

# THE NOTION OF DIACRITICS IN MODERN ENGLISH GRAPHOLOGY

VILÉM FRIED

*Caroline University, Prague*

Professor Vachek's pioneering contributions to the analysis of written language (or, as he prefers to call it, the written norm of language) in contrast with the spoken language (or spoken norm), the first of which he published as early as in the 1930's, have been acknowledged by linguistic scientists all over the world. May this modest paper dealing with a minor problem in the analysis of some graphemic factors in ModE. spelling be a tribute to the eminent Czech linguist and an expression of our profound admiration for his extensive scholarly work.

J. Vachek has so far summed up his research into problems of the written language in his study 'Two Chapters on Written English.'<sup>1</sup> He emphasizes that the written norm of a language should not be identified with orthography which is, so to say, a bridge leading from the spoken language to its written counterparts, whereas in the opposite direction, from the written language to the spoken language, we deal with pronunciation. We would prefer here to use the term 'reading aloud'. A spelling system (or the graphology of a language) may not be adequate in both directions, J. Vachek refers to the two qualities that a spelling system should have to serve both purposes, viz. the qualities of 'surveyability' and 'learnability', the former suits the reader and the latter suits the writer. There are not many spelling systems that meet the needs of both. Significantly the American Rudolf Flesch says in his often quoted book *Why Johnny can't read, and What we can do about it*: 'Many years ago, when I was about fifteen, I took a semester's course in Czech. I have since forgotten everything about the language itself, but I still remember how the letters are pronounced, plus the simple rule that all words have the accent on the first syllable. Armed with this knowledge, I once surprised a native of Prague by reading aloud from a Czech newspaper. "Oh, you know Czech?", he asked. "No, I don't understand a word of it", I answered, "I can only read".'<sup>2</sup> This is understandable if we take into account that Czech spelling is made up of 25 unmarked graphemes (including one digraph) and 15 graphemes marked with a diacritic, i.e. 40 graphemes altogether, practically all of which are phonemically univalent. This greater number of graphemes makes reading easier in Czech than in English though English has only 26 single graphemes at the disposal of its spelling, the great majority of which, however, are phonemically multivalent.<sup>3</sup> If we apply Vachek's statement about the two qualities that a spelling system should have in order to serve both the reader and the writer to ModE. spelling, we may look at its seemingly high irregularity and often beemoaned chaotic state from two points of view, i.e. from the point of the writer and from the point of the reader. Both, no doubt, face great difficulties in learning to write or to read English. But we may ask: 'Are the difficulties arising from this state of irregularity the same both for the reader and the writer?' It seems

that that question was answered in numerous treatises published in recent years which showed that there are definitely regular features—pattern structures—in present-day English spelling that, if disclosed, could make the task of learning to read English easier and more successful.<sup>4</sup> ‘Learning to read means developing a considerable range of high-speed recognition responses to specific sets of patterns of graphic shapes’, says Ch. C. Fries in the introduction to his book. Accordingly there are only a small number of spellings of words that do not fit into one of the spelling patterns and they present primarily problems for the writer. Fries elaborates this idea further and says (185): ‘The high-speed automatic recognition responses which readers must acquire differ quite clearly both in kind and in quantity from the productive skills which writers must achieve. This difference seems especially clear in respect of unstressed syllables.’ The reader of ModE. spelled texts must acquire the habit of rapid recognition of comparatively few arbitrary contrastive sets of spelling patterns. These patterns cannot be viewed only in the light of a simple letter-to-sound relationship, but must be seen in their complexity and should not be related only to phonology, but also to morphology and even to syntax. It means that the reader should learn to respond to a set of graphemes mostly representing words or at least syllables.

These pattern structures of ModE. spelling need not immediately convey lexical meaning. As A. McIntosh so ably showed, a written text normally carries a double load of meaning, but not necessarily always both, one is normal meaning (lexical) and the other load bears information which enables us to read out the text aloud.<sup>5</sup> Indeed an adequate spelling system should always enable the reader to read aloud with fair correctness any written utterance even if he does not necessarily understand its lexical meaning.<sup>6</sup> English spelling is sufficiently structured into typical patterns to enable us to set up a class of structures which are not in recognized use, but are acceptable in patterns and thus are legible (i.e. may be read aloud), e.g. *vake*, *cark*, *orage*, but not *crasgle*, *crk*, *avskle* which are not legible because the patterns do not exist.

