

LINGUISTICS

GRAPHOTACTICS OF THE OLD ENGLISH 'ALEXANDER'S LETTER TO ARISTOTLE'

ROBERT D. STEVICK

University of Washington

ABSTRACT

This paper describes empirical evidence in the Old English MS text of 'Alexander's Letter to Aristotle' for study of syntactic, morphotactic, and phonotactic aspects of a text representative of late West Saxon Old English – the same linguistic system in which *Beowulf* is preserved. It is written by the same scribe who copied more than half of that superb poem. That evidence lies in graphotactic features, where spacings occur within letter-strings, and how wide they are in linear measure; these features are obliterated in printed editions on which nearly all historical linguistics of English has been based. The person who penned these texts was following conventions of writing then shared (and subsequently modified), which recorded linguistic features that later textual conventions left out. Here is extensive and consistent empirical evidence for syllable division, for example, which agrees with scholarly inferences about syllable division, but with some small exceptions; here is evidence about the prosodies of reflexive constructions (*þā wolde wē ūs ge-restan*, for example); here is evidence about the hierarchic clustering of particle + prefixed verb (e.g., *on be-cwōme*). A new edition of this text in electronic format will be published on my Graphotactics website. Full linguistic and historical analysis of this electronic text can then be undertaken by anyone with basic computer facilities.

Standard editions of the Old English translation of 'Alexander's Letter to Aristotle' (British Library Cot Vit A.XV, fols. 107r-131v) print the words, which come predictably packaged between equalized blank spaces. Editions of *Beowulf* and two other prose texts, in the same manuscript and written by the same hand, do the same. Not only that, the words also have been compressed so that no spaces are allowed inside them. That is, the editions disregard the spacings in the manuscript, assuming them to be irrelevant, if not arbitrary or capricious. Yet these were produced by Anglo-Saxons who, to their credit and to our good fortune, had not embraced 'canonical word separation' and instead used spacing to record something more than lexical demarcations. Much is to be

learned about English and its history, I believe, from the graphotactics of certain vernacular texts written by native speakers of English before the early eleventh century. 'Alexander's Letter' is a prime example.

I use the term *graphotactics* to designate patterns of letter-string formation. The early English texts are later than the Latin texts first segmented by Irish readers and writers, whether the segments parsed text per cola et commata, or for words or small combinations of them. That initial segmentation by small markings was succeeded by spacings between letter-strings, a method picked up in England (and elsewhere) and used regularly in the earliest extant writings in English. This method of data structuring then reached a unique stage of development in some tenth and early eleventh century texts.

This transitional use of spacings has its clearest illustration in a swatch from one of the manuscript texts of Ælfric's *Grammar* of Latin, British Library MS Royal 15.B.xxii, f. 65r; see Fig. 1. The *Grammar* is composed in English, describing a language foreign to native speakers of English. It is meaningless to refer to 'normal word spacing' in a text of this kind. Some word sequences are written without intervening space, some words have internal spacing, and the spacing between letter strings that correspond to words is anything but normative. Below the facsimile is the same text edited to represent the spacings for their locations, their morphological contexts, and their magnitudes. Arabic numerals represent the relative measures of the spacings. Further:

space numeral space	represents separation of words
hyphen numeral hyphen	represents morphemic separation within words
— numeral —	represents separation at other than morpheme (and word) boundary

Overall, it will be obvious that the spacings, where they occur, and their relative magnitudes, correspond in part to the sequence of syntactic structures at the sentence level and the phrasal level, and some at the syllable level. Pointing is ancillary, and so is use of majuscule letters and rubrication. The graphotactics can guide a spoken realisation of the text, and seems to have been carefully composed to do just that.

