Correlative Indexes. VIII. Subject-Indexing *vs.* Word-Indexing

By CHARLES L. BERNIER and E. J. CRANE
Chemical Abstracts, The Ohio State University, Columbus, Ohio
Received October 16, 1961

"When I use a word, it means just what I choose it to mean, neither more nor less," Lewis Carroll, "Alice Through The Looking Glass."

Subject-indexing means indexing the subject reported by the author rather than the words he uses. The indexer, not the author, chooses precisely what subject-index words mean.

Searchers usually seek information not on symbols, but on what they signify. In arranging for a race at the fairground, horses are entered; their names are used, but these names win no races.

Although the author deals with subjects and discusses them in his own words, only a very few of these are needed in a subject index; sometimes none of his words is suitable. The author uses many words that introduce, define, explain, and give data about the subject or subjects. *Nearly all of these words are unsuitable for subject-index headings.* Effective subject-index entries usually have fewer than 1% of the number of words in a chemical research paper indexed, and the index words frequently are different from those of the paper.

Five different terms, for example, act as subject-index headings for the average abstract in *Chemical Abstracts*. The average abstract has about 133 words in it. Thus, about 4% of the total number of words (some repeaters) in abstracts (but not necessarily the same words) are used satisfactorily as subject-index headings that function, together with modifying phrases, as complete keys to the chemical information in the abstract. How is it possible for such a limited number of terms to be adequate in unlocking all of the new information from abstracts, and thus from the original documents that average 20-40 times longer than the abstracts? The answer lies, in part, in the word "key." The subject index is designed to be a key to the information, not to be the information itself. Index-heading terms usually represent more or less general areas of information in the document, but do not give specifically the information itself. The modifying phrases (modifications is the term used by the indexing staff of *Chemical Abstracts*), such as those printed in light-faced type in the subject indexes of *Chemical Abstracts*, specify fairly definite parts of the general areas designated by the headings and thus help the index user effectively to select references under headings; *i.e.*, modifications save the time that would be wasted in looking up undistinguished, irrelevant, page references.

The subject-index heading, modifying phrase, and page reference taken together constitute an index entry. The average entry in the subject index of *Chemical Abstracts* consists of about five words including the bold-faced heading, and the column reference counted as a word. Thus, it requires an average of about 25 words effectively to subject-index an abstract and corresponding original document. That so few words are needed in a good subject index is fortunate. In fields other than chemistry the number of words may differ, but probably not widely. On the average fewer than 1% of the words of a chemical document are required to subject-index it. Thus, more than 99% of the words that the author uses are useless and even undesirable in a subject index. They are useless to the index user because they do not guide him to the new information reported by the author; they are undesirable because they dilute the useful entries to make for confusion and time-wasting reading.

A book entitled "Chemists' Handbook" would be subject-indexed under "Chemistry" since it would carry data on substances of interest to chemists; it would be word-indexed under "Chemists." Another book entitled "Who's Who in Chemistry" would be subject-indexed under "Chemists", but word-indexed under "Chemistry." This simple example, offered to illustrate how many words used by the author can be inadequate in a subject index, has many more subtle counterparts. The experienced subject indexer thinks in terms of meanings instead of words and indexes accordingly.

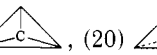
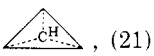
It has been suggested that titles be supplemented before publication so as to bring in suitable index words; however, this would be awkward, space-consuming, and difficult. It would call for the attention of index experts at the manuscript stage, which would delay publication and expand the total indexing effort. Furthermore, good, thorough indexes are based on the full information of abstracts and papers, not on their titles only.

New indexers often have difficulty in understanding the difference between subject- and word-indexing. This difficulty is somewhat similar to the difficulty that may be experienced by students of grammar in learning what is meant by the "subject" of a sentence.

