

SUMMATION BY CHAIRMAN

DOSTERT: You have noticed on the program that a snare has been set for the chairman, and that I have the distinction of being the first chairman, and therefore I am the first one to react to that little snare which is called "summation by the chairman". This is a very dangerous thing to do for two reasons. One is that the speakers will feel that their statements have been distorted in the summation, and this is not the way to make friends; the other is that the chairman will project his own biases and not give a summation but rather try to argue a different case. I will try to avoid both of these pitfalls. The way to avoid the first one is to say that the papers were so lucid, so orderly, so ably presented that it would be presumptuous of a chairman to attempt a summation. So I have gotten out of that one. As for the second one, I think on the basis of what has been said this morning that there are certain general comments which are warranted. It is quite obvious that there is widespread concern with the problems of methodology, that is to say, of finding rational, efficient, effective ways to tackle the many-faceted difficulties that machine translation presents. In the early years--the heroic years--when experience had not chastised many of us, there were some very blind and categorical a priori viewpoints which time has taken care to eliminate, and now there seems to be a greater disposition to be guided by the data of slow methodical empirical experience. It seems to me there is one thing, though, that emerges from the pattern of concern and preoccupation in the field of methodology. I do not want to prejudge what will be said later, nor am I unaware that many able workers in the field will not completely agree with what I am about to say. I think it appears, at this reading at least, that the research seems to be widely text-focused. I realize that I am not referring to what is to be said later but from what has been said this morning. It does appear that a focus on actual texts combined with a measure of theoretical formulation, blended with the empirically derived data from the text, is--at this stage at least--rather widely recognized as valid. A second process of the methodology is the cyclical approach; that is to say, the improvement of the linguistic formulation on the

## Session 1: CURRENT RESEARCH

basis of continuing re-examination of machine-produced text. Also, there seems to be a growing acceptance of the important future role of the revisor. I would rather call him revisor than posteditor because it seems to me that the person who will take the machine output is called upon to do more than edit. It is true that he is called upon to revise the text in order to make it more acceptable and practical to the user, but the job that he is called upon to do in cooperation with the researchers involves a good deal of effort and knowledge which transcends that of mere editing. I submit this idea for consideration because at the end of my remarks I will suggest something concerning the problem of emerging terminology in our field. There is also a growing concern about devising efficient revising techniques, and this is a problem that many groups will be facing soon. If we can have an exchange of ideas at an early stage, and profit from one another's experience, I believe it will be to the benefit of the over-all research effort. At Georgetown we have in mind this summer to take some 40 or 50 translators in various government services and to organize a summer seminar for 8 weeks to give them certain orientations on how to work on machine-produced text, so that the linguists and the programmers will derive maximum benefit from their revising of those texts. We do not by any means claim to have devised a course for the orientation of revisors, and we have a completely open mind for any suggestion as to techniques which might prove plausible. The second general point is that it becomes increasingly obvious that the problem of semantics is the blank wall before which all of us find ourselves and which creates the greater measure of perplexity. Here again it will be very helpful in facilitating communication if we develop a set of terms that will enable us to completely understand what we mean. It seems to me, and I submit this with some reticence because these are merely preliminary notions, that we could think of the semantic problems at possibly three different levels. One would be the level at which the semantic ambiguity is partly resolvable by structural data. I say partly, but not totally--which leaves another area, then, of subgrammatical-class-definitions which, blended with existing grammatical data, would perhaps move us toward a solution of certain types of semantic ambiguities. The second level would be the analysis of text for those grammatical items which, though in part

## Session 1: CURRENT RESEARCH

grammatically definable, do not come to a proper elucidation by grammatical data exclusively. I have been working with the French word "en", and I have tried to study its behavior. There are a number of readily discernible patterns. But there are also a number of patterns which are not readily classifiable. I would say that the first pattern would fall in the category of partly structurally resolvable ambiguities, and then those which do not fall in such categories fall into the hamper of unresolved semantic problems in respect to this particular particle. I believe that the problem of what you might call lexical polysemia appears to me to be less formidable than the problem of non-lexical polysemia. It seems to me that it is, or should be, more feasible to establish monovalence in a multivalence situation when the problem is strictly lexical, than it would be for certain elusive particles which are not subject to the sort of monovalence establishment which certain other lexical data may be. I think that Professor Harper made a very significant statement when he pointed to the fact that the Soviet Union went through a first phase in their machine translation research--the empirical and popularization phase--which I believe was just as indispensable to them as it has proved to be to us. I completely agree that we are moving now in a second phase where the results will be based on possibly more rigorous, more knowledgeable efforts and certainly upon a greater amount of exchange of information.