
a number of these four. We can already make it of practical
value.

MR. BAR-HILLEL: I would say they have no prac-
tical importance. And it should never pay to have a special
rule to point it out.

MR. REIFLER: There is no special rule. We drop
a rule there.

MR. BAR-HILLEL: But it would mislead to translate,
"This book." instead of having, "This book is his book."

MR. REIFLER: You misunderstood that. We've
reduced the number of instructions.

MR. BULL: I would make a suggestion, that now
that Mr. Forrester is here, I think we should let him talk
and make him discuss his business and Huskey's together.

MR. FORRESTER: I have not prepared a paper. I
seem to be out of order here in that regard, and I would

rather find out what kind of questions you would have in mind, and would like to ask, that might be in my area of interest. Is there some way that he can give me a few clues?

MR. BULL: I think I can summarize several questions that you want to know. How much stuff can be stored in the various parts of the machine?

MR. YNGVE: What is the relative size of the high-speed to the low-speed?

MR. BULL: Let's run through the one that we've got.

MR. BAR-HILLEL: It's not a dream.

MR. WIESNER: Yesterday or the day before Dr. Bar-Hillel proposed a logical syntactical analysis which the machine would have to perform. It might be interesting to have him speak on it for a couple of minutes, and then you could take off from there.

MR. BAR-HILLEL: I can't tell what this means in type of memory.

MR. WIESNER: There are two problems, one is memory and the other is the machines's ability to do logical operations.

MR. BAR-HILLEL: I see the following problem at the moment. No pre-editor, no reduction to micro or ideoglossary, or something that can translate everything without pre-editing the source language, which I take to be German.

The maximum goal--no pre-editor, but a post-editor, and the
source language is German, and the target language doesn't
matter. This is the maximum goal, and I don't think we
shall have more than that.

Now this requires the following situation. We have
allowed 450,000 lexical units. Each lexical unit in German
has to be multiplied by something which I don't know--no
linguist apparently knows what is the multiplicity of the
derivatives of the lexical units in German. I can give you
my own estimate which will be around three for German.
Through this process of derivation a few will coincide. In
German "leben" can be either plural or singular.

We come to about one to one and a half million running
words. Therefore, to be able to do this, what I call "opera-
tion syntax," we need a word index of so many entries of one
to one and a half million entries, which will tell with
respect to each entry or frequency to the German word "leben"
that it can be either this or that or that.

        MR. BULL: You're getting too detailed.

        MR. OSWALD: No, he's not.

        MR. BAR-HILLEL: We would have up to two-tenths of
a second to get each entry.

        MR. YNGVE: How many average letters per word, if
I understand this one to one and a half million words?

        MR. BAR-HILLEL: If you don't take into account

savings, twenty bits per word, and the access time, as I
see it at this moment, is about ten minutes.

This is only the first stage; then we need a unit which
will do certain operations upon whatever is entered here. I
do not see any great difficulties because here we will have
not more than a few hundreds of different entries on the
side. This will be coded syntactical functions, and I don't
believe you will have more than a few hundred of them. The
operation will be nothing more than certain matching opera-
tions. As soon as this is done the outcome can be done.

Now the next step would be a bi-lingual dictionary.
Here we need only a bi-lingual dictionary about this size
with 450,000 words. Here I tend to believe that a second
or two might do. So both a smaller capacity and less. This
is a bad thing to do because here I need a higher capacity
and higher access time.

MR. FORRESTER: Can you support for me your state-
ment of two-tenths of a second here and one second on your
dictionary access? But I can't imagine you are using a
machine of this sort and turning out a project that is at
all competitive with a human translation, with speeds of that
sort.

MR. BAR-HILLEL: I believe you won't arrive...

MR. WIESNER: Before you can turn out a word a
second, it's too slow.

MR. FORRESTER: It's much less expensive to teach language.

MR. BAR-HILLEL: It's too slow for translation?   I don't think so.

MR. RICHARDS: One thing was that in the first place this is a research operation.   They can't get translators at all.

MR. WIESNER: Another point that was put out, it appears to me at the moment that MT is so much more costly in terms of what we're talking now than a human translator.

MR. FORRESTER: You can get human translators if you make it a financially successful profession to the extent that this will be.

MR. BAR-HILLEL: I think you are mistaken.

MR. WIESNER: I think there is no point in arguing this economic thing.   We have a limited amount of time.

MR. HUSKEY: Dostert told me last night that there is a satisfactory supply of human translators.   That there are plenty of people on the Civil Service lists--there is no real shortage.

MR. BULL: I think there was an error in the notion of human translators.   When they set up the simultaneous translators in the U. N., he had three weeks to get translators--they were not available.   So then you come to the next stage, can we get human translators?

MR. WIESNER: You give me what it costs to build the machine and to maintain it, and I can train hundreds of human translators.