Among the patterns of the graphemic structure of ModE. graphology are those that occur most frequently and should therefore be recognized first in order to develop the necessary reading skill. They include the monosyllabic words of the type C—V—C (consonantal or consonantal digraph or polygraph—C, vowel letter or vowel digraph or polygraph—V). We may contrast C<sub>1</sub>—V—C: C<sub>2</sub>—V—C or C—V<sub>1</sub>—C: C—V<sub>2</sub>—C or C—V—C<sub>1</sub>: C—V—C<sub>2</sub> or C<sub>1</sub>—V<sub>2</sub>—C: C<sub>2</sub>—V<sub>2</sub>—C etc. This basic pattern C—V—C should further be contrasted with the pattern C—V—C + *e*. Fries maintains that this second set that uses the final letter <*e*> that differentiates it from the pattern C—V—C is equally highly productive.<sup>7</sup> The ‘silent’ <*e*> therefore performs an important function in ModE. graphology. This is only one of its many functions which deserves careful interpretation and analysis. One of the possible interpretations of ‘silent’ <*e*> in this particular function of contrasting the phonic load of the pattern C—V—C + *e* against that of C—V—C is to call it a diacritical grapheme.

The suggestion of interpreting some of the English graphemes in a given position as having the function of diacritics or of diacritical letters was put forward by some scholars in recent years. The term diacritics is usually used only in the sense of diacritical marks in writing, i.e. signs or similar devices that are placed above or below a letter in order to indicate different sounds of this letter, or they indicate accents, diaeresis, etc. If we interpret the term diacritics in the narrow sense of the word, no doubt the system of ModE. graphology appears to be rather poor in this respect and J. Nosek’s general characterization of the ModE.

written norm is quite appropriate on this point.<sup>8</sup> But the use of the term diacritics can be easily extended to graphemes. W. N. Francis gives the following definition: 'When a grapheme has no phonemic reference of its own but serves instead to indicate or limit the phonemic reference of a nearby grapheme, we may say that it is being used as a diacritic or distinguishing marker. The most versatile of the graphemes of the English system in this regard is silent <e>... The most common use of <e> as a diacritic is following a single consonantal, which itself follows a single vocalic, in a syllable bearing at least tertiary stress... The grapheme <e> also serves as a diacritic with certain ambiguous consonantals and consonantal combinations. Thus it may indicate the /s/ value of <c> or the /j/ value of <g> etc. It may also combine its two diacritical functions as in the pair *lace* /læk : leis/.<sup>9</sup> According to W. N. Francis also other graphemes beside <e> may also serve as diacritics of various sorts. Thus <i> may function like <e> to distinguish the /j/ reference of <g> as in *vestigial*. This assertion is in disagreement with the author's definition of a diacritical grapheme; the letter <i> denotes here its own phonemic reference. His further application of the term 'diacritical grapheme' in the case of <h> in letter combinations *ch*, *sh*, *th* is rather doubtful. The grapheme <h> does not indicate here phonemic change in the phonic values of <t>, <s>, <c> in the sense of structure or pattern of structure as in the case of 'silent' <e>. Wherever the digraphs *th*, *ch*, *sh* appear they seem to be inseparable or indivisible and cannot be contrasted with graphemic structures in which <t>, <s>, <c> appear separately in the same sense of regularity as, let us say, in the pairs of the type of *man* : *mane*. Items like *mishap*, *mishandle* should be considered marginal, signalling apart from phonic meaning also grammatical (morphemic) meaning, to use A. McIntosh's arguments. While the diacritical grapheme <e> in all cases where the letter *e* has this function seems to indicate a phonemic feature or a change of phonemic values (if we consider the graphemic structure pattern without *e* as the unmarked member of the contrasting pair) of the same basic character, viz. the lengthening or diphthongization of the syllabographic vowel grapheme, we cannot say the same about the changes indicated by the grapheme <h> in respect to <t>, <s>, <c> because the corresponding phonemes /s/, /ç/, /ð/, /θ/ have hardly any phonemic relevance in common. If we were to include in this series also the digraph *ph*, which we can no longer include among graphemic foreignisms in a synchronic interpretation of ModE. graphology we have yet another argument against interpreting <h> as a diacritical grapheme in ModE. Neither would we find it appropriate so to interpret <h> in the digraph *gh*, which is, no doubt, one of the most difficult graphemic devices in ModE. spelling to be analysed as a diacritic although in Middle English writing the letter *h* seems to have acquired an exact, specific diacritical function, viz. to denote the voiceless character of the sound referred to by the first element of the digraph that stood for /χ/.<sup>10</sup>

