

Session 7: THE DICTIONARY

INTRODUCTION

SWANSON: I am going to assume, since the function of an introducer is a relatively new invention, that a certain amount of freedom to experiment is permissible. Introducing the subject matter, I think, may be somewhat less important than a slightly different but closely related task. I suspect that a newcomer to this field who has innocently walked into the lively discussions of this conference would come to the conclusion that it is a good deal less important to understand what is being said than it is to know who is on which side. So, it seems reasonable to think that the introducer would fulfill a most useful function if he would identify at least one or two of the "schools" and establish a few general criteria that distinguish them. Perhaps the more sophisticated introducer could then follow with an account of the directions and shifts in direction in which a preponderant amount of influence happens to be flowing as the verbal duels are joined. For it is, after all, this whole question of the flow of influence that determines who wins or loses. In any event, I shall not really try to fill all of these ambitious functions, but perhaps what I suggest might one day influence the behavior of future introducers.

Let me adopt a modest objective here of simply identifying one of the major arenas on the subject of dictionaries. Two of its stout defenders can be identified as Professor Reifler and Dr. Gilbert King. I am not choosing this particular contest because I wish to debate it pro or con, but rather because it can be used to raise most of the interesting questions that can be asked about dictionaries. The point of view in question can be defined by saying: "Let us put as much as we can into the dictionary and in this way solve by lookup techniques as many problems of machine translation as possible". The question then is how large the lexical units should be. Certainly they must stand some chance of being repeated often enough in the literature to be useful, but obviously they cannot be as large as are most sentences. It is not my purpose here to discuss the appropriate length of lexical unit; but we may presume that there is some optimal size, and we may presume that complete paradigms of fully inflected word forms are included in the dictionary. This general point of view has the solid virtue of being assailable on only three rather weak grounds:

Session 7: THE DICTIONARY

(1) it is aesthetically unattractive; (2) the researchers, after a few months, are likely to be waist-deep in punched cards; and (3) the cost is extravagant (if we assume experiments with existing general-purpose machinery). This latter objection, though probably of little concern to the scholar, raises a point of view which I think offers a sound perspective for viewing the whole question of dictionaries, and that is: the cost per word of translation that eventually results. It is important to take into account the cost per word of the total automatic translation process, since some ways of organizing dictionaries, though efficient for lookup, make awkward the implementation of subsequent syntactic and semantic rules.

In case my advice were solicited by someone starting work on automatic translation, the first thing I would recommend he do (after choosing up sides, of course) would be to experiment with a fully inflected dictionary and thus evade all of those problems that can be solved by a lookup technique; in this way he must face certain quite difficult problems early which he would otherwise not encounter until too late to perform a graceful retreat. I suspect that this point of view will be clearly presented by Professor Reifler, our first speaker in this session.

The second speaker, Mr. Sherry, I would judge from the title of his talk, represents a different school of thought and a different approach, since he clearly is not going to assume that fully inflected forms are stored in the dictionary. Judging from the titles published in the program and some interesting discussions I have had with Dr. Lamb, I would say that the latter two speakers have probably been too busy being clever to worry about which side they are on.

Thank you, I return the session to your chairman, Dr. Brown.