MR. BULL: Let's discuss it academically as a research project.

MR. BAR-HILLEL: How do you think a human translator can translate more than a second per word? When I hold here two-tenths of a second, this is the maximum which we will allow the machine to produce at half the rate of a human translator, because I need additional time for the human post-editor. Now I believe it is a good estimate, because I believe the human translator translates a page of a book in at least five minutes.

MR. WIESNER: I don't think you should worry about what a human translator does here.

MR. FORRESTER: We can't divorce this discussion from economic things. You can lay down what you would like to do.

MR. WIESNER: We've discussed this before. I don't think we should rehash the whole thing.

MR. BAR-HILLEL: The minimum program, after a certain type of pre-editing, would require about five thousand lexical units. The remaining twenty thousand, which will occur much less frequently, in one second. This can be even more than that because, if on the page we won't find

more than two or three of these; then it really doesn't matter if the time would be even more than that. So that this would require a good pre-editing which would take probably three-quarters, not much less than the time that a good translator would do this translation. So in the whole output the gain is still very considerable.

MR. REIFLER: This is a gross overstatement.

MR. BULL: Let's not argue that.

MR. FORRESTER: I think that this points out that the real problem in the machine is the storage and the memory function. The computic speed part of the operation I don't think is serious if you have access to the right sort of information. This is why you immediately come up against a distinction, a difference, a qualitative difference, I think, in what the machine does and what the human mind does. There's all this talk about the mechanical brain, but you look at the present machines--the ones that are now being developed or the ones that exist--and you find that they are not well adapted to several things that the eye and the mind apparently does very well. They are not well adapted to the recognition of shapes, which is aside from our present discussion, and they are not well adapted to the extraction of information on the basis of the kinds of rules, the kinds of information, that you want in the process of translation. That is the correlation, the pulling

out of memory, on the basis of a stimulus that you get from
a word. You don't have a scanning process in your mind--
not that you can put a conscious finger on it in any way--
you have an immediate correlation, a relationship between
the input and what you pick from your memory.that thus far
doesn't exist to anyway near the same degree in the machine
that you have available. In the machine you are rather forced
to arrange your dictionary in what I might call a linear way.
You arrange them alphabetically or some memory system, and
you arrange them in some other system, and this is not the
way that we appear to get things when we read one line and
think in terms of a second language. It is a fact, I think,
that the human translation is probably as fast as the high-
est speed machines that we now have doing these extractions
from storage. Especially if you had some of the more subtle
distinctions, the context of the paper, the kinds of things
that go into a good translation. Let me give you some
numerical...

MR. BULL: You pick up somewhere between five
hundred and a thousand pieces of information per second when
you're listening to me talk right now.

MR. WIESNER: That depends on how you measure that.
You could, by one deviation, if you accept a sufficiently
large dictionary and base your measure of information on the
statistics of three words a second out of a dictionary of a

hundred thousand, you probably come out with that kind of a number.

MR. BULL: By the time you did the analysis of what I'm saying and got the information out of it, you would be making a selection between that many items per second.

MR. WIESNER: There are many ways to measure this. There are physiological ways.

MR. FORRESTER: The kind of storage that you're going to use, even limiting ourselves to this linear array idea, the kind of storage, let's review the qualitatively different things there. The kinds of storage that we have available for machine use now, I'll start with the highest speed ones and work down. We have in machines today a low capacity small number of bits. Low capacity being in the order of ten thousand, in tens of thousands-put it that way. Then you probably take the next step down to magnetic drum equipment, which runs the half million level. Incidentally at a cost of $50,000, perhaps, a million. Those of access types are still faster than you're talking there of a fiftieth of a second.

MR. WIESNER: I think there is a mistake on the two-tenths of a second.

MR. FORRESTER: You will use in the machine a procedure that's at least as good as your scanning dictionary to find a word, but even on that basis it's not awfully fast.

It's faster than the two-tenths of a second.

Let us say that we are talking about storing this on the equivalent of what we now have in the way of magnetic drums. I think when you still go to the still lower cost units like tape, you will probably get even too slow to meet even these modest requirements for speed. So you're talking about magnetic drums when you're talking about present day equipment. Magnetic drums are in the order of a half-million digits. So when you start with that and existing machines, you're talking about making experiments. Existing machines will have that order of storage, the factor of two or four, until you start putting in multiple units.

Now if you want to store a million and a half...

MR. BAR-HILLEL: Two hundred and fifty drums.

MR. FORRESTER: That first problem came out to be a hundred million bits, which is two million drums, which is close to a thousand vacuum tubes apiece, and this is not...

MR. REIFLER: You can get those drums from England through the courtesy of Dr. Booth.