Another illustration is drawn from the Parker Chronicle, a segment dealing with Ælfred's wars with the Danes, at a turning point of those fateful events; see Fig. 2. It can be read aloud interpreting the spacings as analogs to timing. Or the marble-slab method can be used to dissect the corpus in a series of cuts correlated with measure of spacing, starting with the largest measures. A rather good parsing of the passage is produced by this means. A translation with graded separation marks based on the spacing variations is included in the figure.

beroran . TBIGINTA DIUISIONES . GRAMMATICÆ
Gramma . ongræsc . is littera onlēden . ARTIS .
 7 on englisc stæf . 7 gramatica is stæf cræft . Se cræft
 geopenað 7 gehylt . leden spræce . 7 nan mann næfð
 leden boca andgite befullon buton he þone cræft
 canne . Se cræft is eall ra bōclicra cræfta ord
 fruma . 7 grund weall . Grammaticus . is se ðe can
 þone cræft gramaticam befullan . 7 se cræft hæfð
 þritig to dāl . þæt forme to dāl is . uox . stemn . þæt
 oðer . littera . / stæf . þæt þridde is . sillaba . We ha
 we ha writon on forewerdre þyssere bec . . .

British Library MS Royal 15.B.xxii, fol. 65r.

Text shown above, edited with graphotactic information.

Gram¹ma . ⁷ on⁰ grēcisc . ⁴ is⁰ littera⁴ on¹ lēden / and⁰ on¹ englisc⁴
 stæf . ⁵ and⁰ grammatica⁴ is² stæf²-cræft . ⁷ Se⁰ cræft / ge¹-openað⁴ and⁰
 ge¹-hylt³ lēden¹-spræce . ⁵ and¹ nān¹ man² næfð / lēden¹-bōca⁶ andgit³
 be¹-fullon⁶ būton² hē² þone² cræft / cunne . ⁶ Se⁰ cræft⁴ is¹ eall³ra³
 bōc¹-licra⁴ cræfta⁵ ord¹-fruma . ⁶ and⁰ grund²-weall . ⁷ Grammaticus . ³ is¹
 sē⁰ ðe⁰ cann / þone³ cræft⁴ gramaticam³ be⁰-fullan . ³ and¹ se⁰ cræft³
 hæfð / þritig⁴ tō¹-dāl . ⁷ þæt² forme¹ tō¹-dāl³ is . ² uox . ⁶ stemn⁷ þæt¹
 oðer . ³ littera . / stæf⁷ þæt² þridde³ is . ³ sillaba . ⁷ stæf-gefeg⁷ Be⁰ þissum
⁴ þrim² tō¹-dālum³ wē / ā⁰-writon⁴ on⁰ fore²-werdre³ þyssere² bec . ⁷ ...

Gramma in Greek is *littera* in Latin, and in English *staff*; and *grammatica* is *staff-craft* [i.e., the art of letters]. That craft reveals and rules the Latin language, and no one has understanding of latin-books entirely unless he knows that craft [i.e., grammar]. The craft is the beginning and foundation of all book-like arts. *Grammaticus* is that one who knows the art of grammar entirely. And that craft has thirty divisions. The first division is *uox* '(vocal) sound.' The second *littera* 'letter.' The third is *sillaba* 'syllable.' We have written about these three divisions in the early part of this book. . . .

Fig. 1 Graphotactics of a portion of Ælfric's *Grammar* of Latin.

The extreme case of parallels occurs in the occasional dittography.

118r.1 ⁴ hīe ² ūtan ² wre¹pedon .³

118r.2 / hīe ² ūtan ² wre²pe¹don ²

118r.14-16

² 7 ¹ ic ¹ swīðe ³ wundra¹de ³ þā ⁰ ge-/sælignesse ⁴ þære ² eorðan ¹

7 ⁰ ic ¹ swīðe ³ wund/ rade ⁴ þā ⁰ ge-³sælignesse ⁴ þære ² eorðan ¹

Textual information encoded in spacings like these has been little studied and less understood, a predictable result of English (and Germanic) philology having developed mainly on the basis of information encoded in printed editions of Ælfric's *Grammar*, the contents of the Nowell Codex, and all the other Anglo-Saxon texts. The conventions of printing did not materially affect the information encoded in the alphabetical elements of the original texts: the words, the syntax marked by inflexions and word-order, the morphology, the segmental phonology. On the other hand, they obliterated all the information that seems to be encoded in the spacings between letter-groups, chiefly phonotactics and the suprasegmentals of the speech being represented.

