

Graphic Linguistics and its Terminology

R. A. Crossland

DURING the past thirty years great advances have been made towards making the study of language a science, but leading linguists have been mainly concerned with spoken language. There has been a certain tendency to suggest that the study of written documents should always be subsidiary to that of some spoken idiom, or even that it is bound to be less scientific than that of spoken idioms, and perhaps not a proper part of "linguistics" at all.¹

These suggestions should be opposed. "Linguistics" should include the study of written languages as well as that of spoken; the former study can and should be as scientific as the latter, and it needs its own terminology which should be basically independent of that of the study of spoken languages. Much confusion, and some mistrust, if not antagonism, among linguists would seem to have resulted from lack of agreed distinct terminologies for the two studies, which might well be called respectively phonic and graphic or epigraphic linguistics.²

The problems of graphic linguistics are probably best approached through consideration of what writing is. A script may be defined as a system of visual symbols whose purpose is to convey the thought of one individual or group to another. Writing is often treated as a means of representing a spoken utterance or utterances by visual symbols, but this is not its primary purpose, except where phonetic or phonemic transcription in linguistic work is concerned. Representation of actual, contemplated or imagined utterance is a particular mecha-

nism for conveying meaning by graphic signals, one whose convenience lies in the small number of signs required. The adoption of a particular form of it, alphabetic writing, in Western Europe, has led to its being widely regarded as the normal and natural mechanism, and some of those who have discussed the analysis of systems of writing have tended to write as if they were all more or less satisfactory systems of phonemic transcription of utterances. This attitude leads to or supports the view that the study of written documents should always be subsidiary to the study of some spoken idiom, or as an extreme to the idea that "texts" are not "language".³ One must leave to psychologists the question whether it is possible to read or write without some thought of phonic⁴ realization, whether based on a known spoken idiom or not. But it can hardly be denied that the users of a system of graphic communication may develop for it conventions of vocabulary and grammar which differ from those of any spoken language which they use, or on which the system was originally based. A group of texts showing similar conventions of grammar and vocabulary may reasonably be termed a "written language".⁵

Most of this will probably be accepted by the majority of those concerned with the study of spoken languages, though in some cases with the proviso that the study of written language should be considered a discipline separate from "linguistics" and "philology." Such differentiation, however, has the disadvantage of tending to dissociate the study of the spoken form of a lan -

1. Cf. W.S. Allen, "Phonetics and Comparative Linguistics", Archivum Linguisticum 3, (Glasgow), 126-36.

2. Choosing between graphic and epigraphic here involves a problem common when technical terminology is devised, whether to use the term which is etymologically the most natural, in spite of its currency in non-technical language in another sense. For epigraphic, cf. A.F.L. Beeston, Transactions of the Philological Society, 1951, 1-26, where it means 'of the inscriptions'.

3. Cf. Allen, op.cit., pp.132, 136.

4. As phonetic is now generally used of description of utterances or segments of utterances according to the manner of their articulation, a more general term to cover all studies concerned with spoken language is required, and phonic seems suitable. The use of phonics proposed by J.R. Firth, Trans. of the Phil. Soc., 1951, 84, has not become widespread.

5. Or a "written dialect", if its relation to another group with closely similar conventions is under consideration.

guage from that of the written, where both forms exist, a development particularly undesirable in the case of semantic studies. "Linguistics" should include the analysis and study of the mechanism of both spoken and written languages, while "philology" should be used of studies of the content of written texts, in particular for historical or literary ends. This usage is in fact normal in American English, and corresponds to German use of Sprachwissenschaft and Philologie. "Philology" and "graphic linguistics" will overlap to some extent, especially in semantic studies, but there is a clear distinction between the two in purpose.

Graphic linguistic study, as well as phonic, may reasonably be called "descriptive" or "structural" if its procedures are appropriate. An analysis of the conventions of a class of texts may be termed "descriptive" if it is not shaped by a preconceived notion of what they should be; "structural," if it aims at determining significant oppositions.