MR. FORRESTER: There has been on the other hand, let's look at what's been happening to computing machines in the last few years. There's been an improvement in performance of machines that's very close to one decimal order per year. This is a very rapid rate of improvement. And if you're talking about twenty years hense, I don't think we

can interpolate it for three I am sure. What happens after
that I don't know.

What I'm doing when I say that we've had a performance
improvement, I'm taking a product, a storage capacity and
speed, and cost, and putting it in the right top or bottom
side of the fraction to be indicative of what we need. Cost
is the denominator, and that fraction has improved about a
factor of ten per year for the years since the war. That
has been done by improving speed. Mostly the biggest factor
has been the improvement of speed. There has been some
improvement in storage capacity, but in the last two or
three years there hasn't been much change in the high-speed
internal storage.

In other words, the machines have not been getting
simpler nor much more complicated. The next three factors
of ten should come by a little more improvement in speed,
not a lot, by lower cost in complexity, and by considerable
improvement in reliability--which may be the first order of
importance here. So that actually what we see in the next
three years is not going to change the picture a lot--not a
lot when we're talking about a discrepancy in several orders
of what we have and what we'd like to have.

Now I think you're interested in the cost per digit of
storage.

MR. WIESNER: I think it's unfortunate for you to

concentrate on that higher set of numbers that **Dr.** Bar-Hillel has put out, because many people here believe that we're somewhat closer to the lower limit. I think you can do a better job on that. You might say that one or two or three might do some kind of a job. One is too low.

MR. FORRESTER: Magnetic drum storage costs around fifteen cents a digit a bit. You can vary these by some small factor one way or the other depending on who you talk to. It would cost about $1.75 per unit. The electrostatic tube would be $8 per digit and would come out to here around 160. Then, you see, we've had a trend here where this has been going up, and that's been going down. The unfortunate thing is that it's got to where it's priced the product out of the market.

MR. BAR-HILLEL: You cannot utilize the performance of the electronic tube? It does such a high capacity of work, and you don't use it?

MR. FORRESTER: It does not follow in most applications, and it does not follow here.

MR. WIESNER: You can't afford to use flip-flops.

MR. BAR-HILLEL: Why?

MR. FORRESTER: This is it. It comes out about the same here. As a matter of fact, most of the storage tubes are somewhere near this range. The vacuum tube flip-flop was one of the first, and the cost per digit on that is

going to be excessive. It's not even considered at the present time in storage of large quantitites because it's \$50 each. So I would say that it isn't even to be considered except for very small amounts of high-speed storage that is used in, say, the computing process within the machine as temporary storage running only a few words of storage in a machine. I'm only talking up to ten thousand digits. When you get to the hundreds of thousands, you have this available.

Magnetic arrays are coming up. They are in the research stage in many places, and I think long about 1956, you will find that we are getting figures here to the order of a dollar. In other words, we start back down in cost per digit. We will still be going up or down improving in the cost per unit performance. In other words, we will be getting better performance and lower cost.

If you want to know how much the machine costs, this is the figure. This column is the thing that counts if you want to know what you are getting for your money. The only thing that I've put on the board that is at all of interest in storing this large amount of information is that first unit.

There are people who have other proposals for ways of storing large amounts of information--puting it on film, on magnetic cards, and things. I haven't given it very much

thought. I don't think they are very close from an engineering standpoint or very useful.

MR. WIESNER: You gave a figure of $5,000 for a half-million bits.

MR. FORRESTER: For the drum.

MR. HUSKEY: With respect to the maximum requirement.. This is a two-tenths average, and it would end up with one second or two. You are getting a very different picture.

MR. FORRESTER: It would be something like .005.

MR. BAR-HILLEL: If this one and a half million, we could store 1,400,000 on the tape and the remainder, 100,000, since this 1,400,000 will be consulted so rarely, it's still there.

MR. FORRESTER: There is a possibility if you divide your vocabulary. If this minimum vocabulary still has the things which are very infrequently consulted, there is a possibility. But you'll have to bear in mind that this magnetic tape units give you twenty inches per second, talking about thousand foot reels, so that when you do have to consult it, the access time will run to a few minutes.

MR. BULL: About 20 per cent of your volume would have to be consulted.

MR. YNGVE: If you go through with a single head, or let's say six heads, so that you have the words letter by

letter, and if you have a coincidence type arrangement to pick out, and you always start from one end; then if you give an over-all frequency law, such as frequency times rank as a constant; then you divide the vocabulary by the logarithm vocabulary and it's about one-tenth of the way down.

MR. FORRESTER: On that basis you are really going to search.

MR. YNGVE: You can combine that with any other gimmick, too, you see. But you search one-tenth, and then, perhaps, a fast rewind to the beginning.

MR. BAR-HILLEL: Any combination of using different types of memories and a certain ingenuity in ordering them should be able to reduce the things highly.