"What is the author reporting?" is a question that is answered by naming one or more subjects. In every rational exposition, the author is writing about something. The "things" written about are the subjects. The detection of subjects and their translation into index language is the function of the subject indexer. Index language ("indexese" for fun) is a standardized form of writing that eliminates scattering of like information in the index. Scattering occurs in all kinds of poor indexes. It is most difficult to avoid in subject indexes. Index quality usually can be roughly measured by the amount of scattering, other mistakes being much less common. This measurement is tedious and requires considerable skill and knowledge. Standardization of index language reduces or eliminates scattering of like information. Standardization does not require a list of subject headings or subheadings. To a large extent it is possible to derive subject-index headings by rule rather than by list. This is done in subject-indexing *Chemical Abstracts*, for example. Use of rules saves the indexers' time needed to consult lists, and avoids the obsolescence and cost of lists. The specific rules useful in subject-indexing will be discussed later.

Scattering of like information in an index can occur because of synonymy, by the kind of additional word-indexing that leaves out implied subjects, by lack of under-

standing of the need for true subject-indexing, and by the genus-species relationship (see below).

Word-indexing causes scattering of like information mainly because of synonymy. Even as simple a chemical compound as methane could under varying circumstances be described as (1) methane, (2) marsh gas, (3) firedamp, (4) natural gas, (5) gas, (6) carbon hydride, (7) monocarbon hydride, (8) carbon tetrahydride, (9) monocarbon tetrahydride, (10) hydrogen carbide, (11) tetrahydrogen carbide, (12) hydrogen monocarbide, (13) tetrahydrogen monocarbide, (14) CH_4, (15) HCH_3, (16) H_3CH, (17) H_4C, (18) $H\overset{H}{\underset{H}{C}}H$, (19)  , (20)  , (21) C (notation), (22) C_1 (notation), and (23) 1H (notation).

Scattering caused by variations in forms of expression is controlled by the use of cross references. "See" cross references are made from the other commonly used designations to the heading chosen. An example is, "Vitamin B_1. See *Thiamine*." It has been suggested that a mechanized "thesaurus" (really a cross reference list) be used to standardize headings.[1] This is an interesting possibility for saving time of subject indexers. The number of terms in such a thesaurus for chemistry would be many more than the number of cross references presently used by *Chemical Abstracts* (about 60,000) because existing cross references are limited to the more frequently used synonymous designations.

The so-called genus-species relationship, *e.g.*, metal-gold, bacteria-Mycobacterium, sorption-absorption, and optical instrument-refractometer, are less satisfactorily handled by cross references because there are so many of these relationships. A more satisfactory approach seems to be a thesaurus[1] in book form. The permanence of these relationships makes this form especially attractive. Another approach is to build two indexes—a generic index[2] to handle the genus part; a specific index to handle the species part. These two kinds of indexes are both useful and not competitive, since each serves a different purpose.

True subject-indexing is an art and a science that requires much training and experience. Too often indexing, as of books, is done by inexperienced hands with resultant word-indexing and all of its shortcoming, including scattering and omissions.

Omissions can be exemplified by pointing out that a word index of a statement concerning "the properties of the elements of the 5th group in the periodic table" would lack entries for the specific elements, and another statement about "the saturated hydrocarbons from methane to decane" would have entries for methane and decane, but not for ethane, propane, butane, pentane, hexane, heptane, octane, and nonane.

EXAMPLE

The following sample abstract will further illustrate the difference between word-indexing and subject-indexing of new research results to maximum specificity.

Compounds present in aluminum-rich alloys of the aluminum-molybdenum system. J. W. H. Clare (Aluminum Lab., Ltd., Banbury, Engl.). *J. Inst. Metals*, **89**, 232–4 (1961).—$MoAl_7$ reacts peritectically with liquid at 706° to form $MoAl_{12}$. $MoAl_{12}$ is body-centered cubic with a cell side of 7.575 A.; and $MoAl_7$ is monoclinic with cell dimensions a = 5.12, b = 13.0, c = 13.5 A.; β = 95°.

The words and other symbols of the title and part of this abstract will be taken up in detail, at the risk of being absurd, as a means of stressing the differences between word- and subject-indexing.