That is the reason for preparing an electronic edition of the Old English text of 'Alexander's Letter to Aristotle,' to be available on the worldwide web, at <http://faculty.washington.edu/stevickr/graphotactics>. It is part of a return to a project which was premature for electronic format when begun thirty-five years ago, given the computing equipment then generally available. The materials are now re-edited in format usable by anyone with basic computer facilities. (The new *Beowulf* text has been available since 1999). The full linguistic and historical analysis will have to be done by others, who I hope will find this neglected evidence for English historical linguistics as revealing as it has appear to me to be. Two or three examples will show the kind of work to be done.

First, the phonotactics. Begin with the obligatory break in the letter-strings imposed by the right margin of the text-space. The sequence of letters must be broken every time the writing comes to the right margin. The division may be at a word boundary, or it may be at a root or prefix boundary (less than a word boundary but in any case be a syllable boundary). Or it may be at less than a morpheme boundary. The complete set can be extracted from the new edition by a simple computer program: (1) search for a word, (2) scan it for '/' without a hyphen next to it. The results reflect the general phonotactic features of Old English so consistently that the divisions seem to be trustworthy guides to syllabication for even such words as *fersc-an* and *alex-ander*. That the division of words is rule-governed by phonological features is all the more apparent when we notice that the right margin tends to be ragged.

Now to the suprasegmentals. Begin with the obvious and perhaps trivial instances. Some features are frequent and probably familiar:

- 1) a. Conjunction *ond*, nearly always written as an abbreviation, is usually not set apart from the following letter-string. It is, though, usually preceded by conspicuous linear space. That would be consistent with its being initial and unstressed in a phonological phrase.
- b. The abbreviation for *þæt* is similar.
- c. Relative particle *þe* is similar, except that it is a spelling, hence usually is set apart from the next sequence, but by only minimal spacing. Preceding it, the spacing is little or none when it follows a pronominal, as in *Se þe*, but considerably more when it initiates a clause relative to a NP containing a noun head and a determinative form, as in *tō þēm mere þe wē bī ge-wīcod hæfdon*.
- d. Prepositions are similar to relative particles in several ways.

And so on. Each of these forms has been assumed to have no independent stress-level, leaning on the form that follows it, and the graphotactics correspond regularly to this assumption.

- 2) At the other extremes, sentences usually have wide spacing separating them, many of the boundaries also marked with centered punctuation. These markings seem to reflect segmentation of discourse that is commonly expressed by longer timing between segments. This is an obvious instance of the suprasegmental of timing that marks syntactically self-standing units of the discourse.
- 3) In between these are the complex variations in spacings. Here is where suprasegmental information is less obvious, and also much less trivial.
 - a. Consider a pair of passages with reflexive constructions. Without graphotactic evidence we could only guess at the constituent phrasal structure. Would it be the first, or the second of these?

þa wolde wē | ūs ge-restan .

þa mynton wē | ūs ge-restan .

Or:

þa wolde | wē ūs | ge-restan .

þa mynton | wē ūs | ge-restan .

The graphotactics are consistent in both passages in representing the second clustering pattern clearly. Here they are in full-sentence context:

124v.10-11 ⁹ ðā⁰ wæs⁴ sēo³ þridde⁴ tīd³ þære⁴ nihte⁵
þā⁰ wolde⁹ wē¹ ūs³ ge⁰-restan .⁵ þā² cwōman³ þær² nædran³ eft⁹

125r.16-17 ⁵ þā¹ hit² wæs¹ sēo¹ fīfte³ tīd³ þære⁴ nihte .⁹
þā⁰ mynton³ wē¹ ūs² ge⁰-restan .⁵ ac¹ þā² cwōman⁹ þær³ hwīte⁵
léon³

(Apart from the spacings, note the morphological clue that this text is based on everyday speech in late West Saxon: plural past tense *wolde* vs. *mynton*.)