MR. HUSKEY: Say it takes three minutes for research on this tape. Well, this means that in the three minutes you'll be translating something over two hundred words, so you can only afford one-two hundreds.

MR. BAR-HILLEL: Which is reasonable.

MR. WIESNER: He's putting more in the drums.

MR. HUSKEY: My answer would be to forget the words once in two hundred. Just throw it out.

MR. FORRESTER: I have a couple of slides here that you might like to see, though they don't help the discussion very much.

MR. HELMER: For some proposals two magnetic drums

seem to be all right. I don't see why you say it is out of competition.

MR. FORRESTER: I think we are talking about two different things. I'm talking about doing a good translation, and you are talking about doing a minimum word by word.

MR. WIESNER: We haven't defined what we mean by doing a poor job. It doesn't mean that you are doing a poorer job, it means that your machine is not a general purpose or all-language machine. It's not even a two-language machine without changing its store of information, which might not be too hard to do.

MR. FORRESTER: Don't misunderstand me. As far as being able to do experiments in mechanical translation, the present day machines are all right. In other words, if you cut the problem to fit the available machines, you will probably still have a very significant problem left on your hands. In other words, with present machines you can still set yourself the major task of deciding how a machine ought to do the translation.

MR. WIESNER: I think Dr. Bar-Hillel would go a step further. Somebody said this morning that for $18,000 a month you can rent an IBM machine, which wouldn't be too much poorer, maybe a little better than you would need for this minimum job. So that conceivably, if you were in a situation where you could get the kind of situation that you

need, you could do this. You would have to look very serious-
ly, because for $18,000 a month...

MR. BAR-HILLEL: The cost will be a thousand times
less in three years.

MR. WIESNER: This includes reliability, which
may be the biggest single factor, and storage capacity.
This isn't the situation you have in the transmission of
information. These aren't complete variables.

MR. FORRESTER: The machine is still going to cost
what it cost today, but it's going to cost the same.

MR. BULL: Are we ready at this point to have
Mr. Booth go ahead and put in his "two bits?"

MR. BOOTH: I disagree very fundamentally with
both the computing people who talked today--both in cost,
which may conceivably be a national situation, and possibly
on basis of what you exactly want in a machine. I do this
without any great apology for the simple reason that we have
used the machine, and we also have done the same problem on
the basis of using punch cards equipment. So that we know
it can be done.

I also know what the particular equipment costs, and it
just doesn't compare at all with some of the figures we've
listened to today. We want to consider the cost as a func-
tion of the problem we want to solve. What we want to solve
is the problem of mechanical translation. It's not the

problem of doing some very fast and extremely complicated compilations, it is the simple problem of doing mechanical translation. I maintain that you can do this problem at a far better figure than that which occurs on the left. Well, I've written some figures down here which are quite independent of the ones which Mr. Forrester has brought up. Bits work out at about fifty cents apiece. I'm talking about people who have made these things in England.

MR. FORRESTER: These are all what you would consider a fair commercial price?

MR. BOOTH: They are intended to be. These are mass produced units.

MR. FORRESTER: The attempt here is to make them comparable to one another based on prices that have actually been paid now in this country. I mean the eight dollars, I take anything from four to twelve as being the same thing.

MR. WIESNER: In almost any electrostatic system the tube itself is an incidental part of the cost, compared to the inputs and outputs.

MR. FORRESTER: These are five year figures, and a lot of these tubes don't last very long.

MR. YNGVE: How much equipment is there for fifty cents a bit.

MR. BOOTH: This is the tube. I think someone said that mechanical translation at this stage is a research

project.

MR. WIESNER: I think any serious attempt to build a mechanical translation machine would involve doing it on a commercial quality standards.

MR. BOOTH: I can only say about the British university machines. In fact, I don't think I am incorrect in the statement that the first machine which worked anywhere was the machine at the University of Cambridge. Coming to the question of drums...

MR. FORRESTER: The daliason is more likely to be used. It has lower cost and lower access time.

MR. BOOTH: The figure which applies to the cost of magnetic drums in England is the sort of thing which one would want in the contents of mechanical translation. I consider that a supply ought to be about 250 bits long. This would mean that you could get fifty words of ten alphabetic letters in it. I'm not trying to have variable lengths at this stage, but I'm supposing that an average dictionary would amount to...

MR. HARKIN: Is that a constant length for each word?

MR. BOOTH: No, no.

MR. RICHARDS: You mean fifty characters, don't you?

MR. BOOTH: Yes.

MR. HUSKEY: Two hundred and fifty bits?

MR. BOOTH: Yes, on the basis of the drum unit.
It is now available commercially. I believe essentially
that such units are available. The cost will be for your
entry of 250 bits, ten alphabetic letter words, a total cost
of seventy cents per entry. This is converting the pound
rate of production cost into the dollar standard rate of
exchange.

