

CHAPTER 4

Main Test Results

Efficiency is something which cannot be achieved without effort expended in the appropriate ways, and measurement is one of the ways towards improved efficiency. It is difficult to see how efficiency can be improved without some basis in measurement.

L.T. Wilkins: Social Deviance, page 8.

In the two previous chapters have been outlined the environment in which the tests were carried out and the measures which are used for presenting the results. For those who did not wish to work through these two chapters, a brief summary is now given to assist in the interpretation of the test results presented in this chapter. The simplest method of doing this is to illustrate the various points with an example of the tables which present the test results, as in Figure 4.001, to which reference should be made.

(1) There are four main groups of index languages and these are: identified by roman numerals:

- I Single term index languages (eight languages tested)
- II Simple concept index languages (fifteen languages tested)
- III Controlled term index languages (six languages tested)
- IV Abstract and title searches (four languages tested)

With each index language there are a number of different recall devices. These are identified by arabic numerals, and the full range of these index languages is set out in Figures 2.5, 2.6 and 2.7.

The lower case letter identifies the precision device which is being used. The basic device is coordination, shown by a. The other precision devices are fully explained in Figure 2.8.

(2) At the indexing stage, each term was given a rating in relation to its importance in each particular document, and this permits tests to be made with different levels of exhaustivity of indexing. This is to say that searches can be made on the full indexing, which averaged 31 terms for each document, and this is shown as Exhaustivity 3. Alternatively it can be done on a restricted set of terms, where the average was 25 terms per document; this is shown as Exhaustivity 2. Finally searches can be made on the least exhaustive indexing, where the average was 13 terms per document, and this is shown as Exhaustivity 1.

(3) Various search rules that were used are explained in Figure 2.9. The basic search rule permitted the combination of any terms and is shown by A. Other search rules are explained in the appropriate section of the tables.

- 1 Index Language I 5. a (S. T. Synonyms, Quasi-synonyms. Coordination)
- 2 Exhaustivity of Indexing 3
- 3 Search Rule A
- 4 Document Relevance 1 - 4
- 5 Number of Documents in Collection 1,400
- 6 Number of Questions 221 (Subset 3)
- 7 Number of Relevant Documents 1,590
- 8 Generality Number 5.1

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x y z | | |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|-------|------|-----|
| | Rel. | Non-rel. | | | | x | y | z |
| 1 | 1,548 | (-) | 97.4% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,406 | 114,265* | 88.4% | 1.2%* | 37.099%* | 221 | 44* | 221 |
| 3 | 1,121 | 42,364* | 70.5% | 2.6%* | 13.755%* | 218 | 109* | 220 |
| 4 | 802 | 16,191* | 50.4% | 4.7%* | 5.257%* | 204 | 142* | 212 |
| 5 | 475 | 8,164* | 29.9% | 5.5%* | 2.651%* | 164 | 177* | 197 |
| 6 | 265 | 3,013* | 16.7% | 8.1%* | 0.378%* | 114 | 161* | 164 |
| 7 | 131 | 910 | 8.2% | 12.6% | 0.296% | 79 | 140 | 140 |
| 8 | 50 | 266 | 3.1% | 15.8% | 0.086% | 44 | 105 | 105 |
| 9 | 19 | 56 | 1.2% | 25.3% | 0.018% | 20 | 78 | 78 |
| 10 | 2 | 12 | 0.1% | 14.3% | 0.004% | 6 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.001 SAMPLE TEST RESULTS SHEET

(4) The relevance of the documents to the questions was assessed by the questioners in four grades of relevance (see Vol. I, p.21). Most of the test results are given for documents of all grades of relevance (shown as 1-4), except in Section 4.6 which specifically deals with the effect of relevance.

(5) The main test collection had 1,400 documents, but two other document sets were used. These were of 200 and 350 documents, a characteristic of these smaller sets being that all the documents were concerned with aerodynamics, whereas the main set contained some 300 documents on theory of aircraft structures.

(6) The largest set of questions in the test had 221 questions. Most of the results are based on a subset of 42 questions, all of which were concerned with aerodynamics. Another subset had 35 questions, the characteristic of which was that each question had seven starting terms. Other sets were used in special cases; full details are given in Figure 2.12, and also in the appropriate section of the tables.

(7) The number of relevant documents will vary with the document set, the question set and the relevance grade. This number must be known for calculating the recall ratio.

(8) The generality number is a function of the number of relevant documents and the number of documents in the test collection. With Question Subset 2 (for which there are 198 relevant documents), when the search is made on the 1400 document collection, the generality number is 3.4. When searched on the 200 document collection, the generality number is 23.6. The effect of the increase in generality number is to bring about an apparent improvement in the performance figures. The matter of generality is fully discussed in Chapter 3.

(9) All the test results given in this chapter were based on searches where the coordination level was progressively decreased from the maximum down to a single term. The maximum number of single terms in any question was 15, while the lowest number of terms was 2.

(10) In most situations, for the reason stated in the previous paragraph, the number of questions which can be searched at a given coordination level will be limited by the number of questions having that number of starting terms. This information is given in column z, which shows, for example, that at a coordination level of 6, there are 164 questions which, having six or more starting terms, can be searched at this level.

In certain searches, the number of questions actually searched at a given coordination level was less than the theoretical maximum possible. This was because of the large clerical effort required and the number of questions actually searched is given in column y.

It will be seen from Figure 4.001 that, at a coordination level of 6, only 161 of the possible 164 questions were searched, and from this stage there is an increasing disparity between the figures in column y and z.

Column x shows the number of questions that were able, at any given coordination level, to retrieve any documents, whether relevant or non-relevant. It will be noted, for instance, that it was not until the coordination level had dropped to ten terms was it possible to retrieve a single document; at this level, as can be seen from the figures in columns x and z, of the 52 questions having ten or more starting terms, 6 questions retrieved documents.

(11) In the columns of documents retrieved are shown the total numbers of relevant and non-relevant documents retrieved at the various coordination levels. These figures have been obtained by summing the results for each individual question in the question set.

As mentioned in the previous section, in some cases the searches were not completed at the lower coordination levels. The result is that the figures for non-relevant documents retrieved are estimated by a method described on page 28. All such estimated figures are indicated by an asterisk. However, it should be noted that the figures for the retrieval of relevant documents are always correct.

(12) The actual performance measures presented in the tables are recall ratio, precision ratio and fallout ratio. These are derived from the following:-

- a. Relevant documents retrieved
- b. Non-relevant documents retrieved
- c. Relevant documents not retrieved
- d. Non-relevant documents not retrieved

Recall ratio is $\frac{100a}{a + c}$, that is relevant documents retrieved over the total relevant. All such figures in the tables are correct, but for precision ratio ($\frac{100a}{a + b}$) and fallout ratio ($\frac{100b}{b + d}$), it was in some cases necessary to use the estimated figures discussed in the previous section. Where this has been done, an asterisk is placed against the figure in the table.

TABLE OF RESULTS

The tables of results are presented in nine main sections. Details are given before each section of tables, but the following is a brief resume.

In the first section the results are given for the single term index languages with the largest set of documents and questions. Because of the doubts regarding the totalling method used, the 35 seven-term questions set is given for comparison. This is followed by the 42 aerodynamic questions on the 1400 document collection, with a final table for this set of questions with the 200 document collection. The purpose of this group of tables is first to justify the totalling procedure, then the reduced set of questions and finally the reduced set of documents.

Section 2 gives the results for the 42-question and 200-document sets for the eight single term languages. The tests using precision devices of interfixing and partitioning are given in Section 3, and are followed in Section 4 by comparative results at different levels of exhaustivity. The effect of search rules is shown in Section 5 and the different levels of relevance are tested in Section 6. Results for fifteen simple concepts index languages are presented in Section 7, while Section 8, which deals with the six controlled term index languages, includes results for recall devices, precision devices and search rules. In Section 9 are presented the results of searches on the titles and abstracts,

The tables and plots are numbered according to the above nine sections, 'T' indicating a table and 'P' indicating a plot. In all cases unless otherwise stated, the plots are of recall and precision ratios.

Section 1 Introductory Tables

The first set of results is based on 221 questions (question subset 3) searched on the full collection of 1400 documents. In Tables 4.100T to 4.104T, five of the languages investigated on the single term indexing are presented, namely languages I.1.a, I.2.a, I.3.a, I.5.a, and I.6.a, showing the effect of recall devices. The results for these five languages are presented on a single plot in Figure 4.105P, with a performance curve for each of the languages plotted.

It was very difficult to find a satisfactory method of totalling the results with these 221 questions, because of the large variation in the number of starting terms. This problem was fully discussed in Chapter 3 (pages 51 - 71), but in order to validate the selected method, a subset of 35 questions was selected, the characteristic of which was that each question had seven starting terms, and in this respect was an average set. The results for searches for the same five index languages are presented in Figures 4.110T - 4.114T. The questions are again searched on the 1400 document collection, and a single plot of the five curves is given in Fig. 4.115P.

A further subset of questions, 42 in number (subset 2), is used for the results given in Figs. 4.120T to 4.124T. The same five languages are used, and the 1400 document collection is searched. These questions pose the same problems in totalling as did the 221 questions, since the 42 questions have varying numbers of starting terms, ranging from three to twelve. A single plot of the five languages is given in Fig. 4.125P.

In these three sets of questions, a progression may be observed from the largest set of questions (221) to a smaller set of 35 questions specially selected to minimise the totalling problem and to another small set of 42 questions that has the same problem of totalling, due to the variation in number of starting terms, as the collection of 221 questions. The differences in question sample size may be expected to affect any direct comparison of the three sets of questions, in addition to the totalling method problem. The effect of this can be seen in Fig. 4.130P where the natural language results (Language I.1.a) are compared for the three subsets of questions, the three curves being based on the results in Figs. 4.100T, 4.110T, and 4.120T. Although the comparison is accurate in terms of recall and precision as calculated, the comparison of three different sets of questions brings in a new variation, namely the generality number (G). For the 221 questions G is 5.1, for the 35 questions G is 5.9, and for the 42 questions G is 3.4. The need to allow for this difference in generality has been discussed, and Fig. 4.131P is a graph that allows for this by use of a plot of recall and fallout.

It can now be argued that the performance results based on the smaller sets of questions give valid results. First, there is a close similarity in the relative differences obtained when five recall languages are compared, whichever set of questions is used. This may be seen by comparing Figs. 4.105P, 4.115P and 4.125P where the relative differences between the five languages are very similar. A further comparison is made on a recall fallout plot, in Fig. 4.131P, where generality is allowed for, and the language is held constant as type I.1.a. Some small variation is to be expected when the question sets are altered in size, and when a universal totalling method is to be used, but the subset of 42 questions, on which many of the later results are based, is seen to be representative of both the larger set of 221 questions and of the set with the chosen characteristic of each question having the same number of starting terms.

In the question sets shown so far, the collection size has remained constant at 1400 documents. Most of the later results have been obtained on the 200 document collection (collection subset 1), and the validity of results based on the smaller collection size will be considered next.

Table 4.140T gives the results for a search on language I.1.a made with the 42 questions searched on the 200 document collection. The results from the table are plotted as a performance curve in Fig. 4.140P. Also shown on this plot are the results from Fig. 4.120T, which are based on the same question set and language but searched on the 1400 collection. It would be expected that with a recall/precision plot, the increased generality number would result in a better performance for the searches on the 200 collection as compared to those on the 1400 collection. This is, in fact, the case, for while the recall at each coordination level is seen to be identical, the expected large increase in precision is seen when the 200 collection is searched.

The effect of the change in generality on the precision ratio, as discussed in Chapter 3, is allowed for in Fig. 4.141P, which is a plot of the two curves using recall and fallout ratios, and Fig. 4.142P which plots the two curves on a recall/precision plot with generality adjusted to a constant of 23.6, this number being the generality of the situation in the 200 collection. The result for the 1400 collection is now no longer inferior to the 200 collection, and in fact the situation is reversed. The reason why the 200 collection now has a somewhat worse performance has been investigated in Chapter 3, where it has been shown that the cause of the difference can be adequately explained and allowed for; this was, in fact, done and the result was shown in Figure 3.34T.

The purpose of the introductory results presented is to demonstrate that test results based on a relatively small collection and set of questions do give valid results. The variations observed between the three different

sets of questions and the two collections can all be adequately explained, and can be allowed for by methods of adjustment that have been developed.

LIST OF FIGURES

| | Index Language | No. of Questions | Question Subset | Document Collection | Plots |
|--------|--------------------------------------|---------------------|--------------------|------------------------|--------------|
| 4.100T | I.1.a | 221 | 3 | 1400 | |
| 4.101T | I.2.a | 221 | 3 | 1400 | |
| 4.102T | I.3.a | 221 | 3 | 1400 | |
| 4.103T | I.5.a | 221 | 3 | 1400 | |
| 4.104T | I.6.a | 221 | 3 | 1400 | |
| 4.105P | I.1.a | 221 | 3 | 1400 | Plot 4.100T, |
| | I.2.a | | | | 4.101T, |
| | I.3.a | | | | 4.102T, |
| | I.5.a | | | | 4.103T, |
| | I.6.a | | | | 4.104T. |
| 4.110T | I.1.a | 35 | 1 | 1400 | |
| 4.111T | I.2.a | 35 | 1 | 1400 | |
| 4.112T | I.3.a | 35 | 1 | 1400 | |
| 4.113T | I.5.a | 35 | 1 | 1400 | |
| 4.114T | I.6.a | 35 | 1 | 1400 | |
| 4.115P | I.1.a | 35 | 1 | 1400 | Plot 4.110T, |
| | I.2.a | | | | 4.111T, |
| | I.3.a | | | | 4.112T, |
| | I.5.a | | | | 4.113T, |
| | I.6.a | | | | 4.114T. |
| 4.120T | I.1.a | 42 | 2 | 1400 | |
| 4.121T | I.2.a | 42 | 2 | 1400 | |
| 4.122T | I.3.a | 42 | 2 | 1400 | |
| 4.123T | I.5.a | 42 | 2 | 1400 | |
| 4.124T | I.6.a | 42 | 2 | 1400 | |
| 4.125P | I.1.a | 42 | 2 | 1400 | Plot 4.120T, |
| | I.2.a | | | | 4.121T, |
| | I.3.a | | | | 4.122T, |
| | I.5.a | | | | 4.123T, |
| | I.6.a | | | | 4.124T. |
| 4.130P | I.1.a | 221 | 3 | 1400 | Plot 4.100T, |
| | | 35 | 1 | | 4.110T, |
| | | 42 | 2 | | 4.120T, |
| 4.131P | (same as 4.130P, but recall/fallout) | | | | |

| | Index Language | No. of Questions | Question Subset | Document Collection | Plots |
|--------|---|---------------------|--------------------|------------------------|------------------------|
| 4.140T | I.1.a | 42 | 2 | 200 | |
| 4.141P | I.1.a | 42 | 2 | 1400 200 | Plot 4.120T, 4.140T |
| 4.142P | (same as 4.141P, but recall/adjusted precision) | | | | |

FIGURE 4.100T

Index Language I.1.a. (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1-4

Number of Documents in Collection 1,400

Number of Questions 221 (Subset 3)

Number of Relevant Documents 1,590

Generality Number 5.1

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|-----|-----|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,510 | 159,122 | 95.0% | 0.9% | 51.696% | 221 | 221 | 221 |
| 2 | 1,283 | 58,122 | 80.7% | 2.2% | 18.883% | 221 | 221 | 221 |
| 3 | 946 | 21,933 | 59.5% | 4.1% | 7.125% | 215 | 220 | 220 |
| 4 | 606 | 7,359 | 38.1% | 7.6% | 2.390% | 187 | 212 | 212 |
| 5 | 314 | 2,380 | 19.7% | 11.6% | 0.773% | 131 | 197 | 197 |
| 6 | 154 | 699 | 9.7% | 18.0% | 0.227% | 86 | 164 | 164 |
| 7 | 74 | 216 | 4.7% | 25.5% | 0.070% | 50 | 140 | 140 |
| 8 | 22 | 43 | 1.4% | 33.8% | 0.014% | 18 | 105 | 105 |
| 9 | 8 | 5 | 0.5% | 61.5% | 0.002% | 8 | 78 | 78 |
| 10 | 1 | 0 | 0.1% | 100.0% | 0.000% | 1 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.101T

Index Language I.2.a (S.T. Synonyms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1-4

Number of Documents in Collection 1,400

Number of Questions 221 (Subset 3)

Number of Relevant Documents 1,590

Generality Number 5.1

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,514 | (-) | 95.2% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,313 | 59,734* | 82.6% | 2.2%* | 19.406%* | 221 | 44* | 221 |
| 3 | 981 | 23,654* | 61.7% | 4.0%* | 7.680%* | 216 | 109* | 220 |
| 4 | 644 | 8,850* | 40.5% | 6.8%* | 2.873%* | 192 | 142* | 212 |
| 5 | 355 | 2,946* | 22.3% | 10.4%* | 0.957%* | 139 | 177* | 197 |
| 6 | 169 | 928* | 10.6% | 15.4%* | 0.301%* | 92 | 161* | 164 |
| 7 | 80 | 254 | 5.0% | 24.0% | 0.083% | 55 | 140 | 140 |
| 8 | 24 | 59 | 1.5% | 28.9% | 0.019% | 23 | 105 | 105 |
| 9 | 8 | 8 | 0.5% | 50.0% | 0.003% | 8 | 78 | 78 |
| 10 | 1 | 0 | 0.1% | 100.0% | 0.000% | 1 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.102T

Index Language I.3.a (S.T. Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 221 (Subset 3)

Number of Relevant Documents 1,590

Generality Number 5.1

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,533 | (-) | 96.4% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,338 | 62,765* | 84.2% | 2.1%* | 20.378%* | 221 | 44* | 221 |
| 3 | 1,017 | 24,726* | 64.0% | 3.9%* | 8.028%* | 217 | 109* | 220 |
| 4 | 677 | 9,565* | 42.6% | 6.6%* | 2.530%* | 192 | 142* | 212 |
| 5 | 374 | 3,084* | 23.5% | 10.8%* | 1.001%* | 139 | 177* | 197 |
| 6 | 192 | 1,112* | 12.1% | 14.8%* | 0.361%* | 99 | 164* | 161 |
| 7 | 96 | 333 | 6.0% | 22.4% | 0.108% | 64 | 140 | 140 |
| 8 | 34 | 87 | 2.1% | 28.1% | 0.028% | 28 | 105 | 105 |
| 9 | 13 | 15 | 0.8% | 46.4% | 0.005% | 17 | 78 | 78 |
| 10 | 2 | 0 | 0.1% | 100.0% | 0.000% | 2 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.103T

Index Language I.5.a (S.T. Synonyms, Quasi-synonyms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 221 (Subset 3)

Number of Relevant Documents 1,590

Generality Number 5.1

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,548 | (-) | 97.4% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,406 | 114,265* | 88.4% | 1.2%* | 37.099%* | 221 | 44* | 221 |
| 3 | 1,121 | 42,364* | 70.5% | 2.6%* | 13.755%* | 218 | 109* | 220 |
| 4 | 802 | 16,191* | 50.4% | 4.7%* | 5.257%* | 204 | 142* | 212 |
| 5 | 475 | 8,164* | 29.9% | 5.5%* | 2.651%* | 164 | 177* | 197 |
| 6 | 265 | 3,013* | 16.7% | 8.1%* | 0.378%* | 114 | 161* | 164 |
| 7 | 131 | 910 | 8.2% | 12.6% | 0.296% | 79 | 140 | 140 |
| 8 | 50 | 268 | 3.1% | 15.8% | 0.086% | 44 | 105 | 105 |
| 9 | 19 | 56 | 1.2% | 25.3% | 0.018% | 20 | 78 | 78 |
| 10 | 2 | 12 | 0.1% | 14.3% | 0.004% | 6 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.104T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 221 (Subset 3)
 Number of Relevant Documents 1,590
 Generality Number 5.1

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,557 | (-) | 97.9% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,430 | 116,374* | 89.9% | 1.2%* | 37.783%* | 221 | 44* | 221 |
| 3 | 1,165 | 45,101* | 73.3% | 2.5%* | 14.643%* | 218 | 109* | 220 |
| 4 | 848 | 18,373* | 53.3% | 4.4%* | 5.965%* | 206 | 142* | 212 |
| 5 | 503 | 8,895* | 31.6% | 5.4%* | 2.888%* | 169 | 177* | 197 |
| 6 | 295 | 3,874* | 18.6% | 7.1%* | 1.257%* | 119 | 161* | 164 |
| 7 | 161 | 1,136 | 10.1% | 12.4% | 0.369% | 83 | 140 | 140 |
| 8 | 72 | 344 | 4.5% | 17.3% | 0.112% | 54 | 105 | 105 |
| 9 | 24 | 82 | 1.5% | 22.6% | 0.027% | 25 | 78 | 78 |
| 10 | 8 | 18 | 0.4% | 25.0% | 0.006% | 12 | 52 | 52 |
| 11 | 0 | 1 | 0.0% | 0.0% | 0.0003% | 1 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

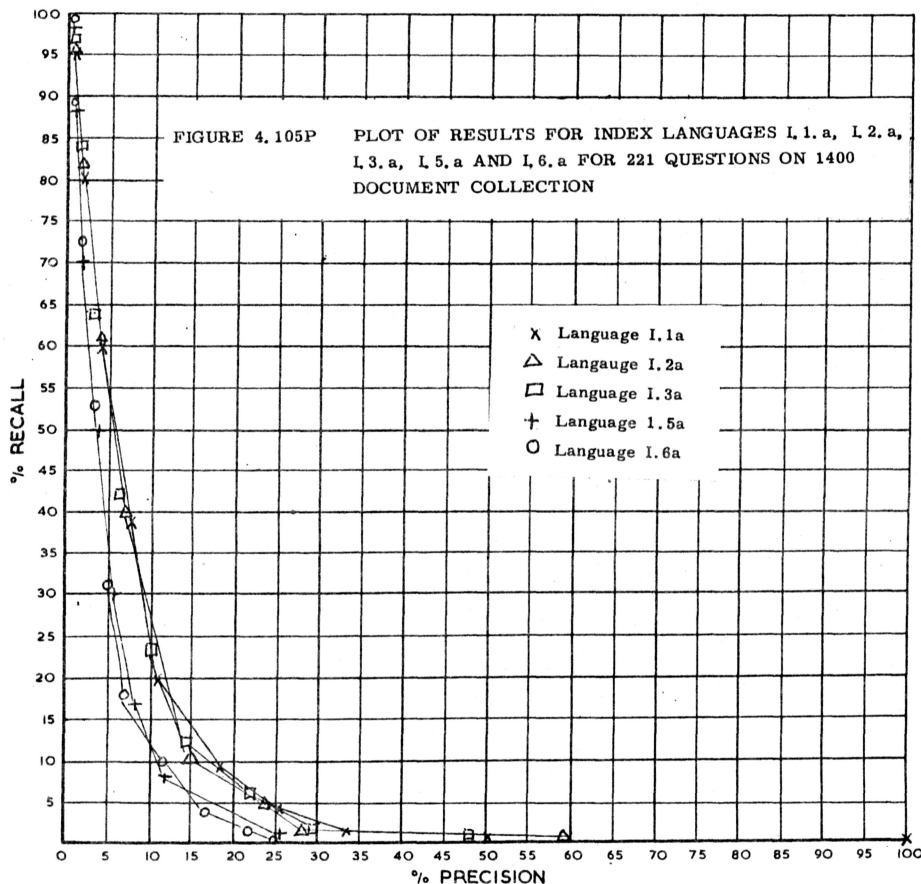


FIGURE 4.110T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 35 (Subset 1)

Number of Relevant Documents 287

Generality Number 5.9

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 268 | 23,681 | 93.4% | 1.1% | 48.329% | 35 | 35 | 35 |
| 2 | 221 | 8,086 | 77.0% | 2.7% | 16.500% | 35 | 35 | 36 |
| 3 | 157 | 2,865 | 54.7% | 5.2% | 5.881% | 34 | 35 | 35 |
| 4 | 94 | 600 | 32.8% | 13.5% | 1.232% | 28 | 35 | 35 |
| 5 | 46 | 147 | 16.4% | 23.8% | 0.281% | 18 | 35 | 35 |
| 6 | 22 | 37 | 7.7% | 37.2% | 0.076% | 7 | 35 | 35 |
| 7 | 8 | 8 | 2.8% | 50.0% | 0.016% | 3 | 35 | 35 |

FIGURE 4.111T

Index Language I.2.a (S.T. Synonyms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 35 (Subset 1)

Number of Relevant Documents 287

Generality Number 5.9

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 269 | (-) | 93.8% | (-) | (-) | 35 | 0 | 35 |
| 2 | 226 | 8,514* | 78.8% | 2.6%* | 17.355%* | 35 | 23* | 35 |
| 3 | 164 | 3,028 | 57.2% | 5.1% | 6.212% | 34 | 35 | 35 |
| 4 | 99 | 650 | 34.5% | 13.2% | 1.334% | 28 | 35 | 35 |
| 5 | 51 | 151 | 17.8% | 25.2% | 0.310% | 19 | 35 | 35 |
| 6 | 23 | 39 | 8.0% | 37.1% | 0.080% | 8 | 35 | 35 |
| 7 | 8 | 8 | 2.8% | 50.0% | 0.016% | 3 | 35 | 35 |

FIGURE 4.112T

Index Language I.3.a (S.T. Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 35 (Subset 1)

Number of Relevant Documents 287

Generality Number 5.9

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 270 | (-) | 94.1% | (-) | (-) | 35 | 0 | 35 |
| 2 | 227 | 9,174* | 79.1% | 2.4%* | 18.518%* | 35 | 23* | 35 |
| 3 | 170 | 3,408 | 59.3% | 4.7% | 6.996% | 35 | 35 | 35 |
| 4 | 103 | 768 | 35.9% | 11.8% | 1.577% | 30 | 35 | 35 |
| 5 | 55 | 184 | 19.2% | 23.0% | 0.378% | 21 | 35 | 35 |
| 6 | 25 | 44 | 8.8% | 36.2% | 0.090% | 12 | 35 | 35 |
| 7 | 8 | 8 | 2.8% | 50.0% | 0.016% | 3 | 35 | 35 |

FIGURE 4.113T

Index Language I 5.a (S. T. Synonyms, Quasi-synonyms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 35 (subset 1)
 Number of Relevant Documents 287
 Generality Number 5.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 280 | (-) | 97.6% | (-) | (-) | 35 | 0 | 35 |
| 2 | 253 | 17,130* | 88.2% | 1.5%* | 34.959%* | 35 | 23* | 35 |
| 3 | 194 | 7,472 | 67.6% | 2.5% | 15.339% | 35 | 35 | 35 |
| 4 | 125 | 2,086 | 43.6% | 5.6% | 4.282% | 34 | 35 | 35 |
| 5 | 65 | 463 | 22.7% | 12.3% | 0.950% | 30 | 35 | 35 |
| 6 | 35 | 88 | 12.2% | 28.4% | 0.181% | 16 | 35 | 35 |
| 7 | 11 | 18 | 3.9% | 38.0% | 0.037% | 5 | 35 | 35 |

FIGURE 4.114T

Index Language I 6.a (S. T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 35 (Subset 1)
 Number of Relevant Documents 287
 Generality Number 5.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 280 | (-) | 97.6% | (-) | (-) | 35 | 0 | 35 |
| 2 | 255 | 19,264* | 88.9% | 1.3%* | 39.314%* | 35 | 23* | 35 |
| 3 | 202 | 8,070 | 70.4% | 2.4% | 16.566% | 35 | 35 | 35 |
| 4 | 130 | 2,426 | 45.3% | 5.1% | 4.980% | 35 | 35 | 35 |
| 5 | 72 | 571 | 25.1% | 11.2% | 1.172% | 30 | 35 | 35 |
| 6 | 38 | 109 | 13.3% | 25.8% | 0.224% | 18 | 35 | 35 |
| 7 | 13 | 19 | 4.6% | 40.6% | 0.039% | 6 | 35 | 35 |

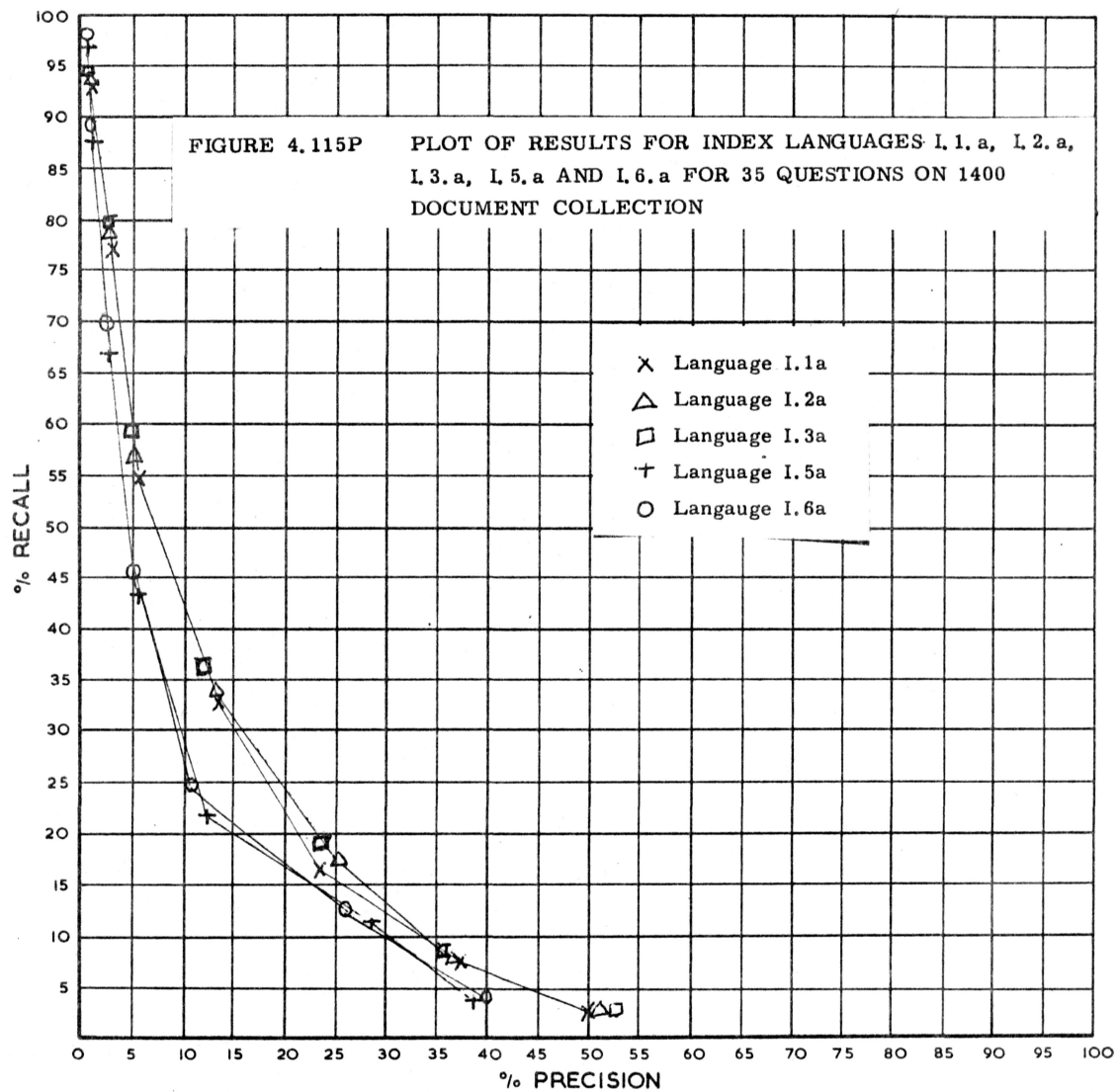


FIGURE 4.120T

Index Language I.1.a (S.T. Natural language. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 3.4

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 189 | 34,127 | 95.5% | 0.6% | 57.068% | 42 | 42 | 42 |
| 2 | 162 | 9,811 | 81.8% | 1.6% | 16.406% | 42 | 42 | 42 |
| 3 | 132 | 3,979 | 66.7% | 3.2% | 6.790% | 42 | 42 | 42 |
| 4 | 88 | 1,039 | 44.4% | 7.8% | 1.772% | 38 | 41 | 41 |
| 5 | 49 | 251 | 24.7% | 16.3% | 0.428% | 25 | 39 | 39 |
| 6 | 25 | 56 | 12.6% | 30.9% | 0.096% | 16 | 33 | 33 |
| 7 | 12 | 11 | 6.1% | 52.2% | 0.019% | 8 | 27 | 27 |
| 8 | 4 | 1 | 2.0% | 80.0% | 0.002% | 2 | 18 | 18 |
| 9 | 0 | 1 | 0.0% | 0.0% | 0.002% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.121T

Index Language I.2.a (S.T. Synonyms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 3.4

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 189 | (-) | 95.5% | (-) | (-) | 42 | 0 | 42 |
| 2 | 165 | 10,476* | 83.3% | 1.6%* | 17.518%* | 42 | 6* | 42 |
| 3 | 134 | 4,614 | 67.7% | 2.8% | 7.873% | 42 | 42 | 42 |
| 4 | 90 | 1,244 | 45.5% | 6.7% | 2.123% | 38 | 41 | 41 |
| 5 | 55 | 320 | 27.8% | 14.7% | 0.546% | 27 | 39 | 39 |
| 6 | 25 | 67 | 12.6% | 27.2% | 0.114% | 16 | 33 | 33 |
| 7 | 12 | 15 | 6.1% | 44.4% | 0.026% | 10 | 27 | 27 |
| 8 | 4 | 1 | 2.0% | 80.0% | 0.002% | 2 | 18 | 18 |
| 9 | 0 | 1 | 0.0% | 0.0% | 0.002% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.122T

Index Language I.3.a (S.T. Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 3.4

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 191 | (-) | 96.5% | (-) | (-) | 42 | 0 | 42 |
| 2 | 166 | 11,268* | 83.8% | 1.5%* | 18,843%* | 42 | 6* | 42 |
| 3 | 139 | 5,054 | 70.2% | 2.7% | 8.624% | 42 | 42 | 42 |
| 4 | 97 | 1,497 | 49.0% | 6.1% | 2.555% | 39 | 41 | 41 |
| 5 | 59 | 419 | 29.8% | 12.3% | 0.715% | 26 | 39 | 39 |
| 6 | 32 | 100 | 16.2% | 24.2% | 0.171% | 17 | 33 | 33 |
| 7 | 14 | 22 | 7.1% | 38.9% | 0.038% | 11 | 27 | 27 |
| 8 | 4 | 6 | 2.0% | 40.0% | 0.010% | 5 | 18 | 18 |
| 9 | 0 | 3 | 0.0% | 0.0% | 0.005% | 2 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.123T

Index Language I.5.a (S.T. Synonyms, Quasi-synonyms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

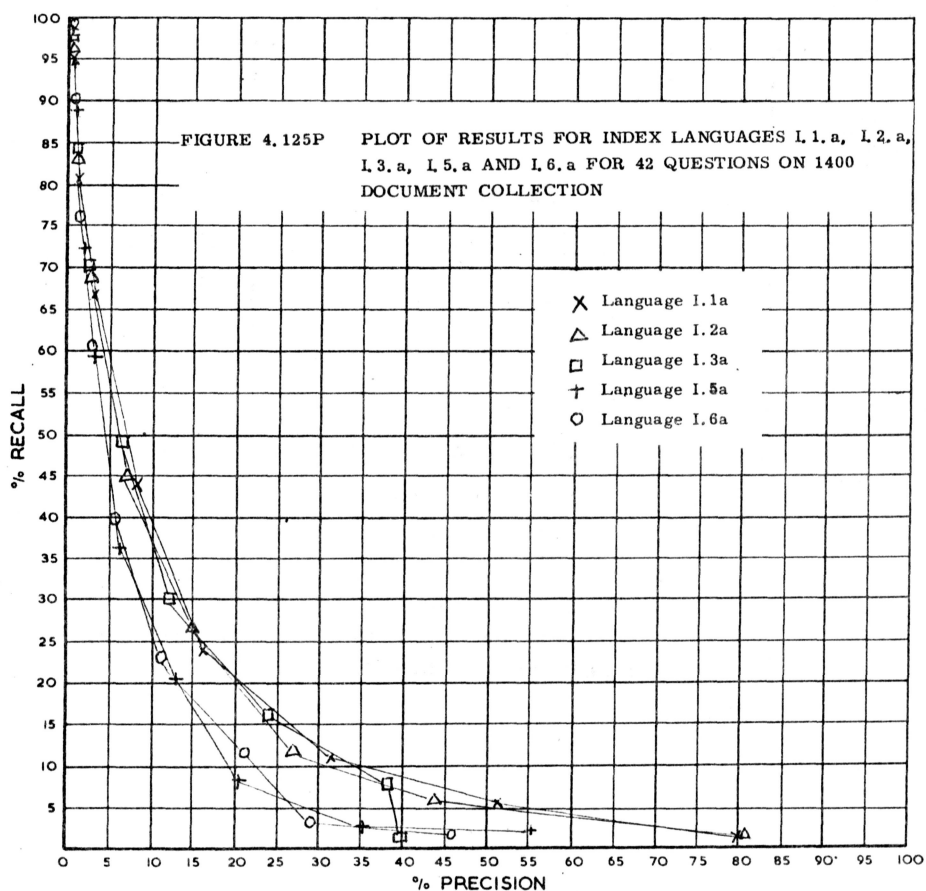
Generality Number 3.4

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 195 | (-) | 98.5% | (-) | (-) | 42 | 0 | 42 |
| 2 | 177 | 13,564* | 89.4% | 1.3%* | 22,682%* | 42 | 6* | 42 |
| 3 | 144 | 9,418 | 72.7% | 1.5% | 16.071% | 42 | 42 | 42 |
| 4 | 117 | 3,248 | 59.1% | 3.5% | 5.542% | 40 | 41 | 41 |
| 5 | 73 | 1,014 | 36.9% | 6.7% | 1.730% | 33 | 39 | 39 |
| 6 | 41 | 275 | 20.7% | 13.0% | 0.469% | 24 | 33 | 33 |
| 7 | 17 | 66 | 8.6% | 20.5% | 0.113% | 16 | 27 | 27 |
| 8 | 6 | 11 | 3.0% | 35.3% | 0.019% | 6 | 18 | 18 |
| 9 | 5 | 4 | 2.5% | 55.6% | 0.007% | 4 | 11 | 11 |
| 10 | 0 | 3 | 0.0% | 0.0% | 0.005% | 2 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.124T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 3.4

| Coord-ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|---------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 195 | (-) | 98.5% | (-) | (-) | 42 | 0 | 42 |
| 2 | 179 | 13,826* | 90.4% | 1.3%* | 23.120%* | 42 | 6* | 42 |
| 3 | 151 | 10,508 | 76.3% | 1.4% | 17.928% | 42 | 42 | 42 |
| 4 | 120 | 3,875 | 60.6% | 3.0% | 6.612% | 40 | 41 | 41 |
| 5 | 79 | 1,299 | 39.9% | 5.7% | 2.217% | 34 | 39 | 39 |
| 6 | 47 | 362 | 23.7% | 11.5% | 0.618% | 24 | 33 | 33 |
| 7 | 27 | 91 | 13.6% | 22.9% | 0.155% | 16 | 27 | 27 |
| 8 | 7 | 17 | 3.5% | 29.2% | 0.029% | 7 | 18 | 18 |
| 9 | 5 | 6 | 2.5% | 45.5% | 0.010% | 5 | 11 | 11 |
| 10 | 0 | 3 | 0.0% | 0.0% | 0.005% | 2 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |



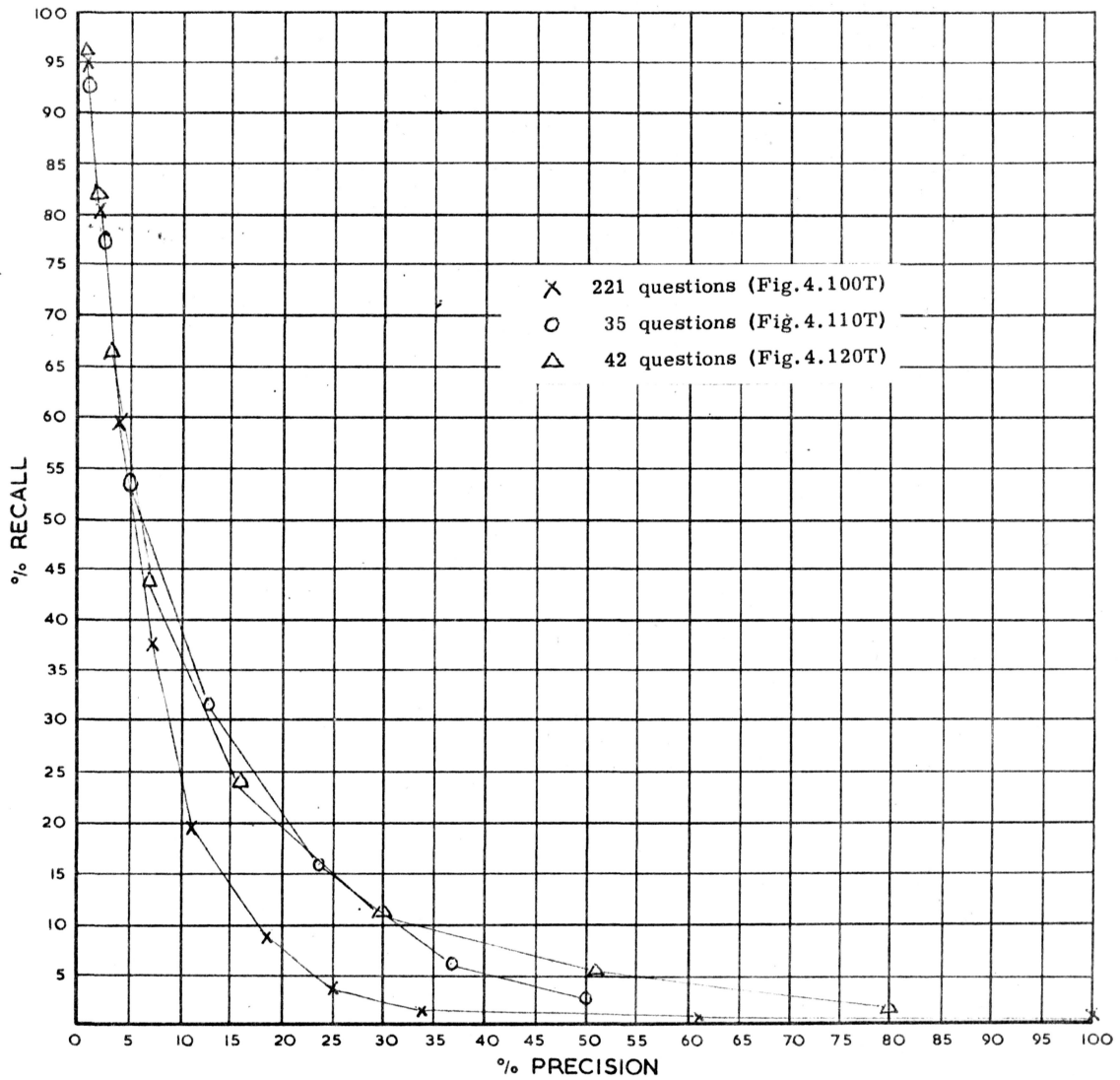


FIGURE 4.130P PLOT OF RESULTS FOR INDEX LANGUAGE I, 1. a FOR
221 QUESTIONS, 35 QUESTIONS AND 42 QUESTIONS ON
1400 DOCUMENT COLLECTION

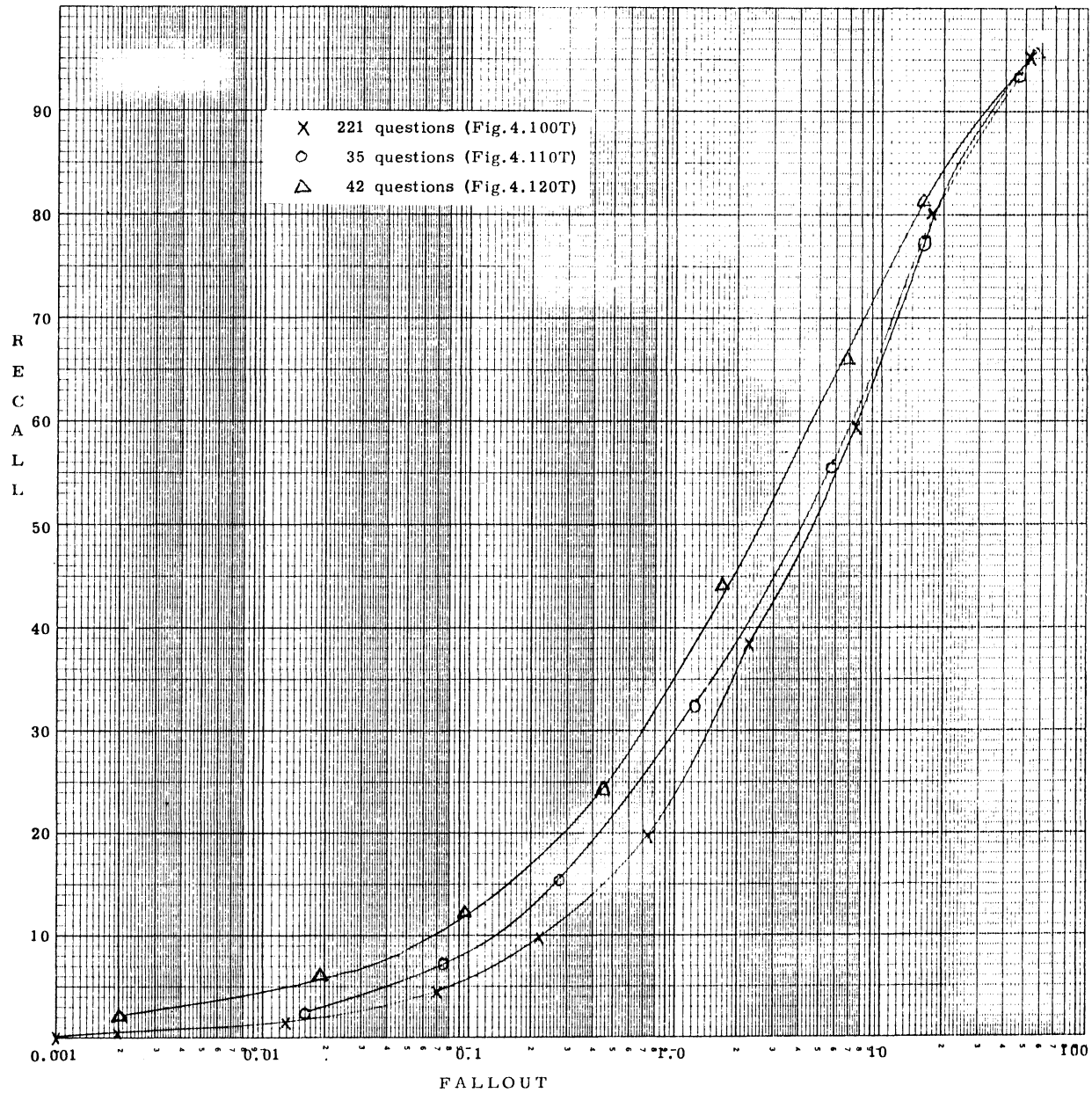
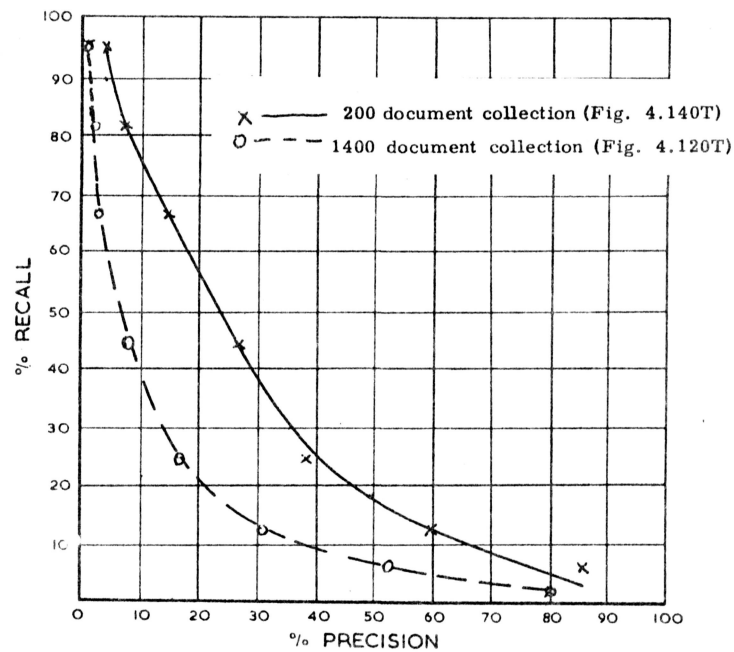


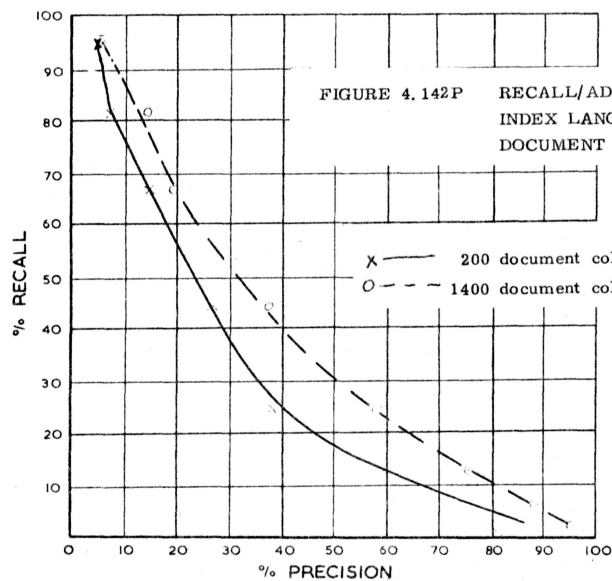
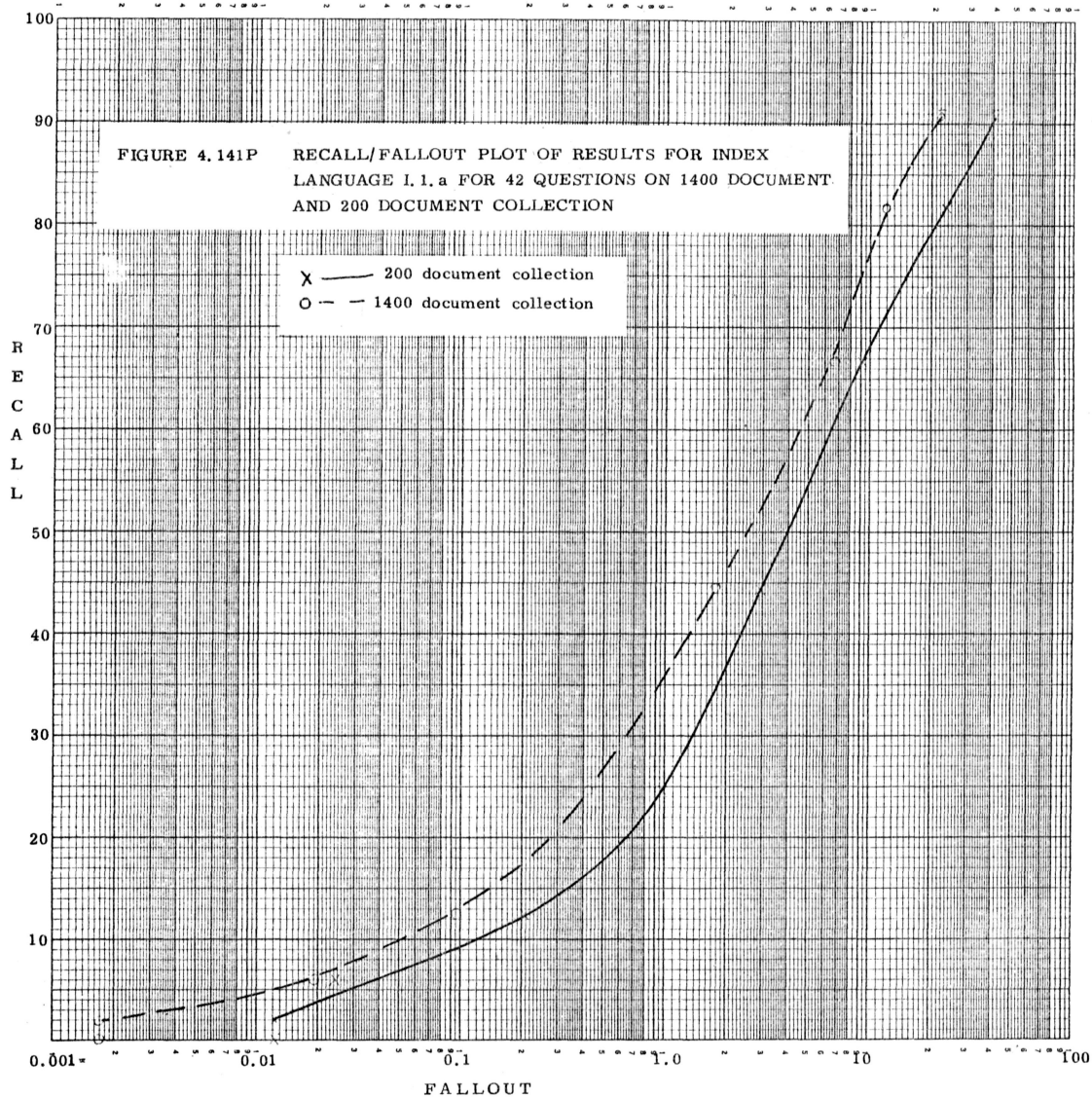
FIGURE 4.131P RECALL/FALLOUT PLOT OF RESULTS FOR INDEX LANGUAGE I.1.a FOR 221 QUESTIONS, 35 QUESTIONS AND 42 QUESTIONS ON 1400 DOCUMENT COLLECTION

FIGURE 4.140T

Index Language I.1.a (S.T. Natural language. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 189 | 4,072 | 95.5% | 4.4% | 49.632% | 42 | 42 | 42 |
| 2 | 162 | 1,929 | 81.8% | 7.7% | 23.519% | 42 | 42 | 42 |
| 3 | 132 | 761 | 66.7% | 14.8% | 9.278% | 42 | 42 | 42 |
| 4 | 88 | 241 | 44.4% | 26.7% | 2.938% | 34 | 41 | 41 |
| 5 | 49 | 80 | 24.7% | 38.0% | 0.975% | 23 | 39 | 39 |
| 6 | 25 | 17 | 12.6% | 59.5% | 0.207% | 15 | 33 | 33 |
| 7 | 12 | 2 | 6.1% | 85.7% | 0.024% | 6 | 27 | 27 |
| 8 | 4 | 1 | 2.0% | 80.0% | 0.012% | 2 | 18 | 18 |
| 9 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |





Section 2 Recall Devices

The tests made on different recall devices on the single term index languages are based on the 42 questions, (subset 2), searched on the 200 document collection. Tables 4.200T to 4.206T give the performance results, and the recall/precision plots given in Figs. 4.200P to 4.206P show the curve for the particular language being tested, together with the curve, (shown by a broken line) obtained by the device-less search in natural language, which was presented in the previous section as Fig. 4.140P.

A combined plot of all the seven devices or their aggregates, together with natural language, is given in Fig. 4.207P.

In presenting these results, all other possible variables of search rules, indexing exhaustivity and document relevance are held constant, and are fixed to the one particular variation indicated at the head of the tables.

LIST OF FIGURES

| | Index Language | No. of Questions | Question Subset | Document Collection | Plots |
|---------|-------------------|---------------------|--------------------|------------------------|---------------|
| 4.200TP | I.2.a | 42 | 2 | 200 | Plot + 4.140T |
| 4.201TP | I.3.a | 42 | 2 | 200 | " " " |
| 4.202TP | I.5.a | 42 | 2 | 200 | " " " |
| 4.203TP | I.6.a | 42 | 2 | 200 | " " " |
| 4.204TP | I.7.a | 42 | 2 | 200 | " " " |
| 4.205TP | I.8.a | 42 | 2 | 200 | " " " |
| 4.206TP | I.9.a | 42 | 2 | 200 | Plot + 4.140T |
| 4.207P | I.1.a | 42 | 2 | 200 | Plot 4.140T, |
| | I.2.a | | | | 4.200T, |
| | I.3.a | | | | 4.201T, |
| | I.5.a | | | | 4.202T, |
| | I.6.a | | | | 4.203T, |
| | I.7.a | | | | 4.204T, |
| | I.8.a | | | | 4.205T, |
| | I.9.a | | | | 4.206T. |

FIGURE 4.200T

Index Language I.2a (S.T. Synonyms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23,6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 189 | 4,193 | 95.5% | 4.3% | 51.147% | 42 | 42 | 42 |
| 2 | 165 | 2,092 | 83.3% | 7.3% | 25.506% | 42 | 42 | 42 |
| 3 | 134 | 837 | 67.7% | 13.8% | 10.205% | 42 | 42 | 42 |
| 4 | 90 | 282 | 45.5% | 24.2% | 3.438% | 34 | 41 | 41 |
| 5 | 55 | 94 | 27.8% | 36.9% | 1.146% | 24 | 39 | 39 |
| 6 | 25 | 23 | 12.6% | 52.1% | 0.280% | 15 | 33 | 33 |
| 7 | 12 | 2 | 6.1% | 85.7% | 0.024% | 6 | 27 | 27 |
| 8 | 4 | 1 | 2.0% | 80.0% | 0.012% | 2 | 18 | 18 |
| 9 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

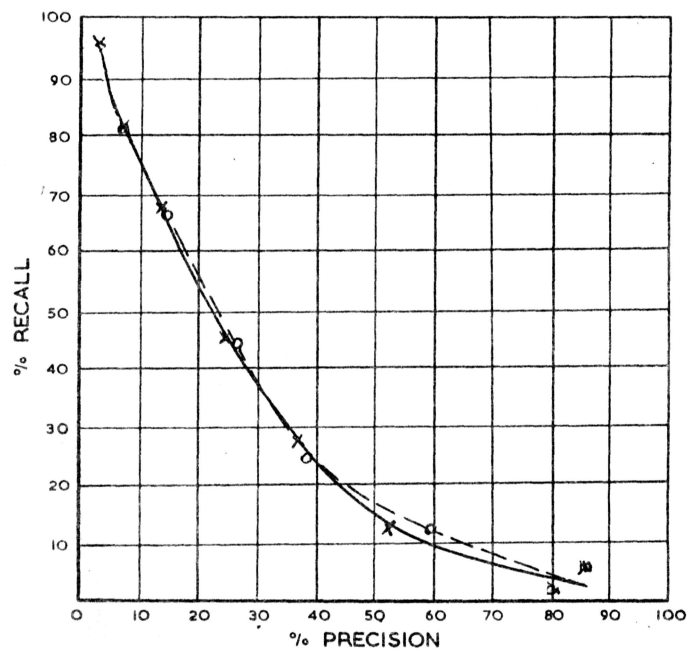


FIGURE 4.201T

Index Language I.3a (S.T. Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord-ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|---------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 191 | 4,416 | 96.5% | 4.1% | 53.826% | 42 | 42 | 42 |
| 2 | 166 | 2,221 | 83.8% | 7.0% | 27.079% | 42 | 42 | 42 |
| 3 | 139 | 913 | 70.2% | 13.2% | 11.131% | 42 | 42 | 42 |
| 4 | 97 | 315 | 49.0% | 23.5% | 3.841% | 35 | 41 | 41 |
| 5 | 59 | 102 | 29.8% | 36.6% | 1.244% | 25 | 39 | 39 |
| 6 | 32 | 28 | 16.2% | 53.3% | 0.341% | 17 | 33 | 33 |
| 7 | 14 | 3 | 7.1% | 32.4% | 0.037% | 8 | 27 | 27 |
| 8 | 4 | 1 | 2.0% | 80.0% | 0.012% | 2 | 18 | 18 |
| 9 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

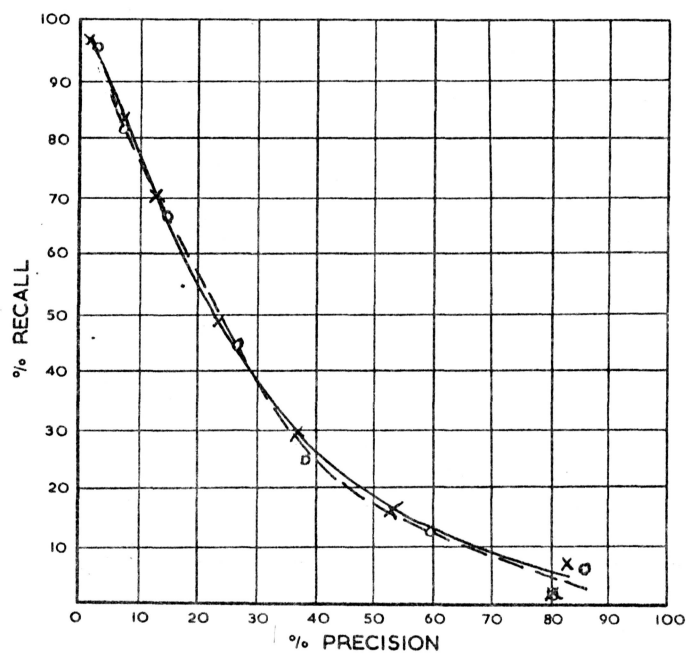


FIGURE 4.202T

Index Language I.5a (S.T. Synonyms, Quasi-synonyms, Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 195 | 6,126 | 98.5% | 3.0% | 74.715% | 42 | 42 | 42 |
| 2 | 177 | 3,565 | 89.4% | 4.7% | 43.465% | 42 | 42 | 42 |
| 3 | 144 | 1,613 | 72.7% | 8.2% | 19.666% | 42 | 42 | 42 |
| 4 | 117 | 660 | 59.1% | 15.1% | 8.047% | 37 | 41 | 41 |
| 5 | 73 | 234 | 36.9% | 23.8% | 2.853% | 30 | 39 | 39 |
| 6 | 41 | 69 | 20.7% | 37.3% | 0.841% | 23 | 33 | 33 |
| 7 | 17 | 14 | 8.6% | 54.8% | 0.171% | 13 | 27 | 27 |
| 8 | 6 | 1 | 3.0% | 85.7% | 0.012% | 3 | 18 | 18 |
| 9 | 5 | 1 | 2.5% | 83.3% | 0.012% | 2 | 11 | 11 |
| 10 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

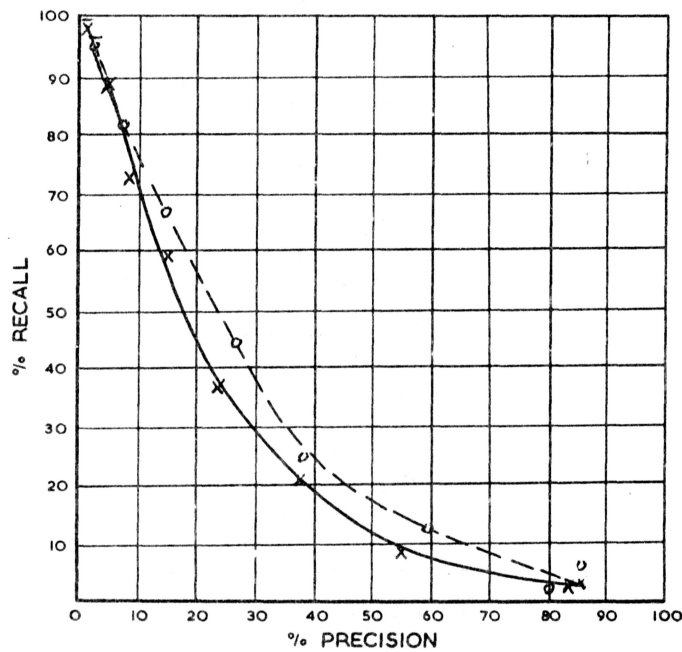


FIGURE 4.203T

Index Language I.6a (S.T. Synonyms, Quasi-synonyms, Word forms, Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 195 | 6,347 | 98.5% | 2.9% | 77.428% | 42 | 42 | 42 |
| 2 | 179 | 3,763 | 90.4% | 4.5% | 45.879% | 42 | 42 | 42 |
| 3 | 151 | 1,785 | 76.3% | 7.8% | 21.763% | 42 | 42 | 42 |
| 4 | 120 | 768 | 60.6% | 13.5% | 9.364% | 37 | 41 | 41 |
| 5 | 79 | 280 | 39.9% | 22.0% | 3.414% | 31 | 39 | 39 |
| 6 | 47 | 87 | 23.7% | 35.1% | 1.061% | 23 | 33 | 33 |
| 7 | 27 | 20 | 13.6% | 57.4% | 0.244% | 15 | 27 | 27 |
| 8 | 7 | 1 | 3.5% | 87.5% | 0.012% | 3 | 18 | 18 |
| 9 | 5 | 1 | 2.5% | 83.3% | 0.012% | 2 | 11 | 11 |
| 10 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

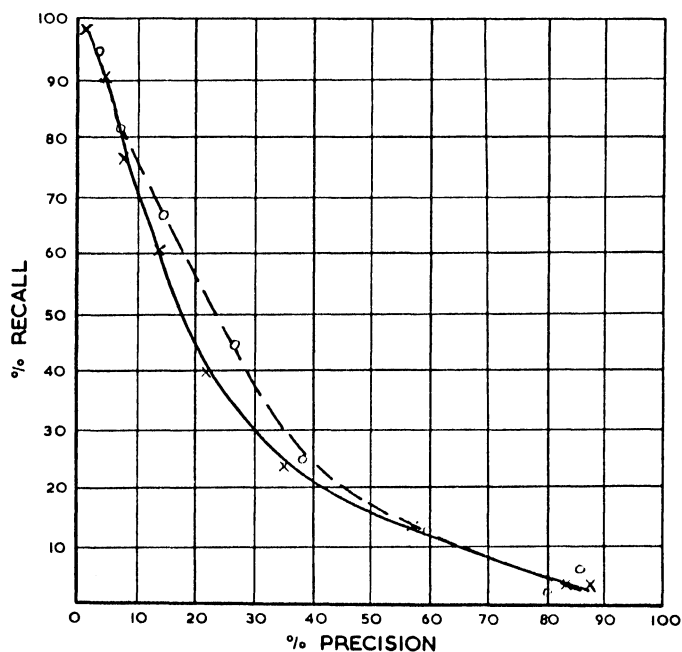


FIGURE 4.204T

Index Language I.7a (S.T. Synonyms, First hierarchical reduction. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 193 | 4,693 | 97.5% | 3.9% | 57.235% | 42 | 42 | 42 |
| 2 | 172 | 2,366 | 86.9% | 6.8% | 28.847% | 42 | 42 | 42 |
| 3 | 143 | 989 | 72.2% | 12.6% | 12.058% | 42 | 42 | 42 |
| 4 | 101 | 353 | 51.0% | 22.2% | 4.304% | 36 | 41 | 41 |
| 5 | 62 | 118 | 31.3% | 34.4% | 1.439% | 27 | 39 | 39 |
| 6 | 31 | 31 | 15.7% | 50.0% | 0.378% | 20 | 33 | 33 |
| 7 | 11 | 4 | 5.6% | 73.3% | 0.049% | 6 | 27 | 27 |
| 8 | 4 | 1 | 2.0% | 80.0% | 0.012% | 2 | 18 | 18 |
| 9 | 1 | 0 | 0.5% | 100.0% | 0.000% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

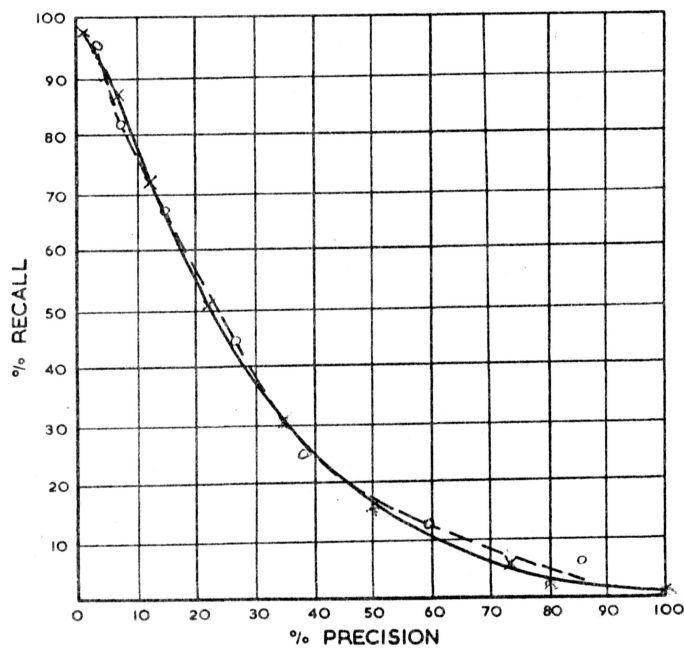


FIGURE 4.205T

Index Language I.8a (S.T. Synonyms, Second hierarchical reduction. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 193 | 4,954 | 97.5% | 3.7% | 60.409% | 42 | 42 | 42 |
| 2 | 173 | 2,551 | 87.4% | 6.4% | 31.102% | 42 | 42 | 42 |
| 3 | 146 | 1,081 | 73.7% | 11.9% | 13.180% | 42 | 42 | 42 |
| 4 | 104 | 418 | 52.5% | 19.9% | 5.096% | 37 | 41 | 41 |
| 5 | 67 | 146 | 33.8% | 31.5% | 1.780% | 27 | 39 | 39 |
| 6 | 33 | 37 | 16.7% | 47.1% | 0.451% | 21 | 33 | 33 |
| 7 | 12 | 7 | 6.1% | 63.2% | 0.085% | 8 | 27 | 27 |
| 8 | 5 | 1 | 2.5% | 83.3% | 0.012% | 2 | 18 | 18 |
| 9 | 2 | 0 | 1.0% | 100.0% | 0.000% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

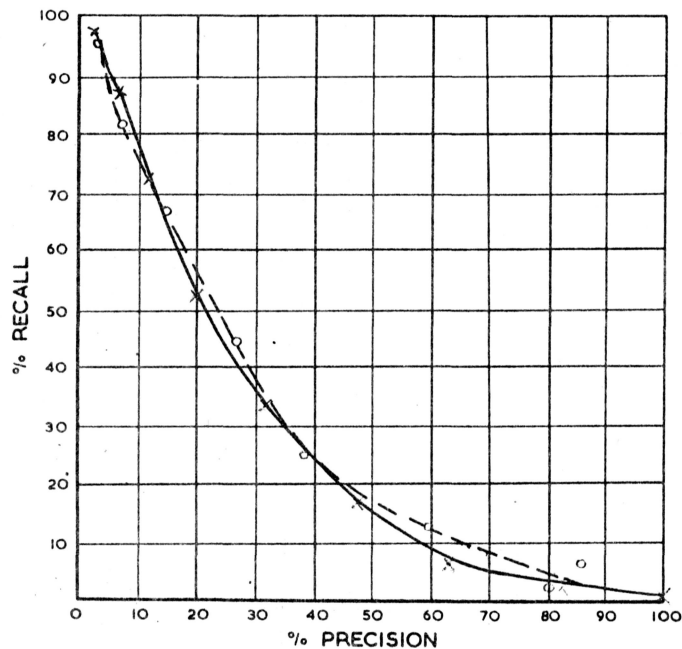


FIGURE 4.206T

Index Language I.9a (S.T. Synonyms, Third hierarchical reduction. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

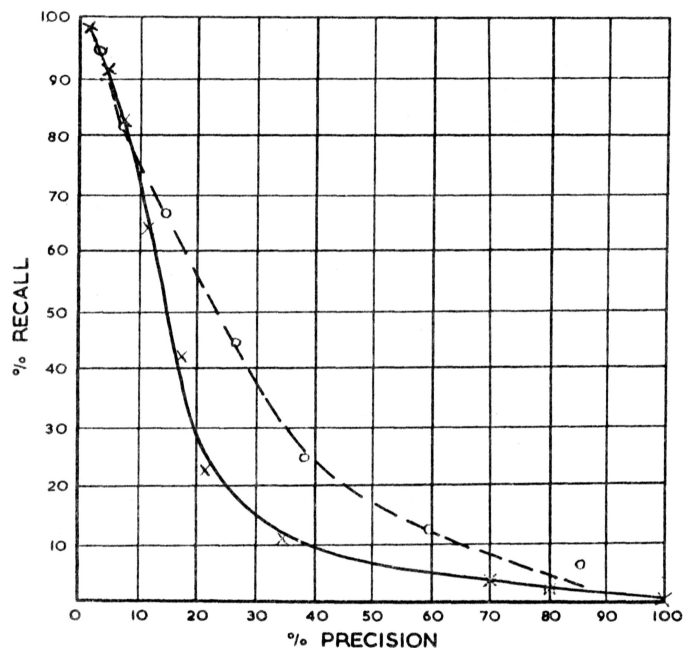
Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 195 | 6,576 | 98.5% | 2.8% | 78.936% | 42 | 42 | 42 |
| 2 | 181 | 3,990 | 91.4% | 4.3% | 48.647% | 42 | 42 | 42 |
| 3 | 163 | 2,034 | 82.3% | 7.4% | 24.799% | 42 | 42 | 42 |
| 4 | 127 | 959 | 64.1% | 11.7% | 11.692% | 39 | 41 | 41 |
| 5 | 83 | 403 | 41.9% | 17.1% | 4.913% | 32 | 39 | 39 |
| 6 | 46 | 171 | 23.2% | 21.2% | 2.085% | 25 | 33 | 33 |
| 7 | 22 | 42 | 11.1% | 34.4% | 0.512% | 14 | 27 | 27 |
| 8 | 7 | 3 | 3.5% | 70.0% | 0.037% | 3 | 18 | 18 |
| 9 | 4 | 1 | 2.0% | 80.0% | 0.012% | 2 | 11 | 11 |
| 10 | 1 | 0 | 0.5% | 100.0% | 0.000% | 1 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |



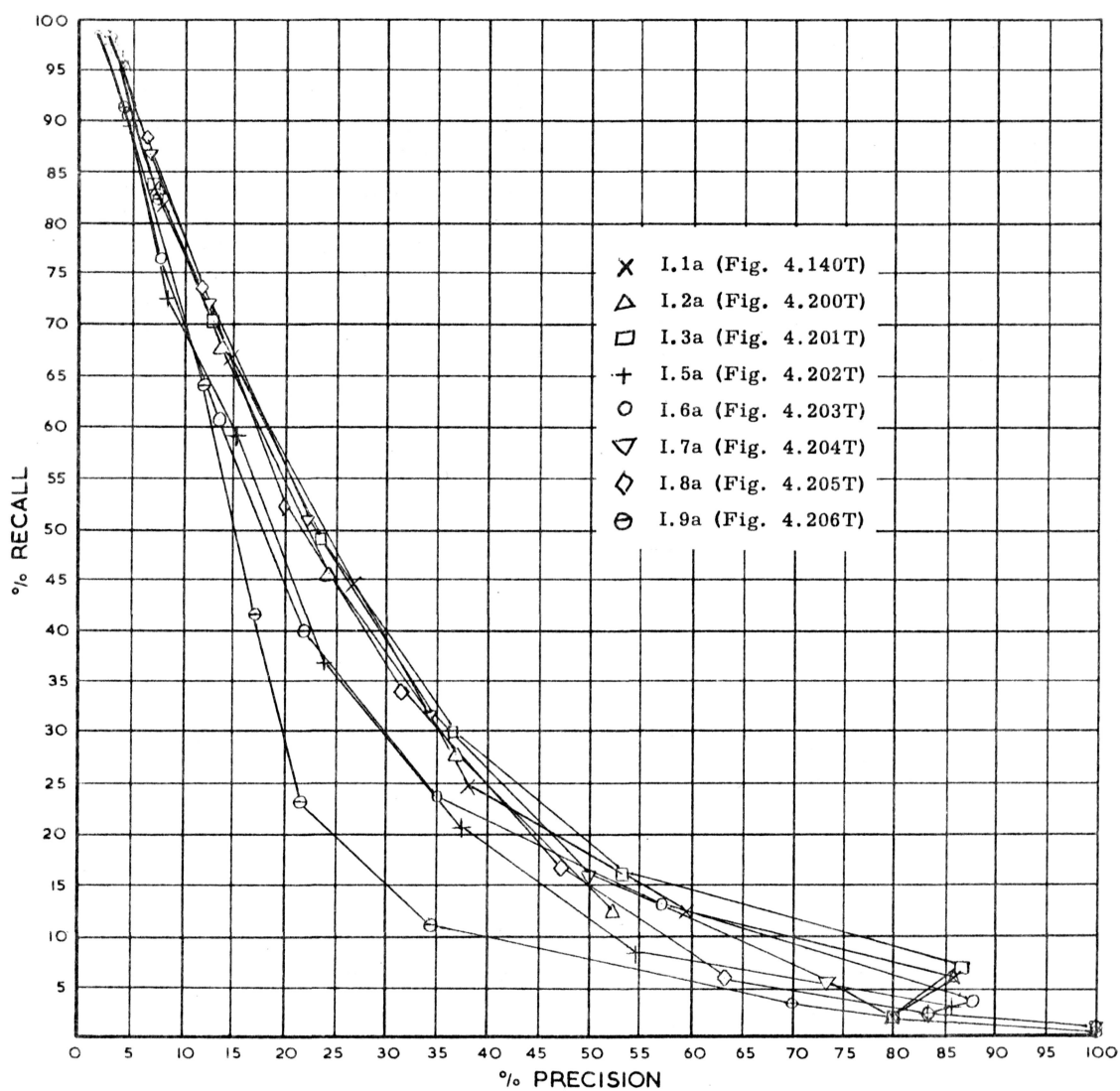


FIGURE 4.207P PLOT OF RESULTS FOR INDEX LANGUAGES I.1a, I.2a, I.3a, I.5a, I.6a, I.7a, I.8a and I.9a for 42 QUESTIONS ON 200 DOCUMENT COLLECTION.

