Book-indexers can often benefit from the experience of other practitioners of their craft. We are pleased to be able to print these contributions by J. ARTHUR GREENWOOD on different aspects of indexing.

ALPHABETIZATION: VERBATIM OR LITERATIM?

Although I have strong (not unconquerable) prejudices in favour of word-by-word indexing, I see no hope of eliminating one of the two methods.

1. Word-by-word has the advantage of familiarity; the reference book most often consulted by the layman (i.e. the telephone directory) is alphabetized word-by-word.

2. Believing that the formulas word-by-word and letter-by-letter mean almost nothing to the layman, I practise to supplement them by ‘New York precedes Newark’ or ‘Newark precedes New York’.

3. Word-by-word alphabetizing can be done fairly readily by using so-called sort programmes on computers, if the collating sequence can be so adjusted that hyphen precedes A, and ampersand (if allowed in index entries) follows Z. This is not an important argument; the convenience of the user is the supreme law, which must not be overridden to accommodate the indexer or his computer.

4. Word-by-word indexing requires some cross-references when a catchword has plausible solid, hyphenated, and two-word forms. The number of cross-references is matter for judgment; the safe rule is to insert cards for all plausible forms not used, and at the end to delete any cross-references that adjoin the form used. A user, in contrast to an optical scanning machine, has a little peripheral vision.

5. Letter-by-letter indexing requires one important exception: in indexing personal names, the family name governs the alphabetization, and the Christian name is used only to break ties. This rule may be left tacit in instructing the user of an index, but must be spelled out in briefing an indexer. The 1969 Chicago Manual of style, paragraphs 18.34-18.71, converted to letter-by-letter from the word-by-word style of the 1949 edition, neither states nor exemplifies this rule, so that an indexer enjoined to ‘follow Chicago’ would be justified in submitting an index containing the sequence

    Clark, Edwin
    Clarke, Jeremiah
    Clark, Eugene

6. The New York Telephone Company’s treatment of double-barrelled surnames is confusing: the current Manhattan directory contains the sequence

    JONES Beulah Etta
    JONES-BEY Zoradea A
    JONES Billie

7. In indexing an encyclopaedia or similar work, if the article titles have been arrayed letter-by-letter, the index must be alphabetized letter-by-letter, and vice versa. To require the hapless reader to bear in mind one convention in the text and the opposite in the index is to ensure confusion and invite disuse of the book.

TYPOGRAPHY OF THE INDEX IN BOOKS REPRODUCED FROM TYPESCRIPT

In the discussion that has appeared in The Indexer about whether the indexer should produce a set of pasted-down slips or alphabetized pack of cards, and stop, or whether he should proceed to make a typewritten index, with indents unmistakably shown (and of course include his labour as a typist in his bill), no one has suggested that the indexer should be his own typographer, yet who knows better than he the exact effects required on the printed page if the reader is to succeed in consulting the index? The notes below suggest that the indexer can and should, and sketch this indexer’s experience.
with the typographical effects obtainable on the typewriter.

1. The salient property of the typewriter as a composing machine is that white space at the end of a line or paragraph, or at the foot of a page, comes gratis. The operator returns the carriage, or removes the page, and has done; with no need to measure or assemble spacing material.

2. That a typewriter cannot produce a square right margin is no drawback in setting an index: cf. Chicago Manual of style, paragraph 18.82. The definition of columns can be improved by carrying short overruns to the right margin.

3. The practice of routinely using I.e. initials reserves cap. initials for distinctive use. This can be helpful in setting a subject index on a machine without italics.

4. It is good form to use a figure 1 distinct from the letter l, preferably a 1 with a pointed top. The possibilities of actual confusion are two: page one [1] of text with page fifty [I] of front matter; 1 with the abbreviation I. for line. The confusion between zero, cap. O, and l.c. o, which plagues the mathematical typist, is not a problem in indexing.

5. In demy octavo, A5, and larger sizes a double-column index is practicable; in demy quarto and A4, triple-column. With short entries, consisting largely of two words and a page number, three columns can be squeezed on to a medium octavo page.

6. To distinguish several classes of entries in one alphabet, the following diacritics are available: roman and italic; light face and black face*; lower case and capitals; underlining. I have used these in combination to distinguish six classes; it is easy to arrange schemes for eight or twelve classes, but difficult to imagine either subject matter requiring such schemes, or a reader who would trouble to learn the key. When the roman/italic contrast is wanted, it can be imple-

*This is not a racial slur. I have ventured to invent the generic name black face because bold face is pre-empted as the specific name of one style of black face, viz. Monotype No. 53.

The Indexer 7 (I) Spring 1970 25
designed as a compromise, to produce pass-
able print at either 10 or 12 pitch; on a
10-pitch machine, italic caps have a passable
appearance.

10. Underscores. The essential point is
not to use the single underscore as a substi-
tute for italics and the double underscore
as a substitute for caps., but to treat the un-
derscores as independent distinctions, applic-
able to any style of type available. In
indexing Edgar A. Poe's *The rationale of
verse* I used roman type (Selectric Delegate)
for subject matter; italic for first lines of verse
quotations; roman underscored once for
English words as objects of discussion; italic
underscored once for foreign words as ob-
jects of discussion; roman underscored twice
for authors cited; italic underscored twice
for magazines cited. The double underscore
has been cut for most typewriters; on the
Selectric it is found on the element with
superior figures and mathematical signs. A
wavy underscore has been cut for some
typewriters—apparently not for the Selectric.