MR. FORRESTER: What was the explanation there?

MR. RICHARDS: His drum is cheap.

MR. FORRESTER: I didn't follow your arithmetic.

MR. BAR-HILLEL: Two hundred and fifty bits for
seventy cents?

MR. WIESNER: You get four bits for the price of
250.

MR. BOOTH: The drum is priced at about $1,200.
Computing machines are about $2,800. You do not have to
have an addition to your drums, you can extend the lengths
of the drums as much as you like. In principle you can
extend the length of the drum. In so doing you are required
to supply certain connecting units to hitch the drum to the
computer, if you use a computer for doing it.

Now there are two general ways of connecting drums on
to the machines. One is expensive, and that is electronic.
The other is a coding switch, which may operate in a

milli-second.

      MR. FORRESTER: You're changing the rules on me.
We started by discussing machines of present-day types.

      MR. BOOTH: This is the way our machine works.
You can't say that just because your machine is of present-
day type that mine isn't. (Laughter)

      MR. WIESNER: That's one way of keeping score.

      MR. BOOTH: Having been attacked, I had to answer.

      MR. HARKIN: Your machine was designed first,
therefore, it's ancient history. (Laughter)

      MR. BOOTH: When I was here in 1937, Mr. Forrester
had part of a machine started, and I had nothing at all in
the laboratory. Now it works out that the cost of fitting
into your drum doesn't involve the amplifier and switching
maker to go with it. The cost works out on the scale of
which I want to make these estimates, at about twenty-five
cents and on to seventy cents, which I want for 250 bits.
So I round this figure out and call it one dollar.

      This is still reasonably cheap, I feel, but this is not
a high-speed sort--it won't do things in twenty milli-seconds.
This is again not the whole answer to the problem. One-
fiftieth of a second is the rate at which we always run and
get out an entry. So accepting that speed limitation, we
think we can do wonderfully for one entry.

      Now engineering wise at the moment these things,

including the computing service, can be made at that particular price. I think that the other figure which I worked out here, which I think Dr. Bar-Hillel knows, about the 100,000 entries.

MR. BAR-HILLEL: This is something in between.

MR. BOOTH: You in fact get 1,000 entries for four inches in every drum length. It works out then that you want ten drums forty inches long and twenty inches in diameter. Now these figures are not the size of the drums which we use, but engineering wise I think the twenty-inch drum would be feasible, forty inches long. But I don't think the five-inch drum would be. I'm not even sure that the twenty-inch drum is.

So you want ten drums forty inches by twenty inches with your switching equipment, which is small compared to the size of the drum, and it would be at the rate of a dollar per entry. So your cost for 100,000 words would be $100,000, excluding the computer service. We've hitched the device, with very slight modifications, to the Apex type computer, and the cost of an Apex type computer in England today is $5,000.

MR. WIESNER: This is still a word for word translation?

MR. BOOTH: Yes, using a microglossary.

MR. BAR-HILLEL: This is done by the machine itself?

This splitting up is done by the machine? I mean the split-
ting up to find out what is the stem which can be done by a
human pre-editor can be prepared on the drum itself? It is
in this case done by the machine, so the machine does the
splitting up?

MR. BOOTH: That's right. You get the input from
a Teletype. When the word gets into the machine it is
stored in a register, and it sends out what is the initial
stem which is found in the dictionary. It may be that the
word is in such a place, it may be, however, that you can
only do part of it. In which case you can patch the stem
and store the extreme translation, and then go on and find
out the tense and so on of the verb from a consideration of
the endings--again by comparison. So you really have your
microglossary with a micro-ending, which I don't think would
be micro at all. This would be a complete language.

MR. HUSKEY: You're not rearranging the word order?

MR. BOOTH: We are not supposing to rearrange the
word order.

MR. BAR-HILLEL: Have you seen his paper?

MR. HUSKEY: No.

MR. HARKIN: How does the machine recognize the
root stem?

MR. BOOTH: By the dictionary, as far as it will
match.

MR. BAR-HILLEL: May I ask for how much per month would you rent such a machine?

MR. BOOTH: I couldn't tell you. I have never rented a machine. I can tell you that this particular type of machine is going to be marketed. I think the cost per year is likely to be of the order of a thousand pounds per year.

MR. RICHARDS: Roughly half the cost.

MR. BOOTH: Don't tell them that I said this.

MR. WIESNER: You're guessing that this machine, well, let me stretch what you said a little bit. You said your storing is going to cost $100,000. Let's say the total machine might cost $200,000 in England. This would mean a million dollar machine. I don't know whether this is right, but let's assume it for a minute and see what this means to Harkin.

MR. RICHARDS: To get your monthly cost in this country, I would say you divide the cost in this country by somewhere between ten and forty.

MR. BOOTH: That would be one hundred thousand bucks a month.