Recent work in phonic linguistics has established a terminology for phonetic and phonemic description of spoken languages, and recently suggestions have been made for a similar terminology to be used in analysis of written languages.⁶ None has yet become generally accepted, however, and those proposed seem unsatisfactory in so far as they are based mainly on the partly phonemic, alphabetic scripts⁷ of Western Europe and are not easily applicable to scripts of other types. The analyses which they imply are in some cases not purely graphic, as they reflect the function of the written signs or the conventions of their combination in representing phonic features of spoken languages.

The terminology now most used in Britain in describing spoken languages permits description at three levels: phonetic description of a single

utterance, phonetic description of a number of utterances, and phonemic description, which may be defined for present purposes as description on the basis of contrasts significant to normal users of the language in question. Distinction is made, for example, between a sound which seems to require definition as "the audible result of a single emission or intake of breath or closure or opening of speech organs by a particular speaker on a particular occasion"; a sound-class - any group of sounds, as just defined, which an investigator associates, perhaps provisionally, in analyzing the phonetic structure of a language, for example, on grounds of phonic similarity or occurrence in similar contexts; and a phoneme, which for convenience may be defined as a sound-class differentiated functionally from others.⁸

It has been recognized that graphic linguistics needs a set of terms similar to sound, sound-class and phoneme in the technical language of phonic linguistics. It would seem to need at least a term for a sign, modification of a sign or feature of arrangement in a particular segment of a particular document; one for a group of similar signs, modifications or features classed together, provisionally or permanently, in graphic analysis; and one for any such group which appears to contrast significantly with another or with zero. Graph or sign suggests itself for the first, graph-class or sign-class for the second, and grapheme for the third. To illustrate the use of these proposed terms, a in a particular written word; for example, class, in

6. See D. Abercrombie, "What is a 'letter'?", Lingua 2, 54-68; P. Diderichsen, "Nye bidrag til en analyse af det danske skriftsprogs struktur", Selskab for Nordisk Filologi, Arsberetning for 1951-52, (Copenhagen), 6-22; E. Pulgram, "Phoneme and Grapheme: a parallel". Word 7, 15-20; H.J. Uldall, "Speech and Writing", Acta Linguistica 4, 11-6; J. Vachek, "Some remarks on writing and phonetic transcription", Acta Ling. 5, 86-93. Diderichsen's article seems particularly important.

7. Cf. Pulgram, Word 7, 15; ". . . . each alphabet has a certain number of classes of symbols" (my underlining).

8. In passing, the choice of sound as a term for the first concept in the publications of most members of the London University School of Oriental and African Studies seems unfortunate. The creation of new terms in technical language is preferable to use of current ones with new artificially restricted meanings. Moreover, sound has long been used in philological and linguistic literature with an accepted sense: the range of "sounds" (in the restricted sense just mentioned) which normal speakers of a language known only from written documents are thought to have produced in pronouncing - "giving phonetic realization to" - a word-segment represented by a given phonic grapheme (cf. "the sound f in Lat. filius", the meaning of which is clear enough). However, a term for the restricted concept to whose expression some would limit sound in the technical language of linguistics is certainly needed. Perhaps phone would serve; cf. Pulgram, Word 7, 15.

this present text, would be described as a graph: all small a's of similar formation in a document or group of documents as a graph-class. Only full examination of how a script is employed in documents under consideration--analysis of its structure, that is to say--will indicate which graph-classes should be termed graphemes. For example, graphic analysis of a sufficient number of documents in modern English would lead to three varieties of written A being distinguished as graph-classes: a, a, A. Structural analysis would probably require the first two being considered to form, together, a single grapheme, since, except in special texts, such as phonetic transcriptions, they never contrast significantly in the same document. Capital A would probably have to be considered a grapheme in written English. Its occurrence at the beginning of sentences may be considered not to involve significant contrast with small a, since sentence division is indicated by the full stop. But there are cases where the use of capital or small a initially is the only graphic indication whether a person, place or group of persons or places is referred to, or some more extensive concept: cf. the Archers and the archers.