11. In a text, the best placement for page
numbers to facilitate reference is outside top;
this is convenient for both right- and left-
handed readers. In an extended index—say,
seven printed pages or more—the place of
honour, as in a dictionary, belongs to the
catchwords. The page number can then be
moved economically to the inside top, where
it instructs the printer and binder, and per-
mits reference to a specific page of the index
si opus sit, without distracting the reader
from the catchwords. The printer should be
warned that this arrangement is deliberate,
and not an accidental misassignment of odd
numbers to versos and even numbers to
rectos.

12. The form of catchword I prefer is two
lines, top left of left-hand page, giving
inclusive contents of two facing pages, in-
cluding any entry carried forward from the
preceding recto; two lines, top right of right-
hand page, identically repeating the above.
On ordinary turn-overs the indentation of the
first line on the verso is sufficient warning
that the reader has entered the middle of an
index entry; but the insertion [MORE] at
the foot of the recto is a useful warning that
the entry continues overleaf. The flag
[MORE] could have been used to advantage
(but was not) in the index to vols. 1-31 of
*Annals of mathematical statistics*, where
the entry normal distribution extends over
pp. 436-449.

13. Suppressing the comma between the
index entry and the page number calls for an
oversized space as separator: say three times
the normal word space. Thus, in Monotype,
set an em quad; on the 'Executive' type-
writer, a 6-unit space; on other typewriters,
tap the space bar three times.

14. Selectric types remain legible at re-
ductions down to 50 per cent; 58 per cent is
the smallest size I find comfortable. This
affords 178 characters to the square inch set
solid, or 119 characters with 1½ line spacing.
The greek types available for the Selectric
are conspicuously smaller than the latin
types; I should not advise a reduction below
64 per cent. The only cyrillic type I have
seen on the Selectric has a very large face,
and so in a mixed Russian-English index the
camera setting can be governed by the latin
type used.

15. The following arbitrary signs may be
found useful in setting an index:

*Asterisk.* Letter Gothic type (Selectric)
has a particularly clear-printing asterisk.

*Single dagger, paragraph, and section* are
on the Selectric element with superior
figures. This element also contains
down and up arrows, which might be
used as signs of sub- and superordina-
tion in complicated indexes.

*Tic tac toe* (used in U.S. for pounds
weight). On typewriters with U.S. lay-
out; the sign on Selectric italic looks
right with most types.¹

*Inverted exclamation mark*: on types with
Spanish layout.

*Turned point*: roll platen down one tooth
and type a full point. The result is
unmistakable, but sets a little too high
to look pleasant.

¹ A sign similar to a thin music sign for sharp, but
italic.

*The Indexer 7 (I) Spring 1970*
Double dagger: type the plus mark, roll down one tooth, back space and type another plus mark.

16. Before starting, have your cards or paste-down marked for indents, type faces, and underscoring; and, preferably, cast off into column lengths. Do not attempt to estimate the time required from your speed as a correspondence typist or a manuscript copyist; in particular, remember that a type change on a Selectric takes a minimum of eight seconds.

TABLE OF RHYMES

On the suspicion that an index to terminations will help the reader of a book of poetry or criticism of poetry to locate verses quoted or discussed, I supplied such an index for Poe's *The rationale of verse*. The table avoids any phonetic arrangement because (1) phonetic symbols are awkward for the typewriter; (2) the labour of mastering a phonetic arrangement would deter most readers from using the table; (3) what with poetic licences, the carelessness of composers, and time elapsed between composition and indexing, rhymes as printed afford no certain clue to the pronunciation.

The indexable termination comprises the vowel or diphthong of the long (or accented) syllable and all letters following; the terminations are alphabetized on the rightmost letter, using the penultimate, antepenultimate, etc., letters to break ties. Terminations adjacent in this anagraphic ordering and not distinguished in pronunciation are bracketed into one entry: for example ED/EAD, EAMING/EEMING, AUGHT/OUGHT, ART/EART, AY/EY.

Within each termination rhymes are recorded as found in the verses, except that the combinations head/said and said/head (say) are not kept separate. Rhymes are alphabetized by the alphabetically earliest word containing the indexable termination: thus, when the rhyme be/sea is sought under EA, the alphabetical place of sea governs; under E, be governs. To break ties, the length of a string of rhymes is used (longest string first); ties between strings of the same length are broken on the alphabetically earliest non-identical words in the strings.

STATES AND TERRITORIES, NAMES OF, PUNCTUATION OF

During a desultory discussion of indexing with a colleague in 1965, I handed him the index to a book on meteorology that contained an account of 'Project Prairie Grass', an experiment in atmospheric diffusion carried out near the town of O'Neill, state of Nebraska. His eye fell on the entry

O'Neill, Nebraska

which he read aloud as 'Nebraska O'Neill' before stopping in disbelief at Nebraska as a Christian name. Clearly, if the entry had read

O'Neill, Alberta

he would have had no reason to reject his misapprehensive reading. Again, in indexing a (hypothetical) book *Edgar A. Poe—his life and loves*, how shall the indexer cater for the sleepy or uninstructed reader, oblivious or unaware that Annie Richmond is a woman, Richmond, Virginia a town, and Virginia Clemm a girl?

An obvious remedy is to make a rule that names of states and territories should be bracketed in indexes. Adding a few plausible entries to the three above we reach the sequence:

- Baltimore (Maryland)
- 'Baltimore Mary'
- Clemm, Maria
- Clemm, Virginia
- Lowell, James Russell
- Lowell (Massachusetts)
- Richmond, Annie
- Richmond (Virginia)
- Richmond Examiner [newspaper]

Is the use of brackets around names of states and territories likely to collide with any other purpose for which an indexer requires brackets?