MR. RICHARDS: A million dollars would be $25,000 per month.

MR. BOOTH: This doesn't imply any criticism of impending machines because ours is very slow.

MR. BAR-HILLEL: To what degree is it expected that the rental cost of similar machines will go down?

MR. REYNOLDS: They don't exist yet, I can't answer your question.

MR. BAR-HILLEL: How does the rental go down, assuming that this is the same machine?

MR. RICHARDS: They never go down. By the time they are where they could reduce them, they build a better machine.

MR. REYNOLDS: The IBM prices are actually fixed. They never go up, and they don't go down. The Company policy is to establish, just as Forrester explained a moment ago, an estimated total cost of production amortization, and this is a fixed rate. Performance goes up with successive models--at least we hope so.

MR. BOOTH: That's all I've got to say about it.

MR. BULL: Now I think we ought to go back now and let Mr. Huskey have the floor and let him say anything that he feels appropriate at the moment.

MR. HUSKEY: I don't think there is very much to say. I think the figures that are on here are pretty much in line to what I said before. I tried to summarize this by saying that I think we could do something in the way of a modest job of translating on some present-day equipment, or equipment that will be ready in the next few months, by

restricting goals--which would still be a satisfactory result for specific purposes such as I mentioned. In other words, being able to give a person who is a specialist in the field enough information about a particular article to know whether he wants to follow it up or not. As far as producing elegant translations, I think we have a long ways to go.

The mechanical equipment has to be improved. Not so much just on the sort of theoretical basis, because I think we have all the principles in hand here that you could go ahead here and put together a machine that would do a reasonably elegant situation; except that there is no limit to the cost to do this, and there may be a reliability problem. Although we can expect the reliability to go up later. It is possible.

MR. BULL: Would you like to make an estimate of about when you think that would be?

MR. HUSKEY: Somebody got a hat? (Laughter)

MR. WIESNER: Ten years is always a safe answer.

MR. HUSKEY: I would say five years from the time you asked the question. (Laughter)

MR. RICHARDS: Very good.

MR. BULL: There's a serious sngle there because I don't believe the other half that you're talking about, the linguistic solution of the problem, can be completed for

awhile. We could do it for the micro situation, but if we then get up to the elegant situation, I don't think that can be done for any single language under about ten years--unless you want to throw in an unlimited amount of money. But if it goes on the way it is now; then ten years is probably an optimistic estimate.

MR. WIESNER: How many years can you go on working without some crude machine to work with?

MR. BULL: At least one full stage of something within the next year; then I think we ought to go back to...

MR. RICHARDS: I'd like to use the available machines to do that sort of work. I'd like to compile the statistics on our present machines.

MR. BAR-HILLEL: It might be worse not to use them, but to use punch card machines, so the translation of one sentence will take an hour.

MR. WIESNER: You've got a long job ahead of you before you do that because we have got to get the materials ready.

MR. BAR-HILLEL: I believe you can close this discussion.

MR. BULL: I don't think we've got much more to say now.

MR. LOCKE: Why couldn't ten post-editors be translating ten pages of the output, if you had ten times the

output speed?

MR. HUSKEY: Any machine that we are visualizing at this point is no better than a human being, and as long as this is true, there is no point in multiplying the post-editors. This is something that I would like to get clarified, because the one person that I feel may know more about it than I do has said that there is no shortage of translators.

MR. RICHARDS: Now listen, we have a tremendous group of intelligent people trying to get hold of Russian documents, and every single one of them goes down to Oak Ridge. At present there is no shortage of money.

MR. BULL: I think they have somebody read the language and extract.

MR. WIESNER: That stuff that came from Germany was walled up at Wright Field.

MR. BULL: As long as we're talking about it, we don't care what this machine costs. There is just the matter that they are paid $4,500 a year, and there is a shortage of people who will work for that money.

MR. WIESNER: Then, of course, we have this long range situation where one might consider translation as a low-grade performance by human beings.

MR. BULL: There are other low-grade operations, such as cleaning out toilets.

MR. RICHARDS: There is going to be a big revolu-
tion in the office. The thing that we can really see coming
in the next ten years is the eliminating of this human book-
keeping. What I say is, let's put those bookkeepers to work
being human dictionaries.

MR. OSWALD: If you had twenty translators working
for a year--and that isn't very many translators--it would
cost you $100,000 a year.

MR. WIESNER: This is just their salary.

MR. OSWALD: This is what it cost to buy something
to do it. And furthermore they are highly replaceable, and
they contribute nothing but the job done. When they go away,
they are gone. A hundred thousand dollars a year for a
lousy twenty translators. You see, I begin to think the
machine is cheap. I begin to think that your million dollar
machine is a wonderful buy. You get a permanent gadget
which will do a tremendous amount of work.

