

[From: *Nature*, November 2, 1963 vol.200]

"INVISIBLE LUNATIC"

Chinese Character Indexes

Ching-yi Dougherty, Sydney M. Lamb and Samuel E. Martin. Vol. 1: Telegraphic Code Index. Pp. xxxi + 364. Vol. 2: Romanization Index. Pp. 3G4. Vol. 3: Radical Index. Pp. 364. Vol. 4: Total Stroke Count Index. Pp. 3G4. Vol. 5: Four Corner System Index. Pp. 364. (Berkeley and Los Angeles: University of California Press; London: Cambridge University Press, 1963.) 220s. per set.

THE production of this five-volume index of Chinese

characters, by a group of people whose main reputation lies in the field of mechanical translation of language, marks an important event in the field of systematic linguistics, and invites a careful examination of the purpose which these volumes may serve, and of their fitness for this end.

The first of the volumes contains a twelve-page introduction which explains the purpose of the work, which is considered to lie under five headings. First, to facilitate the ready conversion of Chinese characters into code symbols and vice versa; secondly, to assemble all the Chinese characters in modern use; thirdly, to establish standard pronunciations for each character; fourthly, to clarify relationships between characters; and fifthly, to serve as a focal point for cross-reference between the different systems at present used in arranging Chinese characters.

An examination of the volumes, and of the introductory material, shows that in the last three of these objectives the authors have to a large extent succeeded in their work. The first two purposes, however, seem to me to be far less well achieved. The contents of the volumes are as follows: Volume 1, the characters listed in sequence by the telegraphic code number; Volume 2, the same, in the alphabetical order of Romanization of pronunciation; Volume 3, the traditional Chinese order by radical number and residual stroke counts; Volume 4, in a sequence according to the total number of strokes; Volume 5, in sequence by the four-corner system.

Now it is true that once a computer programme has been written to deal with Chinese characters the representation of those characters by means of a telegraphic code number achieves a useful saving in time. Unfortunately, the major difficulty of Chinese to English translation lies in the recognition of the characters themselves and, speaking as a non-expert in Chinese, I regard the main difficulty in work within this field to be that of the recognition of the characters themselves. In the present work the recognition criteria are: Romanization, radical number, total stroke number, and sequence by the four-corner system. Those, however, require an expert knowledge of Chinese idiographs in order to obtain the relevant numbers and thus to be able to find the character within the index provided. It would surely have been easy for the authors to write a short factual account of the methods which are used by Sinologists to obtain the characterization numbers and, in particular, the sequence of numbers demanded by the four-corner system. Even supposing that this operation were regarded as too difficult for the non-expert, the volumes would at least have been useful for English-Chinese transcription had they been accompanied by an extra volume listing the Chinese characters in alphabetical order of their English meaning or meanings. In other words, an English to Chinese dictionary. It might be argued that the five volumes so far provided are large and expensive enough to inhibit further activity in the field, but this in itself leads to a criticism of the format of the volumes themselves since the repetition of the Chinese characters and of the associated descriptive material could have been avoided if, accompanying the first volume, that is, the telegraphic dictionary, the remainder consisted merely of a concordance in which

the listings were given and the reader referred back to the telegraphic index for the full particulars including the idiographic symbol.

The second of the objectives, that of assembling all the Chinese characters in modern use, is more difficult an assessment and suggested the title of this review. There is an old (and probably apocryphal) story that a certain computer was asked to translate "Out of sight, out of mind" into Chinese and that a second machine later performed the reverse process, producing as a result the words of the title of this review. The point of the story is that, although telegraphic entries Nos. 4035, 4045 and 4082 are associated with madness and imbecility, neither the word lunatic nor invisible seems to appear in the list. The volumes contain many references to scientific terms of a somewhat recondite character; for example, on the first page the elements actinium and einsteinium are listed with their associated idiograph. In a similar manner there are many terms from organic chemistry, but it would appear to me that far more important symbols and concepts are omitted so that the claim to have assembled all the Chinese characters in modern use appears somewhat curious unless the Chinese themselves have exercised a remarkable selectivity in the addition of modern scientific terms to their vocabulary. As it is, some 10,487 characters appear to have been listed, including one or two (namely, telegraphic code numbers 9155 and 9156) which have no Roman equivalent and no English meaning or meanings, and therefore seem to be of somewhat dubious utility.

It is almost certainly an unreasonable hope that such large and opulent volumes will be soon reprinted; however, in view of the fact that the information concerned is stated to be held in a punch card file there would be no real problem in adding an additional volume giving the information which is suggested here, especially as it is stated that the arrangement and rearrangement of the information contained in the present volumes only took a few days.

The present collection makes a valuable start in giving to students of mechanical linguistics information on a language the principles and content of which are too little understood; but it requires considerable expansion before it becomes a really authoritative work. A. D. BOOTH