Words, terms, formulas, and numbers that might be selected to prepare a word index (or for symbols and numbers, an index of equivalent type) of this abstract are: (1) compounds, (2) aluminum-rich alloys, (3) alloys, (4) aluminum, (5) molybdenum, (6) system, (7) $MoAl_7$, (8) reacts, (9) peritectically, (10) liquid, (11) 706°, (12) form, (13) $MoAl_{12}$, (14) body or body-centered cubic, (15) cubic, (16) cell, (17) side, (18) 7.575 A., (19) monoclinic, (20) dimensions, (21) a, (22) 5.12 A., (23) b, (24) 13.0 A., (25) c, (26) 13.5 A., (27) β, (28) 95°.

Context can be associated in the word index with these terms or numbers to provide greater selectivity, as in a concordance. For example, "compounds" can be associated with "present in aluminum-rich alloys" to give a word-index entry. Alternatively especially coined modifying phrases can be associated with each of the above 28 terms to give entries in a word index.

Instead of selecting terms or words used by the author to give a word index, suppose that we select the most specific terms representing new research results reported by the author to give subject-index entries. J. W. H. Clare was, in brief, studying the existence and crystal structure of $MoAl_7$ and $MoAl_{12}$ in the Mo-Al system and in Al-rich alloys.

A detailed examination of each term in the title and abstract of the paper by Clare can show whether or not it is suitable for use in a subject index and can give insight into the rationale of subject-indexing.

Clare was not studying "compounds" in general (*e.g.*, organic, inorganic, complex, or even all intermetallic); he did not report that his results applied to all types of compounds. Thus, the term "compounds" is too general for a specific subject index. Should the subject indexer have chosen "Compounds" as a subject-index entry or part of one, he would have been pushing the author of the paper far beyond the limits of the reported experimental evidence and conclusions. Thus, the index entry would have been incorrect. To be more specific, Clare did not study compounds in general, but only two compounds, namely, $MoAl_7$ and $MoAl_{12}$; thus his experimental results do not apply to all compounds and, more important, he did not claim that his results applied to all compounds. Had the author reported that his results did apply to all compounds, or probably applied to all compounds, or might apply to all compounds, then "Compounds" would be needed as a heading. Had he reported that his results applied to a class of compounds, *e.g.*, intermetallic compounds, then it would be necessary to index that class of compounds in a general manner. The two required subject-index entries related to the word "compounds" are, "**Aluminum compounds**, with molybdenum, $MoAl_7$ and $MoAl_{12}$", and "**Molybdenum compounds**, with aluminum, $MoAl_7$ and $MoAl_{12}$." Some variation of these entries could have been used, as "**Aluminum-molybdenum compounds**, $MoAl_7$ and $MoAl_{12}$," and "**Molybdenum-aluminum compounds**, $MoAl_7$ and $MoAl_{12}$." The first two entries have been found to lead to less scattering than the last two in a subject index to all of chemistry.

Taking the next word, we find that Clare was not studying "present" or any chemical research result

uniquely designated by this word that indicated "together with." Thus, neither "present" nor a synonym, nor other term designating this concept, can be used as heading in a specific subject index without incorrectly indexing subjects the author reported.

The preposition "in" is the next word. The author, in this case, was not studying "in" as a subject. He could have been. For example, he could study the meanings and use of the preposition "in" in contrast to other propositions that might be used with the word "systems." Such a study, although improbable in the fields of chemistry and metallurgy, would be more likely in the fields of English and linguistics. Had the author studied the use of "in" even incidentally, it would have been necessary to make a subject-index entry starting with "in," in order to subject-index all subjects reported in the paper.

"Aluminum" is the next word. The author is studying a system of aluminum with molybdenum that can be indexed with a subject-index entry, "**Aluminum**, system: Mo—." Also the complementary entry "**Molybdenum**, system Al—" is needed. Substitute entries expressing the same subject could have been "**Aluminum-molybdenum system**" and "**Molybdenum-aluminum system**," for example. The former way of handling the entries has been found to lead to less scattering of like information in a subject index to all of chemistry.

