140. Without a thesaurus random keywords will be presented to the system resulting in an open-ended file in the computer with no connections being made or relationships established between synonyms. To make these relationships known to the system would require operator intervention. The unstructured file of descriptors can become a drawback in retrieval. A completely unstructured list of keywords from which to choose will result in 'spotty' retrievals, that is, retrievals which do not identify all items of the relevant material because different keywords not associated with the subject by the searcher have been used as descriptors. In a complex subject environment in which there is no specialist terminology the lack of a thesaurus can drastically reduce effectiveness of the system, since all word associations and grouping must be done by the user each time he uses the system. With a large, uncontrolled keyword file this can become a major task.

141. Other foreign affairs related system developers have recognized this fact, witness the U.S. State Department and Swiss Foreign Office thesauri. It thus appears that the proper question is not "Whether to?" but "How to?" develop a thesaurus.

142. One answer to this latter question could lie in making use of the experience gained in other similar systems. To this end we have obtained copies of the U.S. State Department and Swiss Foreign Office thesauri (the latter in French), each comprising some five to seven thousand words.

143. Further study may show that our solution lies in another approach, since the adoption of someone else's thesaurus would necessitate amendments to suit the Canadian situation. If the Department should develop its own on-line thesaurus from operating experience no structure for keywords would exist at the start of operation of the new system. Upon material being indexed keywords would be entered into the computer-resident files where they would either create a completely new entry or would be grouped with earlier occurrences of the same word. Over a period of time the rate of increase of new words would decline and eventually stabilize. The file would of course be common to all analysts in the system.

Centralized Computer Indexing

144. The various technical possibilities having been described it is now possible to examine how they could be applied in the Departmental system. The first obvious application to consider would be the substitution of computer indexing for the deficient Kard-veyer system. This approach would likely be somewhat more economical of personnel than either the partial decentralization described in paragraphs 122 and 123 or the "Preferred System" itself. However, it would be remote from the users (as is the present system) and the many demonstrable advantages of Bureau Information Control Centres would be unattainable. There would be some support for Level III and IV activity but the development of practical procedures to provide systematic input for the subsidiary systems would be impeded. While implementation of centralized computerized indexing could probably proceed without any highly visible adverse effects it would probably turn out to be an expensive disappointment for the reason that its benefits would also not be very visible to users, with the result that a powerful potential capability would tend to be much under-utilized. This situation holds to a certain extent at the moment in the U.S. State Department.