MR. HUSKEY: When after you've bought this machine,
it still costs you on the order of $100,000 a year.

MR. BULL: Of course, the question is how much it
translates.

MR. RICHARDS: I don't think that the human brain
if becoming more expensive, there is more of a premium on a
fancy brain.

MR. OSWALD: You can buy any kind of brain for

five thousand bucks.

MR. WIESNER: You don't have to change that much. Make it seven thousand, and the machine is still cheap.

MR. BULL: The best machine we have at the present moment can do the same as only one translator, and it costs you four hundred to five hundred thousand dollars a year. You've bought twenty translators for one hundred thousand dollars.

MR. BAR-HILLEL: The machine will run twenty-four hours a day.

MR. BULL: But your maintenance cost is there.

MR. OSWALD: The computer people are thinking of high-speed operations which the machine can perform so much faster than human beings. The human translator is a heck of a slow gadget.

MR. LOCKE. If he's good.

MR. OSWALD: He's very slow.

MR. BAR-HILLEL: We have not a machine which could improve through learning. May I ask if there has been a discussion at to how it gets into the machine? Do you transcribe it and put it into the machine?

MR. WIESNER: We keep talking about five thousand dollar translators and so on, and one of the bottlenecks in this translation business is bi-lingual people who are experts--this is quite a different kind of specialist. Now

mono-lingual experts in either the target language or the source language are easy to come by. Now when you get into the highly specialized problem, I think the things we've been saying about the ease you come by translators and the relative cost just breaks down. At no cost can you get the people who can translate all the scientific Russian that you want at this time.

MR. LOCKE: We have people at M. I. T. who are translating these things.

MR. BULL: Could we now reorganize our thinking on what we're after now? We don't want to solve who is better, my father or your father. We want to know at this point if we can have a machine that will do something on a modest scale? We have apperently arrived at the conclusion that if we are willing to be modest enough, the present machines will do something that will make "Life" magazine happy.

MR. BAR-HILLEL: I think you could go up to "Scientific American."

MR. BULL: Now then I don't see any point from now on to discuss human cost versus machine cost. Let's go over to the business of how you go about organizing the job. The present research should be aimed at manufacturing this very modest result.

MR. LOCKE: Dr. Booth said the computer may not be the best way to do this. Computer components would be

perhaps, the best way in translating this thing, since it's
going to take us quite awhile to prepare for the machine.
I wonder if this more desirable machine, more effective ma-
chine, might not be ready by the time we're ready?

MR. BOOTH: No, no.

MR. BULL: It's a chicken and egg deal.

MR. LOCKE: The ultimate cost of a more ideal
machine ought to be taken into consideration.

MR. BOOTH: I think we can do anything that you
want done.

MR. LOCKE: For research purposes, we ought to do
with what we've got.

MR. OSWALD: Which takes the thing to where Bill
wants to take us.

MR. BULL: The linguist cannot do the research
until you know the type of machine that it's going to
operate.

MR. BAR-HILLEL: We can only work not one system,
but four or five different systems.

MR. OSWALD: What's needed is a combination of
linguists or a couple of linguists and a couple of men so
that every time the linguist comes up with an idea, they
iron out the edges and find out what fits.

MR. RICHARDS: The machine makes the difficulties,
at least it exposes them.

MR. BAR-HILLEL: The machine will decide who is right.

MR. BOOTH: The machine does what you can do with a pencil and paper.

MR. RICHARDS: That's right.

MR. BAR-HILLEL: Has the performance of the machine itself no influence on the changing of the program?

MR. BOOTH: Not at all.

MR. RICHARDS: It merely uses ultimate parts for an already prepared program.

MR. BULL: We're in a dog fight. All I heard was that the machine wound up and went zoom. Now where were we? Does **somebody** have anything really significant that he wants the machine to iron out at this point?

MR. RICHARDS: Can I tell the gang how a digital machine works? There is a very important difference between a human being learning the inside, and the way a computing machine finds its way through a maze. The machine must have, before it starts, a memory capacity sufficient to encompass all the possible decisions it may have to make. This means at least the same number of words in storage as the number of words incoming.

MR. BULL: Can we go back to the place that we've got half a dozen of these machines? What we want to do at the present moment is to get one or two or any number that

you went to do this. Now if somebody has a proposal that
envisages no dreams, no fancy translations, which is com-
parable to what we think. In other words, if we put a bat-
tery of pre-editors on this end and a battery of post-editors
on the other end, it doesn't have much to do with the prob-
lem, it just proves what we know the machine can do. Does
it take the work away from these people? If it can't do that,
we might as well stop at this point--we're just playing
games.if we're down to that situation. What can we do now,
from the point of view of making the next real step, to have
a piece of paper come out that we can understand?

MR. BOOTH: If I say to you, "Tommorow I will go
to my computer and put your translation problem on, give me
my microglossary." Where is it?