Section 3 Precision Devices

Two subsets of questions have been used to investigate precision devices. Subset 4 consisted of 19 questions all having seven starting terms, while Subset 5 had 17 questions all having eleven starting terms. For reasons explained in Chapter 2, these tests made on precision devices required a different search rule (B) to what has been used so far. This rule required that the main term in a concept must be present if the qualifying term was to be accepted, e.g. with the concept 'Hinged flap', 'Hinged' would not be accepted unless 'Flap' was also present. Therefore it is necessary first to present the results of searches made on these two subsets with search rule B, so that later comparison can be made. Each subset is tested with Index Language I.1.a and I.6.a.

There now follow the results of tests with (b) partitioning, (Figs. 4.310T - 4.313T), (c) interfixing, (Figs. 4.320T - 4.323T) and (a) partitioning plus interfixing (Figs. 4.330T - 4.333T), with both subsets of questions presented separately. For each precision device results are given when two recall languages are used, I.1 Natural Language and I.6. the aggregate of synonyms, quasi-synonyms and word forms.

Because of the large clerical effort required to obtain these results, the precision ratios were not obtained at the lower levels of coordination. This, combined with the relatively small sets of questions used, limits the usefulness of any graphical presentation, for the results show little consistency. Therefore in the plots 4.340P, 4.341P, 4.342P and 4.343P, only a single generalised curve has been drawn for the four sets of data that are presented in each curve. The relative positions of the various symbols give an indication of the performance of the particular system.

LIST OF FIGURES

| Precision Device | Index Language | No. of Questions | Question Subset | Document Collection |
|------------------|----------------|------------------|-----------------|---------------------|
| 4.300T | I.1.a | 19 | 4 | 1400 (Search |
| 4.301T | I.1.a | 17 | 5 | 1400 Rule B) |
| 4.302T | I.6.a | 19 | 4 | 1400 |
| 4.303T | I.6.a | 17 | 5 | 1400 |
| 4.310T | PARTITIONING | I.1.b | 19 | 1400 |
| 4.311T | " | I.1.b | 17 | 1400 |

List of Figures (Cont)

| | Precision Device | Index Language | No. of Questions | Question Subset | Document Collection | Plots |
|--------|--------------------------------|-------------------|---------------------|--------------------|------------------------|--|
| 4.312T | PARTITIONING | I.6.b | 19 | 4 | 1400 | |
| 4.313T | " | I.6.b | 17 | 5 | 1400 | |
| 4.320T | INTERFIXING | I.1.c | 19 | 4 | 1400 | |
| 4.321T | " | I.1.c | 17 | 5 | 1400 | |
| 4.322T | " | I.6.c | 19 | 4 | 1400 | |
| 4.323T | " | I.6.c | 17 | 5 | 1400 | |
| 4.330T | PART. + INT. | I.1.d | 19 | 4 | 1400 | |
| 4.331T | " | I.1.d | 17 | 5 | 1400 | |
| 4.332T | " | I.6.d | 19 | 4 | 1400 | |
| 4.333T | " | I.6.d | 17 | 5 | 1400 | |
| 4.340P | PART., INT., & PART. + INT. | I.1.a,b, c,d | 19 | 4 | 1400 | Plot 4.300T, 4.310T, 4.320T, 4.330T. |
| 4.341P | " | I.1.a,b, c,d | 17 | 5 | 1400 | Plot 4.301T, 4.311T, 4.321T, 4.331T. |
| 4.342P | " | I.6.a,b, c,d | 19 | 4 | 1400 | Plot 4.302T, 4.312T, 4.322T, 4.332T. |
| 4.343P | " | I.6.a,b, c,d | 17 | 5 | 1400 | Plot 4.303T, 4.313T, 4.323T, 4.333T. |

FIGURE 4.300T

Index Language I.1a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule B

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 19 (Subset 4)

Number of Relevant Documents 131

Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 113 | (-) | 86.3% | (-) | (-) | 19 | 0 | 19 |
| 2 | 100 | (-) | 76.3% | (-) | (-) | 19 | 0 | 19 |
| 3 | 68 | 826* | 51.9% | 7.6%* | 3.105%* | 19 | 7* | 19 |
| 4 | 37 | 127 | 28.2% | 22.6% | 0.480% | 15 | 19 | 19 |
| 5 | 16 | 9 | 12.2% | 64.0% | 0.034% | 6 | 19 | 19 |
| 6 | 11 | 0 | 8.4% | 100.0% | 0.000% | 3 | 19 | 19 |
| 7 | 4 | 0 | 3.1% | 100.0% | 0.000% | 2 | 19 | 19 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.301T

Index Language I.1a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule B

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 17 (Subset 5)

Number of Relevant Documents 109

Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 101 | (-) | 92.7% | (-) | (-) | 17 | 0 | 17 |
| 3 | 88 | (-) | 80.7% | (-) | (-) | 17 | 0 | 17 |
| 4 | 66 | 1301 | 60.5% | 4.8% | 5.489 | 17 | 17 | 17 |
| 5 | 51 | 456 | 46.7% | 10.0% | 1.925% | 17 | 17 | 17 |
| 6 | 37 | 178 | 33.9% | 17.2% | 0.751% | 15 | 17 | 17 |
| 7 | 21 | 75 | 19.2% | 21.9% | 0.317% | 12 | 17 | 17 |
| 8 | 10 | 0 | 9.2% | 100.0% | 0.000% | 7 | 17 | 17 |
| 9 | 5 | 0 | 4.6% | 100.0% | 0.000% | 4 | 17 | 17 |
| 10 | 0 | 0 | | | | 0 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.302T

Index Language I.6a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 19 (Subset 4)
 Number of Relevant Documents 131
 Generality Number 4.8

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 120 | (-) | 91.6% | (-) | (-) | 19 | 0 | 19 |
| 2 | 113 | (-) | 86.3% | (-) | (-) | 19 | 0 | 19 |
| 3 | 93 | 2,457* | 71.0% | 3.6%* | 9.237%* | 19 | 7* | 19 |
| 4 | 58 | 765 | 44.3% | 7.0% | 2.890% | 19 | 19 | 19 |
| 5 | 29 | 110 | 22.1% | 20.9% | 0.416% | 15 | 19 | 19 |
| 6 | 17 | 29 | 13.0% | 37.0% | 0.110% | 7 | 19 | 19 |
| 7 | 8 | 7 | 6.1% | 53.3% | 0.026% | 3 | 19 | 19 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.303T

Index Language I.6a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 17 (Subset 5)
 Number of Relevant Documents 109
 Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 105 | (-) | 96.3% | (-) | (-) | 17 | 0 | 17 |
| 3 | 97 | (-) | 88.9% | (-) | (-) | 17 | 0 | 17 |
| 4 | 85 | (-) | 77.9% | (-) | (-) | 17 | 0 | 17 |
| 5 | 68 | 1,495 | 62.3% | 4.4% | 6.310% | 17 | 17 | 17 |
| 6 | 52 | 719 | 47.7% | 6.7% | 3.035% | 17 | 17 | 17 |
| 7 | 35 | 211 | 32.1% | 14.2% | 0.891% | 16 | 17 | 17 |
| 8 | 18 | 116 | 16.5% | 13.4% | 0.490% | 14 | 17 | 17 |
| 9 | 7 | 25 | 6.4% | 21.8% | 0.106% | 7 | 17 | 17 |
| 10 | 1 | 6 | 0.9% | 14.3% | 0.025% | 1 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.310T

Index Language I.1b (S.T. Natural language. Coordination, Partitioning)

Exhaustivity of Indexing 3

Search Rule B

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 19 (Subset 4)

Number of Relevant Documents 131

Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 113 | (-) | 86.3% | (-) | (-) | 19 | 0 | 19 |
| 2 | 93 | (-) | 71.0% | (-) | (-) | 19 | 0 | 19 |
| 3 | 60 | 496* | 45.8% | 10.7%* | 10.876%* | 19 | 7* | 19 |
| 4 | 26 | 75 | 19.8% | 25.7% | 0.283% | 12 | 19 | 19 |
| 5 | 9 | 5 | 6.9% | 64.3% | 0.019% | 2 | 19 | 19 |
| 6 | 7 | 0 | 5.3% | 100.0% | 0.000% | 2 | 19 | 19 |
| 7 | 4 | 0 | 3.1% | 100.0% | 0.000% | 2 | 19 | 19 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.311T

Index Language I.1b (S.T. Natural language. Coordination, Partitioning)

Exhaustivity of Indexing 3

Search Rule B

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 17 (Subset 5)

Number of Relevant Documents 109

Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 97 | (-) | 89.0% | (-) | (-) | 17 | 0 | 17 |
| 3 | 79 | (-) | 72.5% | (-) | (-) | 17 | 0 | 17 |
| 4 | 55 | (-) | 50.5% | (-) | (-) | 17 | 0 | 17 |
| 5 | 35 | 228 | 32.1% | 15.4% | 0.962% | 15 | 17 | 17 |
| 6 | 20 | 44 | 18.3% | 31.3% | 0.186% | 11 | 17 | 17 |
| 7 | 10 | 0 | 9.2% | 100.0% | 0.000% | 6 | 17 | 17 |
| 8 | 6 | 0 | 5.5% | 100.0% | 0.000% | 4 | 17 | 17 |
| 9 | 0 | 0 | | | | 0 | 17 | 17 |
| 10 | 0 | 0 | | | | 0 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.312T

Index Language 1.6b (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination, Partitioning)
 Exhaustivity of Indexing 3
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 19 (Subset 4)
 Number of Relevant Documents 131
 Generality Number 4.9

| Coord-ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|---------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 120 | (-) | 91.6% | (-) | (-) | 19 | 0 | 19 |
| 2 | 109 | (-) | 83.2% | (-) | (-) | 19 | 0 | 19 |
| 3 | 79 | 1,164* | 60.3% | 6.4%* | 4.375%* | 19 | 7* | 19 |
| 4 | 39 | 321 | 29.8% | 10.8% | 1.213% | 16 | 19 | 19 |
| 5 | 17 | 30 | 13.0% | 36.2% | 0.113% | 11 | 19 | 19 |
| 6 | 8 | 5 | 6.1% | 61.5% | 0.019% | 4 | 19 | 19 |
| 7 | 5 | 0 | 3.8% | 100.0% | 0.000% | 3 | 19 | 19 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.313T

Index Language 1.6b (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination, Partiti
 Exhaustivity of Indexing 3
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 17 (Subset 5)
 Number of Relevant Documents 109
 Generality Number 4.6

| Coord-ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|---------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 104 | (-) | 95.4% | (-) | (-) | 17 | 0 | 17 |
| 3 | 90 | (-) | 82.6% | (-) | (-) | 17 | 0 | 17 |
| 4 | 72 | (-) | 66.1% | (-) | (-) | 17 | 0 | 17 |
| 5 | 54 | 702 | 49.5% | 7.1% | 2.963% | 16 | 17 | 17 |
| 6 | 32 | 226 | 29.4% | 12.4% | 0.954% | 14 | 17 | 17 |
| 7 | 17 | 37 | 15.6% | 31.5% | 0.156% | 10 | 17 | 17 |
| 8 | 7 | 19 | 6.4% | 26.9% | 0.080% | 4 | 17 | 17 |
| 9 | 2 | 0 | 1.8% | 100.0% | 0.000% | 1 | 17 | 17 |
| 10 | 1 | 0 | 0.9% | 100.0% | 0.000% | 1 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.320T

Index Language I.1c (S.T. Natural language. Coordination, Interfixing)

Exhaustivity of Indexing 3

Search Rule B

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 19 (Subset 4)

Number of Relevant Documents 131

Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 113 | (-) | 86.3% | (-) | (-) | 19 | 0 | 19 |
| 2 | 92 | (-) | 70.2% | (-) | (-) | 19 | 0 | 19 |
| 3 | 51 | 358* | 38.9% | 12.5%* | 1.346%* | 19 | 7* | 19 |
| 4 | 25 | 54 | 19.1% | 31.6% | 0.204% | 12 | 19 | 19 |
| 5 | 8 | 0 | 6.1% | 100.0% | 0.000% | 2 | 19 | 19 |
| 6 | 4 | 0 | 3.1% | 100.0% | 0.000% | 1 | 19 | 19 |
| 7 | 3 | 0 | 2.3% | 100.0% | 0.000% | 1 | 19 | 19 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.321T

Index Language I.1c (S.T. Natural language. Coordination, Interfixing)

Exhaustivity of Indexing 3

Search Rule B

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 17 (Subset 5)

Number of Relevant Documents 109

Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 97 | (-) | 89.0% | (-) | (-) | 17 | 0 | 17 |
| 3 | 82 | (-) | 75.2% | (-) | (-) | 17 | 0 | 17 |
| 4 | 62 | (-) | 56.9% | (-) | (-) | 17 | 0 | 17 |
| 5 | 48 | 296 | 44.0% | 14.0% | 1.249% | 17 | 17 | 17 |
| 6 | 26 | 118 | 23.9% | 18.1% | 0.498% | 14 | 17 | 17 |
| 7 | 12 | 25 | 11.0% | 32.4% | 0.106% | 8 | 17 | 17 |
| 8 | 7 | 0 | 6.4% | 100.0% | 0.000% | 4 | 17 | 17 |
| 9 | 1 | 0 | 0.9% | 100.0% | 0.000% | 1 | 17 | 17 |
| 10 | 0 | 0 | | | | 0 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.322T

Index Language I.6c (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination, Exhaustivity of Indexing 3 Interfixing)
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 19 (Subset 4)
 Number of Relevant Documents 131
 Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 120 | (-) | 91.6% | (-) | (-) | 19 | 0 | 19 |
| 2 | 107 | (-) | 81.7% | (-) | (-) | 19 | 0 | 19 |
| 3 | 72 | 1.019* | 55.0% | 6.6%* | 3.850%* | 18 | 7* | 19 |
| 4 | 38 | 312 | 29.0% | 10.9% | 1.179% | 17 | 19 | 19 |
| 5 | 14 | 25 | 10.7% | 35.9% | 0.094% | 9 | 19 | 19 |
| 6 | 4 | 5 | 3.1% | 44.4% | 0.019% | 2 | 19 | 19 |
| 7 | 3 | 0 | 2.3% | 100.0% | 0.000% | 1 | 19 | 10 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.323T

Index Language I.6c (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination, Inte Exhaustivity of Indexing 3
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 17 (Subset 5)
 Number of Relevant Documents 109
 Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 106 | (-) | 97.2% | (-) | (-) | 17 | 0 | 17 |
| 3 | 96 | (-) | 88.1% | (-) | (-) | 17 | 0 | 17 |
| 4 | 75 | (-) | 68.8% | (-) | (-) | 17 | 0 | 17 |
| 5 | 60 | 793 | 55.0% | 7.0% | 3.347% | 17 | 17 | 17 |
| 6 | 38 | 268 | 34.9% | 12.4% | 1.131% | 15 | 17 | 17 |
| 7 | 18 | 62 | 16.5% | 22.5% | 0.262% | 12 | 17 | 17 |
| 8 | 9 | 58 | 8.3% | 13.4% | 0.245% | 5 | 17 | 17 |
| 9 | 4 | 8 | 3.7% | 33.3% | 0.034% | 2 | 17 | 17 |
| 10 | 1 | 0 | 0.9% | 100.0% | 0.000% | 1 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.330T

Index Language I.1d (S.T. Natural language, Coordination, Partitioning, Interfixing)
 Exhaustivity of Indexing 3
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 19 (Subset 4)
 Number of Relevant Documents 131
 Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 113 | (-) | 86.3% | (-) | (-) | 19 | 0 | 19 |
| 2 | 88 | (-) | 67.2% | (-) | (-) | 19 | 0 | 19 |
| 3 | 50 | 223* | 38.2% | 18.3%* | 0.842%* | 19 | 7* | 19 |
| 4 | 16 | 36 | 12.2% | 30.8% | 0.136% | 10 | 19 | 19 |
| 5 | 7 | 0 | 5.3% | 100.0% | 0.000% | 3 | 19 | 19 |
| 6 | 4 | 0 | 3.1% | 100.0% | 0.000% | 1 | 19 | 19 |
| 7 | 3 | 0 | 2.3% | 100.0% | 0.000% | 1 | 19 | 19 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.331T

Index Language I.1d (S.T. Natural language. Coordination, Partitioning, Interfixing)
 Exhaustivity of Indexing 3
 Search Rule B
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 17 (Subset 5)
 Number of Relevant Documents 109
 Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 96 | (-) | 88.1% | (-) | (-) | 17 | 0 | 17 |
| 3 | 76 | (-) | 69.7% | (-) | (-) | 17 | 0 | 17 |
| 4 | 53 | (-) | 48.6% | (-) | (-) | 17 | 0 | 17 |
| 5 | 32 | 217 | 29.4% | 12.9% | 0.916% | 14 | 17 | 17 |
| 6 | 16 | 74 | 14.7% | 17.8% | 0.312% | 10 | 17 | 17 |
| 7 | 8 | 0 | 7.3% | 100.0% | 0.000% | 5 | 17 | 17 |
| 8 | 5 | 0 | 4.6% | 100.0% | 0.000% | 3 | 17 | 17 |
| 9 | 0 | 0 | | | | 0 | 17 | 17 |
| 10 | 0 | 0 | | | | 0 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.332T

Index Language 1.6d (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination,
Exhaustivity of Indexing 3 Partitioning, Interfixing)
Search Rule B
Document Relevance 1 - 4
Number of Documents in Collection 1,400
Number of Questions 19 (Subset 4)
Number of Relevant Documents 131
Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 120 | (-) | 91.6% | (-) | (-) | 19 | 0 | 19 |
| 2 | 102 | (-) | 77.7% | (-) | (-) | 19 | 0 | 19 |
| 3 | 64 | 534* | 48.9% | 10.7%* | 2.017%* | 19 | 7* | 19 |
| 4 | 27 | 148 | 20.6% | 15.4% | 0.559% | 12 | 19 | 19 |
| 5 | 9 | 11 | 6.9% | 45.0% | 0.042% | 7 | 19 | 19 |
| 6 | 4 | 5 | 3.1% | 44.4% | 0.019% | 2 | 19 | 19 |
| 7 | 3 | 0 | 2.3% | 100.0% | 0.000% | 1 | 19 | 19 |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.333T

Index Language 1.6d (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination, Partit:
Exhaustivity of Indexing 3 Interf:
Search Rule B
Document Relevance 1 - 4
Number of Documents in Collection 1,400
Number of Questions 17 (Subset 5)
Number of Relevant Documents 109
Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | (-) | 99.1% | (-) | (-) | 17 | 0 | 17 |
| 2 | 104 | (-) | 95.4% | (-) | (-) | 17 | 0 | 17 |
| 3 | 86 | (-) | 78.9% | (-) | (-) | 17 | 0 | 17 |
| 4 | 67 | (-) | 61.5% | (-) | (-) | 17 | 0 | 17 |
| 5 | 45 | 416 | 41.3% | 9.8% | 1.756% | 15 | 17 | 17 |
| 6 | 23 | 141 | 21.1% | 14.0% | 0.595% | 13 | 17 | 17 |
| 7 | 12 | 25 | 11.0% | 32.4% | 0.106% | 7 | 17 | 17 |
| 8 | 6 | 19 | 5.5% | 24.0% | 0.080% | 3 | 17 | 17 |
| 9 | 2 | 0 | 1.8% | 100.0% | 0.000% | 1 | 17 | 17 |
| 10 | 1 | 0 | 0.9% | 100.0% | 0.000% | 1 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

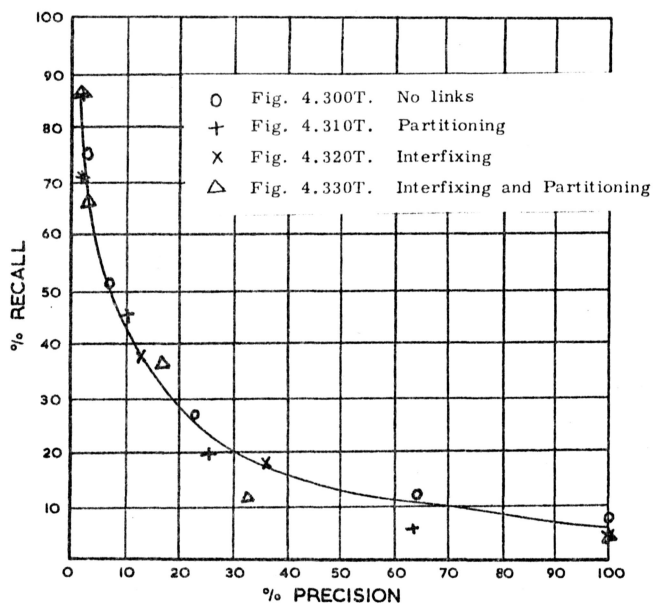


FIGURE 4.340P EFFECT OF LINKING DEVICES FOR INDEX LANGUAGE I.1 WITH 19 QUESTIONS

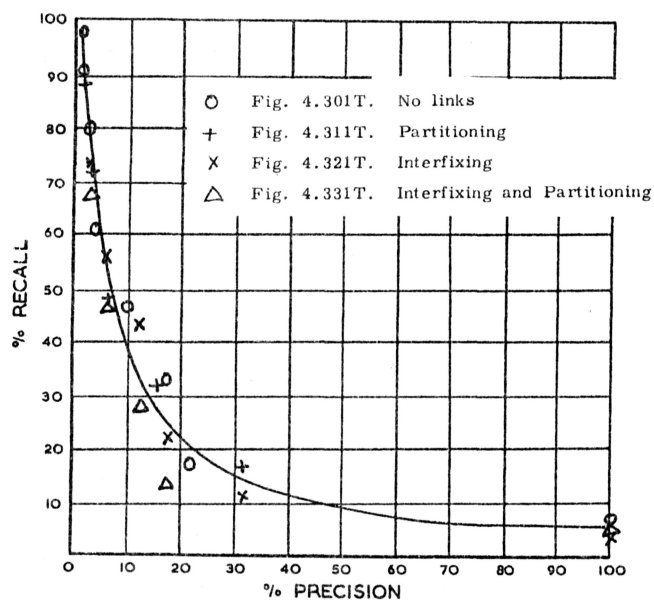


FIGURE 4.341P EFFECT OF LINKING DEVICES FOR INDEX LANGUAGE I.1 WITH 17 QUESTIONS

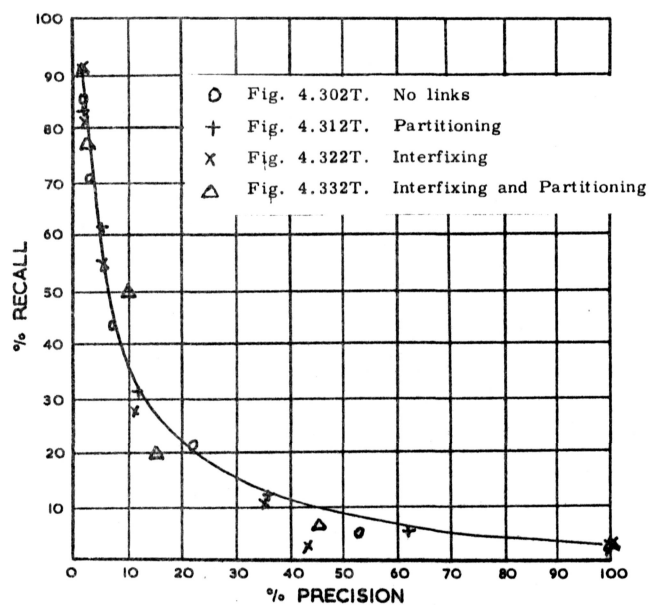


FIGURE 4.342P EFFECT OF LINKING DEVICES FOR INDEX LANGUAGE I.6 WITH 19 QUESTIONS

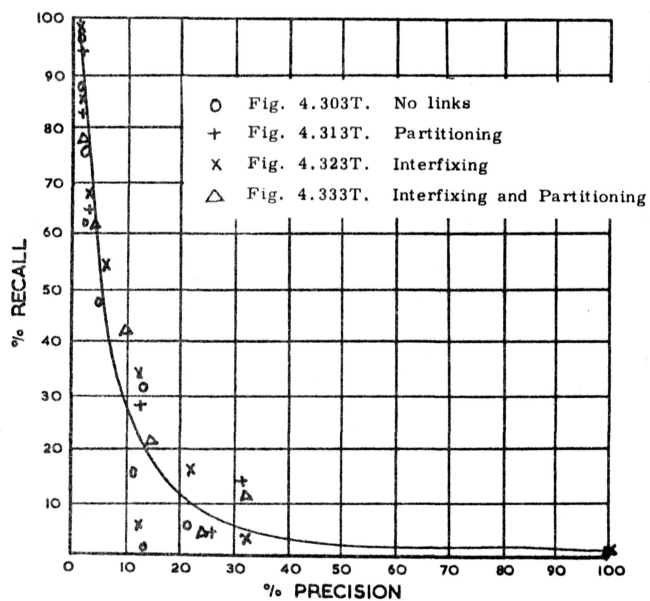


FIGURE 4.343P EFFECT OF LINKING DEVICES FOR INDEX LANGUAGE I.6 WITH 17 QUESTIONS

Section 4 Indexing Exhaustivity

Results when the three different levels of indexing exhaustivity are compared are given in Figs. 4.400T - 4.405T and 4.410T - 4.415T. Three sets of questions are presented separately, being the 221 questions (subset 3) searched on the 1400 collection, the 35 seven starting term questions (subset 1) searched on the 1400 collection, and the 42 questions (subset 2) searched on the 200 collection. Results for index languages I.1.a and I.6.a are given, and the three different performance curves for each question subset and language are given in Figs. 4.420P - 4.425P. No adjustment has been made in these graphs for the differing generality number amongst the different plots, since only the relative position of the curves on each plot is important.

LIST OF FIGURES

| | Exhaustivity | Index Language | No. of Questions | Question Subset | Document Collection | Plots |
|--------|--------------|----------------|------------------|-----------------|---------------------|---------------------------------|
| 4.400T | 2 | I.1.a | 221 | 3 | 1400 | |
| 4.401T | 1 | I.1.a | 221 | 3 | 1400 | |
| 4.402T | 2 | I.1.a | 35 | 1 | 1400 | |
| 4.403T | 1 | I.1.a | 35 | 1 | 1400 | |
| 4.404T | 2 | I.1.a | 42 | 2 | 200 | |
| 4.405T | 1 | I.1.a | 42 | 2 | 200 | |
| 4.410T | 2 | I.6.a | 221 | 3 | 1400 | |
| 4.411T | 1 | I.6.a | 221 | 3 | 1400 | |
| 4.412T | 2 | I.6.a | 35 | 1 | 1400 | |
| 4.413T | 1 | I.6.a | 35 | 1 | 1400 | |
| 4.414T | 2 | I.6.a | 42 | 2 | 200 | |
| 4.415T | 1 | I.6.a | 42 | 2 | 200 | |
| 4.420P | 3, 2, 1 | I.1.a | 221 | 3 | 1400 | Plot 4.100T 4.400T 4.401T |
| 4.421P | 3, 2, 1 | I.1.a | 35 | 1 | 1400 | Plot 4.110T 4.402T 4.403T |
| 4.422P | 3, 2, 1 | I.1.a | 42 | 2 | 200 | Plot 4.140T 4.404T 4.405T |
| 4.423P | 3, 2, 1 | I.6.a | 221 | 3 | 1400 | Plot 4.104T 4.410T 4.411T |
| 4.424P | 3, 2, 1 | I.6.a | 35 | 1 | 1400 | Plot 4.114T 4.412T 4.414T |
| 4.425P | 3, 2, 1 | I.6.a | 42 | 2 | 200 | Plot 4.203T 4.414T 4.415T |

FIGURE 4.400T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 2

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 221 (Subset 3)

Number of Relevant Documents 1,590

Generality Number 5.1

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,480 | (-) | 93.1% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,227 | 49,372* | 77.2% | 2.4%* | 16.040%* | 221 | 44* | 221 |
| 3 | 886 | 18,764* | 55.7% | 4.5%* | 6.092%* | 210 | 109* | 220 |
| 4 | 552 | 6,813* | 34.7% | 7.5%* | 2.212%* | 177 | 142* | 212 |
| 5 | 279 | 2,301* | 17.5% | 10.8%* | 0.747%* | 121 | 177* | 197 |
| 6 | 131 | 514* | 8.2% | 20.3%* | 0.167%* | 76 | 161* | 164 |
| 7 | 58 | 133 | 3.6% | 30.4% | 0.043% | 41 | 140 | 140 |
| 8 | 18 | 28 | 1.1% | 39.1% | 0.009% | 14 | 105 | 105 |
| 9 | 8 | 0 | 0.5% | 100.0% | 0.000% | 6 | 78 | 78 |
| 10 | 1 | 0 | 0.1% | 100.0% | 0.000% | 1 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.401T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 1

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 221 (Subset 3)

Number of Relevant Documents 1,590

Generality Number 5.1

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,373 | (-) | 86.4% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,055 | 23,164* | 66.4% | 4.4%* | 7.525%* | 221 | 44* | 221 |
| 3 | 655 | 7,823* | 41.2% | 7.7%* | 2.540%* | 199 | 109* | 220 |
| 4 | 331 | 2,169* | 20.8% | 13.2%* | 0.704%* | 154 | 142* | 212 |
| 5 | 148 | 477* | 9.3% | 23.7%* | 0.155%* | 88 | 177* | 197 |
| 6 | 67 | 134* | 4.2% | 33.2%* | 0.043%* | 49 | 161* | 164 |
| 7 | 29 | 38 | 1.8% | 43.3% | 0.012% | 25 | 140 | 140 |
| 8 | 7 | 9 | 0.4% | 43.8% | 0.003% | 9 | 105 | 105 |
| 9 | 2 | 0 | 0.1% | 100.0% | 0.000% | 2 | 78 | 78 |
| 10 | 0 | 0 | | | | 0 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.402T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 2

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 35 (Subset 1)

Number of Relevant Documents 287

Generality Number 5.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 263 | (-) | 91.7% | (-) | (-) | 35 | 0 | 35 |
| 2 | 211 | 6,823* | 73.6% | 3.0%* | 13.924%* | 35 | 23* | 35 |
| 3 | 141 | 2,442 | 49.2% | 5.4% | 5.013% | 34 | 35 | 35 |
| 4 | 85 | 488 | 29.7% | 15.0% | 1.002% | 25 | 35 | 35 |
| 5 | 43 | 94 | 15.0% | 31.4% | 0.193% | 17 | 35 | 35 |
| 6 | 23 | 20 | 8.0% | 53.5% | 0.041% | 6 | 35 | 35 |
| 7 | 8 | 8 | 2.8% | 50.0% | 0.016% | 3 | 35 | 35 |

FIGURE 4.403T

Index Language I.1.a (S.T. Natural language. Coordination.)