Martin Joos, too, upholds the idea of introducing the notion of diacritical graphemes in an analysis of ModE. graphology.<sup>11</sup> He mentions 'the function of silent <e> as a diacritic of vowel-sounds traditionally called long'. In addition to the possible diacritical graphemes enumerated by N. Francis, M. Joos refers to 'the letter <r> which may be regarded, to any appropriate extent, as a diacritic like the hats extending the power of the alphabet without extra letters.' We consider this an important observation to which we would like to return in a later section of this paper.

The notion of diacritics has been applied to a great extent in the very interesting Russian monograph on the graphology of ModE. by V. I. Balinskaja.<sup>12</sup> The author

distinguishes diacritical signs attached to letters which help to enlarge the number of the letters of the alphabet to correspond more adequately to the phonemics of a given language, and letter diacritics which are letters of the alphabet used to denote a certain phonic meaning of other letters of the same alphabet. Such diacritical letters either precede or follow the letter the phonemic valence of which they help to determine. In this way a great number of letters used in the ModE. alphabet may be charged with the function of serving as diacritics, e.g. preceding ⟨w⟩ denotes [ɹ:] in *war* as compared with *bar*, ⟨u⟩ determines the reading of [k] in *cup* etc. Hence V. I. Balinskaja assumes that silent final ⟨e⟩ is not the only factor that acts as a diacritic in structure patterns like *mane*, *ice* etc. She speaks here about the alternating digraphs of the type *be*, *ce*, *de*, *fe*, *ge*, *ke*, *le*, *me*, *ne*, *pe*, *se*, *te*, [ve], *ze*. These digraphs as units perform the function of diacritics at the end of stressed syllabographs denoting their 'long' pronunciation. The 'silent' final ⟨e⟩ is inseparably linked with the preceding consonantal letter and only in this way influences the immediately preceding vowel letter and fully exploits the diacritical function of the whole digraph. The alternating digraphs or occasional trigraphs (*the*, *ste*, *ble*) perform the same function in denoting the length of the immediately preceding vowel letters as do final position digraphs, as e.g. *gh*, *th*, *st* or letter combinations as e.g. *ss*, *nd*. V. I. Balinskaja did not include the digraph *re* among the alternating diacritical digraphs in which the second part is silent final ⟨e⟩. She maintains that there is no analogy between let us say *cane* and *care*. In the first case we have a combination of graphemes ⟨c⟩ + ⟨a⟩ + ⟨ne⟩, in the second case we have ⟨c⟩ + ⟨are⟩ (i.e. a trigraph denoting one specific phoneme), whereas there is analogy between the items *fine* and *find* (⟨f⟩ + ⟨i⟩ + ⟨nd⟩: ⟨f⟩ + ⟨i⟩ + ⟨ne⟩). When summing up the various functions of silent ⟨e⟩ the author also speaks about its diacritical function in the grapheme clusters ⟨ble⟩, ⟨cre⟩ (*table*, *acre*) denoting the disyllabic character of the word in question. In the letter combinations *se*, *ge* the diacritical function of ⟨e⟩ is within the digraph itself denoting the pronunciation [s], [ʒ] or [dʒ]. While V. I. Balinskaja pays great attention to the so-called mixed letter groupings (clusters) composed of vowel and consonantal letters in combination with *r* and devotes a special chapter to these combinations, of which she lists 37 combinations divided into five main types (pp. 286—313), she does not seem to be inclined to consider the function of the grapheme ⟨r⟩ in these clusters to be of particular importance in the sense of what she calls 'letter diacritics'. This is due to the fact that in her view, as we mentioned above, any letter of the alphabet may perform this diacritical function as long as by its presence the phonic valence of a preceding or following grapheme is affected. Here the author is underestimating the actual diacritical impact of the grapheme ⟨r⟩ immediately following a stressed vowel letter or a stressed vowel polygraph. The structure patterns in ModE. graphology which consist of vowel graphemes or vowel polygraphs plus ⟨r⟩ represent a very definite set of patterns of frequent occurrence that the English reader must be able to recognize quickly. They definitely contrast with those grapheme patterns in which the syllabographic vowel is not followed by ⟨r⟩, e.g. *can*, *car*—*cane*, *cape*, *care*—*bid*, *bird* etc. They are very numerous in ModE. written texts and, in a way, quite productive. The stimulus for rapid recognition of these patterns is the presence of the grapheme ⟨r⟩ which is the dominating factor in them and which has no direct phoneme correspondence (at least in British standard pronunciation) and therefore in our view performs the function of a diacritical grapheme.