MR. FORRESTER: If somebody comes with a very
simple problem that a dozen people can see right through and
know exactly what's to be done, they'll be six months trying
to get it down. This is a crystal clear problem and has no
relationship to translation as far as I am concerned.

MR. BAR-HILLEL: Within six months somebody will
be able to give us a machine?

MR. WIESNER: That is until you're ready to use
the machine.

MR. FORRESTER: This was six months that it would
take the fellow who had the problem to get , ready to use

the machine.

MR. BAR-HILLEL: No one person would probably be able to do it himself.

MR. WIESNER: I think you have a fundamental decision to make right at the beginning. There are two possible courses of action that you could take. On the one hand hire Bar-Hillel and large numbers of logicians. This is one way, and I think the harder way of going at the problem than if you move over to the other other side of the table and say, "Let us pick a particular field in which there is some interest in doing translation." Let's make it a small enough field so we think we are going to come up with a very restricted microglossary and make a corresponding general vocabulary to go with it, and make a word for word translation without the syntax. That, it seems to me, is a simpler problem.

MR. BAR-HILLEL: There's no reason why we shouldn't do both if we had enough money and manpower.

MR. WIESNER: If you do both, you're just deciding to go down two courses. I say you have to discuss which one or both should you do.

MR. OSWALD: The glossary is more important. For many languages the glossary is more important than the syntax. I think one of the blind alleys which we have been moving down is working with German. I think that Booth and

Richards have quite readily demonstrated that as long as you
work with French or Spanish or Italian, your problem of syn-
tax almost vanishes. But one reason that we began these
syntactic operations was because it was the most complex,
and we wanted to find out if we could work with that. I am
quite confident that if you would produce a microglossary
from one of the languages whose word order is reasonably
close to that of English, that without having to do any
syntactical analysis, you could turn out a job in a remark-
ably small time.

MR. BOOTH: I would say six months with one person.

MR. OSWALD: Let's try to get two people for
awhile. I can do it in less than that, probably in two
weeks, with plenty of guys and an IBM machine.

MR. BULL: It's a good enough place to start work-
ing.

MR. BAR-HILLEL: In addition to the microglossary
we need five thousand most frequent German words arranged in
their frequency order.

MR. BOOTH: I don't know if the computer can put
those on in five months.

MR. OSWALD: You didn't ask me that, you asked me
something else. This is the reason we need these kind of
things. The minute our figures get too high, you say stop.

MR. BAR-HILLEL: I need one drum full. How much

can you put in one drum?

MR. FORRESTER: Five hundred thousand digits on a drum with a factor of two.

MR. BAR-HILLEL: Two thousand? O.K.

MR. HUSKEY: There is another point here. For the purposes of this deal, you don't care how long it takes to do it so long as you can do it on IBM cards. There is no question about the equipment existing, and as long as there is enough money to pay for it, it can be done.

MR. RICHARDS: Suppose that you were to take a selection of twenty or forty or sixty pages of text like brain surgery, suppose you were to get the kinds of pre-editor that you want, one who is familiar with German but with no knowledge of surgery or English. I don't think you ought to simulate such a person. Translate it and get a post-editor if necessary. Although I think that is unnecessary in these days; and then draw up a set of rules. Then you get a girl secretary, give her the proper dictionary, the German and English volume--whatever you people would select--give her a form where she can enter each. This is an actual printed form with a block above and a block below for each chunk of the source document. After it's pre-edited, she enters the chunks, principally single words or pieces of words, in these blocks and writes below them what she gets out of the printed dictionary--according to the

simple rules that you gave here.

MR. OSWALD: She won't get anything. You have got to make a dictionary.

MR. RICHARDS: But, anyhow, give here a short set of rules and this proposed dictionary, and ask her to do it simulating a machine. Then you can have those punched on IBM cards, and just print them out and study them. I think that would be a mighty interesting subject.

MR. BAR-HILLEL: Nobody has sufficient imagination to foresee everything which might happen.

MR. RICHARDS: I'd like to see that.

MR. WIESNER: This has been going on here. This is what he's been doing.

MR. BAR-HILLEL: That proves that for most languages it is not sufficient.

MR. FORRESTER: In doing that you should make sure that someone who has had contact with machine coding approaches this blindly, because it's awfully easy where you're doing what the machine is not going to do.

MR. YNGVE: I should say that you code the original language into nonsense; then you give it to a person to go through so that she or he doesn't know the meaning, and, therefore, will follow the rules.

MR. REYNOLDS: What does this prove that you can't prove by punching it up on IBM cards and running it through

a sorter?

MR. BULL: Let me see if we can carry this on to the next step.

... At this point the conference continued on with a discussion on future programming and ways and means to raise money to carry on the work of mechanical translation ...