Exhaustivity of Indexing 1

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 35 (Subset 1)

Number of Relevant Documents 287

Generality Number 5.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 237 | (-) | 82.6% | (-) | (-) | 35 | 0 | 35 |
| 2 | 184 | 3,532* | 64.1% | 4.9%* | 7.208%* | 35 | 23* | 35 |
| 3 | 111 | 1,045 | 38.7% | 9.6% | 2.145% | 34 | 35 | 35 |
| 4 | 54 | 168 | 18.9% | 24.3% | 0.345% | 22 | 35 | 35 |
| 5 | 22 | 19 | 7.7% | 53.7% | 0.039% | 6 | 35 | 35 |
| 6 | 10 | 2 | 3.5% | 83.3% | 0.004% | 5 | 35 | 35 |
| 7 | 5 | 2 | 1.8% | 71.4% | 0.004% | 3 | 35 | 35 |

FIGURE 4.404T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 2

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 186 | 3,401 | 93.9% | 5.2% | 40.488% | 42 | 42 | 42 |
| 2 | 157 | 1,605 | 79.3% | 8.9% | 19.568% | 42 | 42 | 42 |
| 3 | 120 | 582 | 60.6% | 17.1% | 7.096% | 40 | 42 | 42 |
| 4 | 81 | 164 | 40.9% | 33.1% | 2.000% | 32 | 41 | 41 |
| 5 | 39 | 46 | 19.7% | 45.9% | 0.561% | 20 | 39 | 39 |
| 6 | 22 | 10 | 11.1% | 68.8% | 0.122% | 12 | 33 | 33 |
| 7 | 10 | 1 | 5.1% | 90.9% | 0.012% | 3 | 27 | 27 |
| 8 | 3 | 0 | 1.5% | 100.0% | 0.000% | 2 | 18 | 18 |
| 9 | 0 | 0 | | | | 0 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.405T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 1

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 175 | 2,426 | 88.4% | 6.7% | 28.880% | 42 | 42 | 42 |
| 2 | 132 | 942 | 66.7% | 12.3% | 11.485% | 42 | 42 | 42 |
| 3 | 98 | 240 | 49.5% | 29.0% | 2.926% | 35 | 42 | 42 |
| 4 | 50 | 63 | 25.3% | 44.2% | 0.768% | 25 | 41 | 41 |
| 5 | 24 | 17 | 12.1% | 58.5% | 0.207% | 13 | 39 | 39 |
| 6 | 15 | 2 | 7.6% | 88.2% | 0.024% | 6 | 33 | 33 |
| 7 | 6 | 0 | 3.0% | 100.0% | 0.000% | 3 | 27 | 27 |
| 8 | 1 | 0 | 0.5% | 100.0% | 0.000% | 1 | 18 | 18 |
| 9 | 0 | 0 | | | | 0 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

FIGURE 4.410T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 2
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 221 (Subset 3)
 Number of Relevant Documents 1,590
 Generality Number 5.1

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,530 | (-) | 96.2% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,389 | 98,743* | 87.4% | 1.4%* | 32.089 %* | 221 | 44* | 221 |
| 3 | 1,104 | 36,132* | 69.4% | 3.0%* | 11.739 %* | 215 | 109* | 220 |
| 4 | 782 | 13,974* | 49.2% | 5.3%* | 4.540 %* | 203 | 142* | 212 |
| 5 | 467 | 6,701* | 29.3% | 6.5%* | 2.178 %* | 158 | 177* | 197 |
| 6 | 261 | 2,549* | 16.4% | 9.3%* | 0.828 %* | 111 | 161* | 164 |
| 7 | 128 | 663 | 8.1% | 16.2% | 0.215% | 74 | 140 | 140 |
| 8 | 53 | 197 | 3.3% | 21.2% | 0.064% | 40 | 105 | 105 |
| 9 | 19 | 48 | 1.2% | 28.4% | 0.016% | 18 | 78 | 78 |
| 10 | 5 | 7 | 0.3% | 41.7% | 0.002% | 6 | 52 | 52 |
| 11 | 0 | 1 | 0.0% | | 0.0003% | 1 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.411T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 1
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 221 (Subset 3)
 Number of Relevant Documents 1,590
 Generality Number 5.1

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|-----|------|-----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 1,455 | (-) | 91.5% | (-) | (-) | 221 | 0 | 221 |
| 2 | 1,226 | 54,263* | 77.1% | 2.3%* | 17.618%* | 221 | 44* | 221 |
| 3 | 861 | 15,346* | 54.2% | 5.3%* | 4.982%* | 212 | 109* | 220 |
| 4 | 520 | 4,903* | 32.7% | 9.6%* | 1.462%* | 186 | 142* | 212 |
| 5 | 261 | 2,214* | 16.4% | 10.6%* | 0.719%* | 124 | 177* | 197 |
| 6 | 122 | 586* | 7.7% | 17.2%* | 0.190%* | 77 | 161* | 164 |
| 7 | 54 | 141 | 3.4% | 27.7% | 0.046% | 43 | 140 | 140 |
| 8 | 20 | 38 | 1.3% | 34.5% | 0.012% | 22 | 105 | 105 |
| 9 | 6 | 7 | 0.4% | 46.2% | 0.002% | 7 | 78 | 78 |
| 10 | 1 | 1 | 0.1% | 50.0% | 0.0003% | 2 | 52 | 52 |
| 11 | 0 | 0 | | | | 0 | 32 | 32 |
| 12 | 0 | 0 | | | | 0 | 15 | 15 |
| 13 | 0 | 0 | | | | 0 | 8 | 8 |
| 14 | 0 | 0 | | | | 0 | 4 | 4 |
| 15 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.412T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 2
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 35 (Subset 1)
 Number of Relevant Documents 287
 Generality Number 5.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 276 | (-) | 96.2% | (-) | (-) | 35 | 0 | 35 |
| 2 | 247 | 15,946* | 86.1% | 1.5%* | 32.542%* | 35 | 23* | 35 |
| 3 | 187 | 6,567 | 65.2% | 2.6% | 13.481% | 35 | 35 | 35 |
| 4 | 114 | 1,742 | 39.8% | 6.1% | 3.675% | 32 | 35 | 35 |
| 5 | 60 | 397 | 21.0% | 13.1% | 0.815% | 25 | 35 | 35 |
| 6 | 37 | 62 | 12.9% | 37.4% | 0.127% | 16 | 35 | 35 |
| 7 | 12 | 13 | 4.2% | 48.0% | 0.027% | 5 | 35 | 35 |

FIGURE 4.413T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 1
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 35 (Subset 1)
 Number of Relevant Documents 287
 Generality Number 5.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 257 | (-) | 89.6% | (-) | (-) | 35 | 0 | 35 |
| 2 | 215 | 8,416* | 75.0% | 2.5%* | 17.176%* | 35 | 23* | 35 |
| 3 | 142 | 2,901 | 49.5% | 4.7% | 5.955% | 34 | 35 | 35 |
| 4 | 74 | 539 | 25.8% | 12.1% | 1.106% | 30 | 35 | 35 |
| 5 | 32 | 85 | 11.2% | 27.4% | 0.174% | 14 | 35 | 35 |
| 6 | 12 | 5 | 4.2% | 70.6% | 0.010% | 6 | 35 | 35 |
| 7 | 6 | 2 | 2.1% | 75.0% | 0.004% | 3 | 35 | 35 |

FIGURE 4.414T

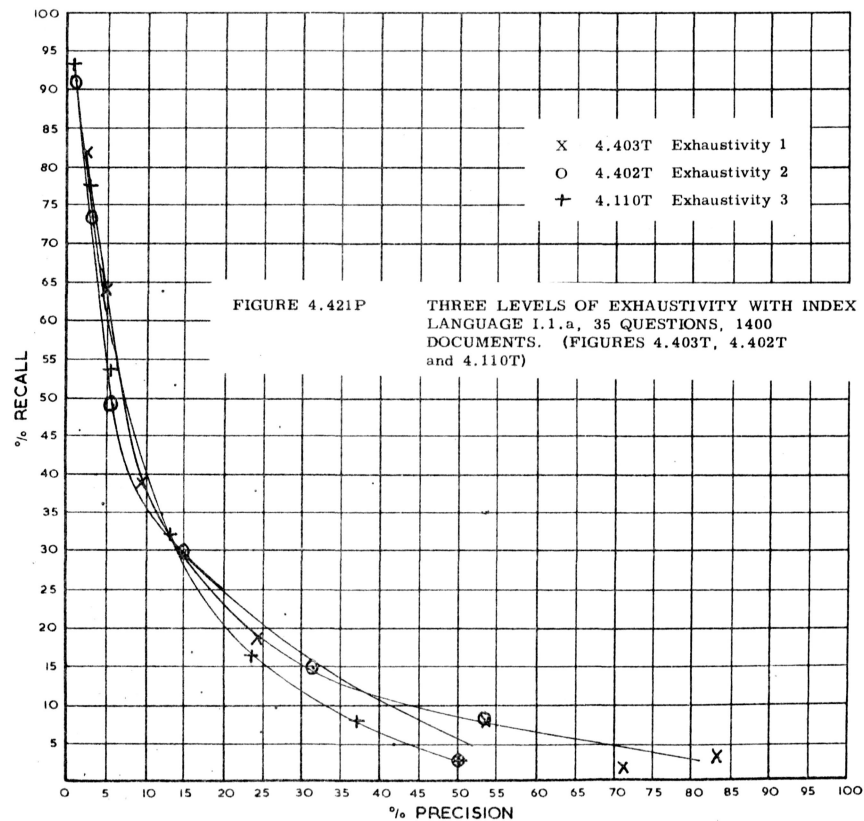
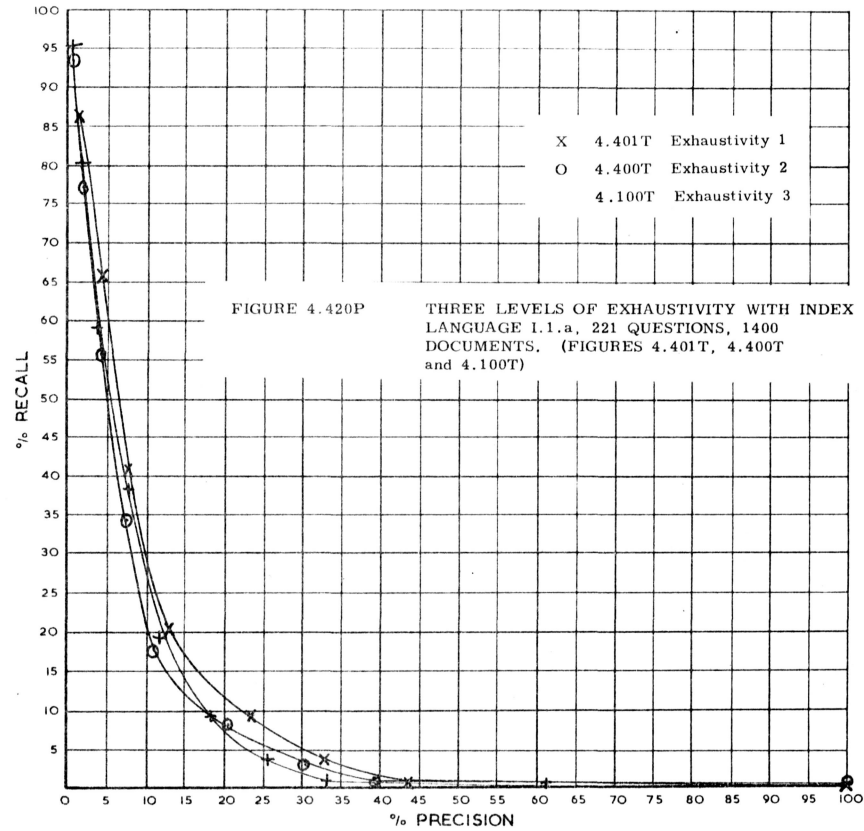
Index Language 1.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 2
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

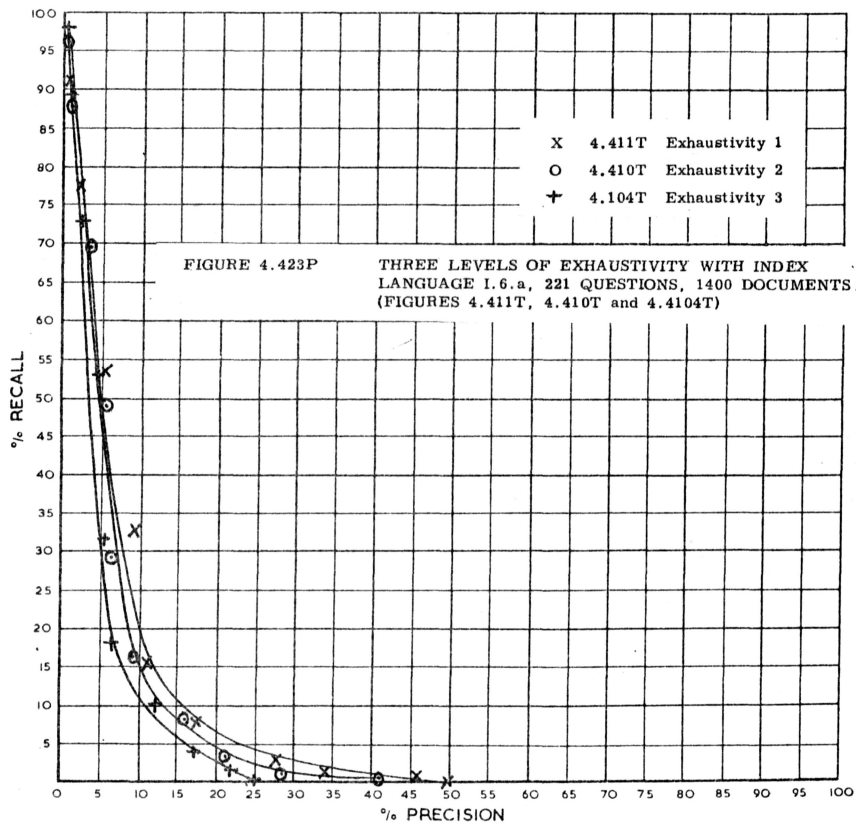
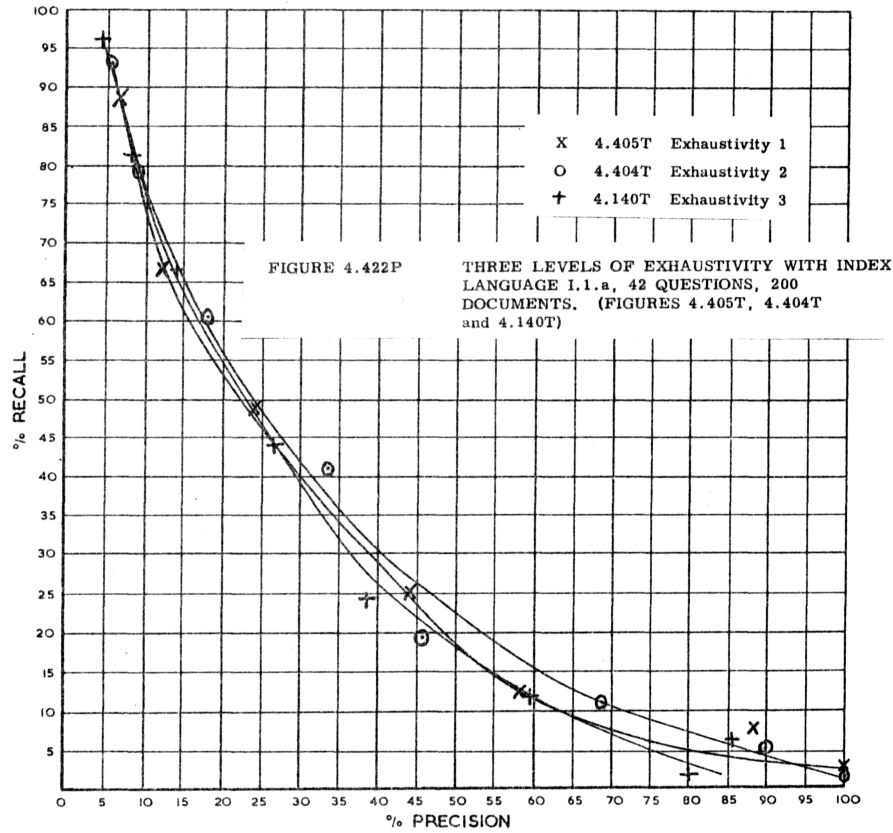
| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 193 | 5,693 | 97.5% | 3.4% | 67.773% | 42 | 42 | 42 |
| 2 | 175 | 3,253 | 88.4% | 5.1% | 39.661% | 42 | 42 | 42 |
| 3 | 145 | 1,431 | 73.2% | 9.2% | 17.447% | 41 | 42 | 42 |
| 4 | 113 | 544 | 57.1% | 17.2% | 6.633% | 36 | 41 | 41 |
| 5 | 64 | 176 | 32.3% | 26.7% | 2.146% | 30 | 39 | 39 |
| 6 | 42 | 40 | 21.2% | 51.2% | 0.488% | 20 | 33 | 33 |
| 7 | 21 | 13 | 10.6% | 61.8% | 0.158% | 12 | 27 | 27 |
| 8 | 5 | 1 | 2.5% | 83.3% | 0.012% | 2 | 18 | 18 |
| 9 | 4 | 0 | 2.0% | 100.0% | 0.000% | 2 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

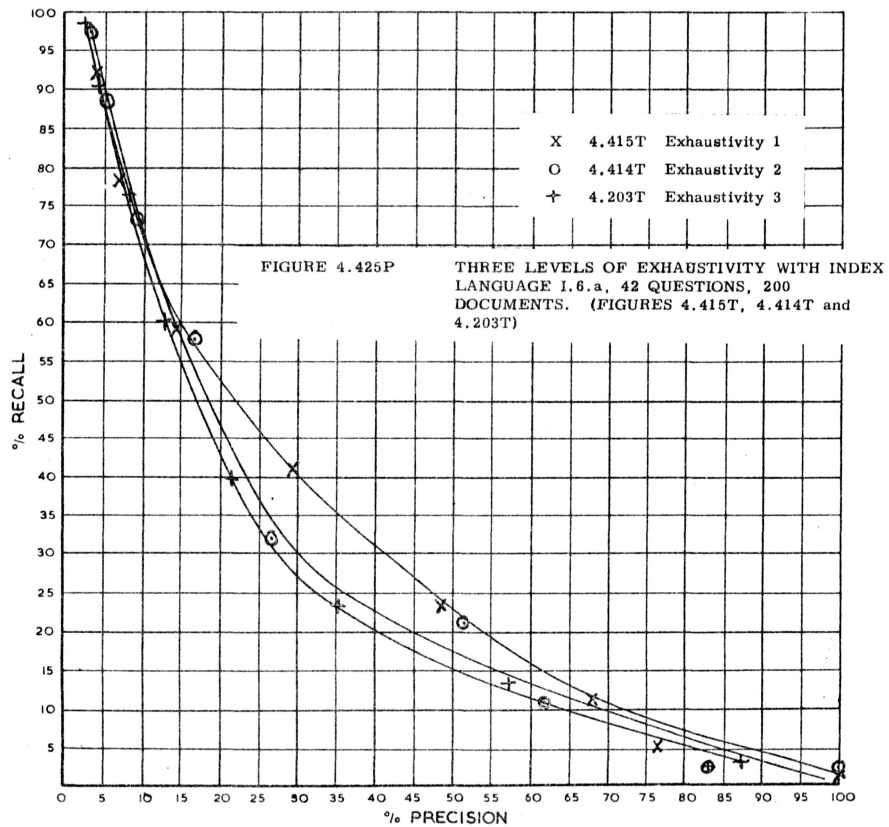
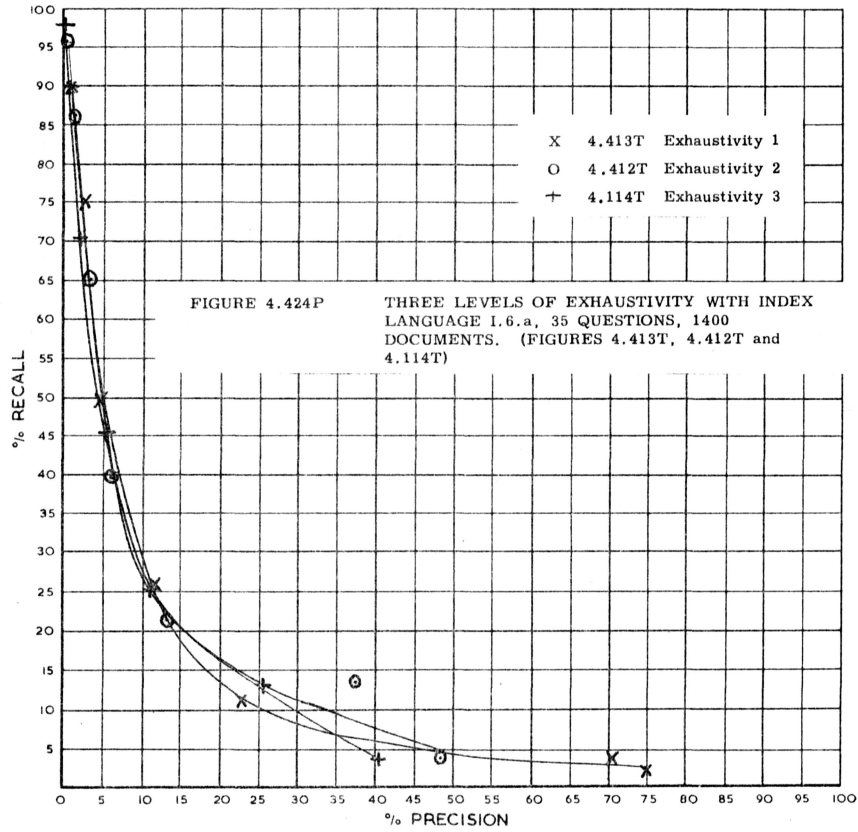
FIGURE 4.415T

Index Language 1.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)
 Exhaustivity of Indexing 1
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 183 | 4,127 | 92.4% | 4.2% | 49.130% | 42 | 42 | 42 |
| 2 | 155 | 2,118 | 78.3% | 6.8% | 25.823% | 42 | 42 | 42 |
| 3 | 117 | 685 | 59.1% | 14.6% | 8.352% | 39 | 42 | 42 |
| 4 | 81 | 180 | 40.9% | 29.9% | 2.317% | 32 | 41 | 41 |
| 5 | 47 | 51 | 23.7% | 48.0% | 0.622% | 22 | 39 | 39 |
| 6 | 24 | 11 | 12.1% | 68.6% | 0.134% | 14 | 33 | 33 |
| 7 | 10 | 3 | 5.1% | 76.9% | 0.037% | 5 | 27 | 27 |
| 8 | 2 | 0 | 1.0% | 100.0% | 0.000% | 1 | 18 | 18 |
| 9 | 2 | 0 | 1.0% | 100.0% | 0.000% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |







Section 5 Search Rules

The first search rule investigated is type B, which, as explained in Chapter 2, was used to provide a minimal amount of intellect prior to making the test of precision devices. The sets of questions (subsets 4 and 5) and languages (I.1.a and I.6.a) that were used to test precision devices are investigated. Figs. 4.500T - 4.503T give tables of results for search A, this being the basic type of search used in presenting the results on, for instance, the recall devices. This particular series of results has to be given because of the different set of questions which was used for testing Search B; the results of these latter searches have already been presented in Figs. 4.300T to 4.303T in Section 3. Graphs comparing the curves of search A with search B are given in Figs. 4.504P to 4.507P. Search Rule C demanded a selection of the original starting terms taken from the question. This eliminated such terms as Problem, Influence, Comparison, etc., and the result was to reduce the number of starting terms in every question. Whereas originally, in this subset, each question had seven starting terms, when search rule C was in force, the number varied from two to six.

Search Rule D retained the terms selected for search rule C but imposed restrictions concerning the combinations of terms that would be accepted at every coordination level, thereby eliminating non-sensical combinations.

For search rules C and D a set of 20 questions (subset 6), all taken from the 35 questions having seven starting terms (subset 1) is used, searched on the 1400 collection. Languages I.1.a and I.6.a are tested and Figs. 4.510T and 4.511T give the results with search rule A.

Results for searches C and D were totalled by method 1B, (see chapter 3, page 62), since the ordinary strict coordination level method gave unsatisfactory results due to the small number of questions and variation in starting terms that resulted from the rules used in the searches. Figs. 4.512T and 4.513T give the results of search C, and Figs. 4.514T and 4.515T the results of search D; because of the different totalling methods the tables of results have to include the additional data which shows the reduced number of documents regarded as relevant at the higher coordination levels, and also the resultant change in generality number. Two plots Figs. 4.516P and 4.517P present the result for searches A, C and D with the two index languages.

Although the comparison between search A on the one hand and searches C and D on the other hand might seem to be influenced by the different totalling method used, all possible methods were tried and that used was found to be the most satisfactory.

| | Search Rule | Index Language | No. of Questions | Question Subset | Document Collection | Plot |
|--------|-------------|-------------------|---------------------|--------------------|------------------------|-------------|
| 4.500T | A | I.1.a | 19 | 4 | 1400 | |
| 4.501T | A | I.1.a | 17 | 5 | 1400 | |
| 4.502T | A | I.6.a | 19 | 4 | 1400 | |
| 4.503T | A | I.6.a | 17 | 5 | 1400 | |
| 4.504P | A & B | I.1.a | 19 | 4 | 1400 | Plot 4.300T |
| | | | | | | 4.500T |
| 4.505P | A & B | I.1.a | 17 | 5 | 1400 | Plot 4.301T |
| | | | | | | 4.501T |
| 4.506P | A & B | I.6.a | 19 | 4 | 1400 | Plot 4.302T |
| | | | | | | 4.502T |
| 4.507P | A & B | I.6.a | 17 | 5 | 1400 | Plot 4.303T |
| | | | | | | 4.503T |
| 4.510T | A | I.1.a | 20 | 6 | 1400 | |
| 4.511T | A | I.6.a | 20 | 6 | 1400 | |
| 4.512T | C | I.1.a | 20 | 6 | 1400 | |
| 4.513T | C | I.6.a | 20 | 6 | 1400 | |
| 4.514T | D | I.1.a | 20 | 6 | 1400 | |
| 4.515T | D | I.6.a | 20 | 6 | 1400 | |
| 4.516P | A, C & D | I.1.a | 20 | 6 | 1400 | Plot 4.510T |
| | | | | | | 4.512T |
| | | | | | | 4.514T |
| 4.517P | A, C & D | I.6.a | 20 | 6 | 1400 | Plot 4.511T |
| | | | | | | 4.513T |
| | | | | | | 4.515T |

FIGURE 4.500T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 19 (Subset 4)

Number of Relevant Documents 131

Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 118 | 12,700 | 90.1% | 0.9% | 47.744% | 19 | 19 | 19 |
| 2 | 100 | 4,294 | 76.3% | 2.4% | 16.143% | 19 | 19 | 19 |
| 3 | 72 | 1,426 | 55.0% | 5.0% | 5.360% | 19 | 19 | 19 |
| 4 | 39 | 188 | 29.8% | 17.2% | 0.710% | 15 | 19 | 19 |
| 5 | 18 | 16 | 13.7% | 52.9% | 0.060% | 8 | 19 | 19 |
| 6 | 10 | 3 | 7.6% | 76.9% | 0.011% | 2 | 19 | 19 |
| 7 | 4 | 0 | 3.1% | 100.0% | 0.000% | 2 | 19 | 19 |

FIGURE 4.501T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 17 (Subset 5)

Number of Relevant Documents 109

Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 108 | 14,710 | 99.1% | 0.7% | 61.807% | 17 | 17 | 17 |
| 2 | 105 | 8,030 | 96.3% | 1.3% | 33.739% | 17 | 17 | 17 |
| 3 | 91 | 3,652 | 83.5% | 2.4% | 15.344% | 17 | 17 | 17 |
| 4 | 74 | 1,559 | 67.9% | 4.5% | 6.550% | 17 | 17 | 17 |
| 5 | 52 | 618 | 47.7% | 7.8% | 2.609% | 17 | 17 | 17 |
| 6 | 39 | 237 | 35.8% | 14.1% | 1.000% | 15 | 17 | 17 |
| 7 | 22 | 100 | 20.2% | 18.0% | 0.422% | 14 | 17 | 17 |
| 8 | 13 | 22 | 11.9% | 37.1% | 0.093% | 10 | 17 | 17 |
| 9 | 6 | 4 | 5.5% | 60.0% | 0.017% | 5 | 17 | 17 |
| 10 | 1 | 0 | 0.9% | 100.0% | 0.000% | 1 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |

FIGURE 4.502T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word Forms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 19 (Subset 4)
 Number of Relevant Documents 131
 Generality Number 4.9

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 126 | (-) | 96.2% | (-) | (-) | 19 | 0 | 19 |
| 2 | 114 | 9,568 | 87.0% | 1.2% | 35.969% | 19 | 19 | 19 |
| 3 | 98 | 3,962 | 74.8% | 2.5% | 14.895% | 19 | 19 | 19 |
| 4 | 61 | 1,142 | 46.6% | 5.1% | 4.314% | 19 | 19 | 19 |
| 5 | 33 | 192 | 25.2% | 14.7% | 0.725% | 16 | 19 | 19 |
| 6 | 18 | 30 | 13.7% | 37.5% | 0.113% | 8 | 19 | 19 |
| 7 | 8 | 1 | 6.1% | 88.9% | 0.004% | 3 | 19 | 19 |

FIGURE 4.503T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms, Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 1,400
 Number of Questions 17 (Subset 5)
 Number of Relevant Documents 109
 Generality Number 4.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 109 | (-) | 100.0% | (-) | (-) | 17 | 0 | 17 |
| 2 | 108 | 13,865 | 99.1% | 0.8% | 58,256% | 17 | 17 | 17 |
| 3 | 101 | 7,264 | 92.7% | 1.4% | 30,521% | 17 | 17 | 17 |
| 4 | 91 | 4,124 | 83.5% | 2.2% | 17,328% | 17 | 17 | 17 |
| 5 | 73 | 2,353 | 67.0% | 3.0% | 9.932% | 17 | 17 | 17 |
| 6 | 54 | 1,026 | 49.5% | 5.0% | 4.331% | 17 | 17 | 17 |
| 7 | 37 | 361 | 34.0% | 9.3% | 1.524% | 17 | 17 | 17 |
| 8 | 23 | 136 | 21.1% | 14.5% | 0.574% | 16 | 17 | 17 |
| 9 | 9 | 41 | 8.3% | 18.0% | 0.173% | 10 | 17 | 17 |
| 10 | 2 | 12 | 1.8% | 14.3% | 0.051% | 4 | 17 | 17 |
| 11 | 0 | 0 | | | | 0 | 17 | 17 |

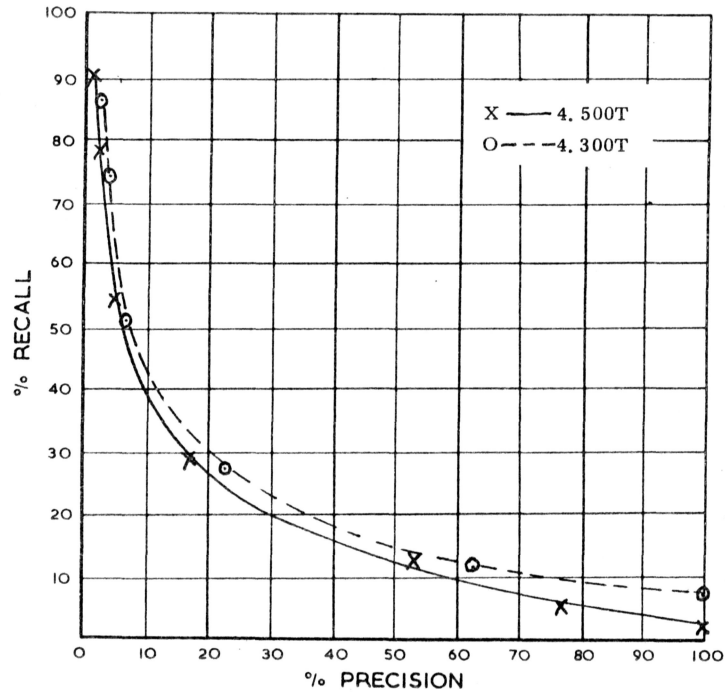


FIGURE 4.504P SEARCH RULES A and B, 19 QUESTIONS, INDEX LANGUAGE I.1.a (FIGURES 4.500T and 4.300T)

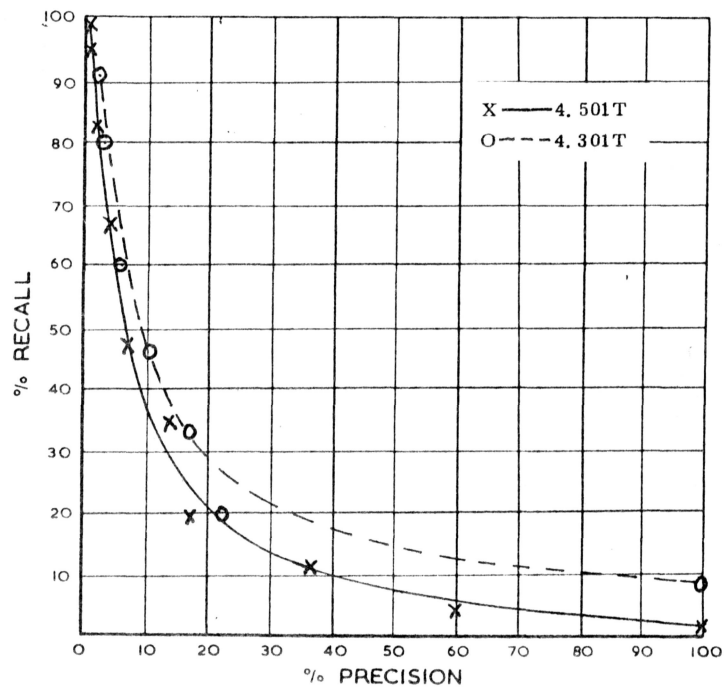


FIGURE 4.505P SEARCH RULES A and B, 17 QUESTIONS, INDEX LANGUAGE I.6.a (FIGURES 4.501T and 4.301T)

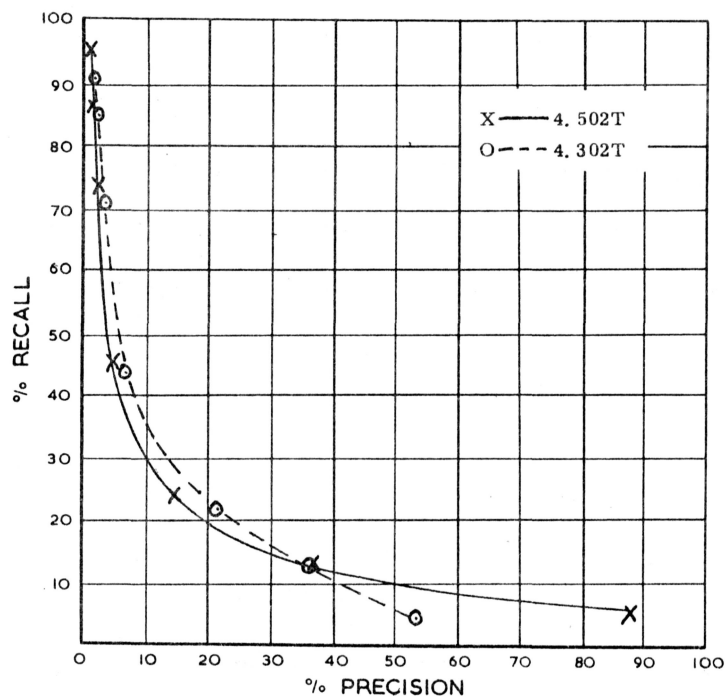


FIGURE 4.506P SEARCH RULES A and B, 19 QUESTIONS, INDEX LANGUAGE 1.1.a (FIGURES 4.502T and 4.302T)

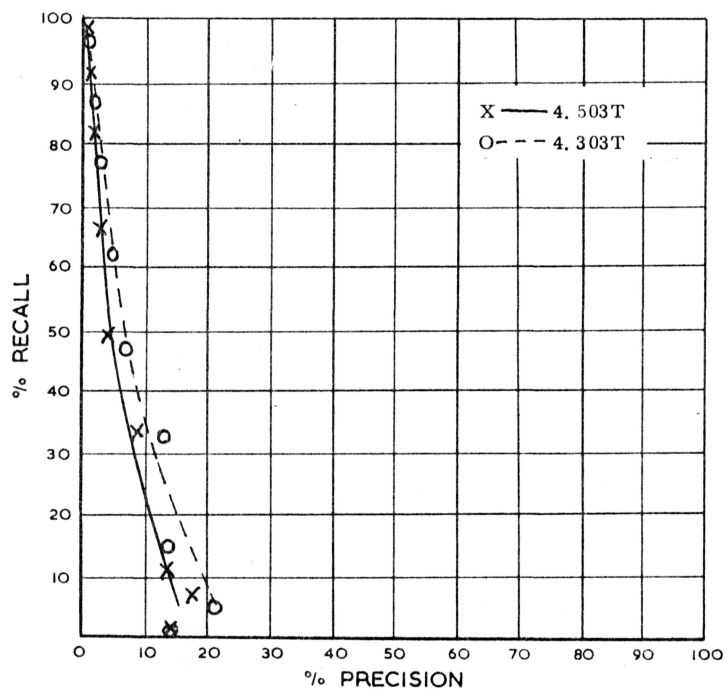


FIGURE 4.507P SEARCH RULES, A and B, 17 QUESTIONS, INDEX LANGUAGE 1.6.a (FIGURES 4.503T and 4.303T)

FIGURE 4.510T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 20 (Subset 6)

Number of Relevant Documents 147

Generality Number 5.3

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 133 | 14,225 | 90.5% | 0.9% | 50.804% | 20 | 20 | 20 |
| 2 | 114 | 4,704 | 77.6% | 2.4% | 16.808% | 20 | 20 | 20 |
| 3 | 82 | 1,630 | 55.9% | 4.9% | 5.852% | 20 | 20 | 20 |
| 4 | 47 | 230 | 32.0% | 17.0% | 0.826% | 16 | 20 | 20 |
| 5 | 21 | 24 | 14.3% | 46.7% | 0.086% | 9 | 20 | 20 |
| 6 | 10 | 3 | 6.8% | 76.9% | 0.011% | 2 | 20 | 20 |
| 7 | 4 | 0 | 2.7% | 100.0% | 0.000% | 2 | 20 | 20 |

FIGURE 4.511T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 20 (Subset 6)

Number of Relevant Documents 147

Generality Number 5.3

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 142 | (-) | 96.6% | (-) | (-) | 20 | 0 | 20 |
| 2 | 129 | 8,088* | 87.8% | 1.6%* | 29.038%* | 20 | 13* | 20 |
| 3 | 109 | 4,657 | 74.1% | 2.3% | 16.720% | 20 | 20 | 20 |
| 4 | 69 | 1,276 | 46.9% | 5.1% | 4.581% | 20 | 20 | 20 |
| 5 | 36 | 226 | 24.5% | 13.7% | 0.811% | 16 | 20 | 20 |
| 6 | 19 | 32 | 12.9% | 37.3% | 0.115% | 8 | 20 | 20 |
| 7 | 8 | 1 | 5.4% | 88.9% | 0.004% | 3 | 20 | 20 |

FIGURE 4.512T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule C

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 20 (Subset 6)

Number of Relevant Documents see table

Generality Number see table

| Coord- ination Level | Relevant Documents | Generality Number | Documents Retrieved | | Recall Ratio | Precision Ratio | Fallout Ratio | x | y | z |
|----------------------------|-----------------------|----------------------|------------------------|----------|-----------------|--------------------|------------------|----|-----|----|
| | | | Rel. | Non-rel. | | | | | | |
| 1 | 147 | 5.3 | 128 | (-) | 87.1% | (-) | (-) | 20 | 0 | 20 |
| 2 | 147 | 5.3 | 96 | 1,809* | 65.3% | 5.0%* | 6.495%* | 20 | 13* | 20 |
| 3 | 119 | 5.0 | 57 | 533 | 47.9% | 9.7% | 2.251% | 17 | 17 | 17 |
| 4 | 94 | 4.8 | 26 | 63 | 27.7% | 29.2% | 0.323% | 9 | 14 | 14 |
| 5 | 43 | 4.4 | 7 | 8 | 16.3% | 46.7% | 0.082% | 2 | 7 | 7 |
| 6 | 21 | 5.0 | 1 | 2 | 4.8% | 33.3% | 0.048% | 1 | 3 | 3 |

FIGURE 4.513T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule C

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 20 (Subset 6)

Number of Relevant Documents see table

Generality Number see table

| Coord- ination Level | Relevant Documents | Generality Number | Documents Retrieved Rel. Non-rel. | Recall Ratio | Precision Ratio | Fallout Ratio | x | y | z |
|----------------------------|-----------------------|----------------------|---|-----------------|--------------------|------------------|----|-----|----|
| 1 | 147 | 5.3 | 140 (-) | 95.2% | (-) | (-) | 20 | 0 | 20 |
| 2 | 147 | 5.3 | 113 3,323* | 76.9% | 3.3%* | 11.930%* | 20 | 13* | 20 |
| 3 | 119 | 5.0 | 68 1,544 | 57.1% | 4.2% | 6.519% | 17 | 17 | 17 |
| 4 | 94 | 4.8 | 38 299 | 40.4% | 11.3% | 1.533% | 11 | 14 | 14 |
| 5 | 43 | 4.4 | 12 33 | 27.9% | 26.7% | 0.338% | 6 | 7 | 7 |
| 6 | 21 | 5.0 | 2 2 | 9.5% | 50.0% | 0.048% | 1 | 3 | 3 |

FIGURE 4.514T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule D

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 20 (Subset 6)