From what we have said so far about the possibility of applying the notion of

grapheme diacritics in an analysis of ModE. graphology follows that we believe that this notion should be primarily applied to 'silent' ⟨e⟩ in certain positions of a written syllable or word and to ⟨r⟩ in certain positions of a written syllable or word. While in the case of diacritical ⟨e⟩ we may refer to it as to a non-contagious diacritical grapheme denoting the phonemic valence of the not-immediately preceding syllabic vowel grapheme, in the case of diacritical ⟨r⟩ we have a contagious diacritical grapheme denoting the immediately preceding syllabic vowel grapheme. In cases like *care*, *cure*, *mere* etc. both grapheme diacritics ⟨e⟩ and ⟨r⟩ perform their functions. We do not think it opportune to segment ModE. graphemic structures in such a way as to detect in them grapheme clusters or polygraphs that would include final 'silent' ⟨e⟩ and 'silent' ⟨r⟩. Grapheme combinations of this kind are not 'fixed, recurrent, indivisible typified groups of (two, three or more) letters which stand mostly for one phoneme each (as ie, oo, ch, teh, gh, sc etc.)'.<sup>13</sup> The diacritical graphemes indicate in them certain phonemic changes that occur when the vowel graphemes in the syllables that are usually under stress are read aloud. These changes are qualitatively related and in many ways regular, i.e. they represent graphemic structure patterns.

One of the characteristic features of ModE. graphology is the frequent occurrence of what are commonly called 'silent letters'. These are graphemes that have zero phonic valence or no phonic reference or 'meaning' to use the terms suggested by A. McIntosh. These silent graphemes, however, perform other functions in order to enforce the quality of 'surveyability' of English written texts, they are distinguishing markers denoting differences in lexical meaning, in particular with homophones. in morphemic segmentation, in word or syllable boundaries. The really 'silent' letters, that is graphemes that perform no function whatsoever, may be detected only very rarely and are mostly spelling foreignisms, e.g. *debt*, *rheumatic* etc. Let us take a closer view of the many functions that are performed by the 'silent final ⟨e⟩'.<sup>14</sup> These functions are usually divided into those with no phonic reference and those with some phonic reference. The first group includes those cases in which ⟨e⟩ indicates the end of a word, e.g. *live*, *love*, *breeze*, *some*, *swore*, *-gue*, *-cue*, *-ture*, *-sure* etc., the cases in which it prolongs the written word because there is a very definite tendency in ModE. spelling against two-letter words, e.g. *doe*, *toe*, *die* etc., the cases in which it is part of a morpheme, e.g. *fined* against *find*, *bard* against *barred* etc., the cases in which it is a distinction marker of homophones or a prophylaxis against homographs, e.g. *tens* against *tense*, *hears* against *hearse*, *cleans* against *cleanse*, *dying* against *dyeing* etc. To the second group belong the cases in which 'silent' ⟨e⟩ after a single consonantal grapheme or occasionally after a digraph (*th*, *st*) performs a diacritical function in reference to the preceding usually stressed syllabic vowel grapheme, e.g. *provide*, *mistake*, *mode* etc., in a way it also performs here the function of denoting the disyllabic appearance of the written word, especially if we contrast *take* against *taking*, *provide* against *providing* etc.; this brings us to the next case of this group in which 'silent' ⟨e⟩ clearly emphasizes the disyllabic character of the word and may in this way also perform a diacritical function, e.g. *trouble*, *nestle* etc.; in the cases where it denotes the phonemic value of the preceding *g* or *c* its diacritical functions seems to us doubtful and grapheme clusters like *ge*, *gi*, *gy*, *ci*, *ce*, *cy* against *gu*, *ga*, *go*, *g* + consonantal grapheme, *ca*, *co*, *cu* and *c* + consonantal grapheme rather seem to speak in favour of an interpretation in the sense of digraphs; the same holds good for the function of the grapheme ⟨e⟩ in combinations with *th*, *s* in which it helps to denote the voiced consonant (e.g. *breathe*, *raise* etc.).