- b. Then consider constructions such as on *be-cwōm-*. This is in the line of historical development of phrasal verbs in English. These had conspicuous development in the fourteenth century, and in the nineteenth and twentieth centuries have virtually flooded the language.

stand up	stand off	look up	look into	
throw up	throw away	throw in	keep up	keep down
stand up to	get away with	put up with	keep up with	

But they had to begin somewhere. Listed below are the instances of *be-cwōm-* (and *be-cuman*) culled from 'Alexander's Letter.' Those instances which have *on* preceding the verb form are shown in italics. Two are especially informative. At fol. 124r and fol. 125r, at the end of a line of writing there is room for *on* and *be*. However, both times the line division is after *on*, apparently intended to avoid canceling the graphotactic patterns of the phrasal construction. These are the occurrences of *be-cwōm-*.

111r .⁵ Ond² swā[?] mid² mī'ne² werode³ on⁻¹-sunde³ in¹ patriacen³
þ² lond / wē⁰ be⁻³-cwōman³ mid² golde⁴

111v .⁶ ðā² be⁻¹-cwōman³ wē³ on¹ þa³ lond⁻³-ge⁻²-mæro / me²do³ 7⁰
persa .⁶

111v þ¹ wē² tō² þæm¹ londe / 7⁰ tō¹ þære⁵ stōwe⁵ be⁻¹-cwōman .⁴

114r .⁶ ðā¹ be⁻¹-cwōman³ wē⁰ syð²þan³ tō¹ þæm³ wudum⁴ in¹die³

.⁵ Ond² wē⁴ þā¹ eft³ in⁰ fasia¹cen² þæt / [115v] lond⁴ be⁻¹-cwōman³

116r þ⁰ ūs¹ic² þær² on¹ be⁻¹-cwōme .⁶

116v 7⁰ þā² fȳr³ þe¹ ūs² þær² in⁰ þæm / wīcum³ on¹ be⁻²-cwōman²

117v / frīneð⁴ hwæt¹ gōdes³ oþðe³ yfles³ him¹ be⁻¹-cu/man³ sceal .⁵

118r .⁵ Ðā⁰ be⁻²-cwōm² ic³ on / caspīam³ þ¹ lond³

118v ³ þȳ⁰ læs / wē¹ on⁰ ðā³ be⁻¹-cwōmon .⁴

119r ³ þ⁰ mē² þā² earfeðu / be⁻¹-cwōman ;⁴

121v .⁵ *þ⁰ ūsic² ðonne³ sem¹ninga / hwelce³ earfe²ðo⁴ on¹ be⁻²-cwōme .⁵

122r .⁴ ðā³ be⁻²-cwōm / sum[?] on⁻⁰-gris⁻²-līc[?] wīse⁴ on² hīe .³

123r 7⁰ þ⁰ wē¹ ge⁻³-nōg⁰ raðe³ tō⁰ þæm¹ be⁻¹-cwōman /

124r ⁴ gif² ūs¹ on⁰ niht³ un⁻⁰-cūðes³ hwæt³ on / be⁻²-cwōme⁵

125r ³ þāra¹ þinga³ þe² ūs¹ on / be⁻¹-cwōmon³ swā³

monigra³ ge⁻⁰-swenc¹nissa³ 7¹ earfeðo .⁵

125v ² 7⁰ ge⁻¹-swenc²nissum² þe² ūs¹ on¹ be⁻³-cwōm= /

128v be⁻²-cuman³ in⁰ mace¹dōniam⁴ tō¹ olimphi¹ade⁵ mīnre⁴ mē¹der

³ 7⁰ mīnum² ge⁻¹-swus¹trum .⁵

128v þ⁰ ic² eft² cwic³ ne / mōste³ in⁰ mīn¹ne³ ēpel⁴ be⁻⁰-cuman .⁴

Synoptically now. With such frequent spacing dividing the letter-strings, with the specific linguistic positions at which they occur, with the variability in measure of the spacings, and with consistent contexts of the patterns of variants, with all this 'busyness' the manuscript text can hardly be a hasty or careless production. I believe it is a deliberate, committed, serious composition of written text, the last place to expect textual arbitrariness.

Now, if spacing features in the 'Alexander's Letter' (as well as in *Beowulf*) do have extensive correlations with syntactic and morphotactic and syllabic patterns, what can we infer? The most obvious and natural inference to draw from these correlations is that the spacing features in the written text provide an analog to some prosodic features of the spoken text: they would have been derived from them, and their purpose was in turn to cue the reading for segmentation of the syllable string into meaningful constructions and to guide the appropriate linkages of those constructions within sentences, not to mention separation of sentence units (or of discourse units, more likely, that approximate sentence units).