Number of Relevant Documents see table

Generality Number see table

| Coord- ination Level | Relevant Documents | Generality Number | Documents Retrieved Rel. Non-rel. | Recall Ratio | Precision Ratio | Fallout Ratio | x | y | z |
|----------------------------|-----------------------|----------------------|---|-----------------|--------------------|------------------|----|----|----|
| 1 | 147 | 5.3 | 97 2,213 | 66.0% | 4.5% | 7.945% | 20 | 20 | 20 |
| 2 | 147 | 5.3 | 73 653 | 49.7% | 10.1% | 2.344% | 20 | 20 | 20 |
| 3 | 119 | 5.0 | 42 163 | 35.3% | 20.5% | 0.688% | 15 | 17 | 17 |
| 4 | 94 | 4.8 | 20 14 | 21.3% | 58.8% | 0.072% | 8 | 14 | 14 |
| 5 | 43 | 4.4 | 7 4 | 16.3% | 63.6% | 0.041% | 2 | 7 | 7 |
| 6 | 21 | 5.0 | 1 2 | 4.8% | 33.3% | 0.048% | 1 | 3 | 3 |

FIGURE 4.515T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule D

Document Relevance 1 - 4

Number of Documents in Collection 1,400

Number of Questions 20 (Subset 6)

Number of Relevant Documents see table

| Coord- ination Level | Relevant Documents | Generality Number | Documents Retrieved Rel. Non-rel. | Recall Ratio | Precision Ratio | Fallout Ratio | x | y | z |
|----------------------------|-----------------------|----------------------|---|-----------------|--------------------|------------------|----|----|----|
| 1 | 147 | 5.3 | 102 2,668 | 69.4% | 3.7% | 9.579% | 20 | 20 | 20 |
| 2 | 147 | 5.3 | 81 791 | 55.1% | 9.3% | 2.840% | 20 | 20 | 20 |
| 3 | 119 | 5.0 | 48 189 | 40.3% | 20.3% | 0.798% | 15 | 17 | 17 |
| 4 | 94 | 4.8 | 25 21 | 26.6% | 54.3% | 0.108% | 11 | 14 | 14 |
| 5 | 43 | 4.4 | 10 6 | 23.3% | 62.5% | 0.061% | 4 | 7 | 7 |
| 6 | 21 | 5.0 | 1 2 | 4.8% | 33.3% | 0.048% | 1 | 3 | 3 |

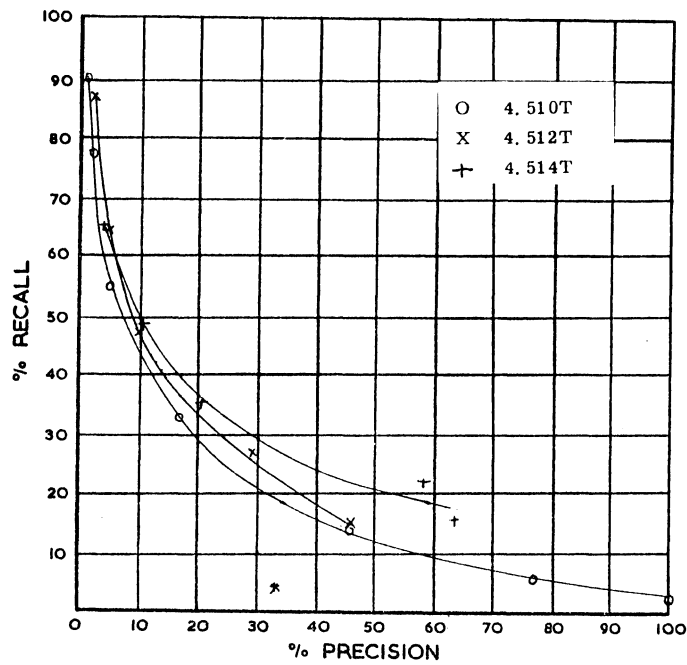


FIGURE 4.516P SEARCH RULES A, C, and D, 20 QUESTIONS, INDEX LANGUAGE I.1.a (FIGURES 4.510T, 4.512T and 4.514T)

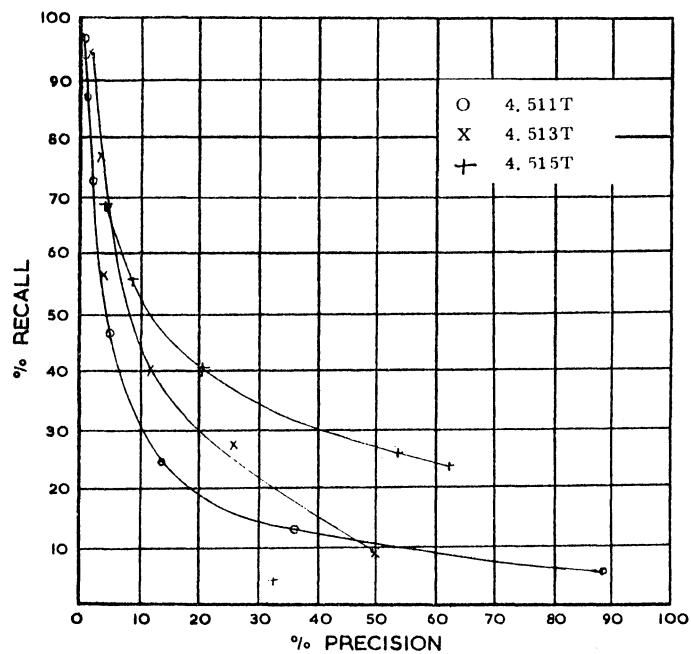


FIGURE 4.517P SEARCH RULES, A,C and D, 20 QUESTIONS, INDEX LANGUAGE I.G.a (FIGURES 4.511T, 4.513T and 4.515T)

Section 6 Document Relevance

The four grades of document relevance were first tested on the 35 questions (subset 1), and results are given for languages I.1.a and I.6.a. Figs. 4.600T - 4.605T give the tables of results, and Figs. 4.606P - 4.607P the plots of performance curves. The change in the different relevance decisions causes a change in the generality number, which is shown in each table, and therefore recall/fallout plots are presented. Although the set of questions used was typical, the test results of the two higher grades of relevance (1-2 and 1) suffer from the fact that not all questions had any relevant documents of these grades. 27 questions had relevant documents when grade 1-2 was tested, but only 11 questions when grade 1 was tested. All 35 questions, however, were kept in the set and the non-relevant documents retrieved by these questions were still counted. It can be noted that the change in document relevance from low to high merely transfers some of the relevant documents from the relevant retrieved to the non-relevant retrieved category, while the total documents retrieved always stays constant.

Because of the small total of documents having relevance 1 documents, a further test was made on a set of fifty questions, for which the criterion of selection was that each question must have a relevance 1 document. These were tested on the 1400 document collection with index language I.1.a and the results are shown in Figs. 4.610T - 4.613T, with a recall/fallout plot as Figure 4.614P.

LIST OF FIGURES

| | Relevance | Index Language | No. of Questions | Question Subset | Document Collection | Plots |
|--------|-----------|----------------|------------------|-----------------|---------------------|-------------|
| 4.600T | 1-3 | I.1.a | 35 | 1 | 1400 | |
| 4.601T | 1-3 | I.6.a | 35 | 1 | 1400 | |
| 4.602T | 1-2 | I.1.a | 35 | 1 | 1400 | |
| 4.603T | 1-2 | I.6.a | 35 | 1 | 1400 | |
| 4.604T | 1 | I.1.a | 35 | 1 | 1400 | |
| 4.605T | 1 | I.6.a | 35 | 1 | 1400 | |
| 4.606P | 1-4 | I.1.a | 35 | 1 | 1400 | Plot 4.110T |
| | 1-3 | | | | | 4.600T |
| | 1-2 | | | | | 4.602T |
| | and 1 | | | | | 4.604T |
| 4.607P | 1-4 | I.6.a | 35 | 1 | 1400 | Plot 4.114T |
| | 1-3 | | | | | 4.601T |
| | 1-2 | | | | | 4.603T |
| | and 1 | | | | | 4.605T |
| 4.610T | 1-4 | I.1.a | 50 | 9 | 1400 | |
| 4.611T | 1-3 | I.1.a | 50 | 9 | 1400 | |
| 4.612T | 1-2 | I.1.a | 50 | 9 | 1400 | |
| 4.613T | 1 | I.1.a | 50 | 9 | 1400 | |
| 4.614P | 1-4 | I.1.a | 50 | 9 | 1400 | Plot 4.610T |
| | 1-3 | | | | | 4.611T |
| | 1-2 | | | | | 4.612T |
| | 1 | | | | | 4.613T |

FIGURE 4.600T

Index Language 1.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 3

Number of Documents in Collection 1,400

Number of Questions 35 (subset 1)

Number of Relevant Documents 212

Generality Number 4.3

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 194 | 23,755 | 91.5% | 0.6% | 48.520% | 35 | 35 | 35 |
| 2 | 156 | 8,151 | 73.6% | 1.9% | 19.259% | 35 | 35 | 35 |
| 3 | 112 | 2,914 | 52.8% | 3.7% | 5.965% | 34 | 35 | 35 |
| 4 | 69 | 625 | 32.5% | 9.9% | 1.281% | 28 | 35 | 35 |
| 5 | 34 | 150 | 16.0% | 18.5% | 0.307% | 18 | 35 | 35 |
| 6 | 17 | 43 | 8.0% | 28.3% | 0.088% | 7 | 35 | 35 |
| 7 | 6 | 10 | 2.8% | 37.5% | 0.020% | 3 | 35 | 35 |

FIGURE 4.601T

Index Language 1.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 3

Number of Documents in Collection 1,400

Number of Questions 35 (Subset 1)

Number of Relevant Documents 212

Generality Number 4.3

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 205 | (-) | 96.7% | (-) | (-) | 35 | 0 | 35 |
| 2 | 186 | 19,333* | 87.7% | 1.0%* | 39.455%* | 35 | 23* | 35 |
| 3 | 146 | 8,126 | 68.9% | 1.8% | 16.656% | 35 | 35 | 35 |
| 4 | 93 | 2,463 | 43.9% | 3.6% | 5.048% | 35 | 35 | 35 |
| 5 | 50 | 593 | 23.6% | 7.8% | 1.215% | 30 | 35 | 35 |
| 6 | 28 | 119 | 13.2% | 19.0% | 0.244% | 18 | 35 | 35 |
| 7 | 10 | 22 | 4.7% | 31.3% | 0.045% | 6 | 35 | 35 |

FIGURE 4.602T

Index Language 1.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 2

Number of Documents in Collection 1,400

Number of Questions 35 (8 questions had no relevant documents)

Number of Relevant Documents 79

Generality Number 1.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 73 | 23,876 | 92.4% | 0.3% | 48.726% | 35 | 35 | 35 |
| 2 | 55 | 8,252 | 69.6% | 0.7% | 20.176% | 35 | 35 | 35 |
| 3 | 43 | 2,983 | 54.4% | 1.4% | 6.089% | 34 | 35 | 35 |
| 4 | 30 | 664 | 38.0% | 4.3% | 1.357% | 28 | 35 | 35 |
| 5 | 16 | 177 | 20.3% | 8.3% | 0.361% | 18 | 35 | 35 |
| 6 | 9 | 51 | 11.4% | 15.0% | 0.104% | 7 | 35 | 35 |
| 7 | 3 | 13 | 3.8% | 18.8% | 0.027% | 3 | 35 | 35 |

FIGURE 4.603T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 2

Number of Documents in Collection 1,400

Number of Questions 35 (8 questions had no relevant documents)

Number of Relevant Documents 79

Generality Number 1.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 78 | (-) | 98.7% | (-) | (-) | 35 | 0 | 35 |
| 2 | 70 | 19,449* | 88.6% | 0.4%* | 39.691%* | 35 | 23* | 35 |
| 3 | 53 | 8,210 | 67.1% | 0.6% | 16.801% | 35 | 35 | 35 |
| 4 | 37 | 2,510 | 46.8% | 1.4% | 5.149% | 35 | 35 | 35 |
| 5 | 24 | 610 | 30.4% | 3.7% | 1.265% | 30 | 35 | 35 |
| 6 | 17 | 130 | 21.5% | 11.6% | 0.266% | 18 | 35 | 35 |
| 7 | 6 | 26 | 7.6% | 18.8% | 0.053% | 6 | 35 | 35 |

FIGURE 4.604T

Index Language I.1.a (S.T. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1

Number of Documents in Collection 1,400

Number of Questions 35 (24 questions had no relevant documents)

Number of Relevant Documents 18

Generality Number 0.4

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 17 | 23,932 | 94.4% | 0.1% | 48.840% | 35 | 35 | 35 |
| 2 | 16 | 8,291 | 88.9% | 0.2% | 16.920% | 35 | 35 | 35 |
| 3 | 12 | 3,014 | 66.7% | 0.4% | 6.145% | 34 | 35 | 35 |
| 4 | 8 | 686 | 44.4% | 1.2% | 1.401% | 28 | 35 | 35 |
| 5 | 5 | 188 | 27.8% | 2.6% | 0.384% | 18 | 35 | 35 |
| 6 | 2 | 58 | 11.1% | 3.3% | 0.118% | 7 | 35 | 35 |
| 7 | 2 | 14 | 11.1% | 12.5% | 0.028% | 3 | 35 | 35 |

FIGURE 4.605T

Index Language I.6.a (S.T. Synonyms, Quasi-synonyms, Word forms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1

Number of Documents in Collection 1,400

Number of Questions 35 (24 questions had no relevant documents)

Number of Relevant Documents 18

Generality Number 0.4

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|-----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 18 | (-) | 100.0% | (-) | (-) | 35 | 0 | 35 |
| 2 | 16 | 19,503* | 88.9% | 0.1%* | 39.802% | 35 | 23* | 35 |
| 3 | 13 | 8,259 | 72.2% | 0.2% | 16.861% | 35 | 35 | 35 |
| 4 | 10 | 2,546 | 55.6% | 0.3% | 5.198% | 35 | 35 | 35 |
| 5 | 9 | 636 | 50.0% | 1.4% | 1.298% | 30 | 35 | 35 |
| 6 | 6 | 141 | 33.3% | 4.1% | 0.288% | 18 | 35 | 35 |
| 7 | 3 | 29 | 16.7% | 9.4% | 0.059% | 6 | 35 | 35 |

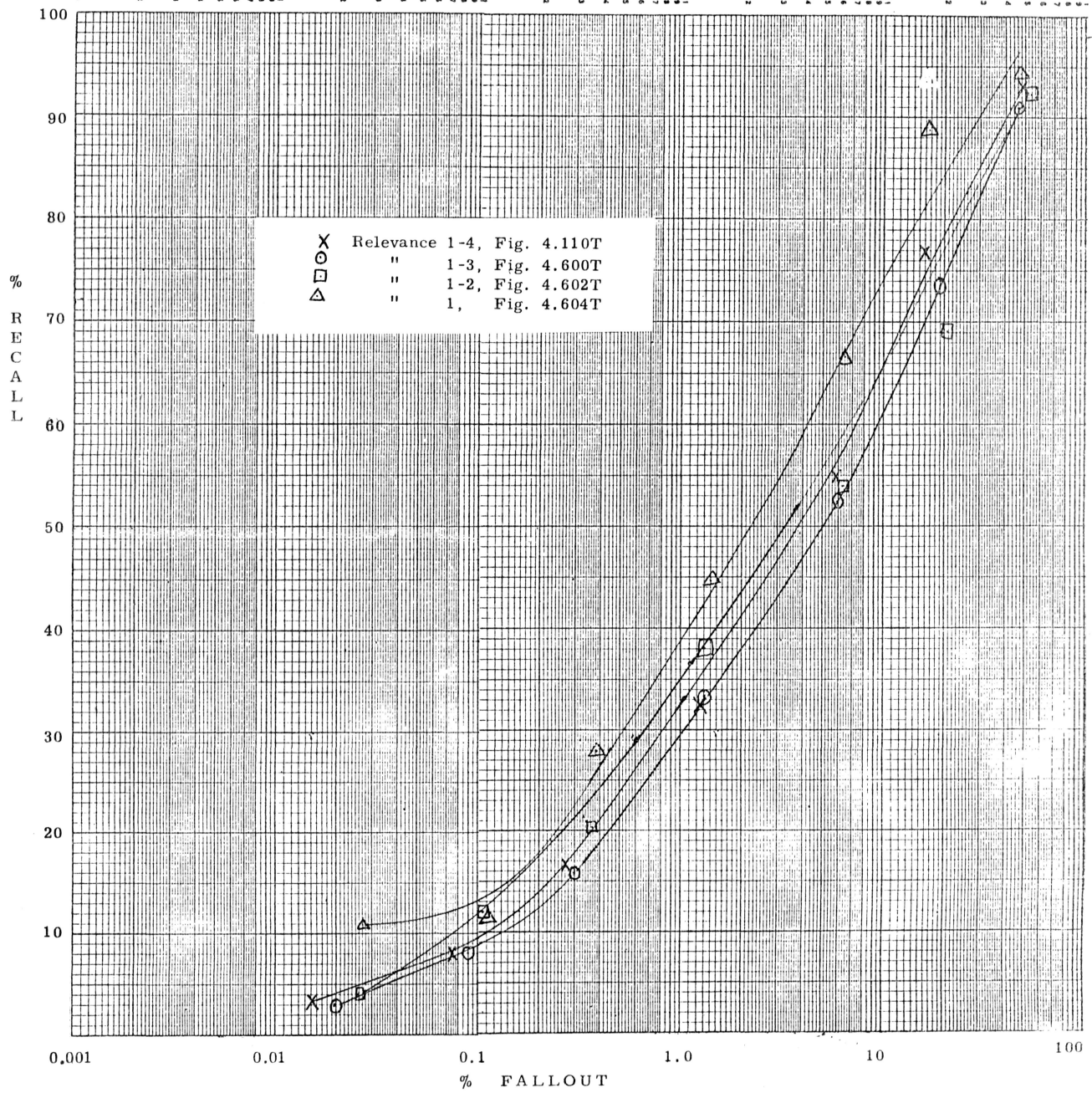


FIGURE 4.606P RECALL-FALLOUT PLOT SHOWING EFFECT OF CHANGES IN RELEVANCE ON 42 QUESTIONS WITH 1400 DOCUMENT COLLECTION, INDEX LANGUAGE I.1.a

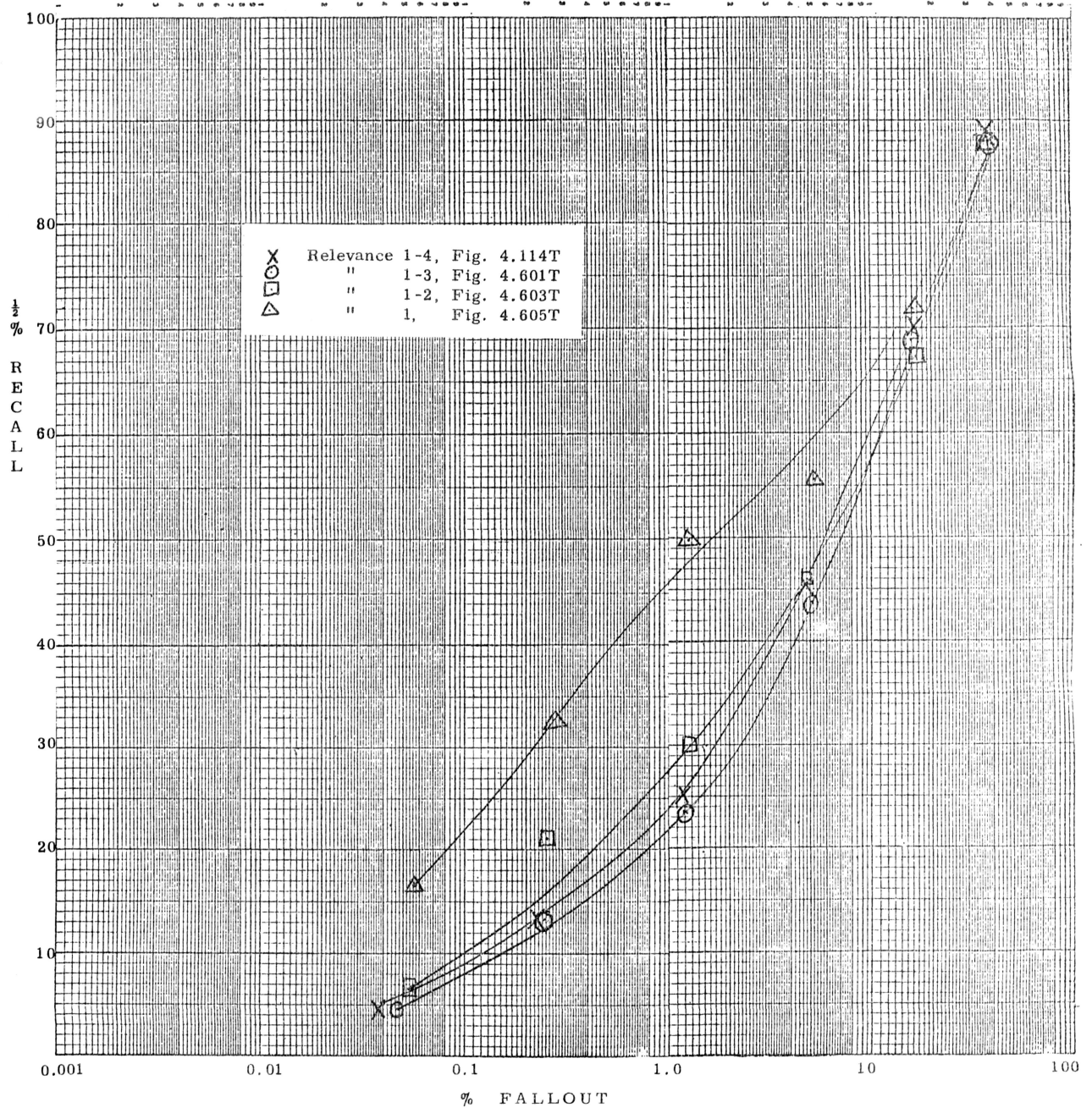


FIGURE 4.607P RECALL-FALLOUT PLOT SHOWING EFFECT OF CHANGES IN RELEVANCE ON 42 QUESTIONS WITH 1400 DOCUMENT COLLECTION, INDEX LANGUAGE I.6.a.

FIGURE 4.610T

Index Language 1.1.a

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1-4

Number of Documents in Collection 1400

Number of Questions 50

Number of Relevant Documents 361

Generality Number 5.2

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 340 | 30,778 | 94.2% | 1.1% | 44.196% | 50 | 50 | 50 |
| 2 | 281 | 11,572 | 77.8% | 2.4% | 16.617% | 50 | 50 | 50 |
| 3 | 176 | 4,717 | 48.8% | 3.6% | 6.759% | 47 | 50 | 50 |
| 4 | 107 | 1,727 | 29.6% | 5.8% | 2.480% | 39 | 46 | 46 |
| 5 | 59 | 717 | 16.3% | 7.6% | 1.015% | 21 | 40 | 40 |
| 6 | 35 | 273 | 9.7% | 11.4% | 0.392% | 12 | 30 | 30 |
| 7 | 19 | 80 | 5.3% | 19.2% | 0.115% | 8 | 23 | 23 |
| 8 | 7 | 17 | 1.9% | 29.2% | 0.024% | 4 | 12 | 12 |
| 9 | 5 | 2 | 1.4% | 71.4% | 0.003% | 3 | 10 | 10 |
| 10 | 1 | 0 | 0.3% | 100.0% | 0.000% | 1 | 8 | 8 |
| 11 | 0 | 0 | | | | 0 | 6 | 6 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.611T

Index Language 1.1.a

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1-3

Number of Documents in Collection 1400

Number of Questions 50

Number of Relevant Documents 297

Generality Number 4.2

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 277 | 30,841 | 93.2% | 0.9% | 44.246% | 50 | 50 | 50 |
| 2 | 235 | 11,618 | 79.1% | 2.0% | 16.651% | 50 | 50 | 50 |
| 3 | 162 | 4,731 | 54.5% | 3.4% | 6.787% | 47 | 50 | 50 |
| 4 | 97 | 1,737 | 32.7% | 5.3% | 2.492% | 39 | 46 | 46 |
| 5 | 49 | 724 | 16.5% | 6.4% | 1.039% | 21 | 40 | 40 |
| 6 | 29 | 279 | 9.8% | 9.2% | 0.400% | 12 | 30 | 30 |
| 7 | 16 | 83 | 5.3% | 16.2% | 0.119% | 8 | 23 | 23 |
| 8 | 6 | 18 | 2.0% | 25.0% | 0.026% | 4 | 12 | 12 |
| 9 | 4 | 3 | 1.3% | 57.1% | 0.004% | 3 | 10 | 10 |
| 10 | 1 | 0 | 0.3% | 100.0% | 0.000% | 1 | 8 | 8 |
| 11 | 0 | 0 | | | | 0 | 6 | 6 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.612T

Index Language I.1.a
Exhaustivity of Indexing 3
Search Rule A
Document Relevance 1-2
Number of Documents in Collection 1400
Number of Questions 50
Number of Relevant Documents 155
Generality Number 2.2

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 146 | 30,918 | 94.2% | 0.5% | 44.267% | 50 | 50 | 50 |
| 2 | 125 | 11,726 | 80.6% | 1.0% | 16.789% | 50 | 50 | 50 |
| 3 | 87 | 4,806 | 56.1% | 1.8% | 6.958% | 47 | 50 | 50 |
| 4 | 58 | 1,776 | 37.4% | 3.2% | 2.543% | 39 | 46 | 46 |
| 5 | 33 | 740 | 21.3% | 4.3% | 1.059% | 21 | 40 | 40 |
| 6 | 21 | 287 | 13.5% | 6.9% | 0.411% | 12 | 30 | 30 |
| 7 | 10 | 89 | 6.5% | 10.2% | 0.127% | 8 | 23 | 23 |
| 8 | 6 | 18 | 3.9% | 25.0% | 0.026% | 4 | 12 | 12 |
| 9 | 4 | 3 | 2.6% | 57.1% | 0.004% | 3 | 10 | 10 |
| 10 | 1 | 0 | 0.6% | 100.0% | 0.000% | 1 | 8 | 8 |
| 11 | 0 | 0 | | | | 0 | 6 | 6 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.613T

Index Language I.1.a
Exhaustivity of Indexing 3
Search Rule A
Document Relevance 1
Number of Documents in Collection 1400
Number of Questions 50
Number of Relevant Documents 95
Generality Number 1.4

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 90 | 30,974 | 94.7% | 0.3% | 44.309% | 50 | 50 | 50 |
| 2 | 81 | 11,770 | 85.3% | 0.7% | 16.837% | 50 | 50 | 50 |
| 3 | 57 | 4,836 | 60.0% | 1.2% | 6.918% | 47 | 50 | 50 |
| 4 | 40 | 1,794 | 42.1% | 2.2% | 2.566% | 39 | 46 | 46 |
| 5 | 24 | 749 | 25.3% | 3.1% | 1.071% | 21 | 40 | 40 |
| 6 | 14 | 294 | 14.7% | 4.5% | 0.420% | 12 | 30 | 30 |
| 7 | 7 | 92 | 7.4% | 7.1% | 0.153% | 8 | 23 | 23 |
| 8 | 3 | 21 | 3.2% | 14.3% | 0.030% | 4 | 12 | 12 |
| 9 | 2 | 5 | 2.1% | 28.6% | 0.007% | 3 | 10 | 10 |
| 10 | 0 | 1 | 0.0% | 0.0% | 0.001% | 1 | 8 | 8 |
| 11 | 0 | 0 | | | | 0 | 6 | 6 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

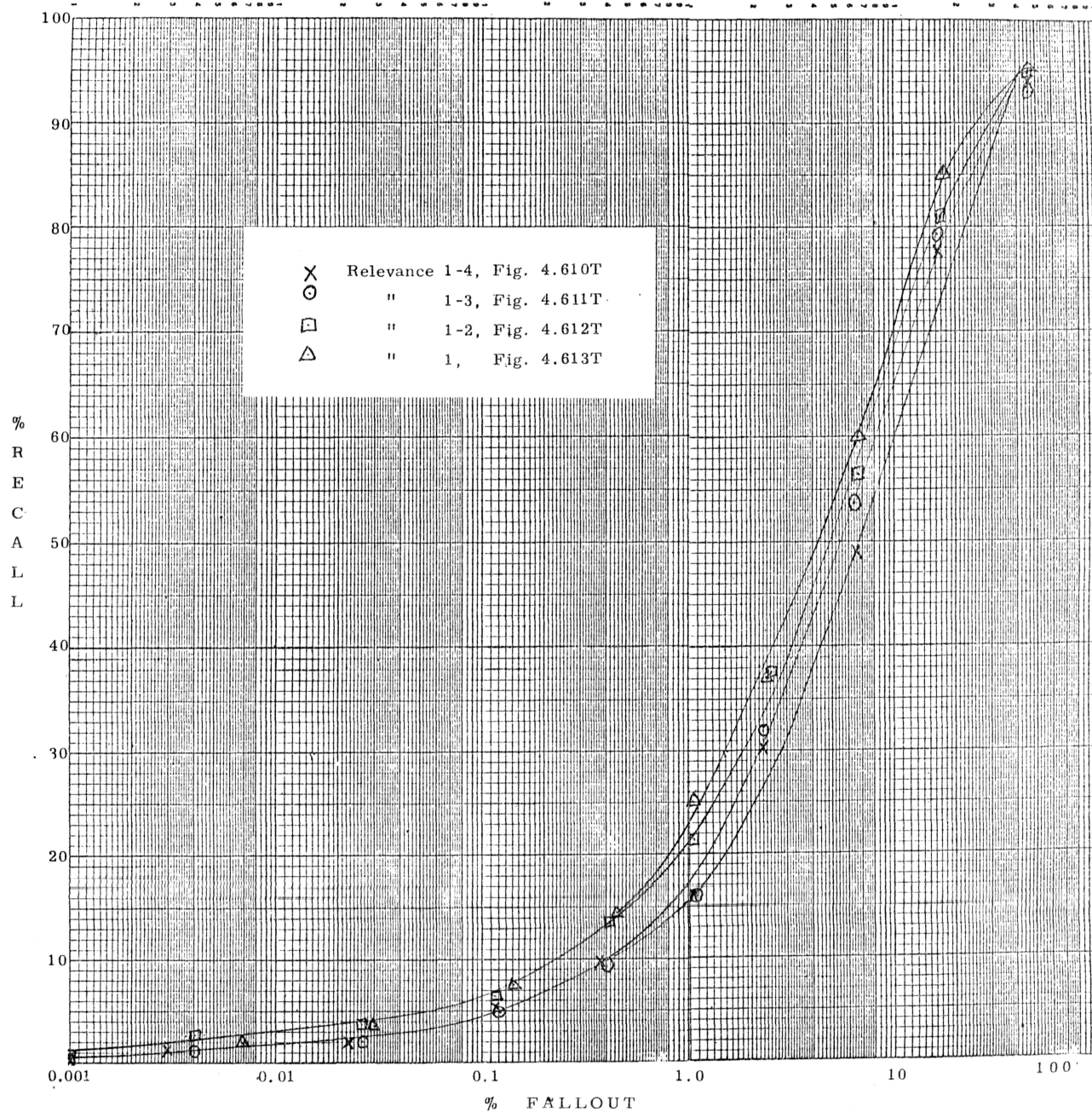


FIGURE 4.614P RECALL-FALLOUT PLOT SHOWING EFFECT OF CHANGES IN RELEVANCE ON 50 QUESTIONS WITH 1400 DOCUMENT COLLECTION, INDEX LANGUAGE 1.1.a

Section 7 Simple Concepts:- Recall Devices

The simple concept languages are tested by the 42 questions (subset 2) searched on the 200 collection (subset 1). Fifteen separate languages are presented in the tables, 4.700T - 4.714T. Index language I.1.a uses the basic natural terms, and index language 1.2.a incorporates the recall device of synonym control. All subsequent index languages are based on index language I.2.a,

Index languages I.3.a, I.4.a, I.5.a, I.6.a, I.7.a and I.8.a investigate different selections of classes based on the schedules as given in Appendix 5.4 of Volume I. Index languages I.9.a, I.10.a and I.11.a represent selections from the alphabetical index as given in Appendix 5.5 of Volume I. Index languages I.12.a, I.13.a and I.14.a investigate the complete classes of the schedules of Appendix 5.4, while index language I.15.a brings in both the alphabetical and the hierarchical languages.

A table presenting the fifteen languages and their relationships appears as Figure 2.6. (page 10).

Seven graphs are included in this section, each plotting two or more of the sets of test results; these graphs fall into three main groups and provide comparison of

- (1) the different hierarchical devices, including the selective and obligatory types, (Figs. 4.715P, 4.716P, 4.718P, 4.719P and 4.721P).
- (2) the alphabetical device, (Figs. 4.717P and 4.721P).
- (3) hierarchical and alphabetical devices (Figs. 4.719P and 4.721P).