The word "rich" meaning "a relatively great quantity" is not the subject of study. It is in the same category as is "present" above.

Alloys are studied, but not alloys in general. Only Al-rich-Mo alloys are studied. The subject indexing of this concept is handled in *Chemical Abstracts* by indexing "**Aluminum alloys**, with molybdenum ..." and "**Molybdenum alloys**, with aluminum ...." Another form of heading could have been chosen, as "**Aluminum-molybdenum alloys**" and "**Molybdenum-aluminum alloys**." However the first way of indexing has been found to lead to better grouping (less scattering).

Often in an abstract of this type it is considered unnecessary to subject-index both the "alloy" and the "system" viewpoints since both overlap to so great an extent. The chemist-user of the index will know that alloys are usually formed in intermetallic systems and that alloys always represent parts of such systems. However, in this case, both "alloys" and "systems" have been indexed since both were studied and reported by the author. Sometimes overlapping indexing of this nature can be avoided by cross references that guide the index user from "alloy" to "system" and *vice versa*. Examples of such cross references are "**Aluminum alloys**. (See also 'systems' under **Aluminum**.)" and "**Aluminum**, systems with metals ... see also '**Aluminum alloys**'." Even with such cross references in use in an index, entries for both "alloys" and "systems" are sometimes needed.

The preposition "of" and the article "the," just as the preposition "in" discussed above, do not designate subjects studied nor are they related to subjects of study.

The name of the author should be entered in an author index. Discussion of author indexes is beyond the scope of this paper. Also, the laboratory where the work was done would be an important candidate for an entry in a place or location index (also outside of the scope of this paper).

The name of the journal itself already has been covered by lists of periodicals and the subject matter in the journal is carried in its tables of contents. Thus nothing, from a subject point of view, needs indexing here.

"$MoAl_7$" needs to be indexed and already has been discussed.

"Reacts" is not a subject or related to one since the author was, according to the abstract, studying the products of the reaction and one reactant, but not the reaction itself. Although a reaction did occur it either was not new or was incidental to the preparation of a studied product. Had the reaction been new and studied, it would have been necessary to index it under "Reactions" or a more specific "reaction" heading.

The next four words, peritectically, with, liquid, and at, in the abstract can be analyzed as potential sources of subjects analogously to some of the terms above. None is a subject for indexing.

The number 706° is not a source of a subject either. The author was not studying the uniqueness or novelty of the number 706. Also he was not concerned uniquely with temperature, as indicated by the degree sign. He did not study the relationship between "706" and temperature. He might have been studying these things. For example the temperature, 706°, might have some special significance in chemistry or metallurgy with this study related to the special significance. There is actually nothing of subject interest about 706, the degree sign, temperature, or their combinations. While this is a reported numerical datum, it is not a subject studied by the author. Thus, no subject-index entries are needed.

Every word, term, and number in the abstract can be analyzed as above to see whether it is related to a subject studied by the author or is suitable for use as a heading in the subject index.

Concepts implied by the author but not expressed in words must be examined for potential subject-index entries.

One additional entry, derived from the body of the abstract, is "Crystal structure." Crystal structures or parts of them were reported for $MoAl_{12}$ and $MoAl_7$. The subject here is the "crystal structure" and not the numerical values.

SUBJECT-INDEX STATISTICS ON *CHEMICAL ABSTRACTS*

For abstracts other than those on synthetic organic chemistry, for which most of the entries are for compounds, about 18% of the entries in the *Chemical Abstracts* subject indexes are made by use of words or synonyms not found in the material indexed. The heading terms of these entries are more generic, more specific, or merely suggestive in indirect ways of terms actually in the papers.[3] The indexer has supplied the index words from his knowledge of the field indexed. For example, a study on a new diaphragm for a Castner cell may not mention the terms "**Chlorine**" nor "**Sodium hydroxide**," yet both are proper subject headings and entries in an index devoted to chemistry. Another example is a study of new techniques for fractioning of air. "**Nitrogen**" and "**Oxygen**" are usually good subject headings here although they may not be mentioned in the abstract indexed. The helium group gases may call for entries.