A principal difficulty of graphic analysis will be to decide whether certain features should be considered independent graphs or graphemes (according to the level of analysis) or not. In the case of most scripts there will be an obvious division into what may be called provisionally unitary graphemes and graphemes of arrangement or modification. The simplest case is offered by a linear phonemic script, which uses gaps to indicate word-division. In this case each letter will be a unitary grapheme representing a segment of a spoken or imagined word. Sequence of unit graphemes from right to left in scripts using the Latin alphabet, will be an arrangement grapheme representing temporal order of enunciation of the segments which they represent. Juxtaposition of unitary graphemes, at less than certain intervals in normal texts, will be an arrangement grapheme indicating that the segments represented constitute a word. Italicizing to indicate emphasis is an example of a modification grapheme. Description of graphemes according to their function in scripts which are only partly phonic in principle will be a good deal more complicated. It might be fairly simple in a fundamentally ideographic script--Chinese is the only example, I think, apart from the earliest Sumerian.⁹

9. The Chinese script is the obvious example. Others are the earliest Sumerian and Egyptian, and the Mayan.

The differentiation of unitary graphemes and graphemes of arrangement or modification should be a fairly simple process. It will often be more difficult to decide whether a particular symbol is to be regarded as an independent grapheme or not. Decisions will have to be made on grounds of ease of recognition, or with regard to the ideas of those who normally use the script in question. For example, it is arbitrary and a matter of convenience whether we analyze the Sanskrit signs usually transcribed -ra, re, -r, (final position only), ri, ru, pa, pe, -p (final only), pi, pu, as eight separate graphemes, or as six, k and p, modified by a grapheme zero (indicating following a), and graphemes representing following i, following u and absence of following vowel. If, in analysis of a linear script, superlinear or sublinear symbols are treated as graphemes, it will presumably be necessary to differentiate them from unitary graphemes and graphemes of arrangement or modification.

Differentiation of graphemes on the basis of the manner of their employment in the script to which they belong is the only proper differentiation in a descriptive study of a written language. Differentiation of graphemes according to the manner in which they are used to represent concepts and their nexus will be necessary when the history of a script or the interaction of written and spoken forms of a language is studied. One may then want to make a distinction, for example, between phonic graphemes, which indicate a concept by indicating more or less accurately its oral realization in a spoken language, and what are generally termed ideograms, but which for the sake of symmetry within the terminology one might better call idea graphemes, concept graphemes or notional graphemes.¹⁰

A complex terminology would be needed to describe e.g. Babylonian cuneiform, which is partly syllabic, partly ideographic.¹¹

From the point of view of mechanical translation, the following seem important:

10. Logogram should only be used of a sign representing a particular word. It would be incorrect, for example, to apply it to the Sumerian sign No. 172 in P.A. Deimel, Sumerisches Lexikon, which represents in different contexts bil, "burn", and izi, "fire". A purely logographic script would be impracticable for most inflected languages. The number of signs required would be prohibitive.

1. Written texts can be scientifically described and analyzed without reference to any spoken form of the language in which they are written or to the spoken language which the script in which they are written originally was devised to represent.

2. Problems of ambiguity resulting from homography in written texts are not likely to be more frequent or more serious than those which result from homophony in a spoken language.

3. No system in regular use will represent the nuances conveyed by emphasis or intonation in a spoken language, but this is not a serious objection to mechanical translation of written documents of the type in use in most modern ci-

vilized countries. In the written forms of many languages, nuances, of the type mentioned, in the spoken forms are conveyed by alternative means, and an individual may quite well express his ideas in the written form of a language, (or even in a dialect or foreign language which he does not speak) more precisely than in the spoken idiom which he normally uses.

4. Although a phonemic text may be regarded as an abstraction of utterances, it is probably better to regard written and spoken forms of a language as different realizations of concepts and their nexus than to regard either as on a higher level of abstraction than the other.

11. A syllabic grapheme may be defined as one representing a phonic segment which those who devised a syllabic or partly syllabic script thought they could distinguish when they attempted to analyze words of the language which they spoke, for graphic representation.