LIST OF FIGURES

| | Index Language | No. of Questions | Document Subset | Document Collection |
|--------|-------------------|---------------------|--------------------|------------------------|
| 4.700T | II.1.a | 42 | 2 | 200 |
| 4.701T | II.2.a | 42 | 2 | 200 |
| 4.702T | II.3.a | 42 | 2 | 200 |
| 4.703T | II.4.a | 42 | 2 | 200 |
| 4.704T | II.5.a | 42 | 2 | 200 |
| 4.705T | II.6.a | 42 | 2 | 200 |
| 4.706T | II.7.a | 42 | 2 | 200 |
| 4.707T | II.8.a | 42 | 2 | 200 |
| 4.708T | II.9.a | 42 | 2 | 200 |
| 4.709T | II.10.a | 42 | 2 | 200 |
| 4.710T | II.11.a | 42 | 2 | 200 |
| 4.711T | II.12.a | 42 | 2 | 200 |

| | Index Language | No. of Questions | Document Subset | Document Collection | Plot |
|--------|---|---------------------|--------------------|------------------------|--|
| 4.712T | II.13.a | 42 | 2 | 200 | |
| 4.713T | II.14.a | 42 | 2 | 200 | |
| 4.714T | II.15.a | 42 | 2 | 200 | |
| 4.715P | II.1.a, II.2.a, II.3.a, II.4.a, | 42 | 2 | 200 | Plots 4.700T, 4.701T, 4.702T, 4.703T. |
| 4.716P | II.2.a, II.6.a, II.7.a, | 42 | 2 | 200 | Plots 4.701T, 4.705T, 4.706T. |
| 4.717P | II.2.a, II.9.a, II.10.a, | 42 | 2 | 200 | Plots 4.701T, 4.708T, 4.709T. |
| 4.718P | II.2.a, II.12.a, II.13.a, II.14.a, | 42 | 2 | 200 | Plots 4.701T, 4.711T, 4.712T, 4.713T. |
| 4.719P | II.5.a, II.6.a, II.7.a, II.9.a, | 42 | 2 | 200 | Plots 4.704T, 4.705T, 4.706T, 4.708T. |
| 4.720P | II.8.a, II.10.a, | 42 | 2 | 200 | Plots 4.707T, 4.709T. |
| 4.721P | II.8.a, II.11.a, II.14.a, II.15.a, | 42 | 2 | 200 | Plots 4.707T, 4.710T, 4.713T, 4.714T. |

FIGURE 4.700T

Index Language II.1.a (Simple concepts. Natural language. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 73 | 459 | 36.9% | 13.7% | 5.597% | 42 | 42 | 42 |
| 2 | 28 | 27 | 14.1% | 50.9% | 0.329% | 18 | 41 | 41 |
| 3 | 10 | 0 | 5.1% | 100.0% | 0.000% | 5 | 38 | 38 |
| 4 | 0 | 0 | | | | 0 | 32 | 32 |
| 5 | 0 | 0 | | | | 0 | 20 | 20 |
| 6 | 0 | 0 | | | | 0 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.701T

Index Language II.2.a (Simple Concepts. Synonyms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 101 | 635 | 51.0% | 13.7% | 7.743% | 42 | 42 | 42 |
| 2 | 44 | 65 | 22.2% | 40.4% | 0.792% | 25 | 41 | 41 |
| 3 | 10 | 2 | 5.1% | 83.3% | 0.024% | 6 | 38 | 38 |
| 4 | 2 | 0 | 1.0% | 100.0% | 0.000% | 2 | 32 | 32 |
| 5 | 0 | 0 | | | | 0 | 20 | 20 |
| 6 | 0 | 0 | | | | 0 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.702T

Index Language II.3.a (Simple Concepts. Species (selected). Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 115 | 1,010 | 58.1% | 10.2% | 12.317% | 42 | 42 | 42 |
| 2 | 60 | 115 | 30.3% | 34.3% | 1.890% | 32 | 41 | 41 |
| 3 | 24 | 8 | 12.1% | 75.0% | 0.098% | 12 | 38 | 38 |
| 4 | 4 | 0 | 2.0% | 100.0% | 0.000% | 3 | 32 | 32 |
| 5 | 0 | 0 | | | | 0 | 20 | 20 |
| 6 | 0 | 0 | | | | 0 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.703T

Index Language II.4.a (Simple concepts. Superordinate (Selected). Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 129 | 1,338 | 65.2% | 8.8% | 16.292% | 42 | 42 | 42 |
| 2 | 67 | 255 | 33.8% | 20.8% | 3.109% | 36 | 41 | 41 |
| 3 | 19 | 40 | 9.6% | 32.2% | 0.488% | 16 | 38 | 38 |
| 4 | 2 | 8 | 1.0% | 20.0% | 0.098% | 4 | 32 | 32 |
| 5 | 0 | 2 | 0.0% | 0.0% | 0.024% | 1 | 20 | 20 |
| 6 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.704T

Index Language II.5.a (Simple concepts. Species and Superordinate (Selected). Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 138 | 1,614 | 69.7% | 7.8% | 19.682% | 42 | 42 | 42 |
| 2 | 81 | 359 | 40.9% | 18.4% | 4.377% | 39 | 41 | 41 |
| 3 | 32 | 60 | 16.2% | 34.8% | 0.732% | 22 | 38 | 38 |
| 4 | 6 | 10 | 3.0% | 37.5% | 0.122% | 5 | 32 | 32 |
| 5 | 0 | 4 | 0.0% | 0.0% | 0.049% | 1 | 20 | 20 |
| 6 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.705T

Index Language II.6.a (Simple Concepts. Co-ordinate (Selected). Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 118 | 1,124 | 59.6% | 9.5% | 13.707% | 42 | 42 | 42 |
| 2 | 57 | 155 | 28.8% | 26.9% | 1.890% | 33 | 41 | 41 |
| 3 | 18 | 26 | 9.1% | 40.9% | 0.219% | 11 | 38 | 38 |
| 4 | 4 | 5 | 2.0% | 44.4% | 0.061% | 4 | 32 | 32 |
| 5 | 0 | 0 | | | | 0 | 20 | 20 |
| 6 | 0 | 0 | | | | 0 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.706T

Index Language II.7.a (Simple Concepts. Co-ordinate and Collateral (Selected). Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 137 | 1,532 | 69.2% | 8.2% | 18.682% | 42 | 42 | 42 |
| 2 | 75 | 318 | 37.9% | 19.1% | 3.877% | 33 | 41 | 41 |
| 3 | 23 | 50 | 11.6% | 31.5% | 0.610% | 14 | 38 | 38 |
| 4 | 6 | 9 | 3.0% | 40.0% | 0.110% | 4 | 32 | 32 |
| 5 | 1 | 0 | 0.5% | 100.0% | 0.000% | 1 | 20 | 20 |
| 6 | 0 | 0 | | | | 0 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.707T

Index Language II.8.a (Simple Concepts. Species, Superordinate, Coordinate and Collateral (Selected). Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 159 | 2,133 | 80.3% | 6.9% | 26.012% | 42 | 42 | 42 |
| 2 | 99 | 805 | 50.0% | 14.1% | 7.376% | 40 | 41 | 41 |
| 3 | 38 | 123 | 19.2% | 23.6% | 1.500% | 23 | 38 | 38 |
| 4 | 14 | 33 | 7.1% | 29.8% | 0.402% | 10 | 32 | 32 |
| 5 | 5 | 8 | 2.5% | 38.5% | 0.098% | 2 | 20 | 20 |
| 6 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.708T

Index Language II.9.a (Simple Concepts. First alphabetical collateral (Selected). Coordination).
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 135 | 1,524 | 68.2% | 8.1% | 18.585% | 42 | 42 | 42 |
| 2 | 75 | 296 | 37.9% | 20.2% | 3.609% | 33 | 41 | 41 |
| 3 | 37 | 45 | 18.7% | 45.1% | 0.549% | 19 | 38 | 38 |
| 4 | 13 | 6 | 6.6% | 68.4% | 0.073% | 8 | 32 | 32 |
| 5 | 1 | 2 | 0.5% | 33.6% | 0.024% | 3 | 20 | 20 |
| 6 | 0 | 0 | | | | 0 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.709T

Index Language II.10.a (Simple Concepts. Second alphabetical collateral (selected). Coordination.)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 182 | 2,892 | 91.9% | 5.9% | 35.288% | 42 | 42 | 42 |
| 2 | 129 | 942 | 65.2% | 12.0% | 11.485% | 32 | 41 | 41 |
| 3 | 79 | 253 | 40.0% | 23.8% | 3.085% | 34 | 38 | 38 |
| 4 | 27 | 44 | 13.6% | 38.0% | 0.536% | 18 | 32 | 32 |
| 5 | 7 | 6 | 3.5% | 53.8% | 0.073% | 6 | 20 | 20 |
| 6 | 1 | 0 | 0.5% | 100.0% | 0.000% | 1 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.710T

Index Language II.11.a (Simple Concepts. Species, superordinate, co-ordinate, collateral, and second alphabetical collateral (selected). Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 183 | 3,317 | 92.4% | 5.2% | 40.451% | 42 | 42 | 42 |
| 2 | 146 | 1,259 | 73.7% | 10.4% | 15.350% | 41 | 41 | 41 |
| 3 | 94 | 408 | 47.5% | 18.72% | 4.974% | 37 | 38 | 38 |
| 4 | 34 | 99 | 17.2% | 25.6% | 1.207% | 19 | 32 | 32 |
| 5 | 14 | 25 | 7.1% | 35.9% | 0.305% | 9 | 20 | 20 |
| 6 | 3 | 1 | 1.5% | 75.0% | 0.012% | 3 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.711T

Index Language II.12.a (Simple Concepts. Species (complete). Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 128 | 1,432 | 64.6% | 8.2% | 17.463% | 42 | 42 | 42 |
| 2 | 72 | 287 | 38.4% | 20.1% | 3.500% | 34 | 41 | 41 |
| 3 | 27 | 32 | 13.6% | 45.8% | 0.390% | 13 | 38 | 38 |
| 4 | 9 | 3 | 4.5% | 75.0% | 0.037% | 5 | 32 | 32 |
| 5 | 0 | 0 | | | | 0 | 20 | 20 |
| 6 | 0 | 0 | | | | 0 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.712T

Index Language II.13.a (Simple Concepts. Species, Superordinate and Co-ordinate (complete). Coordination).
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 168 | 2,490 | 84.6% | 6.3% | 30.365% | 42 | 42 | 42 |
| 2 | 117 | 937 | 59.1% | 11.1% | 11.424% | 40 | 41 | 41 |
| 3 | 49 | 244 | 24.7% | 16.7% | 2.975% | 29 | 38 | 38 |
| 4 | 20 | 54 | 10.1% | 27.0% | 0.658% | 11 | 32 | 32 |
| 5 | 8 | 21 | 4.0% | 27.6% | 0.256% | 3 | 20 | 20 |
| 6 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.713T

Index Language II.14.a (Simple Concepts. Species, Superordinate, Coordinate and Collateral (complete). Coordination).
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 180 | 3,684 | 90.9% | 4.6% | 44.926% | 42 | 42 | 42 |
| 2 | 143 | 2,047 | 72.2% | 6.5% | 24.957% | 40 | 41 | 41 |
| 3 | 66 | 600 | 33.3% | 9.9% | 7.315% | 33 | 38 | 38 |
| 4 | 26 | 179 | 13.1% | 12.7% | 2.182% | 19 | 32 | 32 |
| 5 | 13 | 59 | 6.6% | 18.1% | 0.719% | 7 | 20 | 20 |
| 6 | 0 | 11 | 0.0% | 0.0% | 0.134% | 2 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

FIGURE 4.714T

Index Language II.15.a (Simple Concepts. Species, Superordinate, Coordinate, and Collateral (complete), Second alphabetical Collateral (selected). Coordination).
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 190 | 4,542 | 96.0% | 4.0% | 55.390% | 42 | 42 | 42 |
| 2 | 168 | 2,590 | 84.8% | 6.1% | 31.578% | 41 | 41 | 41 |
| 3 | 117 | 1,020 | 59.1% | 10.3% | 12.436% | 37 | 38 | 38 |
| 4 | 51 | 325 | 25.8% | 13.6% | 3.962% | 26 | 32 | 32 |
| 5 | 21 | 96 | 10.6% | 17.9% | 1.170% | 12 | 20 | 20 |
| 6 | 4 | 15 | 2.0% | 21.1% | 0.183% | 5 | 9 | 9 |
| 7 | 0 | 0 | | | | 0 | 3 | 3 |

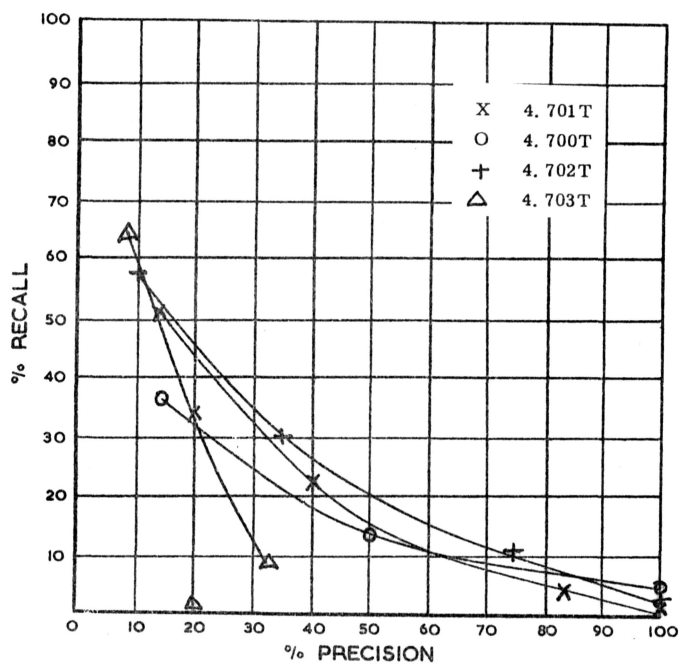


FIGURE 4.715P SIMPLE CONCEPT INDEX LANGUAGES II.1.a, II.2.a, II.3.a, and II.4.a (FIGURES 4.700T - 4.703T)

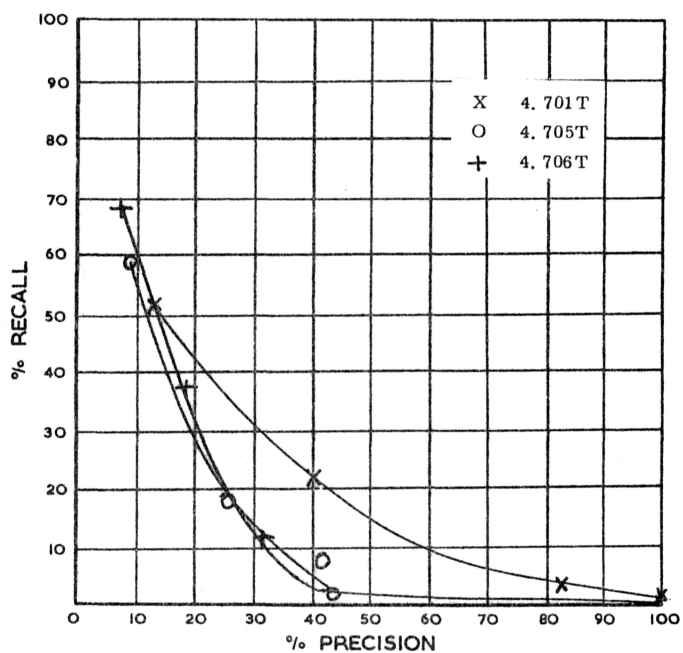


FIGURE 4.716P SIMPLE CONCEPT INDEX LANGUAGES II.2.a, II.6.a and II.7.a (FIGURES 4.701T, 4.705T and 4.706T)

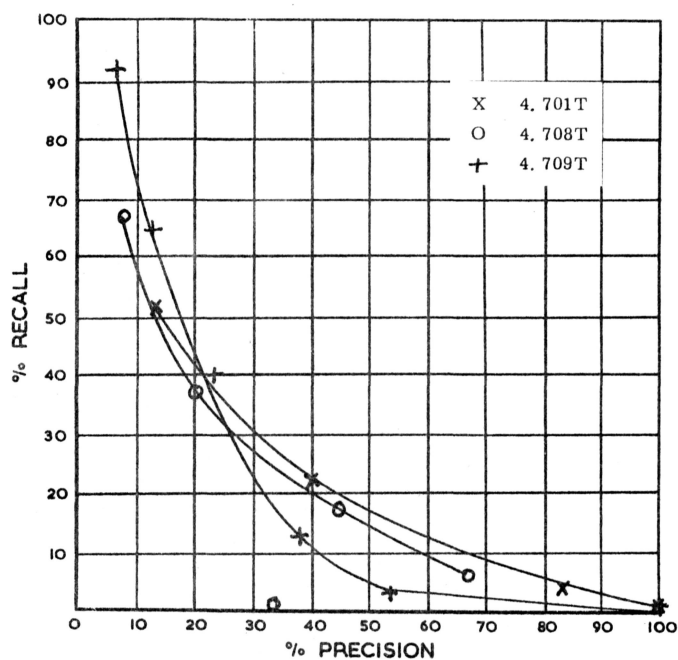


FIGURE 4.717P SIMPLE CONCEPT INDEX LANGUAGES II.2.a, II.9.a, and II.10.a (FIGURES 4.701T, 4.708T and 4.709T)

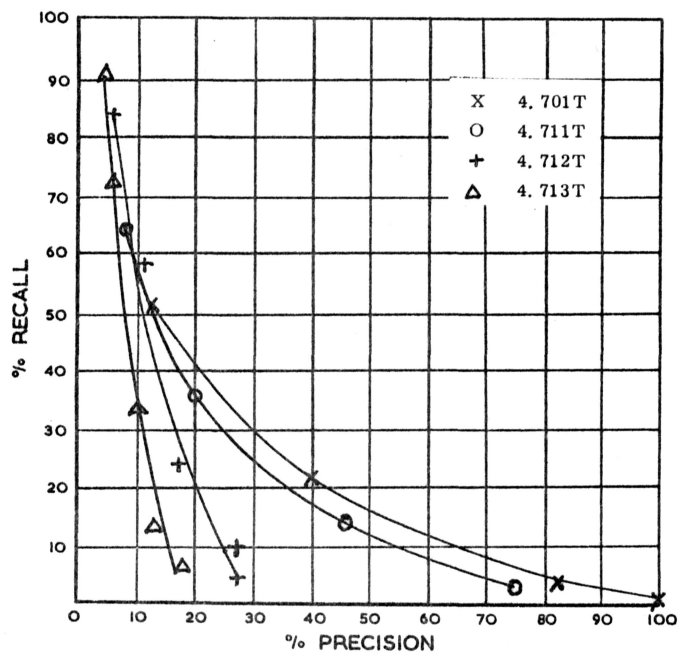


FIGURE 4.718P SIMPLE CONCEPT INDEX LANGUAGES II.2.a, II.12.a, II.13.a and II.14.a (FIGURES 4.701T, 4.711T, 4.712T and 4.713T)

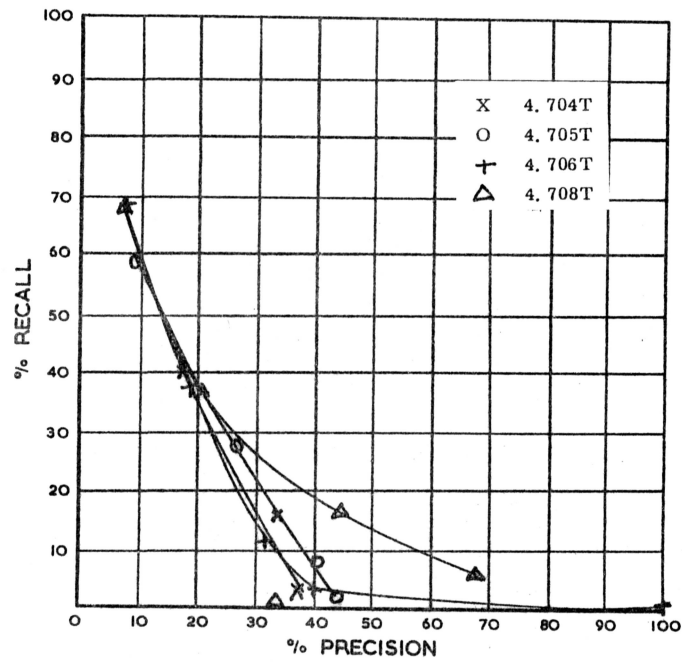


FIGURE 4.719P SIMPLE CONCEPT INDEX LANGUAGES II.5.a, II.6.a, II.7.a and II.9.a (FIGURES 4.704T, 4.705T, 4.706T and 4.708T)

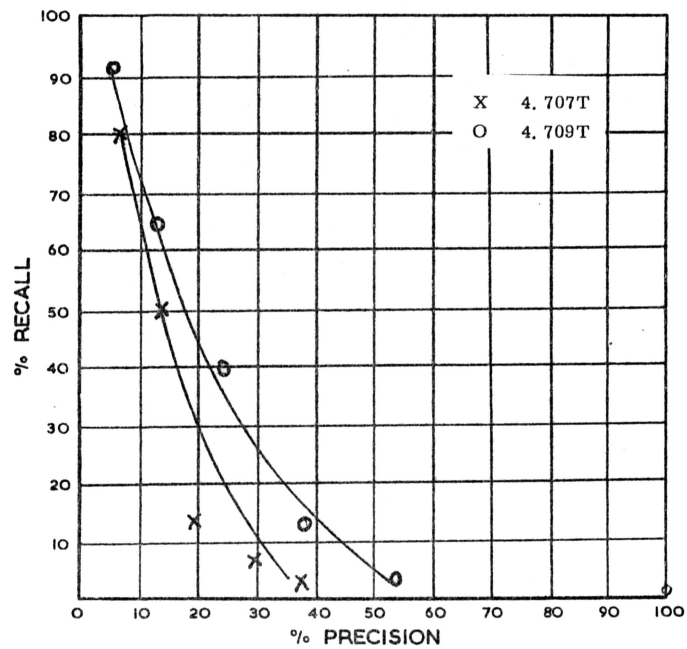


FIGURE 4.720P SIMPLE CONCEPT INDEX LANGUAGES II.8.a, II.10.a (FIGURES 4.707T and 4.709T)

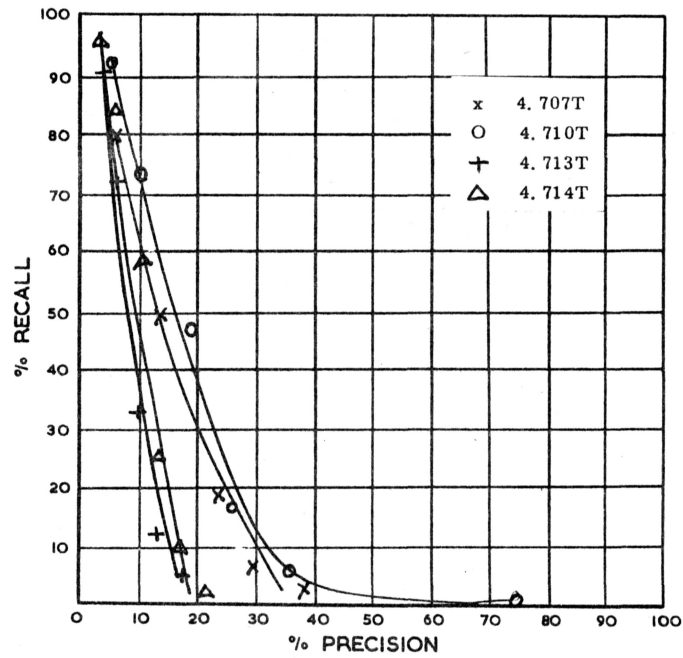


FIGURE 4.721P

SIMPLE CONCEPT INDEX LANGUAGES II.8.a,
II.11.a, II.14.a and II.15.a (FIGURES 4.707T,
4.710T, 4.713T and 4.714T)

Section 8 Controlled terms

Recall devices

The investigations on the controlled term languages were carried out on two sets of questions and two collection sizes: the 42 questions (subset 2) searched on the 200 document collection (subset 1), and the 77 questions (subset 7) searched on the 350 document collection (subset 2). Six recall languages were tested, and the effect of each recall device is seen in the tables and plots given in Figs. 4.800TP - 4.805TP, tested with a search rule A. A plot of the six curves is given in Figure 4.806P. The recall devices are shown in Fig. 2.7 (p.11)

Search Rules

Further comparison of the six index languages is given with the tests made using Search Rule E on both the 42 questions with the 200 document collection and the 77 questions with the 350 document collection. The results are presented in Figs. 4.810TP - 4.815TP and 4.820T - 4.825T, with plots covering the various index languages as Figs. 4.816P and 4.826P. With Search Rule E all the combinations of acceptable terms were selected for each coordination level. Examples of search formulations for a number of questions are given in Appendix 8.1.

Tests with Search Rule F were done on index languages III.5.a and III.6.a with 42 questions on the 200 document collection. This search was superimposed on Search E, and the results are presented at the various coordination levels according to the number of basic terms as apart from related terms. Thus, at a coordination level of 4, the tables (Figs. 4.850T and 4.851T) show the results when all terms were basic terms, where one term was a related term, where two terms were related terms and so on. It is obvious that when all the terms are basic terms, then the results must be the same as for index language III.1.a with Search E; when all the terms are related terms, then the results must be the same as for the corresponding index language with the basic Search E. Therefore in the plot 4.850P, the two main curves represent index languages III.1.a and III.5.a with Search E and in Figure 4.851P the main curves represent index languages III.1.a and III.6.a, again with Search E. The additional results obtained with Search F now produce a series of secondary curves at each coordination level which span the main curves.

Precision Device

The precision device of weighting was tested, in which search questions were weighted and tested on the most exhaustive index language, using the weights assigned in indexing. Figs. 4.830TP and 4.831TP give results for languages III.1.a and III.6.a respectively, and in the plots a comparison is made of the weighted and unweighted searches, carried out with search rule E.

SECTION 8 LIST OF FIGURES

| | Index Language | Search Rule | No. of Questions | Document Collection | Plot |
|---------|--|----------------|---------------------|------------------------|--|
| 4.800TP | III.1.a | A | 42 (2) | 200 | |
| 4.801TP | III.2.a | A | 42 (2) | 200 | |
| 4.802TP | III.3.a | A | 42 (2) | 200 | |
| 4.803TP | III.4.a | A | 42 (2) | 200 | |
| 4.804TP | III.5.a | A | 42 (2) | 200 | |
| 4.805TP | III.6.a | A | 42 (2) | 200 | |
| 4.806P | III.1.a, III.2.a, III.3.a, III.4.a, III.5.a, III.6.a | A | 42 (2) | 200 | 4.800T, 4.801T, 4.802T, 4.803T, 4.804T, 4.805T |
| 4.810TP | III.1.a | E | 42 (2) | 200 | |
| 4.811TP | III.2.a | E | 42 (2) | 200 | |
| 4.812TP | III.3.a | E | 42 (2) | 200 | |
| 4.813TP | III.4.a | E | 42 (2) | 200 | |
| 4.814TP | III.5.a | E | 42 (2) | 200 | |
| 4.815TP | III.6.a | E | 42 (2) | 200 | |
| 4.816P | III.1.a, III.2.a, III.3.a, III.4.a, III.5.a, III.6.a | E | 42 (2) | 200 | 4.810T, 4.811T, 4.812T, 4.813T, 4.814T, 4.815T |
| 4.820T | III.1.a | E | 77 (7) | 350 | |
| 4.821T | III.2.a | E | 77 (7) | 350 | |
| 4.822T | III.3.a | E | 77 (7) | 350 | |
| 4.823T | III.4.a | E | 77 (7) | 350 | |
| 4.824T | III.5.a | E | 77 (7) | 350 | |
| 4.825T | III.6.a | E | 77 (7) | 350 | |
| 4.826P | III.1.a, III.2.a, III.3.a, III.4.a, III.5.a, III.6.a | E | 77 (7) | 350 | 4.820T, 4.821T, 4.822T, 4.823T, 4.824T, 4.825T |
| 4.830TP | III.1.e (weighting) | E | 42 (2) | 200 | + 4.810T |
| 4.831TP | III.6.e (weighting) | E | 42 (2) | 200 | + 4.815T |
| 4.840P | III.1.a | A & E | 42 (2) | 200 | 4.800T, 4.810T |
| 4.841P | III.2.a | A & E | 42 (2) | 200 | 4.801T, 4.811T |
| 4.842P | III.3.a | A & E | 42 (2) | 200 | 4.802T, 4.812T |
| 4.843P | III.4.a | A & E | 42 (2) | 200 | 4.803T, 4.813T |
| 4.844P | III.5.a | A & E | 42 (2) | 200 | 4.804T, 4.814T |
| 4.845P | III.6.a | A & E | 42 (2) | 200 | 4.805T, 4.815T |
| 4.850TP | III.5.a | F | 42 (2) | 200 | + 4.814T |
| 4.851TP | III.6.a | F | 42 (2) | 200 | + 4.815T |

FIGURE 4.800T

Index Language III.1.a (Controlled terms. Basic terms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 174 | 2,884 | 87.9% | 5.6% | 35.170% | 42 | 42 | 42 |
| 2 | 136 | 946 | 68.7% | 12.6% | 11.534% | 42 | 42 | 42 |
| 3 | 79 | 213 | 39.9% | 27.1% | 2.597% | 31 | 41 | 41 |
| 4 | 36 | 49 | 18.2% | 42.4% | 0.597% | 20 | 34 | 34 |
| 5 | 17 | 7 | 8.6% | 70.8% | 0.085% | 6 | 24 | 24 |
| 6 | 5 | 3 | 2.5% | 62.5% | 0.037% | 3 | 13 | 13 |
| 7 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 8 | 8 |
| 8 | 1 | 1 | 0.5% | 50.0% | 0.012% | 2 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

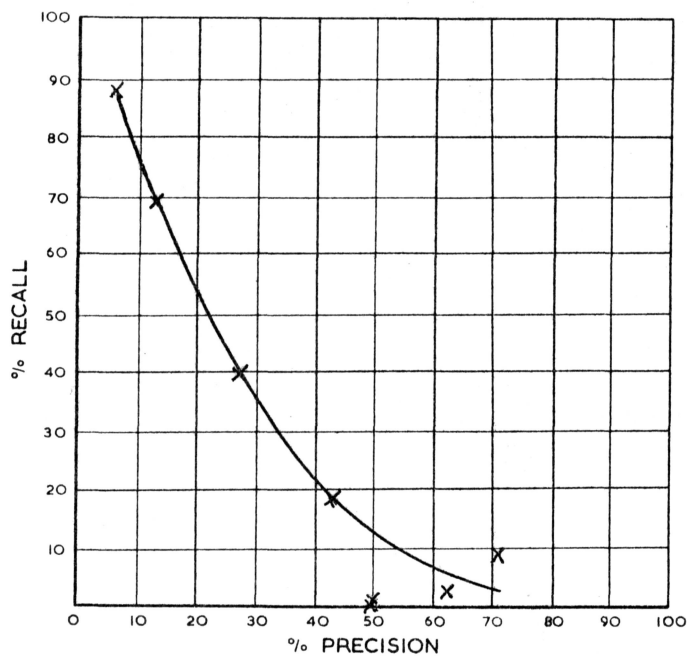


FIGURE 4.800P INDEX LANGUAGE III.1.a

FIGURE 4.801T

Index Language III.2.a (Controlled terms. Narrower terms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 176 | 3.062 | 88.9% | 5.4% | 37.34% | 42 | 42 | 42 |
| 2 | 137 | 1,024 | 69.2% | 11.8% | 12.485% | 42 | 42 | 42 |
| 3 | 83 | 232 | 41.9% | 26.3% | 2.829% | 31 | 41 | 41 |
| 4 | 38 | 53 | 19.2% | 41.8% | 0.646% | 20 | 34 | 34 |
| 5 | 17 | 9 | 8.6% | 65.4% | 0.110% | 6 | 24 | 24 |
| 6 | 6 | 3 | 3.0% | 66.7% | 0.037% | 3 | 13 | 13 |
| 7 | 2 | 2 | 1.0% | 50.0% | 0.012% | 2 | 8 | 8 |
| 8 | 1 | 1 | 0.5% | 50.0% | 0.012% | 2 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

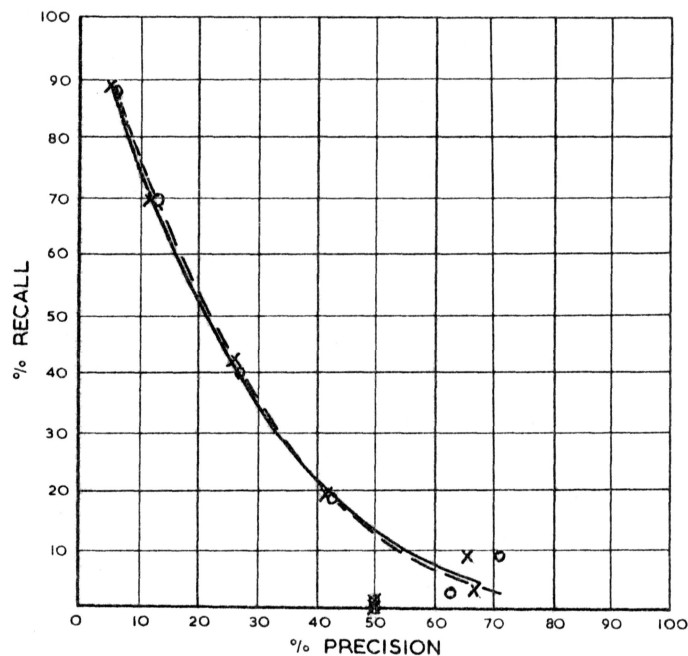


FIGURE 4.801P INDEX LANGUAGE III.2.a
(Index Language III.1.a Broken line)

FIGURE 4.802T

Index Language III.3.a (Controlled terms. Broader Terms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule A
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 178 | 3,565 | 89.9% | 4.7% | 43.475% | 42 | 42 | 42 |
| 2 | 147 | 1,575 | 74.2% | 8.5% | 19.203% | 42 | 42 | 42 |
| 3 | 90 | 364 | 45.5% | 19.8% | 4.438% | 37 | 41 | 41 |
| 4 | 43 | 78 | 21.7% | 35.5% | 0.951% | 22 | 34 | 34 |
| 5 | 18 | 8 | 9.1% | 69.2% | 0.098% | 6 | 24 | 24 |
| 6 | 7 | 3 | 3.7% | 70.0% | 0.037% | 4 | 13 | 13 |
| 7 | 22 | 2 | 1.0% | 50.0% | 0.024% | 4 | 8 | 8 |
| 8 | 1 | 1 | 0.5% | 50.0% | 0.012% | 2 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

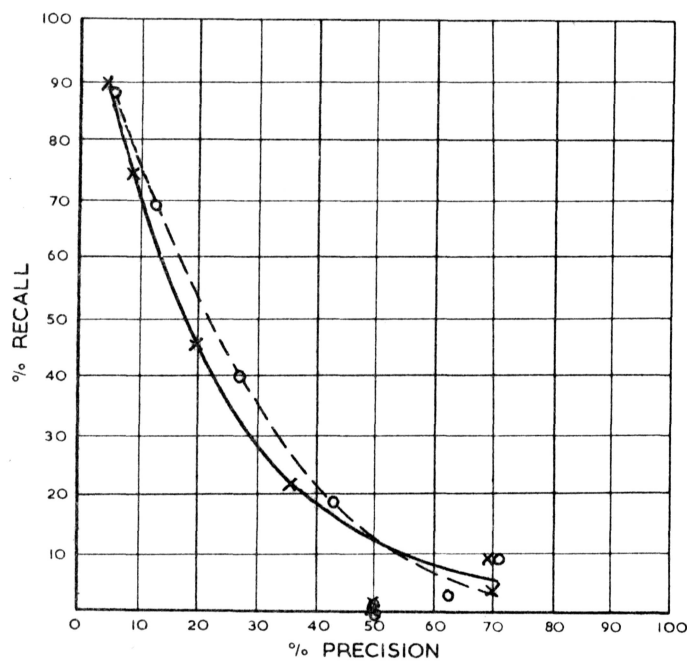


FIGURE 4.802P INDEX LANGUAGE III.3.a
 (Index Language III.1.a Broken line)

FIGURE 4.803T

Index Language III.4.a (Controlled terms. Narrower and Broader terms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 180 | 3,614 | 90.9% | 4.7% | 44.073 | 42 | 42 | 42 |
| 2 | 148 | 1,661 | 74.7% | 8.2% | 20.251% | 42 | 42 | 42 |
| 3 | 94 | 386 | 47.5% | 19.6% | 4.706% | 37 | 41 | 41 |
| 4 | 44 | 82 | 22.2% | 34.9% | 1.000% | 23 | 34 | 34 |
| 5 | 18 | 12 | 9.1% | 60.0% | 0.146% | 8 | 24 | 24 |
| 6 | 8 | 3 | 4.0% | 72.7% | 0.037% | 4 | 13 | 13 |
| 7 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 8 | 8 |
| 8 | 1 | 1 | 0.5% | 50.0% | 0.012% | 2 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

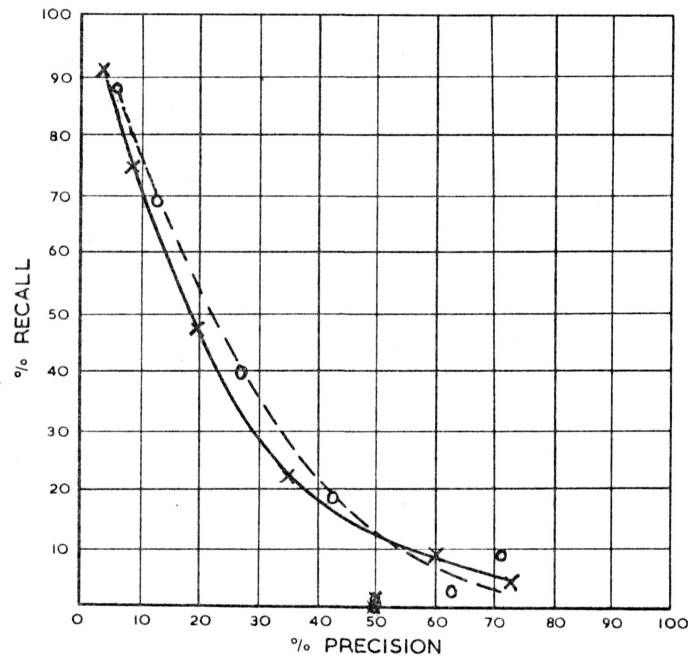


FIGURE 4.803P INDEX LANGUAGE III. 4. a
(Index Language III. 1. a Broken line)

FIGURE 4.804T

Index Language III.5.a (Controlled terms. Related terms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 42 (Subset 2)

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 195 | 5,424 | 98.5% | 3.4% | 66.146% | 42 | 42 | 42 |
| 2 | 184 | 3,044 | 92.9% | 5.7% | 37.113% | 42 | 42 | 42 |
| 3 | 133 | 1,252 | 67.2% | 9.6% | 15.265% | 41 | 41 | 41 |
| 4 | 62 | 420 | 31.3% | 12.9% | 5.121% | 29 | 34 | 34 |
| 5 | 35 | 122 | 17.7% | 22.3% | 1.487% | 19 | 24 | 24 |
| 6 | 12 | 26 | 6.1% | 31.6% | 0.317% | 7 | 13 | 13 |
| 7 | 4 | 7 | 2.0% | 36.4% | 0.085% | 3 | 8 | 8 |
| 8 | 2 | 3 | 1.0% | 40.0% | 0.037% | 2 | 4 | 4 |
| 9 | 2 | 1 | 1.0% | 66.7% | 0.012% | 2 | 3 | 3 |

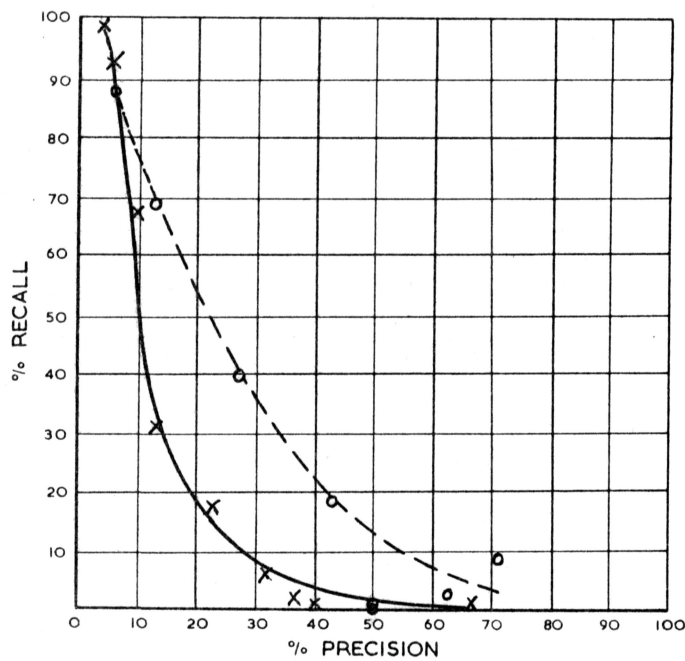


FIGURE 4.804P INDEX LANGUAGE III.5.a
(Index Language III.1.a Broken line)

FIGURE 4.805T

Index Language III.6.a (Controlled terms. Narrower, broader and related terms. Coordination)

Exhaustivity of Indexing 3

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 196 | 5,665 | 99.0% | 3.3% | 69.085% | 42 | 42 | 42 |
| 2 | 187 | 3,483 | 94.4% | 5.1% | 42.465% | 42 | 42 | 42 |
| 3 | 149 | 1,609 | 75.3% | 8.5% | 19.617% | 41 | 41 | 41 |
| 4 | 72 | 527 | 36.4% | 12.0% | 6.425% | 31 | 34 | 34 |
| 5 | 38 | 140 | 19.2% | 21.3% | 1.707% | 20 | 24 | 24 |
| 6 | 13 | 31 | 6.6% | 29.5% | 0.378% | 7 | 13 | 13 |
| 7 | 4 | 7 | 2.0% | 36.4% | 0.085% | 3 | 8 | 8 |
| 8 | 2 | 3 | 1.0% | 40.0% | 0.037% | 2 | 4 | 4 |
| 9 | 2 | 1 | 1.0% | 66.7% | 0.012% | 2 | 3 | 3 |