About 17% of the subject-index entries in a chemical index are derived from synonyms in the non-organic material indexed.[3] For example, "Thiamine" rather than "Vitamin B_1" is chosen as index heading in order to put the entries alphabetically at the place where most index users will probably look first.[4]

For other than organic abstracts in *Chemical Abstracts*, about 34% of subject-index entries are derived from information given in the titles. In about 67% of these cases, the index-heading terms are identical with those in the titles; in the remainder, index-heading terms are not present or are there as synonymous designations.

The remaining 66% of subject-index entries for abstracts other than organic ones are derived from the bodies of the abstracts. For certain kinds of abstracts, such as those of synthetic organic papers, the average percentage of entries derived from the bodies of abstracts is much higher. These data show that complete subject-indexing is not possible from titles of chemical abstracts or papers. Most words in titles are not suitable subject-index headings and most words or their synonymous counterparts for subject-indexing are not found in the titles of such abstracts. From the above facts, it is clear that a concordance (word index) based upon titles of abstracts of chemical papers or the papers themselves would be incomplete from a subject viewpoint.

Title lists keyed by words have value for quick distribution and fast use since time is often a very important element in the obtaining of information. Such lists do not serve adequately for thorough searching. It is true that the usefulness of such a concordance can be extended by taking certain time-consuming steps. If *all* key words in a title concordance are examined, those closely associated with the subject concepts in the minds of the user tend to trigger responses that lead to examination of the context words. If this examination increases the interest of the searcher, then the full title is examined further to see whether effort needed to obtain and examine the paper or an abstract of it is justified. Thus, a title concordance may be more useful than would seem from the above data on index entries. However, it must obviously be incomplete, must have many unnecessary entries, and would not prove suggestive enough to users who lack background in the subjects sought.

RULES FOR INDEXING

Indexing primarily by rule rather than by choosing terms from a list has the advantage given above. Although every subject indexer develops many rules for his specific subfields, there are a relatively few basic rules that apply to all subject-indexing.

1. Only new information or information helpfully organized for use, as in comprehensive reviews, is to be subject-indexed. The old information was indexed when it was new. Old information re-indexed leads the users to material they already know or have access to and wastes both their valuable time and costly index space.

2. Indexing is to the maximum specificity unless a generic index is being built. The reasons for this rule have been discussed above.

3. Subjects outside of the scope of the index often can be handled satisfactorily by generalization. For example, in a subject index to chemistry, "calculus" can be indexed under "Mathematics" provided that a cross reference is used.

4. Unless misleading or ruled out by official standardization, the most commonly used of several synonymous terms is selected to be index heading and cross references are made to this heading from all of the unselected synonymous terms likely to be encountered.

5. Cross references are provided extensively to bring the users' vocabularies into harmony with the vocabulary of the indexes and to lead to sources of related information.

6. Headings used in previous indexing are used in preference to starting new headings from synonymous or closely related terms. This helps searchers through a series of indexes.

7. Modifying phrases (modifications) are provided to make entries more precise and thus to enable the user to select references to sources of information most closely related to his interests.

MODIFICATION-WRITING RULES

While there are many more or less specific rules for the writing of suitable modifications, *Chemical Abstracts* uses only six general rules.[5] The most fundamental rule for modification writing, for use particularly when modifications are alphabetized as in *Chemical Abstracts*, is to put the most significant word first. However, certain arbitrary rules, helpful in the selection of the most significant words, reduce scattering and index-editing time and cost without too-often affecting the general rule to put the most significant word first. These rules are:

1. The modification is started with action words when this can be done so that the heading will read correctly into the beginning of the modification. Starting the modification with action words is avoided if the heading can be read correctly only into the middle or end of the modification.

Examples of approved form:
Argon, adsorption by Cu, ...
Copper, argon adsorption by, ...
Penicillin, enzyme effect on, ...
Zinc, reaction with O,. . . .