Apparently the only means already in hand in the tenth century for writing English vernacular texts were alphabetic symbols in lineal succession and spacings of linguistic segments. We do not find notation for levels of stress or for intonation, such as those that must be supplied by linguists in transcribing the full features of modern spoken languages. The spellings could have been improved some, of course, but not developed to represent any new kind of information. The spacing, on the other hand, could be developed to represent linguistic information that the alphabetic system could not. There was no need for a different kind of system, say, to mark sentence-syntax (for example tree-diagrams, inter-linear symbols for parts of speech or for sequencing, so called syntactic glosses) because the syntax was already encoded in sequence patterns of words, in the valence of the lexical items, and in the grammatical inflections. But variable spacing could clarify constituent structures and even some of their hierarchies, and certainly it was developed in this way for a number of texts. There was no

need either to mark word-stress for native speakers. Phrase-accent, on the other hand, could be signified at least indirectly by variable spacing, in its correlation with constituent structure marking. Varied spacing would be a natural representation of timing features. And to the extent that timing variations correlate with pitch patterns, the need to represent the 'tune' would not have risen to the level of needing separate representation, such as by rising and falling patterns in the letter strings, or annotations. Spacing, with its varied measures, could have represented enough prosodic information to cue the written text for oral interpretation by a native speaker of English, in accordance with the author's composition (or the copyist's understanding).

How much have we learned about the earliest English by studying texts in the radically transformed data structure of modern printed editions? It took me many years to decode the basis for sectional divisions of religious poetic texts in Old English. Printed editions ignored them, or explained them away in embarrassingly ad hoc ways. But they are in fact carefully and coherently placed, in a manner consistent with design of the finest page illuminations and the design of stone crosses and fine metalwork in the Hiberno-Saxon world. In like manner, I believe, the graphotactics of several texts turn out to be carefully and coherently constructed, recording linguistic information that has been overlooked, and which may help us understand more fully, and more accurately, the early structure of English, and its history.

REFERENCES

- Doane, A. N.
 1998 "Spacing, placing, and effacing: Scribal textuality and Exeter Riddle 30a/b", in: Sarah Larratt Keefer – Katherine O'Brien O'Keefe (eds.), 45-65.
- Fulk, R. D.
 1998 "The role of syllable structure in Old English quantitative sound changes", *North-Western European Language Evolution* 33: 3-35.
- Kiernan, Kevin (ed.)
 1999 *The electronic Beowulf*. London: The British Library.
- Larratt Keefer, Sarah – Katherine O'Brien O'Keefe (eds.)
 1998 *New approaches to editing Old English verse*. Rochester, N. Y.: D. S. Brewer.
- Malone, Kemp (ed.)
 1963 *The Nowell Codex: British Museum Cotton Vitellius A.xv Second MS*. (Early English Manuscripts in Facsimile XII). Copenhagen: Rosenkilde and Bagger.
- O'Brien O'Keefe, Katherine
 1998 "Introduction", in: Sarah Larratt Keefer – Katherine O'Brien O'Keefe (eds.), 1-9.
- Orchard, Andy
 2003 *Pride and prodigies: Studies in the monsters of the Beowulf-manuscript*. (Revised edition). Toronto: University of Toronto Press.
- Rypins, Stanley (ed.)
 1924 *Three Old English prose texts in MS. Cotton Vitellius A.xv*. (Early English Text Society, Original Series 161). London: Oxford University Press.

Saenger, Paul

1997 *Space between words: The origin of silent reading*. Stanford: Stanford University Press.

Stevick, Robert D.

1967 "Scribal notation of prosodic features in The Parker Chronicle, Anno 894 [893]", *Journal of English Linguistics* 1: 57-66.

1968 *Suprasegmentals, meter, and the manuscript of Beowulf*. The Hague: Mouton. [faculty.washington.edu/stevickr/graphotactics/beowulf.html]