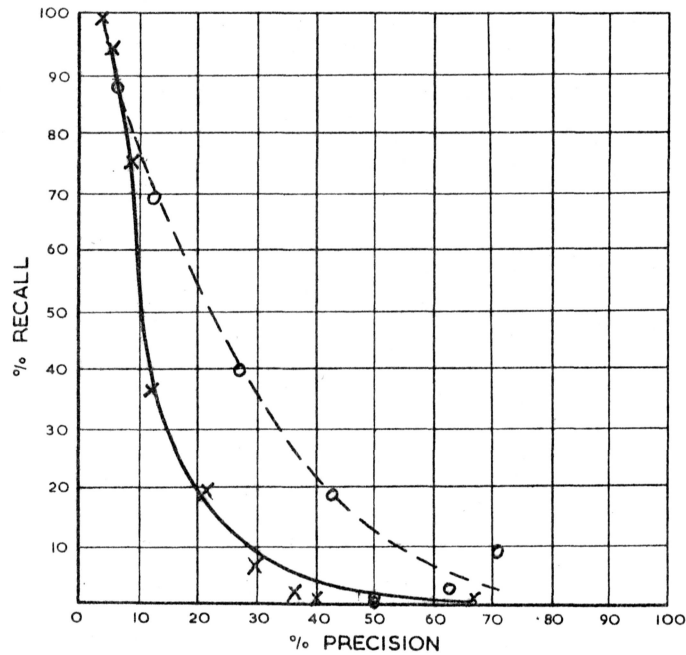


FIGURE 4.805P INDEX LANGUAGE III.6.a
(Index Language III.1.a Broken line)

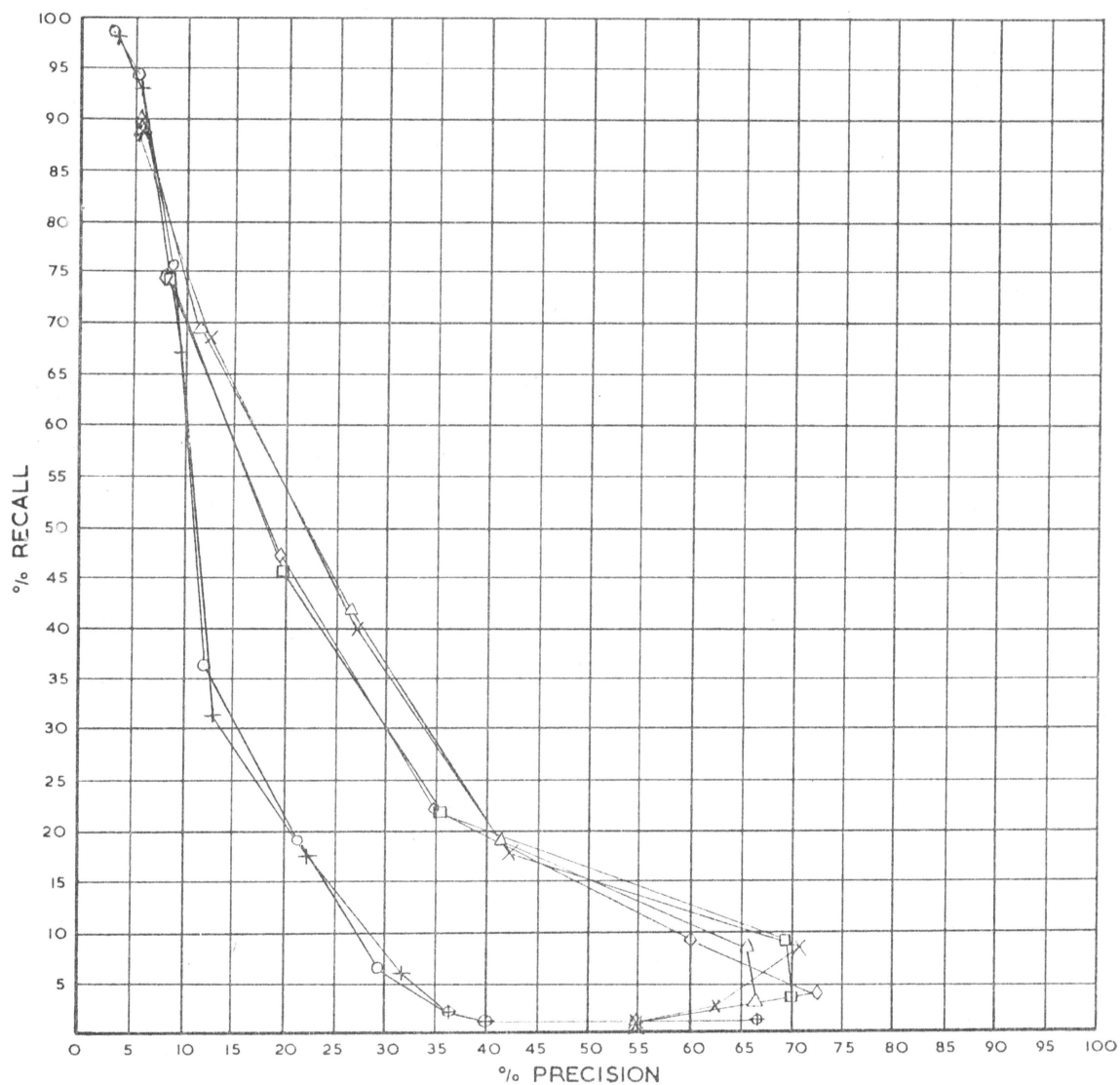


FIGURE 4.806P INDEX LANGUAGES III. 1. a (X), III. 2. a (Δ),
 III. 3. a (□), III. 4. a (◇), III. 5. a (+) and
 III. 6. a (O). SEARCH A (Figures 4.800T-4.805T)

FIGURE 4.810T

Index Language III.1.a (Controlled terms. Basic terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 169 | 1,933 | 85.4% | 8.0% | 23,573% | 42 | 42 | 42 |
| 2 | 114 | 428 | 57.6% | 21.0% | 5.218% | 39 | 42 | 42 |
| 3 | 62 | 92 | 31.3% | 40.3% | 1.122% | 29 | 41 | 41 |
| 4 | 25 | 13 | 12.6% | 65.8% | 0.158% | 15 | 34 | 34 |
| 5 | 10 | 4 | 5.1% | 71.4% | 0.049% | 6 | 24 | 24 |
| 6 | 0 | 2 | 0.0% | 0.0% | 0.024% | 1 | 13 | 13 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 8 | 8 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 4 | 4 |
| 9 | 0 | 0 | 0.0% | 0.0% | 0.012% | 0 | 3 | 3 |

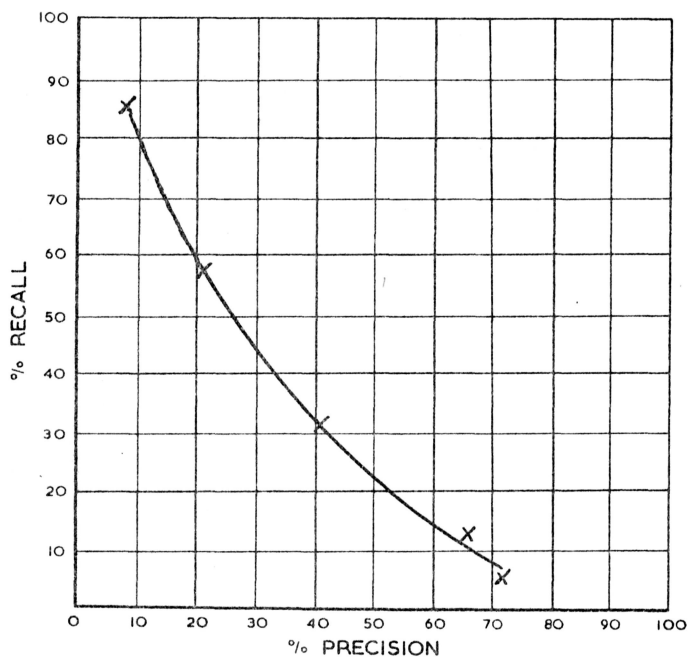


FIGURE 4.810P INDEX LANGUAGE III.1.a SEARCH E
200 DOCUMENTS

FIGURE 4.811T

Index Language III.2.a (Controlled terms. Narrower terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 171 | 2,064 | 86.4% | 7.6% | 25.414% | 42 | 42 | 42 |
| 2 | 116 | 455 | 58.6% | 20.3% | 5.547% | 39 | 42 | 42 |
| 3 | 64 | 100 | 32.3% | 39.0% | 1.219% | 29 | 41 | 41 |
| 4 | 28 | 14 | 14.1% | 66.7% | 0.171% | 15 | 34 | 34 |
| 5 | 11 | 4 | 5.6% | 73.3% | 0.049% | 6 | 24 | 24 |
| 6 | 0 | 2 | 0.0% | 0.0% | 0.024% | 1 | 13 | 13 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 8 | 8 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 4 | 4 |
| 9 | 0 | 0 | 0.0% | 0.0% | 0.012% | 0 | 3 | 3 |

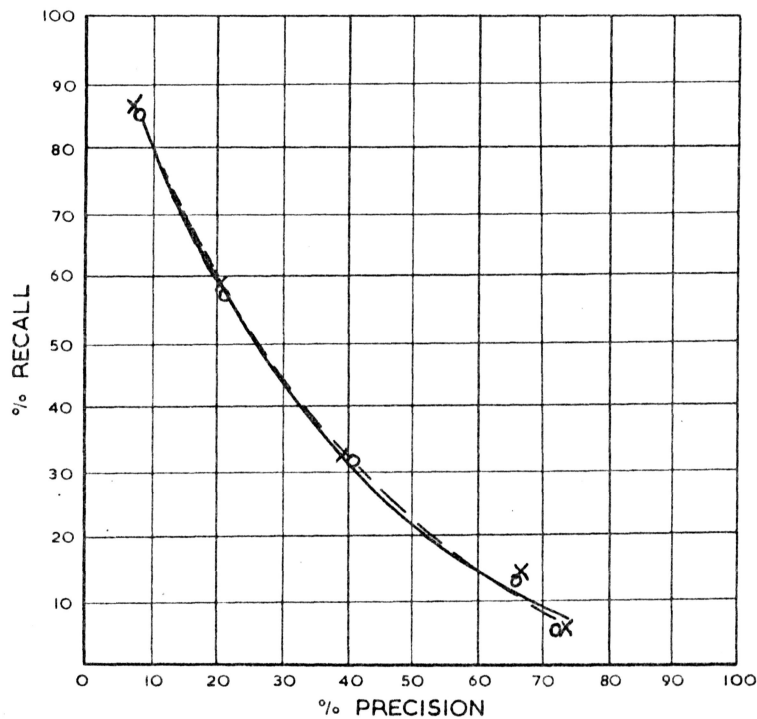


FIGURE 4.811P INDEX LANGUAGE III.2.a SEARCH E
200 DOCUMENTS
(Index Language III.1.a Broken line)

FIGURE 4.812T

Index Language III.3.a (Controlled terms. Broader terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 173 | 2,758 | 87.4% | 5.9% | 33.634% | 42 | 42 | 42 |
| 2 | 117 | 762 | 59.1% | 13.3% | 9.290% | 39 | 42 | 42 |
| 3 | 79 | 172 | 39.9% | 31.5% | 2.097% | 35 | 41 | 41 |
| 4 | 31 | 24 | 15.7% | 56.4% | 0.2926% | 19 | 34 | 34 |
| 5 | 12 | 5 | 6.1% | 70.6% | 0.0609% | 7 | 24 | 24 |
| 6 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 13 | 13 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 8 | 8 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

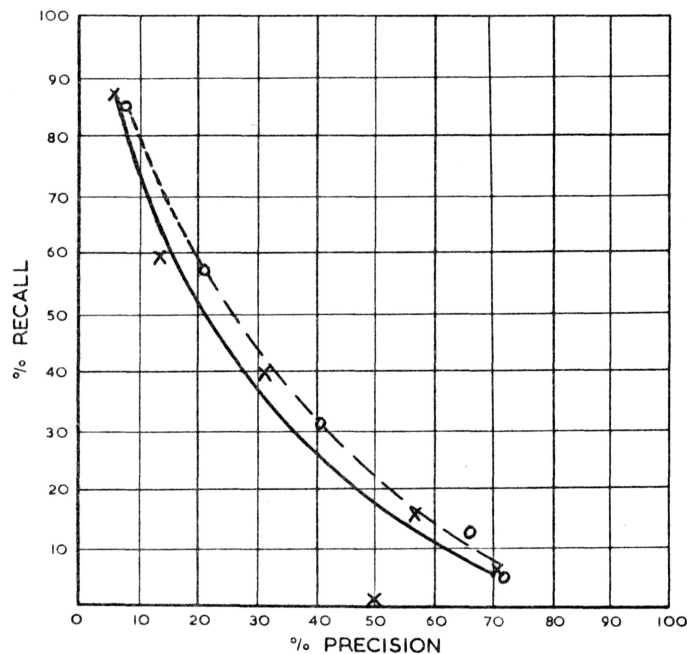


FIGURE 4.812P INDEX LANGUAGE III.3.a SEARCH E
200 DOCUMENTS
(Index Language III.1.a Broken line)

FIGURE 4.813T

Index Language III.4.a (Controlled terms. Narrower and broader terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 175 | 2,854 | 88.4% | 5.8% | 34.804% | 42 | 42 | 42 |
| 2 | 129 | 797 | 65.2% | 13.9% | 9.717% | 39 | 42 | 42 |
| 3 | 81 | 180 | 40.9% | 31.0% | 2.195% | 35 | 41 | 41 |
| 4 | 34 | 24 | 17.2% | 58.6% | 0.2926% | 19 | 34 | 34 |
| 5 | 13 | 6 | 6.6% | 68.4% | 0.073% | 7 | 24 | 24 |
| 6 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 13 | 13 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 8 | 8 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.012% | 1 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

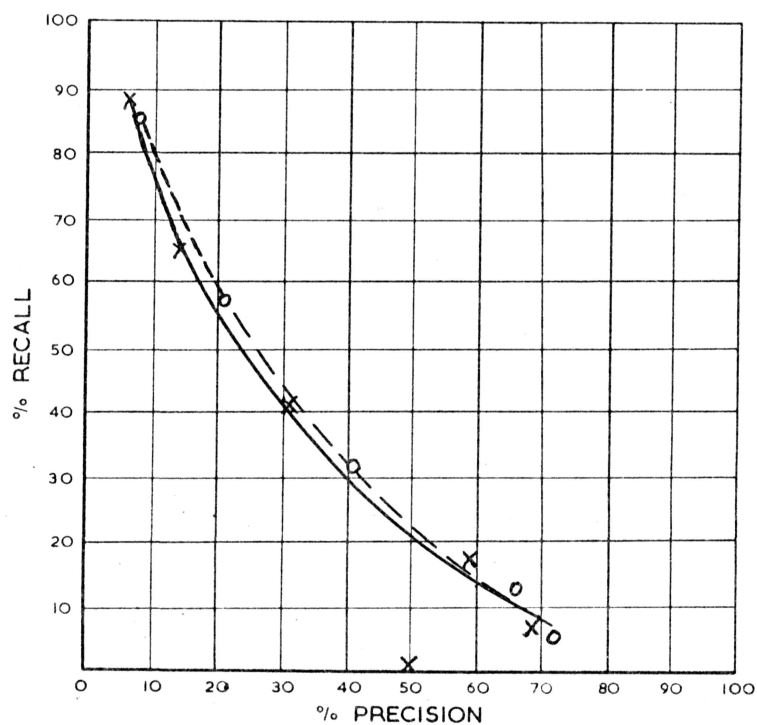


FIGURE 4.813P INDEX LANGUAGE III.4.a SEARCH E
200 DOCUMENTS
(Index Language III.1.a Broken line)

FIGURE 4.814T

Index Language III.5.a (Controlled terms. Related terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 194 | 4,339 | 98.0% | 4.3% | 52.914% | 42 | 42 | 42 |
| 2 | 156 | 1,660 | 78.8% | 8.6% | 20.239% | 42 | 42 | 42 |
| 3 | 104 | 628 | 52.5% | 14.2% | 7.657% | 40 | 41 | 41 |
| 4 | 48 | 155 | 24.2% | 23.6% | 1.890% | 27 | 34 | 34 |
| 5 | 27 | 47 | 13.6% | 36.5% | 0.573% | 17 | 24 | 24 |
| 6 | 11 | 10 | 5.6% | 52.4% | 0.122% | 5 | 13 | 13 |
| 7 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 8 | 8 |
| 8 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 4 | 4 |
| 9 | 2 | 1 | 1.0% | 66.7% | 0.012% | 2 | 3 | 3 |

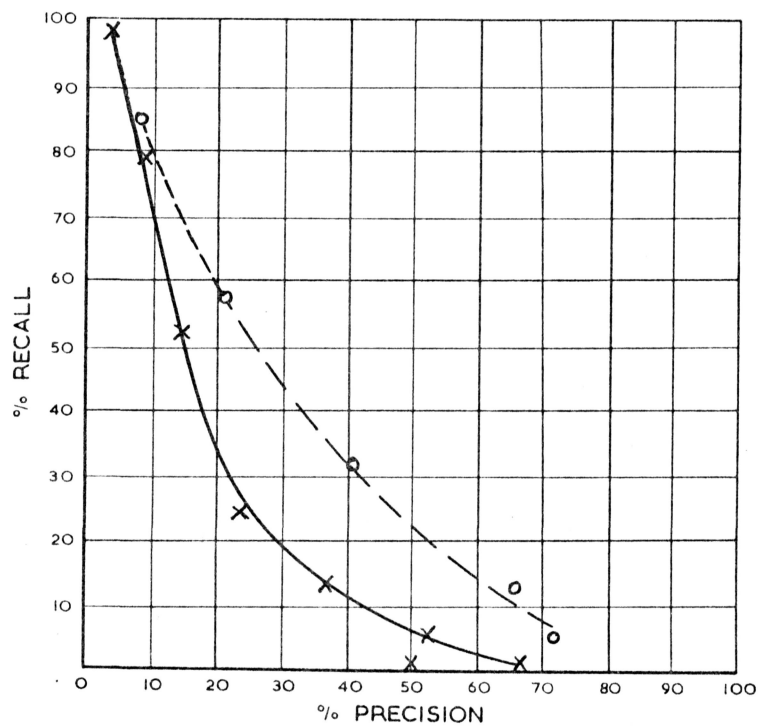


FIGURE 4.814P INDEX LANGUAGE III.5.a SEARCH E
200 DOCUMENTS
(Index Language III.1.a Broken line)

FIGURE 4.815T

Index Language III.6.a (Controlled terms. Narrower, broader and related terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 196 | 4,685 | 99.0% | 4.0% | 57.134% | 42 | 42 | 42 |
| 2 | 169 | 2,092 | 85.4% | 7.5% | 25.506% | 42 | 42 | 42 |
| 3 | 129 | 848 | 65.2% | 13.2% | 10.339% | 42 | 41 | 41 |
| 4 | 60 | 210 | 30.3% | 22.2% | 2.560% | 31 | 34 | 34 |
| 5 | 30 | 55 | 15.2% | 35.3% | 0.671% | 18 | 24 | 24 |
| 6 | 11 | 10 | 5.6% | 52.4% | 0.122% | 5 | 13 | 13 |
| 7 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 8 | 8 |
| 8 | 2 | 2 | 1.0% | 50.0% | 0.024% | 2 | 4 | 4 |
| 9 | 2 | 1 | 1.0% | 66.7% | 0.012% | 2 | 3 | 3 |

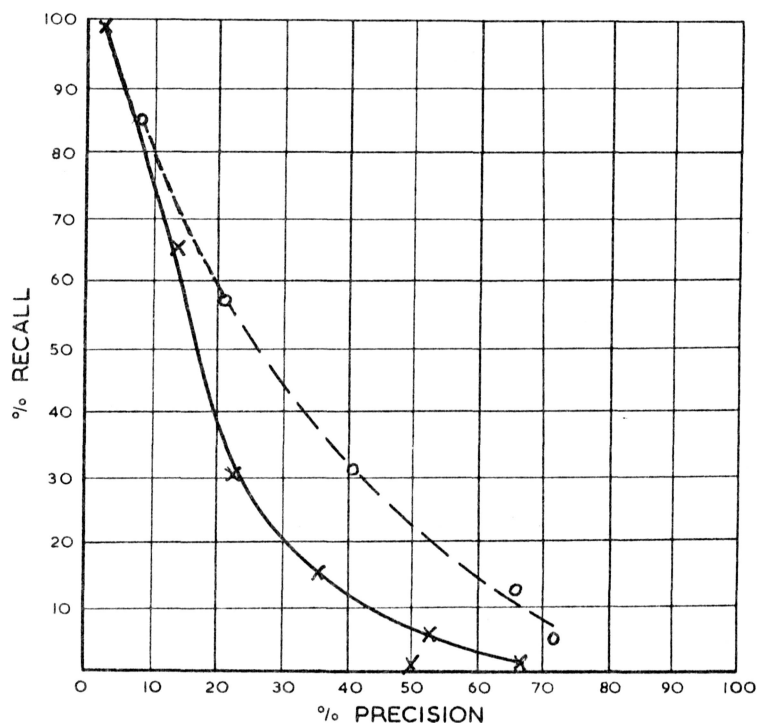


FIGURE 4.815P INDEX LANGUAGE III.6.a SEARCH E
200 DOCUMENTS
(Index Language III.1.a Broken line)

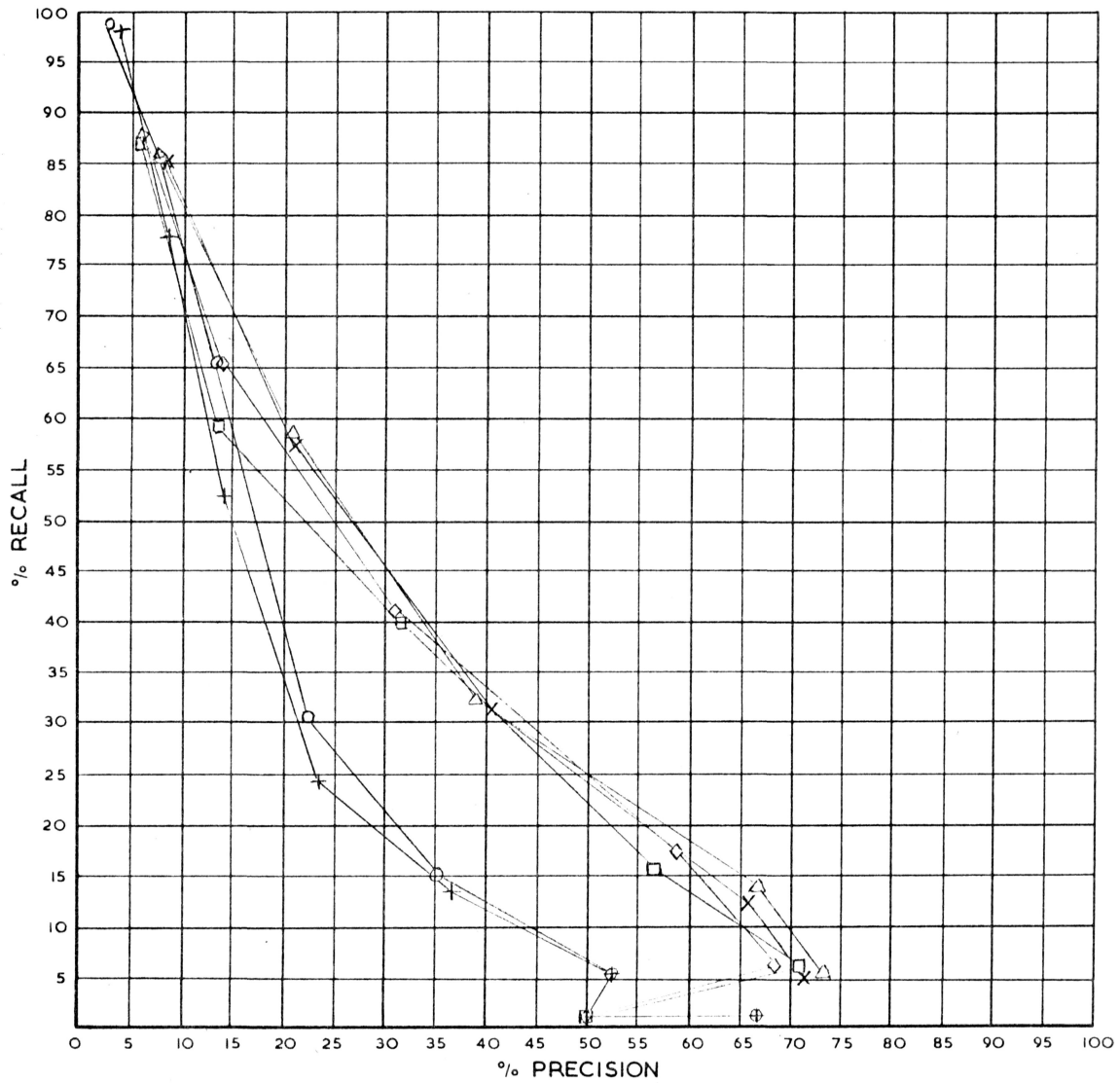


FIGURE 4.816P INDEX LANGUAGES III.1.a (X), III.2.a (A),
III.3.a (P), III.4.a (phi), III.5.a (+),
III.6.a (O). SEARCH E. 200 DOCUMENTS
(Figures 4.810T - 4.815T)

FIGURE 4.820T

Index Language III.1.a (Controlled terms. Basic terms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule E
 Document Relevance 1 - 4
 Number of Documents in Collection 350 (Subset 2)
 Number of Questions 77 (Subset 7)
 Number of Relevant Documents 454
 Generality Number 16.8

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 361 | 6,320 | 79.5% | 5.4% | 23.849% | 77 | 77 | 77 |
| 2 | 231 | 1,413 | 50.9% | 14.1% | 5.333% | 73 | 77 | 77 |
| 3 | 109 | 289 | 24.0% | 27.4% | 1.091% | 54 | 73 | 73 |
| 4 | 40 | 31 | 8.8% | 56.3% | 0.117% | 23 | 63 | 63 |
| 5 | 13 | 6 | 2.9% | 68.4% | 0.023% | 9 | 47 | 47 |
| 6 | 0 | 2 | 0.0% | 0.0% | 0.008% | 1 | 28 | 28 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 17 | 17 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 11 | 11 |
| 9 | 0 | 0 | | | | 0 | 6 | 6 |
| 10 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.821T

Index Language III.2.a (Controlled terms. Narrower terms. Coordination)
 Exhaustivity of Indexing 3
 Search Rule E
 Document Relevance 1 - 4
 Number of Documents in Collection 350 (Subset 2)
 Number of Questions 77 (Subset 7)
 Number of Relevant Documents 454
 Generality Number 16.8

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 377 | 6,762 | 83.0% | 5.3% | 25.516% | 77 | 77 | 77 |
| 2 | 245 | 1,554 | 54.0% | 13.6% | 5.865% | 73 | 77 | 77 |
| 3 | 113 | 335 | 24.9% | 25.2% | 1.264% | 54 | 73 | 73 |
| 4 | 46 | 35 | 10.1% | 56.8% | 0.132% | 25 | 63 | 63 |
| 5 | 13 | 6 | 2.9% | 68.4% | 0.023% | 9 | 47 | 47 |
| 6 | 0 | 2 | 0.0% | 0.0% | 0.008% | 1 | 28 | 28 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 17 | 17 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 11 | 11 |
| 9 | 0 | 0 | | | | 0 | 6 | 6 |
| 10 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.822T

Index Language III.3.a (Controlled terms. Broader terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 350 (Subset 2)

Number of Questions 77 (Subset 7)

Number of Relevant Documents 454

Generality Number 16.8

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 376 | 9,202 | 82.8 | 3.9% | 37.724% | 77 | 77 | 77 |
| 2 | 259 | 2,562 | 57.0% | 9.2% | 9.669% | 73 | 77 | 77 |
| 3 | 138 | 496 | 30.4% | 21.8% | 1.872% | 60 | 73 | 73 |
| 4 | 53 | 65 | 11.7% | 44.9% | 0.245% | 31 | 63 | 63 |
| 5 | 18 | 9 | 4.0% | 66.7% | 0.034% | 14 | 47 | 47 |
| 6 | 2 | 3 | 0.4% | 40.0% | 0.011% | 3 | 28 | 28 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 17 | 17 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 11 | 11 |
| 9 | 0 | 0 | | | | 0 | 6 | 6 |
| 10 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.823T

Index Language III.4.a (Controlled terms. Narrower and Broader terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 350 (Subset 2)

Number of Questions 77 (Subset 7)

Number of Relevant Documents 454

Generality Number 16.8

| Coordination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|--------------------|---------------------|----------|-----------------------|--------------------------|------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 394 | 9,369 | 86.8% | 4.0% | 35.354% | 77 | 77 | 77 |
| 2 | 276 | 2,757 | 60.8% | 9.1% | 10.405% | 73 | 77 | 77 |
| 3 | 213 | 599 | 46.9% | 26.2% | 2.261% | 60 | 73 | 73 |
| 4 | 58 | 73 | 12.8% | 44.3% | 0.276% | 33 | 63 | 63 |
| 5 | 18 | 11 | 4.0% | 62.1% | 0.042% | 14 | 47 | 47 |
| 6 | 2 | 3 | 0.4% | 40.0% | 0.011% | 3 | 28 | 28 |
| 7 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 17 | 17 |
| 8 | 0 | 1 | 0.0% | 0.0% | 0.004% | 1 | 11 | 11 |
| 9 | 0 | 0 | | | | 0 | 6 | 6 |
| 10 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.824T

Index Language III.5.a (Controlled terms. Related terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 350 (Subset 2)

Number of Questions 77 (Subset 7)

Number of Relevant Documents 454

Generality Number 16.8

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 430 | 13,965 | 94.7% | 3.0% | 52.698% | 77 | 77 | 77 |
| 2 | 350 | 5,896 | 77.1% | 5.6% | 22.252% | 77 | 77 | 77 |
| 3 | 211 | 1,943 | 46.5% | 9.8% | 7.333% | 72 | 73 | 73 |
| 4 | 103 | 441 | 22.7% | 18.9% | 1.664% | 53 | 63 | 63 |
| 5 | 40 | 132 | 8.8% | 23.3% | 0.498% | 31 | 47 | 47 |
| 6 | 12 | 27 | 2.6% | 30.8% | 0.102% | 9 | 28 | 28 |
| 7 | 2 | 5 | 0.4% | 28.6% | 0.019% | 4 | 17 | 17 |
| 8 | 2 | 3 | 0.4% | 40.0% | 0.011% | 2 | 11 | 11 |
| 9 | 2 | 1 | 0.4% | 66.7% | 0.004% | 2 | 6 | 6 |
| 10 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.825T

Index Language III.6.a (Controlled terms. Narrower, broader and related terms. Coordination)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 350 (Subset 2)

Number of Questions 77 (Subset 7)

Number of Relevant Documents 454

Generality Number 16.8

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 440 | 15,172 | 96.9% | 2.8% | 57.252% | 77 | 77 | 77 |
| 2 | 370 | 7,383 | 81.5% | 4.8% | 27.865% | 77 | 77 | 77 |
| 3 | 252 | 2,664 | 55.5% | 8.6% | 10.054% | 72 | 73 | 73 |
| 4 | 126 | 630 | 27.8% | 16.7% | 2.378% | 59 | 63 | 63 |
| 5 | 45 | 167 | 9.9% | 20.7% | 0.630% | 35 | 47 | 47 |
| 6 | 13 | 36 | 2.9% | 26.5% | 0.136% | 11 | 28 | 28 |
| 7 | 2 | 7 | 0.4% | 22.2% | 0.026% | 5 | 17 | 17 |
| 8 | 2 | 3 | 0.4% | 40.0% | 0.011% | 2 | 11 | 11 |
| 9 | 2 | 1 | 0.4% | 66.7% | 0.004% | 2 | 6 | 6 |
| 10 | 0 | 0 | | | | 0 | 1 | 1 |

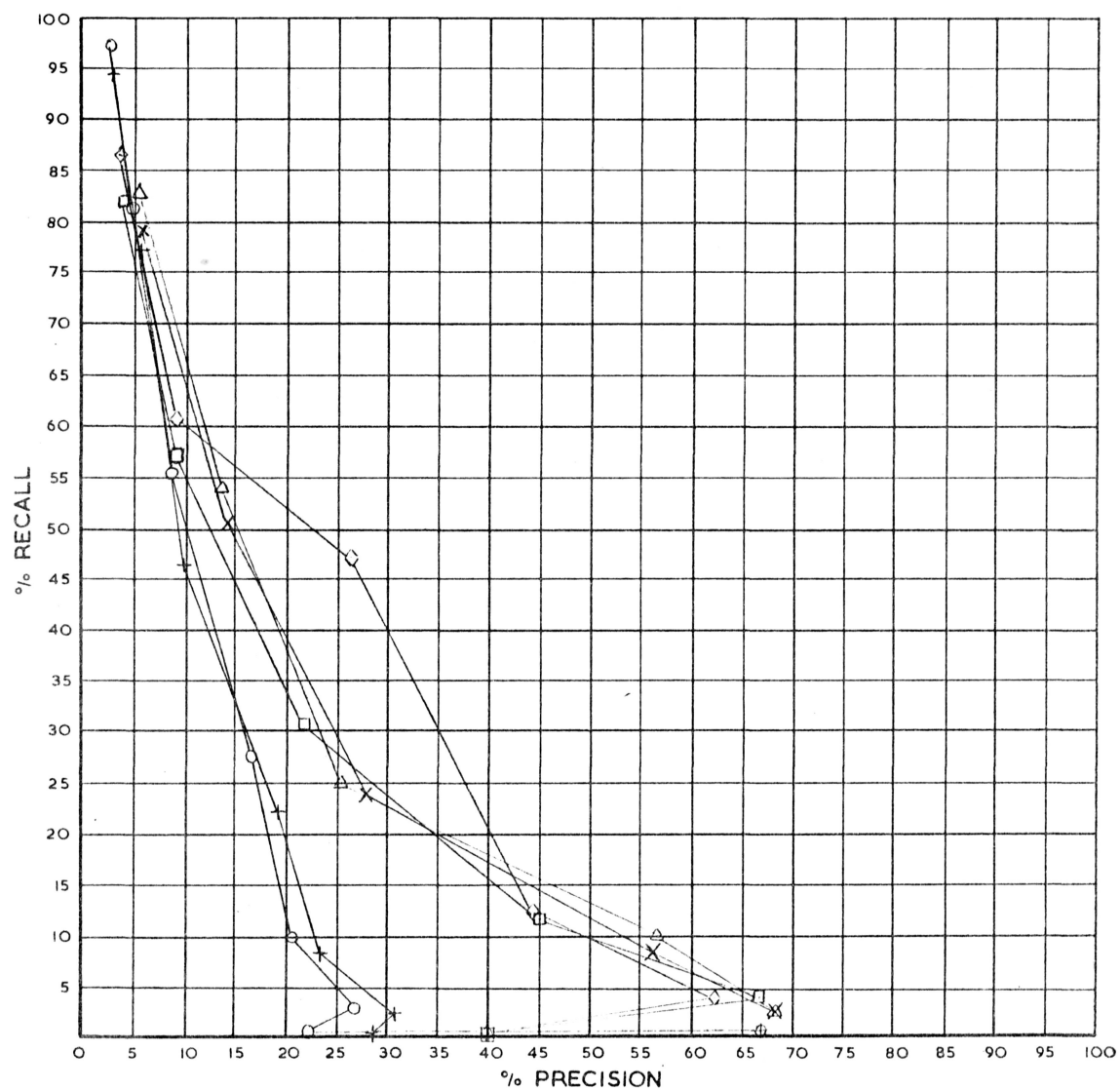


FIGURE 4.826P INDEX LANGUAGES III.1.a (X), III.2.a (Δ), III.3.a (◻), III.4.a (◊), III.5.a (+), III.6.a (O), SEARCH E. 350 DOCUMENTS (Figures 4.820T - 4.825T)

FIGURE 4.830T

Index Language III.1.e (Controlled terms. Basic terms. Coordination, weighting)

Exhaustivity of Indexing 3

Search Rule E

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 155 | (-) | 78.3% | (-) | (-) | 42 | 0 | 42 |
| 2 | 90 | 215 | 45.5% | 29.5% | 2.621% | 36 | 42 | 42 |
| 3 | 39 | 27 | 19.7% | 59.1% | 0.329% | 23 | 41 | 41 |
| 4 | 11 | 4 | 5.6% | 73.3% | 0.049% | 7 | 34 | 34 |
| 5 | 6 | 0 | 3.0% | 100.0% | 0.000% | 3 | 24 | 24 |
| 6 | 0 | 0 | | | | 0 | 13 | 13 |
| 7 | 0 | 0 | | | | 0 | 8 | 8 |
| 8 | 0 | 0 | | | | 0 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

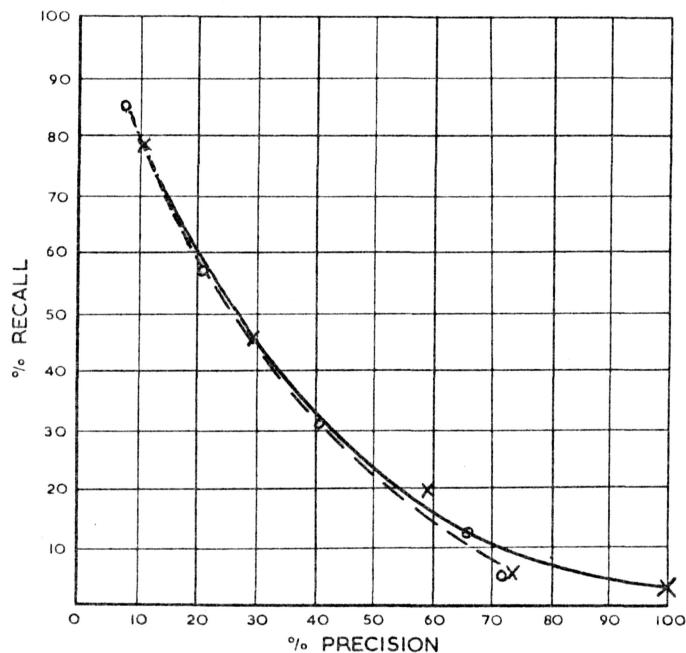


FIGURE 4.830P INDEX LANGUAGE III.1.e SEARCH E
(Index Language III.1.a Broken Line)