Examples of less suitable form:
Argon, copper adsorption of, ...
Copper, adsorption of A by, ...
Penicillin, effect of enzymes on, ...
Zinc, oxygen reaction with, ...

2. Alphabetical order is used wherever possible. The principal exception to this rule is for a natural sequence of operations that is unalphabetical. Multiple entries for the same document under the same heading are not used.

Examples of approved form:
Copper, analysis, detn. in blood, cerebrospinal fluid and gastric juice, ...
Hexane, prepn., distn. and testing of, ...
Copper ores, screening, grinding and flotation of, ...
Air, argon and He in, ...

3. The modification is started with a term used as heading in another part of the index rather than with a synonymous term. (Some exceptions may occur; *e.g.*,

DDT, 2,4-D, *etc.*, are common terms used at the start of modifications although they are not used as index headings. This practice saves space and probably helps the user.)

Examples of approved form:

Rice, thiamine extn. from, ...
Calcium acetate, acetone prepn. from, ...
Corn, DDT detn. on, ...

Examples of less suitable form:

Rice, vitamin B_1 extn. from, ...
Calcium acetate, 2-propanone prepn. from, ...
Corn, 1,1,1-trichloro-2,2,bis-(*p*-chlorophenyl)-ethane detn. on, ...

4. Unnecessary symbols and terms are omitted, especially at the beginnings of modifications. Examples are: "L-," under most amino acids; "surface" under Coating(s); "iron" and "steel" under "Corrosion;" and "sulfate" under "Quinine" in most circumstances. Although "L-" should be omitted under most amino acids, isomer designations are often needed at the beginnings of modifications under many headings.

5. When two or more terms in conjunction are associated on an equal basis with a third term, the modification is started with the third term even though this violates rule 1. This rule prevents ambiguity. This form is used instead of multiple entries under the same heading for the same abstract.

Examples of approved form:

Hexane, analysis for 2-methyl pentane and pentane, ...
Heart, effect of adrenaline and noradrenaline on, ...
Charcoal, adsorption of Ne and Xe by, ...
Peas, deficiencies of Cu and Zn in, ...
Insulin, treatment of diabetes and schizophrenia with, ...

Examples of less suitable form:

Hexane, 2-methylpentane and pentane detn. in, ...
Heart, adrenaline and noradrenaline effect on, ...
Charcoal, neon and Xe adsorption by, ...
Peas, copper and Zn deficiencies in, ...
Insulin, diabetes and schizophrenia treatment with, ...

6. For some headings for special reasons (usually apparent) the above rules apply secondarily to special rules. Examples of such headings in *Chemical Abstracts* are "Oils," under which the kind of oil starts the modification, and "Dyes," under which a form using few propositions and a fixed order (explained in a note at the start) are employed. The first five rules always are applied if the modification structure is not completely determined by the special rule.

Observation of these six rules will save much editing when the index cards are being made ready for printing because fewer modifications will need change and fewer index cards will need transposition.

SPECIFIC RULES

Standardization is achieved in a subject index built by rule by recording each decision that affects the index. In the *Chemical Abstracts* office if the note or rule is of general interest, it is replicated and given to all indexers; if not, it is used by the indexer making the rule to increase his precision of indexing. Indexing to the maximum specificity can be started without any specific rules and without a list of subject headings. Editing of the assembled index entries with the help of inverted cross references[4] removes scattering, standardizes headings, and brings in needed entries from a master file of about 60,000 cross references.