It would be wrong to suggest that the 'silent' <e> has in all these cases only one definite function, i.e. one of the many functions that have been enumerated above. In the analysis of items we would have to state that two or more functions of 'silent' <e> coincide or that one of the functions is sometimes neutralized by another, e.g. in *finéd* <e> has the function of a diacritic and also that of a morpheme marker, in *gavé* it has the function of a diacritic and of indicating the end of the word, in *tablé* the function of a diacritic and of denoting the disyllabic character of the word, etc. In other words the grapheme <e> in ModE. has a large load of potential phonic and linguistic meanings;<sup>16</sup> in the situational context of its occurrences it may have one or two or even three actual linguistic or phonic meanings. When 'silent' <e> performs its diacritical function it forces the reader to recognise automatically and immediately the written syllable or word taken as a whole so that he may respond to it quickly and adequately in the process of reading aloud. The same may be said about the grapheme <r> when functioning as a diacritic. In stating this we would like to re-emphasize the clearly syllabographic or logographic character of ModE. graphology which has been stressed so often before by many scholars who have contributed to a better understanding of the system and structure of ModE. spelling.

#### NOTES

- <sup>1</sup> J. Vachek, 'Two Chapters on Written English', *Brno Studies in English* 1. 7—38 (Prague, 1959)—Cf. also his more recent paper 'K obecným otázkám pravopisu a psané normy jazyka' [On General Problems of Orthography and the Written Norm of Language], *Slovo a slovesnost* 25. 117—26 (Prague, 1964).
- <sup>2</sup> 23 (New York, 1955).
- <sup>3</sup> Cf. Jiří Kraus, 'K některým otázkám pravopisu z hlediska grafematické soustavy češtiny' [Some Problems of Orthography in the Light of Czech Graphology], *Slovo a slovesnost* 25. 51—4 (Prague, 1965). The author draws attention to the usefulness of the terms 'orthoanagraphic' referring to the univalent correspondence of letters to phonemes, thus serving the purpose of writing, and orthoanagnostic, thus serving the purpose of reading. These terms have been introduced by M. Janakiev, 'Osnovy teorii ortografii' [Fundamental Theory of Orthography], *Voprosy jazykoznanija* 6. 47—57 (Moscow, 1963).
- <sup>4</sup> Let us quote here only some of them. Ch. C. Fries, *Linguistics and Reading* 265 (New York, 1965); Axel Wijk, *Regularized English*, *Acta Universitatis Stockholmiensis* VII. 361 (1959); Axel Wijk, *Rules of Pronunciation for the English Language* 160 (Stockholm, 1965) (cf. edition of the same in the series *Language and Learning*, Oxford University Press, London, 1966); C. A. Lefevre, *Linguistics and the Teaching of Reading* 252 (New York, 1964); R. H. Weir and R. Venezky, *Rules to Aid the Teaching of Reading*, Research Project No. 2584 (Stanford Univ., 1965); R. H. Weir and R. Venezky, *A Study of Selected Spelling—to—Sound Correspondence Patterns*, Research Project 3090 (Stanford Univ., 1966).—R. H. Weir, *Formulation of Grapheme-Phoneme Correspondence Rules to Aid in the Teaching of Reading*, Research Project S—039 (Stanford Univ., 1964); W. R. Lee, *Spelling Irregularities and Reading Difficulties in English*, Occasional Publication, 74 (London, 1957); M. Lehnert *Schreibung und Aussprache im Englischen* 43 (Berlin, G.D.R., 1963).
- <sup>5</sup> A. McIntosh, 'Graphology and Meaning', *Archivum Linguisticum* 13. 107—20 (Glasgow, 1961); republished in M. A. K. Halliday, A. MacIntosh, *Patterns of Language* 89—110 (London, 1966).
- <sup>6</sup> Cf. M. Joos's review of *Regularized English* by Axel Wijk in *Language* 36. 250—62 (Baltimore, 1960).
- <sup>7</sup> Fries, *op. cit.* 177. The author mentions the following interesting experiment: 'Small groups of ten or more educated adults were asked to write, first, ten one-syllable "nonsense" words; then, afterwards, ten two-syllable "nonsense" words; and finally, ten three-syllable "nonsense" words. The nonsense words produced were all within the three major sets of spelling patterns, and more than 90% within the two sets mentioned here' (182).
- <sup>8</sup> J. Nosek, 'A Systematic Analysis of Modern English Graphics', *Prague Studies in English*