FIGURE 4.831T

Index Language III.6.e (Controlled terms. Narrower, broader and related terms. Coordination weighting)
 Exhaustivity of Indexing 3
 Search Rule E
 Document Relevance 1 - 4
 Number of Documents in Collection 200 (Subset 1)
 Number of Questions 42 (Subset 2)
 Number of Relevant Documents 198
 Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 192 | (-) | 97.0% | (-) | (-) | 42 | 0 | 42 |
| 2 | 137 | 1,134 | 69.2% | 10.8% | 13.826% | 41 | 42 | 42 |
| 3 | 88 | 345 | 44.4% | 20.3% | 4.206% | 38 | 41 | 41 |
| 4 | 39 | 56 | 19.7% | 41.1% | 0.683% | 24 | 34 | 34 |
| 5 | 20 | 15 | 10.1% | 57.1% | 0.183% | 11 | 24 | 24 |
| 6 | 8 | 2 | 4.0% | 80.0% | 0.024% | 3 | 13 | 13 |
| 7 | 0 | 0 | | | | 0 | 8 | 8 |
| 8 | 0 | 0 | | | | 0 | 4 | 4 |
| 9 | 0 | 0 | | | | 0 | 3 | 3 |

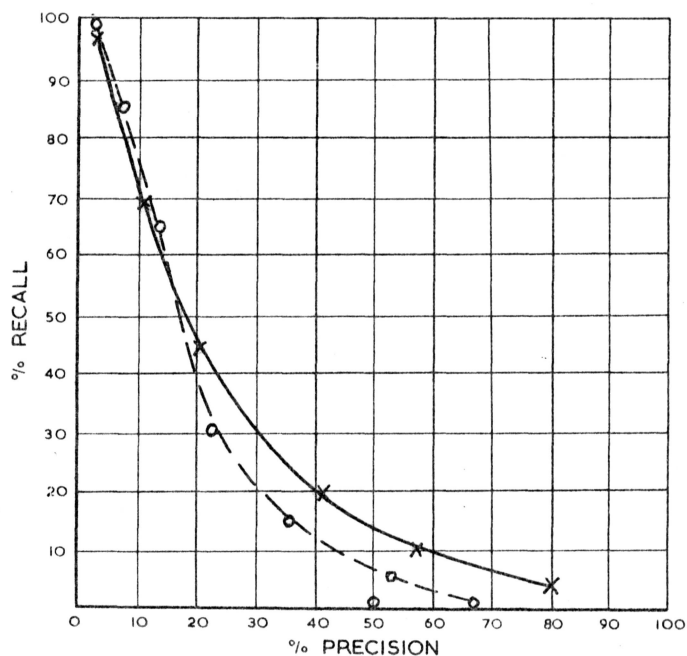


FIGURE 4.831P INDEX LANGUAGE III.6.e SEARCH E
 (Index Language III.6.a Broken line)

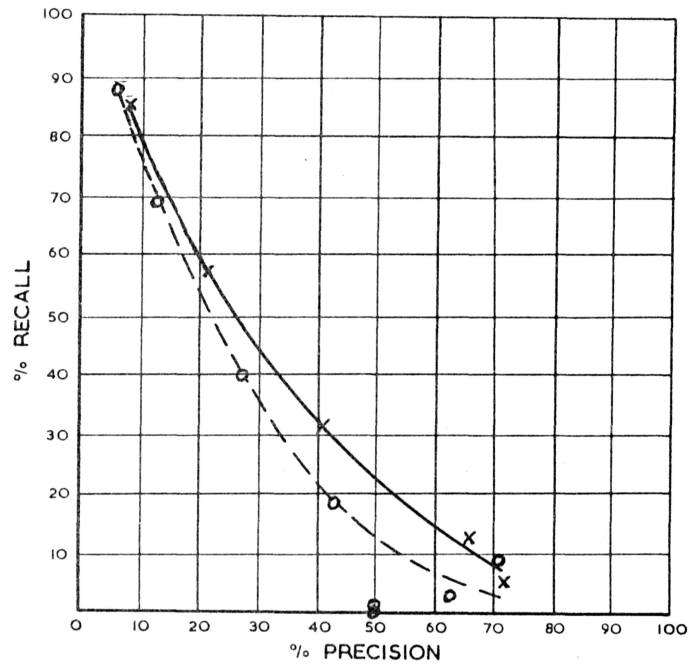


FIGURE 4.840P INDEX LANGUAGE III.1.a. SEARCH A (BROKEN LINE) AND SEARCH E. 200 DOCUMENTS (FIGURES 4.800T and 4.810T)

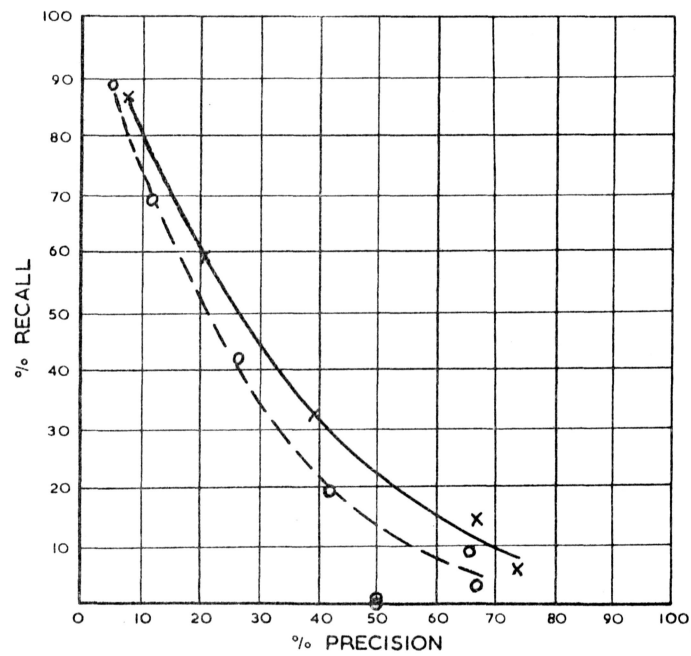


FIGURE 4.841P INDEX LANGUAGE III.2.a. SEARCH A (BROKEN LINE) AND SEARCH E. 200 DOCUMENTS (FIGURES 4.801T and 4.811T)

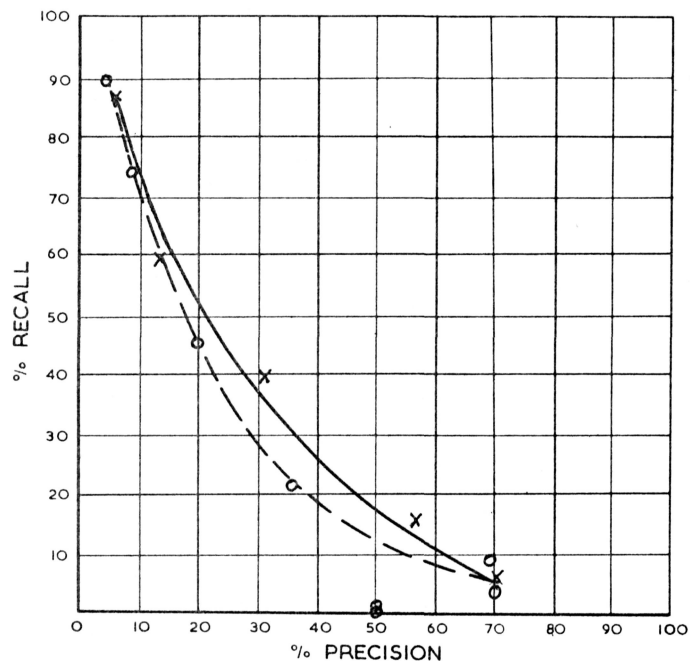


FIGURE 4.842P INDEX LANGUAGE III.3.a. SEARCH A (BROKEN LINE) AND SEARCH E. 200 DOCUMENTS (FIGURES 4.802T and 4.812T)

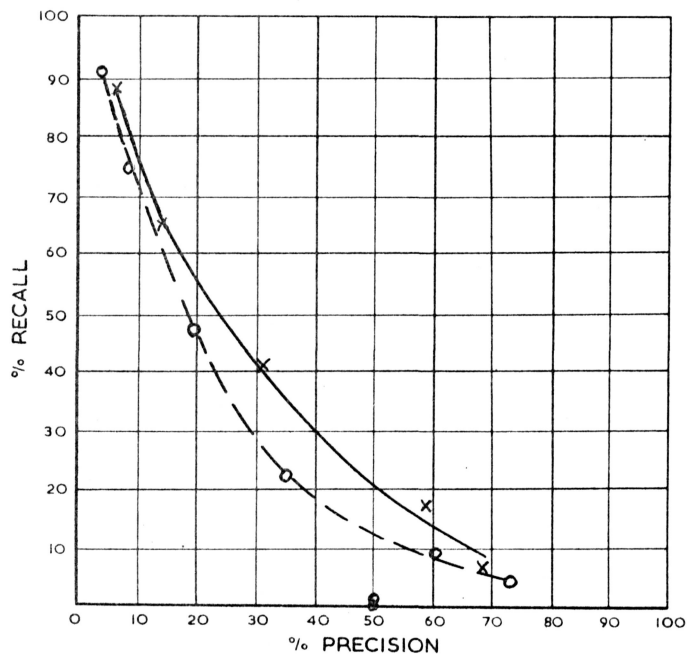


FIGURE 4.843P INDEX LANGUAGE III.4.a. SEARCH A (BROKEN LINE) AND SEARCH E. 200 DOCUMENTS (FIGURES 4.803T and 4.813T)

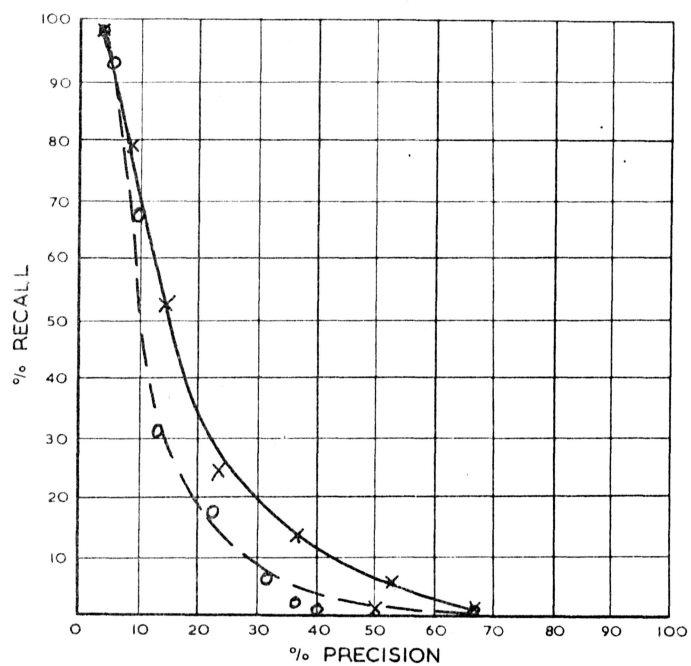


FIGURE 4.844P INDEX LANGUAGE III.5.a. SEARCH A (BROKEN LINE) AND SEARCH E. 200 DOCUMENTS (FIGURES 4.804T and 4.814T)

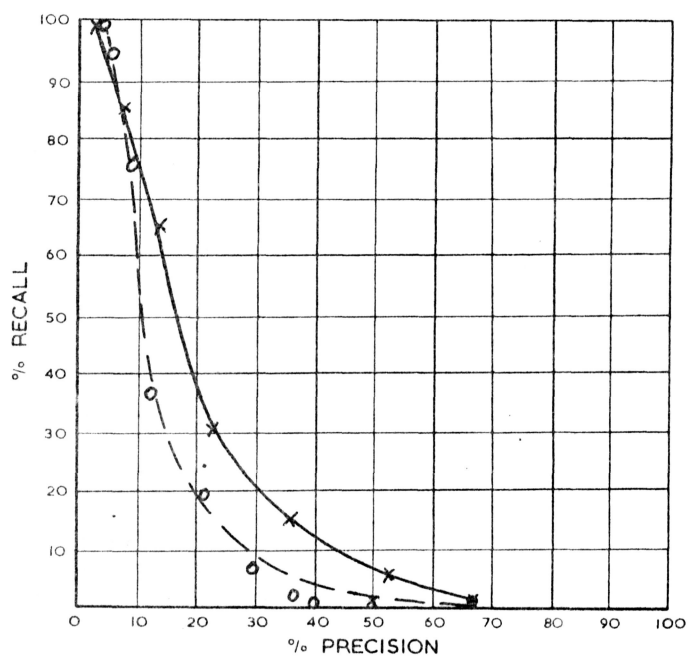


FIGURE 4.845P INDEX LANGUAGE III.6.a. SEARCH A (BROKEN LINE) AND SEARCH E. 200 DOCUMENTS (FIGURES 4.805T and 4.815T)

FIGURE 4.850T

Index Language III.5.a

Exhaustivity of indexing 3

Search Rule F

Document Relevance 1-4

Number of Documents in Collection 200 (subset 1)

Number of Questions 42 (subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Starting Terms | Documents Retrieved | | Recall Ratio | Precision Ratio | Fallout Ratio |
|----------------------------|-------------------|------------------------|----------|-----------------|--------------------|------------------|
| | | Rel. | Non-rel. | | | |
| 2 | 0 | 156 | 1,660 | 78.0% | 8.6% | 20.329% |
| | 1 | 146 | 1,133 | 73.7% | 11.4% | 13.817% |
| | 2 | 114 | 428 | 57.6% | 21.0% | 5.218% |
| 3 | 0 | 104 | 628 | 52.5% | 14.2% | 7.657% |
| | 1 | 100 | 534 | 50.5% | 15.8% | 6.512% |
| | 2 | 89 | 339 | 44.9% | 21.1% | 4.134% |
| 4 | 3 | 62 | 92 | 31.3% | 40.3% | 1.122% |
| | 0 | 48 | 155 | 24.2% | 23.6% | 1.890% |
| | 1 | 47 | 146 | 23.7% | 24.3% | 1.780% |
| 5 | 2 | 47 | 100 | 23.7% | 32.0% | 1.220% |
| | 3 | 44 | 55 | 22.2% | 44.4% | 0.671% |
| | 4 | 25 | 13 | 12.6% | 65.8% | 0.158% |
| 6 | 0 | 27 | 47 | 13.6% | 36.5% | 0.573% |
| | 1 | 27 | 46 | 13.6% | 37.0% | 0.561% |
| | 2 | 27 | 37 | 13.6% | 42.2% | 0.451% |
| 6 | 3 | 25 | 29 | 12.6% | 46.3% | 0.354% |
| | 4 | 20 | 15 | 10.1% | 57.1% | 0.183% |
| | 5 | 10 | 4 | 5.1% | 71.4% | 0.049% |
| 6 | 0 | 11 | 10 | 5.6% | 52.4% | 0.122% |
| | 1 | 10 | 10 | 5.1% | 50.0% | 0.122% |
| | 2 | 10 | 10 | 5.1% | 50.0% | 0.122% |
| 6 | 3 | 10 | 6 | 5.1% | 62.5% | 0.073% |
| | 4 | 10 | 4 | 5.1% | 71.4% | 0.049% |
| | 5 | 6 | 2 | 3.1% | 75.0% | 0.024% |
| 6 | 6 | 0 | 2 | 0.0% | 0.0% | 0.024% |

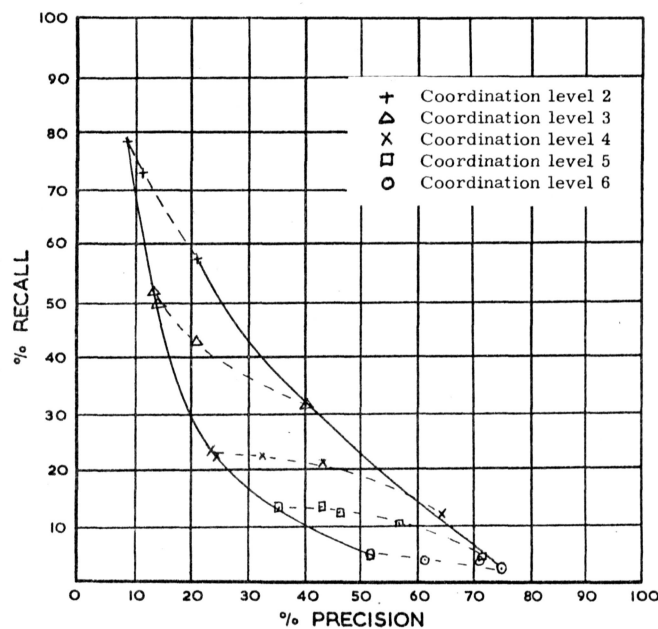


FIGURE 4.851T

Index Language III.6.a

Exhaustivity of indexing 3

Search Rule F

Document Relevance 1-4

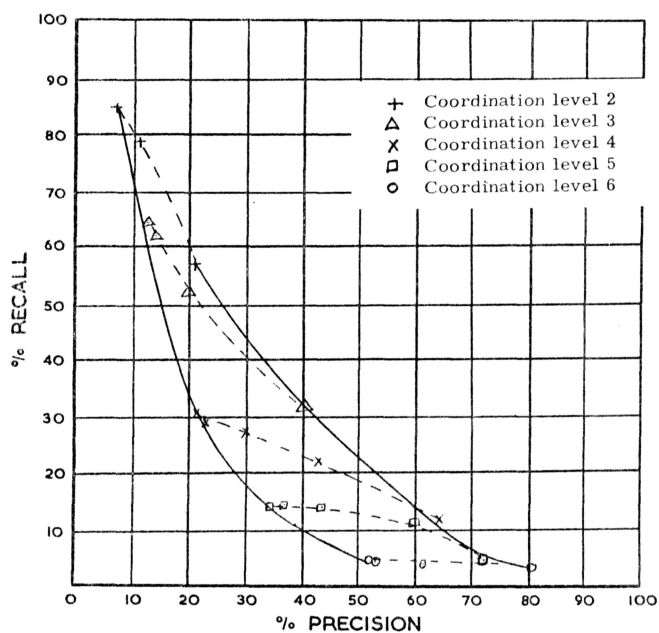
Number of Documents in Collection 200 (subset 1)

Number of Questions 42

Number of Relevant Documents 198

Generality Number 23.6

| Coordination Level | Starting Terms | Documents Retrieved | | Recall Ratio | Precision Ratio | Fallout Ratio |
|--------------------|----------------|---------------------|----------|--------------|-----------------|---------------|
| | | Rel. | Non-rel. | | | |
| 2 | 0 | 169 | 2,092 | 85.4% | 7.5% | 25.506% |
| | 1 | 158 | 1,321 | 79.8% | 10.7% | 16.110% |
| | 2 | 114 | 428 | 57.6% | 21.0% | 5.218% |
| 3 | 0 | 129 | 848 | 65.2% | 13.2% | 10.339% |
| | 1 | 122 | 685 | 61.6% | 15.1% | 8.354% |
| | 2 | 105 | 398 | 53.0% | 20.9% | 4.854% |
| 4 | 3 | 62 | 92 | 31.3% | 40.0% | 1.122% |
| | 0 | 60 | 210 | 30.3% | 22.2% | 2.560% |
| | 1 | 59 | 192 | 29.8% | 23.5% | 2.341% |
| | 2 | 57 | 128 | 28.8% | 30.8% | 1.561% |
| 5 | 3 | 47 | 64 | 23.7% | 42.3% | 0.780% |
| | 4 | 25 | 13 | 12.6% | 65.8% | 0.158% |
| | 0 | 30 | 55 | 15.2% | 35.3% | 0.671% |
| | 1 | 30 | 51 | 15.2% | 37.0% | 0.622% |
| | 2 | 30 | 41 | 15.2% | 42.3% | 0.500% |
| 6 | 3 | 30 | 30 | 15.2% | 50.0% | 0.366% |
| | 4 | 23 | 15 | 11.6% | 60.5% | 0.183% |
| | 5 | 10 | 4 | 5.1% | 71.4% | 0.049% |
| | 0 | 11 | 10 | 5.6% | 52.4% | 0.122% |
| | 1 | 10 | 10 | 5.1% | 50.0% | 0.122% |
| | 2 | 10 | 9 | 5.1% | 52.6% | 0.110% |
| | 3 | 10 | 6 | 5.1% | 62.5% | 0.073% |
| | 4 | 10 | 4 | 5.1% | 71.4% | 0.049% |
| | 5 | 8 | 2 | 4.0% | 80.0% | 0.024% |
| | 6 | 0 | 2 | 0.0% | 0.0% | 0.012% |



Section 8 Abstracts and Titles

This section does not introduce any new index language device, since it deals with searches carried out on the abstracts and titles of the 200 document collection. As such it represents a variation in exhaustivity of indexing, and also, in some cases, different concept indexing to that done by the project staff. The searches are done first in single term natural language and then with word forms confounded.

Figures 4.900T and 4.901T present the results of searches on the titles only, and Figures 4.902T and 4.903T present the results obtained with titles and abstracts. Figure 4.904P is a plot of the four tables.

In order to make a comparison from the viewpoint of exhaustivity, Figures 4.910T and 4.911T present the results of Index Language I.3.a (word forms) at exhaustivity levels of 2 and 1, on the 200 document collection. Plot 4.912P now gives the performance curves for the three levels of exhaustivity with the project indexing, for the search on titles alone and for the search on titles and abstracts.

We were able to do this series of tests by having a print out of a concordance of the titles and abstracts of the 200 documents. This was prepared with the SMART programme, and we are indebted to Prof. Salton for making it available

LIST OF FIGURES

| | Index Language | Exhaustivity | No. of Questions | Question Subset | Document Collection | Plots |
|--------|-------------------|--------------|---------------------|--------------------|------------------------|--------|
| 4.900T | IV.1.a | | 42 | 2 | 200 | |
| 4.901T | IV.2.a | | 42 | 2 | 200 | |
| 4.902T | IV.3.a | | 42 | 2 | 200 | |
| 4.903T | IV.4.a | | 42 | 2 | 200 | |
| 4.904P | IV.1.a | | 42 | 2 | 200 | 4.900T |
| | IV.2.a | | | | | 4.901T |
| | IV.3.a | | | | | 4.902T |
| | IV.4.a | | | | | 4.903T |
| 4.910T | I.3.a | 2 | 42 | 2 | 200 | |
| 4.911T | I.3.a | 1 | 42 | 2 | 200 | |
| 4.912P | IV.2.a | | 42 | 2 | 200 | 4.901T |
| | IV.4.a | | | | | 4.903T |
| | I.3.a | 3 | | | | 4.201T |
| | I.3.a | 2 | | | | 4.910T |
| | I.3.a | 1 | | | | 4.911T |

FIGURE 4.900T

Index Language IV.1.a (Single terms. Natural language. Coordination)

Exhaustivity of Indexing Title only

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 155 | 1,902 | 78.3% | 7.5% | 23.189% | 42 | 0 | 42 |
| 2 | 108 | 463 | 54.5% | 18.9% | 5.645% | 40 | 42 | 42 |
| 3 | 58 | 76 | 29.3% | 43.3% | 0.927% | 32 | 42 | 42 |
| 4 | 26 | 15 | 13.1% | 63.4% | 0.183% | 16 | 41 | 41 |
| 5 | 15 | 3 | 7.6% | 83.3% | 0.037% | 9 | 39 | 39 |
| 6 | 6 | 0 | 3.0% | 100.0% | 0.000% | 3 | 33 | 33 |
| 7 | 3 | 0 | 1.5% | 100.0% | 0.000% | 1 | 27 | 27 |
| 8 | 0 | 0 | | | | 0 | 18 | 18 |
| 9 | 0 | 0 | | | | 0 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.901T

Index Language IV.2.a (Single terms. Word forms. Coordination)

Exhaustivity of Indexing Title only

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 159 | 2,017 | 80.3% | 7.3% | 24.592% | 42 | 0 | 42 |
| 2 | 111 | 512 | 56.1% | 17.8% | 6.242% | 40 | 42 | 42 |
| 3 | 65 | 94 | 32.8% | 40.9% | 1.146% | 34 | 42 | 42 |
| 4 | 29 | 18 | 14.6% | 61.7% | 0.219% | 16 | 41 | 41 |
| 5 | 17 | 4 | 8.6% | 81.0% | 0.049% | 12 | 39 | 39 |
| 6 | 7 | 0 | 3.5% | 100.0% | 0.000% | 3 | 33 | 33 |
| 7 | 3 | 0 | 1.5% | 100.0% | 0.000% | 1 | 27 | 27 |
| 8 | 0 | 0 | | | | 0 | 18 | 18 |
| 9 | 0 | 0 | | | | 0 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.902T

Index Language IV.3.a (Single terms. Natural language. Coordination)

Exhaustivity of Indexing Title and Abstract

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 185 | 6,261 | 93.4% | 2.9% | 76.353% | 42 | 0 | 42 |
| 2 | 161 | 2,158 | 81.3% | 6.9% | 26.311% | 42 | 42 | 42 |
| 3 | 120 | 854 | 60.6% | 12.3% | 10.412% | 40 | 42 | 42 |
| 4 | 75 | 288 | 37.9% | 20.7% | 3.511% | 32 | 41 | 41 |
| 5 | 44 | 97 | 22.2% | 31.2% | 1.183% | 21 | 39 | 39 |
| 6 | 27 | 25 | 13.6% | 51.9% | 0.305% | 13 | 33 | 33 |
| 7 | 12 | 2 | 6.1% | 85.7% | 0.024% | 6 | 27 | 27 |
| 8 | 3 | 0 | 1.5% | 100.0% | 0.000% | 2 | 18 | 18 |
| 9 | 2 | 0 | 1.0% | 100.0% | 0.000% | 1 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.903T

Index Language IV.4.a (Single terms. Word forms. Coordination)

Exhaustivity of Indexing Title and Abstract

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 187 | 6,770 | 94.4% | 2.7% | 82.540% | 42 | 0 | 42 |
| 2 | 166 | 2,654 | 83.8% | 5.9% | 32.358% | 42 | 42 | 42 |
| 3 | 130 | 1,150 | 65.7% | 10.2% | 14.021% | 41 | 42 | 42 |
| 4 | 92 | 432 | 46.5% | 17.6% | 5.267% | 35 | 41 | 41 |
| 5 | 51 | 158 | 25.8% | 24.4% | 1.926% | 28 | 39 | 39 |
| 6 | 32 | 53 | 16.2% | 37.6% | 0.646% | 19 | 33 | 33 |
| 7 | 20 | 13 | 10.1% | 60.6% | 0.158% | 10 | 27 | 27 |
| 8 | 9 | 4 | 4.5% | 69.2% | 0.049% | 7 | 18 | 18 |
| 9 | 3 | 1 | 1.5% | 75.0% | 0.012% | 4 | 11 | 11 |
| 10 | 1 | 0 | 0.5% | 100.0% | 0.000% | 1 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

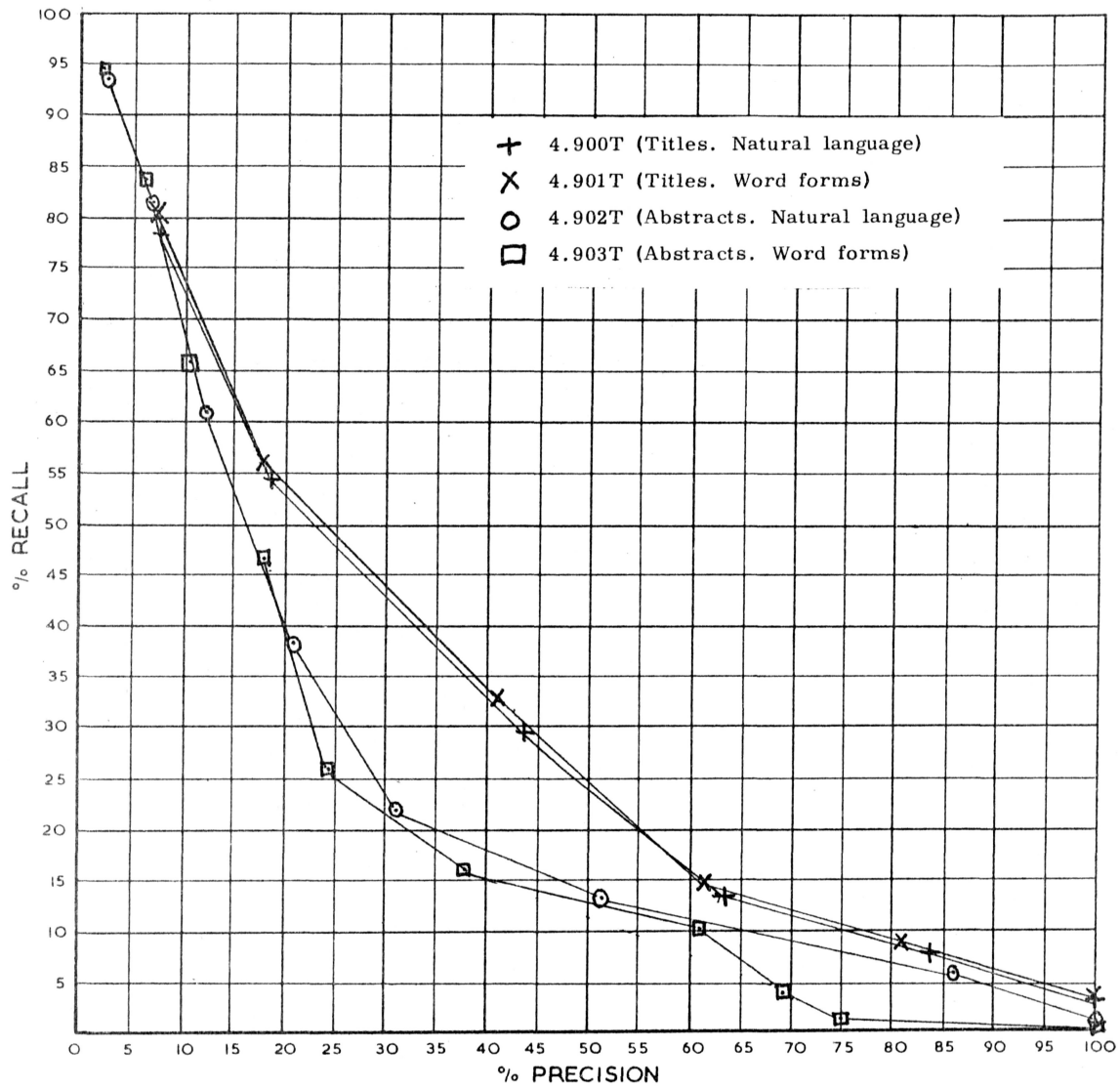


FIGURE 4.904P INDEX LANGUAGES IV.1.a, IV.2.a, IV.3.a and IV.4.a. (Figures 4.900T - 4.903T)

FIGURE 4.910T

Index Language I.3.a (Single terms. Word forms. Coordination)

Exhaustivity of Indexing 2

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 188 | 4,037 | 94.9% | 4.4% | 49.220% | 42 | 42 | 42 |
| 2 | 162 | 1,867 | 81.8% | 8.0% | 22.763% | 42 | 42 | 42 |
| 3 | 128 | 694 | 64.6% | 15.6% | 8.461% | 40 | 42 | 42 |
| 4 | 87 | 208 | 43.9% | 29.5% | 2.536% | 32 | 41 | 41 |
| 5 | 47 | 66 | 23.7% | 41.6% | 0.805% | 22 | 39 | 39 |
| 6 | 30 | 17 | 15.2% | 63.8% | 0.207% | 13 | 33 | 33 |
| 7 | 11 | 2 | 5.6% | 84.6% | 0.024% | 6 | 27 | 27 |
| 8 | 3 | 0 | 1.5% | 100.0% | 0.000% | 2 | 18 | 18 |
| 9 | 0 | 0 | | | | 0 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

FIGURE 4.911T

Index Language I.3.a (Single terms. Word forms. Coordination)

Exhaustivity of Indexing 1

Search Rule A

Document Relevance 1 - 4

Number of Documents in Collection 200 (Subset 1)

Number of Questions 42 (Subset 2)

Number of Relevant Documents 198

Generality Number 23.6

| Coord- ination Level | Documents Retrieved | | Recall Ratio a/a+c | Precision Ratio a/a+b | Fallout Ratio b/b+d | x | y | z |
|----------------------------|------------------------|----------|--------------------------|-----------------------------|---------------------------|----|----|----|
| | Rel. | Non-rel. | | | | | | |
| 1 | 177 | 2,626 | 89.4% | 6.3% | 32.017% | 42 | 42 | 42 |
| 2 | 138 | 1,101 | 69.7% | 11.1% | 13.424% | 42 | 42 | 42 |
| 3 | 103 | 308 | 52.0% | 25.1% | 3.755% | 36 | 42 | 42 |
| 4 | 60 | 80 | 30.3% | 42.9% | 0.975% | 28 | 41 | 41 |
| 5 | 30 | 23 | 15.2% | 56.6% | 0.280% | 14 | 39 | 39 |
| 6 | 17 | 3 | 8.6% | 85.0% | 0.037% | 9 | 33 | 33 |
| 7 | 7 | 0 | 3.5% | 100.0% | 0.000% | 3 | 27 | 27 |
| 8 | 1 | 0 | 0.5% | 100.0% | 0.000% | 1 | 18 | 18 |
| 9 | 0 | 0 | | | | 0 | 11 | 11 |
| 10 | 0 | 0 | | | | 0 | 7 | 7 |
| 11 | 0 | 0 | | | | 0 | 3 | 3 |
| 12 | 0 | 0 | | | | 0 | 1 | 1 |

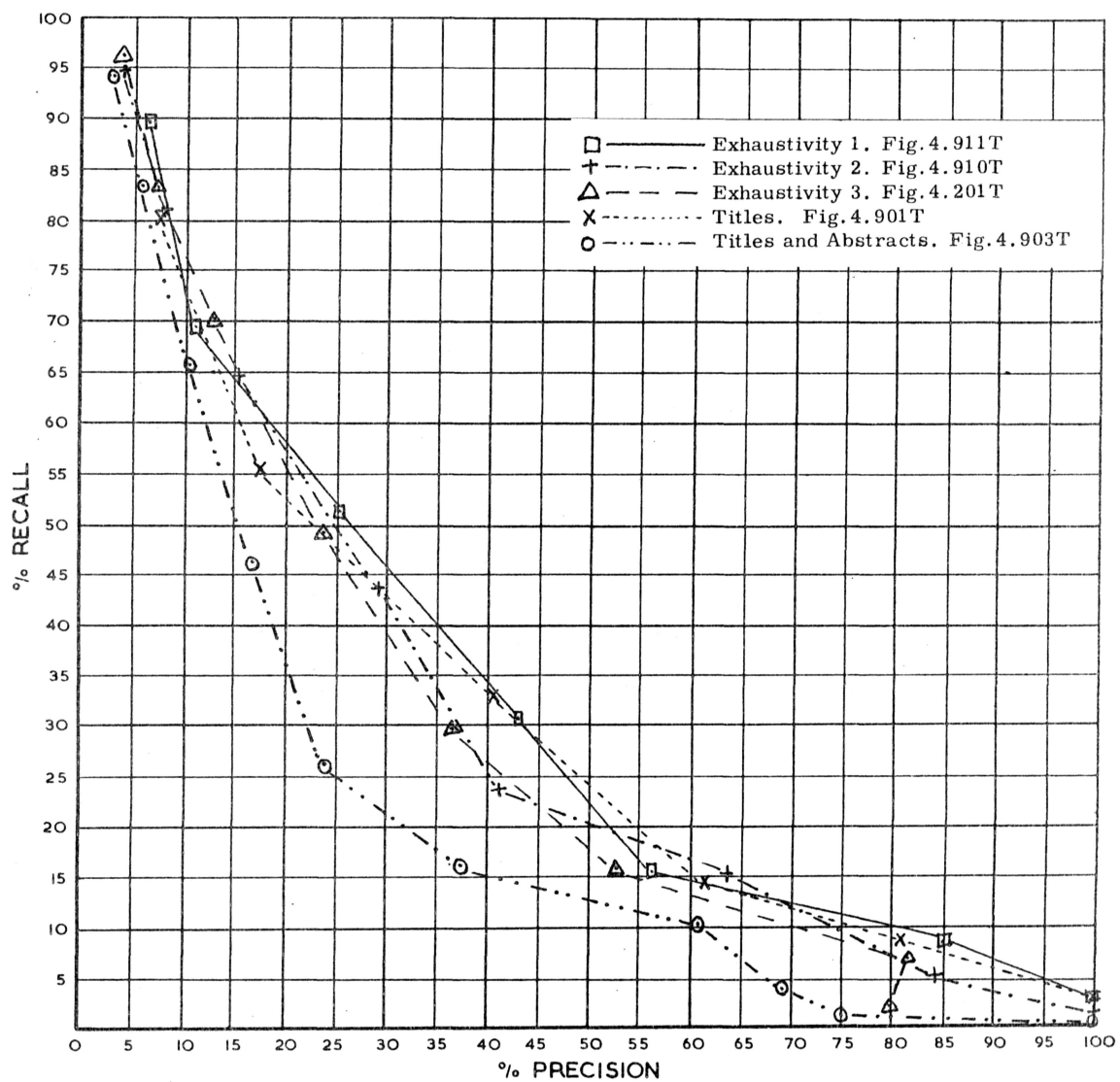


FIGURE 4.912P PLOT OF RESULTS FOR THREE LEVELS OF EXHAUSTIVITY OF INDEXING, FOR TITLES AND FOR ABSTRACTS, ON 42 QUESTIONS, 200 DOCUMENT COLLECTION, INDEX LANGUAGE I.3.a