APPLICATIONS OF SUBJECT-INDEXING

Subject-indexing is useful in guiding searchers to subjects, but not to words in documents. The latter function requires a concordance or word index. The philosophy, rationale, and techniques of subject-indexing are quite independent of the structure and form of the final index. That is, the index can be alphabetical, classified, or correlative;[6] it can be in the form of pamphlet, book, deck of file cards, punched cards, magnetic tape, permuted trope numbers and titles,[7] edge-notched cards, microform, *etc.* For example, in correlative-trope indexing,[2] the general terms of the list are chosen to apply to new subjects reported in the documents to give a correlative trope *subject* index. The general terms could be chosen to relate to words used in the documents, in which case a correlative trope *word* index would result. It would be difficult to overemphasize the importance of the difference between subject- and word-indexing and of the fact that these two kinds of indexing and indexes are independent of the structure and form of the index. Word-indexing by use of a list of trope terms will result in a key to documents that will be largely ineffective in locating subjects. An example (from the abstract above) may help to make vivid this point. The above abstract of the paper by Clare reports that "$MoAl_7$ *reacts* peritectically with liquid . . . ." The trope term "reactions" could have been chosen to index this abstract on the grounds that a reaction occurred between $MoAl_7$ and Al to yield $MoAl_{12}$. This would be word-indexing because Clare did not study the reaction (its kinetics, mechanism, catalysts, *etc.*). He was, according to the abstract, studying the products of the reaction and one reactant, but not the reaction itself.

For indexing this abstract, suitable correlative trope terms from the list given in Appendix A of part VII of this series[7] would be: "Structure," "Transition elements," "Periodic group IIIA," and "Alloys." The term "Structure" indexes the new data reported on crystal structure of $MoAl_{12}$ and $MoAl_7$. "Transition elements" takes care of molybdenum, "Periodic group IIIA" handles aluminum, and "Alloys" subject-indexes the alloys and intermetallic compounds studied.

With the use of the numbers assigned to the terms in the previous paper of this series and with the omission of the page reference, one trope-index entry would be "3, 28, 30, 40 Compounds present in aluminum-rich alloys of the aluminum-molybdenum system."

There would be three other entries with the numbers 28 3 30 40, 30 3 28 40, and 40 3 28 30 preceding the same title and page reference.

CONCLUSIONS

Word indexes, including concordances, are grossly inferior to subject indexes in providing a subject approach

to the material indexed. Only when the index user wants statistical studies on word frequency and usage (quite different from subject approach) are word indexes and concordances of greater value than subject indexes. Thus, word indexes are useful for linguists, archeologists, and cryptographers. For those interested in finding *all* of the information of papers or abstracts, without suffering the limitations of words unorganized for keying, of titles only, and of expressions containing inference, a subject index is far superior as a key in locating such documents. This superiority comes about principally from fewer useless entries in true subject indexes, freedom from scattering of like information, and avoidance of omissions.

It is not the purpose of this paper to discuss subject-indexing *versus* word-indexing in relation to mechanized systems for storing and retrieving information, but it seems appropriate to make one general comment. In the building of such mechanized systems it will be well to bear in mind the differences between subjects and words in providing keys. Mechanized systems based only on words as found in the literature have the serious shortcomings of word indexes as keys to information about subjects.

REFERENCES

(1) C. L. Bernier, and K. F. Heumann, "Correlative Indexes. III. Semantic Relations among Semantemes—The Technical Thesaurus," *Am. Document.*, **8**, 211-220 (1957).

(2) C. L. Bernier, "Correlative Indexes. II. Correlative Trope Indexes, *Am. Document.*, **8**, 47-50 (1957).

(3) C. L. Bernier, "The Indexing Problem," *J. Chem. Doc.*, **1**, No. 3 (1961).

(4) C. L. Bernier and E. J. Crane, "Indexing Abstracts," *Ind. Eng. Chem.*, **40**, 730 (1948).

(5) E. J. Crane, ed., "CA Today—The Production of Chemical Abstracts," American Chemical Society, Washington, D. C., 1958.

(6) C. L. Bernier, "Correlative Indexes. I. Alphabetical Correlative Indexes," *Am. Document.*, **7**, 283-288 (1956).

(7) C. L. Bernier, G. M. Dyson, and H. J. Friedman, "Correlative Indexes. VII. Trope Vocabularies and Trope Indexes for Chemistry," *J. Chem. Doc.*, **2**, 93 (1962).