9.53—67 (Prague, 1961): 'It is a system of visual alphabetic characters, aiming at a (rough) representation of the phonemic system in native words, by means of letters (polygraphs) and excluding diacritics, and it represents higher linguistic units (words and morphemes) globally by means of letter clusters, and often without straight relation to phonemic counterparts.' The absence of diacritical marks in ModE. graphology is even more apparent if we compare it with ModCz. graphology as shown above.

- <sup>9</sup> W. N. Francis, *The Structure of American English*: Chapter 8 ('Writing it down: Graphics') 430-79 (New York, 1958).
- <sup>10</sup> Cf. J. Vachek's paper (note<sup>1</sup>), 25. Nosek's criticism of this point in his review of Vachek's paper (*Philologica Pragensia* 3.117-19 (Prague, 1960) is not justified. In the synchronic interpretation of Middle E. graphology the situation is different from that in Mod.E. graphology and, moreover, Nosek takes a too narrow view of the term diacritics as a means of interpreting writing systems. It may be worth mentioning here that the German present-day reader of New High German writings has no hesitation in reading proper names in which instead of the Umlaut graphemes < ä >, < ö >, < ú > are still used the diagraphs *ae*, *oe*, *ue* (e.g. Goethe, or the typed versions of telegrams) to denote the Umlaut phonemes. The diacritical function of the grapheme letter < e > cannot be doubted here.
- <sup>11</sup> See note<sup>9</sup>.
- <sup>12</sup> V. I. Balinskaja, *Grafika sovremenno anglijskovo jazyka* [The Graphics of Present-Day English] 354 (Moscow, 1964).
- <sup>13</sup> Cf. Nosek's definition of polygraphs (see note<sup>8</sup>).
- <sup>14</sup> Cf. V. I. Balinskaja, *op. cit.* (note<sup>12</sup>) 96 (Chapter 'The final silent letter e'); or A. Wijk, *Regularized English* (see note<sup>4</sup>) 322 (Chapter VII: The Final Silent E); or A. Wijk 'Rules of Pronunciation' (see note<sup>4</sup>) 81 (Chapter VI: The Final Silent E).—It is worth mentioning here A. Wijk's assessment of the importance of 'silent' < e > in ModE. graphology, which could be formulated in the light of our discussion as follows: 'It performs its diacritical function so consistently that in a more "regularized" system of English spelling [i.e. what Wijk has attempted to compose — V. F.] it must be introduced even in these cases where it does not exist in the present spelling to indicate "length", e.g. one should write *finde*, *moste*, etc.'
- <sup>15</sup> We try to follow here along the lines suggested by A. McIntosh in his paper, see note<sup>5</sup>.

## RESUMÉ

### Pojetí diakritik při rozboru grafematické soustavy moderní angličtiny

V několika pojednáních o grafematické soustavě moderní psané angličtiny se operuje pojmem diakritických písmen. Zdůrazňuje se takto „pravidelnost“ dnešní pravopisné soustavy angličtiny zejména z hlediska její funkce vybavovací proti názoru těch, kteří v ní nacházejí převážně „nepravidelnost“ a chaotičnost. Autor příspěvku se domnívá, že je možno plně aplikovat myšlenku „diakritických písmen“ při interpretaci grafematické soustavy dnešní angličtiny na písmena < e > a < r > (tzv. němé < e > a němé < r >) v jedné z funkcí těchto funkčně značně zatížených grafemů.

